

Consumer Expenditure Survey Microdata Users' Workshop, July 2009

Geoffrey D. Paulin

The Consumer Expenditure Survey (CE) is the most detailed source of expenