Consumer Expenditures In The United States:

Survey Description And Distributional Analyses.

Thesia I. Garner and Stephanie Shipp

Household expenditure survey data can be used to provide information concerning the economic status and material well-being of families within and across countries. The purpose of this paper is two-folds first, to describe the survey which is conducted in the United States (U.S.) to collect those data, and second, to present results from recent studies in which these data have been used to examine the distribution of household expenditures. We hope this paper will stimulate grater interest among researchers internationally to examine household expenditure data when addressing distributional issues. The paper is divided into four sections. Section one provides a description of the U.S. Consumer Expenditure Survey (CEX) and procedures. Section two includes results from analyses of the distribution of expenditures across households and the inequality in these expenditures. Section three discusses limitations of the CEX data. Section four is the summary and conclusions.

1. The Consumer Expenditure Survey

Expenditure surveys have boon conducted by the U .S. Bureau of Labor Statistics (BLS since the late 19th century to provide a profile of consumers' spending patterns. On average, these surveys have been conducted at about 10-year intervals. However, since 1980, Consumer Expenditure Survey data have been collected continually in the U.S. Brier to 1980, the most recent expenditure survey data wore collected in 1973. Microlevel data currently available to the public are for the years 1980 to 1987.

The CEX serves two primary purposes. First, it is used to collect data on the

personal spending of consumer units[1] on an ongoing basis; these data are used by a wide range of public and private organizations for economic analysis and planning. Second, the CEX provides the basis for the weights and item selections for the Consumer Price Index (CPI). An unique aspect of the survey data is that detailed demographic information is connected about each consumer unit so that this can be related to family spending patterns and income. The early surveys arose because of the need by U.S. Government policymakers for information relevant for tariff negotiations, and concern over the effects of rapidly rising prices on the living costs of workers. Expenditure surveys have been broadened over the years to collect more than just expenditure data. Additional questions have been added to the survey to obtain information about commodities, income, and family socio-demographic characteristics. Data that once were collected only for selected population groups are now collected for most of the U.S. population and can be used to adress a greater of issues. A wide range of public and private agencies use the CEX data for economic analysis and planning. For example, the data are used to study the welfare of particular segments of the population, such as the aged or low-income families. The U.S. Department of Commerce uses the survey data for revising its benchmark estimates of some of the personal consumption expenditure components of the Gross National Product. Economic policymakers use the data to study the effects of policy changes, such as the impact of various tax proposals on diverse socioeconomic groups. Econometricians find the data useful in constructing economic models. Market researchers find them valuable in identifying consumer trends for various groups of good and services. Businesses often use the data for forecasting.

The CEX data are collected by the Bureau of the Census (Census), U.S. Department of Commerce, under contract to the BLS, an agency within the U.S. Department of Labor. To those of you from countries with central statistical offices, this contractual arrangement may seem unusual. However, in the U.S., government statistics are produced by many departments independently. Data are collected for a specific purpose, or they may be by-products of administrative functions. The Office of Management and Budget provides limited coordination for Federal government agencies collecting survey data. Approval by this office is required before surveys that may place an undue burden on the public can be undertaken.

The U.S. Consumer Expenditure Survey has two components, each with its

---

[1]Consumer Unit: A single person or group of persons in a sample household related by blood, marriage, adoption, or other legal arrangement, or who share responsibility for at least 2 out of 3 major types of expenses-food, housing, and other expenses. The term household or family is used for convenience, although there may be more than one consumer unit in a household and one person families are included. Each consumer unit has a householder or "reference" person; this person is the first member mentioned by the survey respondent when asked by the interviewer to name the person or persons who owns or rents the home. It is with respect to that person that consumer units are classified.

own questionnaire and sample: (1) an Interview panel in which 5,000 consumer units are visited by an interviewer every 3 months over a five quarter period, and (2) a Diary survey in which 5,000 consumer units are asked to maintain a record of expenses for two consecutive 1-week periods. The Interview survey is used to obtain data on the types of expenditures respondents can be expected to recall for a period of 3 months or longer-such as those for property, automobiles, major appliances and travel, and expenditures that occur on a regular basis such as for rent, insurance, and clothing. The Diary survey is designed to obtain data on frequently purchased items that are more difficult to recall over longer time periods. These include food and beverages, tobacco, drugs, personal care products, and services. Expenditures on trips away from home are not included in the Diary. Business related expenditures are not included in either survey.

The current Consumer Expenditure Interview Survey is an ongoing survey with panels of respondents interviewed on a continuous basis; rotating panels of respondents are interviewed quarterly. Each quarter, one-fifth of the units are new to the survey. After being interviewed for five consecutive quarters, each panel is rotated out of the survey. For the initial interview, information is collected on demographic and family characteristics and on an inventory of major durable goods of each consumer unit. Expenditure information is also collected in this interview, using a one-month recall. This information is used for bounding purposes to minimize telescoping errors. These errors, common in retrospective interviews, result from a tendency to report events prior to the reference period of the survey. In the second through fifth interviews, uniform questionnaires with a three-month recall are used to collect out-of-pocket expenditures and expenditures which are made using credit. Expenditures that are reimbursed, such as medical expenses, are subtracted from the out-of-pocket expenditure reported for a good or service. Detailed wage, salary, and employment information is collected in the second and fifth interviews. The fifth interview is used to obtain a financial profile of the unit, including information on assets and liabilities.

For the Diary, as with the Interview, information on the demographic and family characteristics of the consumer unit is collected from a member of the consumer unit by an interviewer during an initial visit to the consumer unit's home. Each consumer unit then receives a Diary that is to be completed over a one-week period. At the end of the week, the interviewer collects this Diary, reviews the entries, and leaves a second diary. Finally, the interviewer returns a third time to collect the second Diary.

For the current survey, data are being collected in 91 urban areas and 18 rural areas of the country. These areas were selected to produce a sample representative of the U.S. population[2]. The sampling frame, the list from which housing units

---

[2] The population of interest is the total U.S. civilian population. Within this framework, the eligible population is composed of all civilian noninstitutional persons (e.g., those living in houses, condominiums, or apartments and all people residing in the following group quarters: boarding

were chosen, was generated from the 1980. Census 100-percent-detail file. Data on new construction permits augmented this basic list; in areas where new construction permits were inadequate or not available, area probability samples were selected. Average response rates computed on a quarterly basis have been fairly consistent since 1980, fluctuating between 86 and 91 percent for the Diary and between 83 and 86 percent for the Interview survey. These rates are computed on a quarterly basis and do not represent participation in the survey over the entire 2-week or 5-quarter period.

Once the data are collected and reviewed, data imputation routines are used to account for missing or inconsistent expenditure entries on a quarterly basis. Imputation includes the use of hot deck procedures, weighting class procedures, and percent distributions. Currently missing income dare are not imputed. Allocation routines are applied where necessary to transform reports of combined expenditures into more detailed categories of expenditures. The data are also "time adjusted" and this results in monthly income and expenditure data. [3]

During processing, weights are created for consumer unit included in the survey sample. The weight is the inverse of the probability of selecting the consumer unit adjusted to account for field subsampling and nonresponse. The weight is designed to reflect the U.S. population distribution by selected consumer unit characteristics, for example, age, sex, and race.

As noted previously, the data from the CEX surveys are made available to the public in various forms. News releases provide a first look at the data with a few tables showing expenditures by socio-demographic characteristics, such as income, age, and region. These are followed by bulletins which contain more extensive text and tables, more characteristics, and more detailed expenditures. Computer printouts are provided when special requests are made for characteristics not usually published or for more detailed expenditure levels. These requests are easily filled with the aid of an online data base management system. Diskettes containing macro level data are provided to those who want to use a personal computer to perform analyses such as percent changes, percent distributions, or comparisons of expenditures of different groups. Microlevel data are available on public-use computer tapes. These tapes contain information on the characteristics of each consumer unit member (for example, age, employment, and earnings), the expenditure data at the lowest level of classification in our system, ownership of durable

---

houses; housing facilities for students and workers; staff units in hospitals and homes for the aged, infirm, or needy; permanent type living quarters in hotels and motels; and mobile home parks. Patients and inmates are eliminated before sampling. The remaining ineligible persons in group quarters and housing units on military reservations are eliminated by later screening. These include: camps, communes; convents; halfway houses; home for the aged, infirm, or needy; transient quarters in hotels or motels; and missions. The collection of data for the rural population was discontinued in 1982 and 1983 for budgetary reasons.

[3] All expenditures are mapped (assigned) to a standard uniform classification code (UCC). The UCC system is used by both CEX and CPI.

goods and vehicles, and selected other nonexpenditure information (for example, the characteristics of housing units, estimated values of homes, assets, and liabilities). A detailed review process is used to insure the confidentiality of respondents. Identifiers such as household name and address and unusual values which might be revealing, such as very high incomes or very high values of homes, are not included in the data files. Release of geographic detail is limited. Since 1984, flags have appeared on the public-use tapes to indicate whether and adjustment, (for example, allocation, imputation, computation, or some combination of these three) was made to the expenditure data.

In addition to actively producing data for public users, economic and statistical research is conducted within the Bureau using CEX data. In the following section, results are present from studies in which published aggregate data and micro-level data have been used.

2. Distributional Analysis of Expenditures

In this section, results are presented from two studies in which the distribution of household expenditures have been examined. The distribution of expenditures is first presented in terms of aggregate shares[4]. Gini coefficients, based on microlevel expenditure data, are presented in the second part of this section[5]. Data for both of these studies are weighted to represent the U.S. population. Results from each of these studies have implications for social , and economic policy related to the distribution of household expenditures across diverse population groups.

A. Aggregate Expenditure Shares

Aggregate expenditure shares are computed for several categories of expenditure items and are compared across subgroups of the population. An aggregate item expenditure share is the ratio of item expenditures for a particular subgroup of consumer units compared to the item expenditures for the total population. In Tables 1 and 2, aggregate expenditure shares for 198? axe presented by income quintiles[6] and then by age groups, respectively. Consumer units in the lowest income quintile group account for 8 percent of the 1987 total aggregate expenditures, less than half of the proportion of the population they represent. The highest income quintile's aggregate expenditure share for all items is 38 percent, which is

---

[4]Results are from ongoing research in the Division of Consumer Expenditure Survey. For a description of the method and earlier results see: Kaneer, Kirk (1986)

[5] see: Garner (1989)

[6] Across-the-board zero income is considered an invalid response; in this case the consumer unit would be identified as an incomplete income reporter. It is important to note that even complete income reporters may not have provided a full accounting of all income from all sources. Thus, it is possible for consumer units not reporting income completely to be considered complete income reporters.

Table 1. QUINTILES OF INCOME BEFORE TAXES: Aggregate annual expenditures, expenditure shares, and consumer unit characteristics, Consumer Expenditure Survey, 1987
(Aggregates in millions of dollars, unless otherwise noted)

Item	Total complete reporting	Lowest 20 percent	Second 20 percent	Third 20 percent	Fourth 20 percent	Highest 20 percent
Number of consumer units (in thousands)	81,070	16,187	16,215	16,215	16,214	16,239
Percent distribution	100.0	20.0	20.0	20.0	20.0	20.0
Aggregate annual expenditures	$2,008,590	8.3	12.7	17.5	23.9	37.6
Food	304,256	10.2	14.8	19.1	23.1	32.9
Food at home	173,733	12.3	17.1	19.2	22.8	28.6
Food away from home	130,523	7.4	11.7	18.8	23.5	38.6
Alcoholic beverages	24,564	8.0	12.5	21.5	22.4	35.6
Housing	607,295	10.2	13.6	17.3	22.7	36.1
Shelter	330,279	10.3	13.5	17.1	22.4	36.6
Owned dwellings	187,110	5.9	8.7	13.6	24.5	47.3
Rented dwellings	108,715	19.1	23.7	24.9	19.8	12.5
Other lodging	34,455	6.8	7.9	11.4	19.2	54.6
Utilities, fuels, and public services	133,441	13.1	16.5	19.4	22.6	28.4
Household operations	29,753	6.7	9.5	13.9	22.0	47.8
Housekeeping supplies	29,834	10.1	15.4	19.9	23.5	31.1
Apparel and services	117,957	7.5	12.9	16.6	25.3	37.7
Transportation	371,949	6.8	12.7	18.1	25.8	36.6
Vehicle purchases	163,437	6.1	12.1	17.9	26.8	37.2
Gasoline and motor oil	71,423	8.4	14.6	20.2	25.4	31.4
Other vehicle expenses	116,011	6.1	12.6	17.6	25.8	37.9
Public transportation	20,997	9.8	12.5	15.0	19.9	42.8
Health care	92,014	12.9	18.7	20.2	21.3	26.9
Entertainment	98,176	6.3	9.9	17.5	24.5	41.8
Personal care products and services	27,969	8.3	14.1	18.7	22.4	36.4
Reading	11,674	8.9	12.8	18.5	23.5	36.6
Education	26,348	14.8	10.3	13.8	16.9	44.0
Tobacco products and smoking supplies	19,051	14.2	18.0	20.8	24.5	22.5
Miscellaneous	48,561	8.0	10.7	18.3	26.6	36.3
Cash contributions	62,910	4.7	8.9	14.2	21.0	51.2
Personal insurance and pensions	195,784	2.0	6.0	14.3	26.7	51.0
Life and other personal insurance	23,835	7.4	10.3	16.7	25.2	40.3
Pensions and Social Security	171,949	1.2	5.4	14.0	26.9	52.5
Consumer unit characteristics:						
Mean income before taxes 1/	27,326	4,611	11,954	20,943	33,276	65,750
Mean income after taxes 1/	24,871	4,494	11,424	19,500	30,373	58,477
Average number of persons in consumer unit	2.5	1.8	2.2	2.6	2.9	3.2
Age of reference person	47.0	51.7	50.7	44.9	43.0	44.8
Average number in consumer unit						
Earners	1.3	0.6	0.9	1.4	1.7	2.1
Vehicles	2.0	0.8	1.5	2.0	2.5	3.0
Children under 18	0.7	0.4	0.6	0.7	0.9	0.9
Persons 65 and over	0.3	0.4	0.5	0.3	0.2	0.1
Percent distribution						
Housing tenure:						
Homeowner with mortgage	37	11	17	32	53	72
Homeowner without mortgage	25	29	34	26	20	14
Renter	38	60	49	41	27	14
Race of reference person:						
Black	10	19	12	9	7	4
White and other	90	81	88	91	93	96
Education of reference person:						
Elementary (1-8)	12	25	17	10	5	2
High school (9-12)	45	45	53	50	45	30
College	43	28	29	40	50	68
Never attended and other	1	1	1	0	0	0
At least one vehicle owned	86	57	83	94	97	97

1/ Components of income are derived from "Complete income reporters" only; see Note 6.

Table 2. AGE OF REFERENCE PERSON: Aggregate annual expenditures, expenditure shares and consumer unit characteristics, Consumer Expenditure Survey, 1987
(Aggregates in millions of dollars, unless otherwise noted)

Item	All consumer units	Under 25	25 - 34	35 - 44	45 - 54	55 - 64	65 - 74	75 and over
Number of consumer units (in thousands)	94,150	7,811	21,345	18,747	13,395	13,080	11,578	8,194
Percent distribution	100.0	8.3	22.7	19.9	14.2	13.9	12.3	8.7
Aggregate annual expenditures	$2,298,578	4.9	22.5	25.7	18.5	14.6	9.5	4.4
Food	344,966	5.0	22.1	25.2	18.0	14.7	10.0	5.0
Food at home	197,621	4.1	21.5	24.1	17.8	15.3	11.1	6.2
Food away from home	147,345	6.1	22.9	26.7	18.2	14.0	8.4	3.4
Alcoholic beverages	27,209	8.9	27.7	24.7	16.5	12.4	7.0	2.2
Housing	712,621	4.7	24.3	25.9	16.9	13.4	9.7	5.2
Shelter	391,099	5.4	26.2	26.4	16.5	11.9	8.6	4.8
Owned dwellings	223,606	1.2	23.4	30.4	18.7	13.6	8.7	4.0
Rented dwellings	126,632	13.7	35.1	19.7	10.8	6.5	7.6	6.6
Other lodging	40,861	2.9	14.1	25.6	22.3	19.8	11.5	3.8
Utilities, fuels, and public services	157,325	4.0	20.4	23.5	17.6	15.8	11.8	6.9
Household operations	34,930	2.8	28.5	29.1	11.2	10.5	10.4	7.5
Housekeeping supplies	32,105	3.6	20.7	22.8	18.4	16.1	12.7	6.0
Apparel and services	136,141	5.1	22.6	27.6	19.1	13.4	8.8	3.3
Transportation	433,090	6.1	22.4	25.7	19.2	15.1	8.8	2.6
Vehicle purchases	190,371	6.9	22.6	26.6	19.4	14.9	7.7	1.8
Gasoline and motor oil	83,605	5.7	22.8	24.8	18.9	15.3	9.2	3.3
Other vehicle expenses	133,411	5.4	22.4	25.5	19.8	14.9	9.0	3.0
Public transportation	25,703	5.7	18.9	23.6	16.3	17.3	13.4	4.8
Health care	106,860	2.5	15.3	19.0	15.7	16.9	18.3	12.2
Entertainment	112,321	5.2	24.7	28.1	17.5	13.2	9.0	2.2
Personal care products and services	31,070	4.2	21.1	24.1	17.7	16.4	10.7	5.8
Reading	13,369	4.0	20.9	24.0	17.1	15.3	13.1	5.7
Education	31,729	15.2	15.4	27.4	28.7	10.2	3.0	0.2
Tobacco products and smoking supplies	21,843	5.9	22.3	23.7	20.5	15.2	9.5	3.2
Miscellaneous	52,912	3.6	22.8	25.2	19.1	14.7	11.2	3.6
Cash contributions	69,765	0.8	11.8	21.1	23.1	19.9	13.8	9.3
Personal insurance and pensions	204,776	3.2	23.6	28.6	21.7	17.0	4.9	1.0
Life and other personal insurance	27,680	1.7	17.4	27.9	19.6	19.1	11.0	3.5
Pensions and Social Security	177,096	3.4	24.6	28.7	22.0	16.7	4.0	0.6
Consumer unit characteristics:								
Mean income before taxes 1/	$27,326	$12,621	$27,835	$36,240	$36,941	$31,038	$18,598	$12,912
Mean income after taxes 1/	24,871	11,693	25,322	32,666	33,064	28,137	17,637	12,280
Average number of persons in consumer unit	2.6	1.8	2.8	3.4	2.9	2.3	1.9	1.5
Age of reference person	47	21.6	29.6	39.1	49.2	59.6	69.1	80.2
Average number in consumer unit:								
Earners	1.4	1.2	1.5	1.8	2	1.4	0.6	0.2
Vehicles	2	1.2	1.9	2.4	2.7	2.3	1.6	0.9
Children under 18	0.7	0.4	1.1	1.5	0.6	0.2	0.1	0
Persons 65 and over	0.3	0	0	0	0	0.1	1.3	1.3
Percent distribution								
Housing tenure:								
Homeowner with mortgage	38	8	40	59	58	38	18	6
Homeowner without mortgage	25	2	5	9	18	44	59	63
Renter	38	90	55	32	24	18	23	31
Race of reference person:								
Black	11	11	11	11	11	11	10	8
White and other	89	89	89	89	89	89	90	92
Education of reference person:								
Elementary (1-8)	12	2	3	5	9	17	24	39
High school (9-12)	45	44	42	41	47	50	51	39
College	43	54	54	54	43	32	24	21
Never attended and other	1	0	0	0	1	1	1	2
At least one vehicle owned	86	70	88	92	92	91	83	63

1/ Components of income are derived from "Complete income reporters" only; see Note 6.

almost twice their 20-percent proportion of the population. For consumer units in the lowest income quintile, which includes many consumer units headed by a student, the only items for which aggregate expenditure shares approximate their population share are rent and education.

When evaluating the data in terms of the expenditure shares of different age groups, distributional and expenditure level differences also emerge. More than a third of aggregate expenditures for rented dwellings are made by consumer units
in wick the reference person is in the 25 to 34 years age group. Only 45 percent of these families are homeowners. In the next age group (35 to 44), homeownership increases to 68 percent. As a result, those age 35 to 44 account for 30 percent of the aggregate for owned dwellings; this one age group spends the largest share for owned dwellings. These families are more likely to have newer and hence larger mortgages as well as higher interest rates than are families in the other age groups.

Consumer units with reference persons younger than 25 years have the lowest aggregate expenditure share for transportation (6 percent) while the 35 to 44 years age group has the largest (26 percent). Seventy percent of all consumer unite in the younger age group own vehicles, while 92 percent of the consumer units in the 35 to 44 years age group own vehicles.

As expected, consumer units in the top two age groups combined (65-74 and 75 and over) have the largest share of aggregate out-of-pocket health and medical expenditures. These families represent 21 percent of the population; however, they account for almost one-third of the aggregate share for out-of-pocket health care expenditures. Although most of these consumer units participate in the federal
Medicare program, they are frequently saddled with large medical expenditures since our federal health care system is not totally comprehensive.

The largest aggregate expenditures share for education is made by families with reference persons age 45 to 54 years. This group spends 29 percent of the education expenditures for the total population. They are followed closely by families age 35 to 44 who account for 27 percent of the aggregate education expenditures. Families in these two age groups are likely to be paying for their children's college education, which can be expensive in the U.S.

Results such as these can be produced from the data BLS publishes. For certain types of policy questions, aggregate shares may be the most appropriate form in which to present the distribution of expenditures for demographic groups.

B. Consumer Expenditures and Inequality:

Gini coefficients can be used to examine the inequality in the distribution of household consumption expenditures. The Gini coefficient is a measure of the inequality of a distribution; it is a summary statistic of inequality derived from the Lorenz Curve and is equal to one minus twice the area under the Curve. The lower the value of the Gini, the less the inequality in the distribution. For example, as expenditures become more equally distributed across consumer units, the value

of the of the Gini declines. A Gini of zero would represent absolute equality in expenditures. Garner[7] used a method developed by Lerman and Yitzhaki[8] to produce Gini coefficients for consumption expenditures; some of these coefficients are reproduced in this discussion.

Three sets of Gini coefficients are presented. First, Gini coefficients based on total annual consumption expenditures are produced and compared to Gini coefficients based on annual income before and after taxes for the total population. Second, expenditure Gini coefficients are produced for demographic subgroups of the population. The demographic subgroups are defined in terms of consumer unit size, household composition, the race and age of the reference person, and region and degree of urbanization of residence. Consumer unit size is defined in terms of five groups, from one person to five or more persons. Household composition is defined in terms of six categories: (1) single, (2) one parent with children, (3) husband and wife only, (4) husband and wife with their own children, (5) other husband and wife consumer units, and (6) other consumer units. Race is defined as black or non-black. Age of the reference person is divided into two groups: less than 65 years of age and greater than or equal to 65 years of age. Region refers to the U.S. Census regions: Northeast, Midwest, South, and West. Degree of urbanization refers to one of the following: a large urban area, a small urban area, or a rural area (farm and non-farm). Third, the overall Gini coefficient is decomposed and individual budget component Gini coefficients are produced. The effects of the inequality within expenditure budget components on overall inequality are identified by decomposing the Gini. To study the decomposition effects, total expenditure are divided into ten budget components: food, shelter, fuel and utilities, household operations, apparel and services, private transportation, public transportation, medical care and services, entertainment, and other expenditures. (See the Lerman and Yitzhaki papers and the Garner paper for a description of the decomposition method).

Microlevel data from the U.S. Consumer Expenditure Interview Survey, collected in January 1987 through March 1988, are used. Since an objective of the study is to produce Gini coefficients based upon income, the sample is limited to those consumer units who were interviewed for four consecutive quarters during the survey period. Thus, Gini coefficients based upon annual expenditures and annual income could be compared for the same time period. Expenditure items are aggregated into the ten budget components noted previously, with expenditure values of zero assigned to items for with consumer units did not report expenditures. Although collected in the Interview, expenditures for life insurance, endowments, annuities, other personal insurance, retirement, pensions, and Social Security are not included as part of total expenditures for the study since expenditures for these items are not considered to be for current consumption. This excluded group of

[7] see: Garner (1989).

[8] see: Lerman-Yitzhaki (1984), Yitzkaki (1983) and Lerman-Yitzhaki (1985)

items accounts for approximately 9 percent of total expenditures reported by the BLS for the U.S. population for 1987 [9]. For the expenditure-income comparison, a subsample of the primary sample is used. This subsample is restricted to include only complete income reporters. Completeness of income reporting is based upon whether the respondent to the Interview provided values for various sources of income, primarily for major sources such as wages and salaries, self-employment income, and Social Security income.

The primary sample included 2071 consumer units participating in four consecutive interviews; the subsample of complete income reporters included 1797 consumer units. The weighted primary sample represents approximately 27,000,000 consumer units. The greatest percentage of sample consumer units are composed of two persons (.30), although the most prevalent household composition type, given the groupings noted previously, is the one composed of a husband, wife, and their own children (.32). The majority of consumer units include a non-black reference pension (.89), and the reference person is most likely to be less than 65 years of age (.75). The largest percentage of consumer units live in the South (.31), while the urbanization category most descriptive of the area of residence is large urban (.42). This distribution is quite similar to the distribution of all consumer units from which the Interview data are collected, although the weighted sample represents less than the total U.S. population.

The overall Gini coefficient based on annual expenditures for the total weighted sample in 1987 is .33. For complete income reporters, the total expenditures Gini coefficient is .327, while the Gini coefficients based on income are .448 for before tax income and .435 for after tax income. The Gini coefficient for 1987 based on income from the Current Population Survey (CPS) is .392 [10]. The difference in the two survey income Gini coefficients is likely to be related to the fact that income is imputed for the CPS but not for the CEX. The lower value for the total expenditures estimate, compared to the income estimates, could be due to the stability of consumption expenditures relative to income. Or, as noted by Kakwani, the difference may be explained in terms of the inequality of savings. "Since income is the sum of expenditures and savings, and since households with higher incomes tend to save a greater proportion of the incomes, inequality of saving will be high. This is an important reason why the inequality of income is higher than that of expenditures[11]. Kakwani, using Australian data, reported a total expenditures Gini to be .256 compared to .331 for disposable income for 1975-76.

Pooling all consumer units with different characteristics into a single distribution could hide significant differences in inequality trends if substantial differences among inequality estimates of consumer units with different demographic profiles

---

[9]see: U.S. Department of Labor (1989)
[10]see: U.S. Census (1987)
[11]see: Kakwani (1986)

exist. Thus, when evaluating the effect on inequality of changes in expenditures, it would be better to produce separate Gini estimates for demographic subgroups to identify which consumer units will be most affected. Results from this study reveal such differences. Within the consumer unit size groups, the greatest degree of inequality is among one person consumer units (G=.345), the lowest is among four person households (G=.265). Among the household composition groups, one parent with children households experience the most inequality of expenditures (G=.348); inequality is lowest among husband and wife with children households (G=.265). Inequality in expenditures is slightly greater for consumer units in which the reference person is 65 years of age or older (G=.344) than it is for those in which the reference person is younger (G=.301). Consumer units living in the West appear to have total expenditures which are more equally distributed (G=.292) than do consumer units living in other regions. Inequality in expenditures is also lower for rural consumer units (G=.309) relative to consumer units living in other areas. Inequality did not differ much by race. These results substantiate the claim made by others that inequality in economic well-being differs by the demographic characteristics of the units 12.

Decomposition of the Gini provides specific information concerning the concentration or inequality in consumption expenditures by budget component, the contribution of the budget component to overall inequality, and information about how marginal changes in particular expenditure can affect overall inequality. Within budget component Gini coefficients are expected to be higher, for the most part, than the overall Gini because, with the exception of food, not all consumer units have expenditures for each of the budget components. For example, approximately 49 percent of the sample does not have expenditures for public transportation, which is reflected in the high within budget component Gini coefficient of .852. This indicates that public transportation expenditures are highly concentrated among consumer units in the sample. In contrast, expenditures for food and for fuel and utilities are the most equally distributed of au budget components, with Gini coefficients of .324 and .286, respectively. These low Gini estimates result from the high proportion of consumer units with these expenditures as well as from their relatively even distribution among those consumer units which do have these expenditures.

The greatest amount of inequality in total consumption expenditures is contributed by shelter expenditures; 24 percent of the total inequality in expenditures for the population sample is accounted for by these expenditures. Expenditures for shelter are highly correlated with total expenditures (R = .765), they seem to be somewhat concentrated (G = .463), and they represent approximately 23 percent of total consumption expenditures. Public transportation expenditures contribute the least to the inequality in total expenditures (2.6 percent), followed

[12] see: Dagum-Grenier (1984), Jorgenson-Dale-Slesnick (1984), Ray (1985), Silber (1989), Kakwani (1986).

by medical care and services expenditures (3.7 percent). For most components, the contribution to inequality is similar to their share of total expenditures.

The relative effects on overall inequality of a marginal increase in a budget component can be determined by comparing its expenditure share to its share of expenditure inequality. Changes in expenditures which would lead to reductions in inequality are associated with expenditure components for which the expenditure shares are greater than are the shares of expenditure inequality. The direction of the marginal relationship indicates the effect at the margin of an increase in a component expenditure on overall inequality. This means that a component would exert a negative effect on inequality if the relative marginal effect is negative. For example, in this study a negative value for food results; this means that a one percent increase in food expenditures, holding all else constant (e.g., prices, taxes), would decrease the inequality in total consumption expenditures at the margin by 0.04 percent. Increases in expenditures for fuel and utilities, private transportation, and medical care and services would also decrease overall inequality at the margin.

3. Data Limitations

Thus far in this paper, we have described the U.S. Consumer Expenditure Survey and have highlighted results from distributional analyses of the CEX data conducted by researchers within the BLS. However, results based on household survey data must be considered with some caution. One must keep in mind the fact that studies based up on household survey data are subject to both sampling and nonsampling errors. These errors can affect the validity and reliability of the data. For the CEX data we have been producing measures of the sampling error for some time, with coefficients of variation ranging from two to five percent for most published expenditure items. More recently we have been actively dedicating greater resources to the study of nonsampling errors. Nonsampling errors can be attributed to many sources, such as definitional difficulties, differences in the interpretation of questions, inability or unwillingness of the respondent to provide correct or complete information, mistakes in conding or recording the data obtained, and other errors of collection, response, processing, coverage, estimation for missing data, and interviewer variability. Nonsampling errors affect the data, and subsequently the results based on these data. For example, underreporting affects the data in two ways. There may be differential underreporting by demographic groups. Younger families, for example, may be less likely to report all of their expenditures than are older families. There may also be differential underreporting by expenditure component; households are more likely to remember large purchases or routine payments but are less likely to remember infrequent and/or small

purchases.

Within the BLS, research is being conducted to identify and quantify some of the nonsampling errors for the purpose of developing new data collection methodologies. Various types of evaluation are being utilized, including comparisons of CEX data with data from other sources, statistical analysis of the data to identify the reasons for underreporting, and research on the cognitive processes respondents use as they answer survey questions.

The CEX data are regularly compared to the Personal Consumption Expenditure (PCE) Component of the National Income Accounts[13]. After accounting for definitional differences, the levels obtained from the Interview component of the survey are generally lower than the PCE, indicating that the expenditure data may be somewhat underreported. A more recent comparison project begun at the BLS is to evaluate and compare the income, assets, and liabilities data from the CEX and data collected using the Federal Reserve Board's Survey of Consumer Finance.

As noted, statistical analyses of the CEX data have been conducted to investigate possible reasons for underreporting and the effects on the data of this underreporting. Researchers within BLS have examined time-in-sample and recall effects in two studies using aggregate data from 1982-1983 and microlevel data from 1984[14]. Recall bias, telescoping, and time-in-sample have been identified as potential concerns in data collection for the Interview. Recall bias results from the selective recollection of past events and is common in retrospective interviews. It is related to memory and tends to increase with longer recall periods and for less salient events. The incorrect reporting of events' time of occurrence, or "telescoping" interferes with an analyses of recall bias since it often affects reporting in a forward direction. Thus, a combination of "recall effects" results. Time-in-sample may result in underreporting since consumer units are potentially visited five times for the Interview and three times for the Diary. Over time, consumer units may lose interest in the survey and thus may subsequently not report their expenditures fully. Research results revealed that recall, more than panel conditioning, contributed to the underreporting of expenditures[15]. The researchers noted that internal telescoping, or erroneous reporting of the expenditure month, would not account solely for the systematic variations in mean expenditures by recall month. When studying the importance of the relationship of consumer unit (CU) characteristics and recall bias BLS researchers reported the most important variables were the size, composition, and income level of the consumer unit[16]. Age and education of the survey respondent were important respondent characteristics. They also noted sensitivity of reporting patterns to changes in interview

---

[13] see: Gieseman (1987)
[14] Silberstein-Jacobs (1989), Silberstein (1989)
[15] see: Silberstein Jacobs (1989).
[16] Silberstein (1989).

week. Reporting patterns were influenced by the length of the reporting period and the respondent's perception of it.

Related research, focusing on the Diary component, examined the record-keeping behavior of Diary respondents. Data collected in a 1984 supplemental research survey indicated that almost 20 percent of the respondents completed at least one of the two diaries by total recall, an obvious threat to the quality of the data received for the Diary Survey [17].

Specific research on the underreporting of income is also being conducted. A comparison of the CEX income data with Current Population Survey data reveals that underreporting of income, particularly for property and non-wage income, may be a problem in the CEX. Analysis of data in the lower income groups shows that a large percentage of consumer units report a level of expenditures that greatly exceeds their reported income. Research on income reporting and the characteristics of consumer units reveals that the age, education, and occupation of the reference person, housing tenure, and region in which the consumer unit resides are significant variables in determining the probability that a consumer unit wall be a complete income reporter[18]. Studies such as these have important implications for developing methods that can be used to reduce the underreporting of income and for developing procedures that can be used to impute income for nonreporters. Implementing an income imputation procedure should improve the data even if it does not entirely correct the problem at the source.

Cognitive research related to the data collection instruments and interview procedures began in the mid-1980's. During this time the BLS received funding from Congress to create a Survey Design Research Center. Shortly thereafter, in January 1987, a (questionnaire Design Advisory Conference was convened at the BLS[19]. As a result of the recommendations received by the Bureau at this conference, combined with results from the underreporting studies, research on the cognitive aspects of the data collection process increased. Emphasis in the past year has been on the use of cognitive laboratory techniques for investigating respondents' thought processes in response to survey questions, and developing alternative questionnaire formats and question phrasing. Attention has also been focused on the demands placed upon respondents in terms of burden. Major form redesigns are planned for 1991 and 1996 based on findings from the cognitive research studies.

4. *Summary and Conclusions*

The U.S. has been conducting expenditure surveys for several decades, begin-

---

[17]see: Tucker (1986)

[18]see: Garner (1987)

[19]see: Biennia-Dippo-Palmisano (1987)

ning in the late 19th century. These data have been used for a variety of purposes over this time, including the production of weights for the CPI and for economic and trend analyses. Included in this paper are results from distributional studies of 1987 expenditure data. First, aggregate expenditure data were used to identify the proportion of expenditures by item that various demographic subgroups in the population spend. Second, microlevel data were used to produce Gini coefficients based on total annual consumption expenditures for the total sample population and for demographic subgroups.

These studies represent only two ways of examining distributions. In the future we plan to add to this foundation by examining the CEX data using other distribution measures, for example, coefficients of variation, a general entropy measure, and Atkinson's measure of inequality. We also anticipate expanding our work to include studies of the economic status and materiel well-being of families cross-nationally, as similar data from other countries become more available. Throughout this work, we will continue to be aware of data limitations as we compare results from the CEX with those from other data sources.

REFERENCES

Biennia, J., Dippo, C. and Palmisano, M. (1987), *Questionnaire Design: Report on the 1987 BLS Advisory Conference,* Washington, D.C.: Bureau of Labor Statistics.

Dagum, C., Grenier, G. (1984). Male-Female Income Distributions in Four Canadian Metropolitan Areas: An Application Using Personal Income Tax Data. *American Statistical Association,* 1984, Proceedings of the Section on Survey Research Methods.

Garner, T. (1987), Income Reporting in the U.S. Consumer Expenditure Survey, Vicki Hampton (ed.), *American Council on Consumer Interest - 33rd Annual Conference, April 1-4, 1987, Denver, Colorado, The Proceedings,* Columbia, Missouri: American Council on Consumer Interests, pp. 282-290.

Garner, T. (1989), Consumer Expenditures and Inequality: An Analysis Using the Gini Coefficient,
Paper presented at the 35th Annual Conference of the American Council on Consumer Interests,
Baltimore,Maryland,March 29-April 1.

Gieseman, R. (1987), The Consumer Expenditure Survey: Quality Control by Comparative Analy*sis, Monthly Labor Review,* March, pp. 8-14.

Jorgenson, D. W. and Slesnick, D.T. (1984), Inequality in the Distribution of Individual Welfare *Advances fin Econometrics,* 3, R.L. Basmann and George F. Rhodes, Jr. eds, Greenwich, Connecticut: JAI Press, Inc., 67-130.

Yitzkaki, S. (1983), On an Extension of the Gini Inequality Index, *International Economic Review,* 24: 617-628.

Kakwani, N. (1986), *Analyzing Redistribution Policies: A Study Using Australian Data,* Cambridge: Cambridge University Press.

Kaneer, K., (1986), Distribution of consumption examined using aggregate expenditure shares, *Monthly Labor Review,* April: pp. 50-53.

Lerman, R. I. and Yitzhaki, S. (1984), A Note on the Calculation and Interpretation of the Gini Index, *Economics Letters,* 15: 363-368.

Lerman, R. I. and Yitzhaki, S. (1985), Income Inequality Effects by Income Source: A New Approach and Applications to the United States, *Review of Economics and Statistics,* 67(1): 151-156. Ray, R. (1985), Prices, Children and Inequality: Further Evidence for the United Kingdom, 196582, *The Economic Journal,* 95: 1069-1077.

Silber, J. (1989), Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality, *The Review of Economics and Statistics*, 71 (1): 107-115.
Silberstein, A. and Jacobs, C. (1989), Symptoms of Repeated Interview Effects in the Consumer Expenditure Interview Survey, *Panel Surveys*, eds. D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh, New York: John Wiley, in press.
Silberstein, A. (1989), Recall Effects in the U.S. Consumer Expenditure Survey, *Journal of Official Statistics*, in press.
Tucker, C. (1986), An Analysis of the Dynamics in the CE Diary Survey, Proceedings of the Survey Research Methods, *American Statistical Association, pp. 18-27*
U.S. Census (1987), Current Population Series, P-60, No. 162, Table 12. Washington, D.C.: U.S. Government Printing Office.
U.S. Department of Labor *(1989),* Comprehensive Picture of Consumer Spending Released by Bureau of Labor Statistics, *USDL:* 89-330, July 6.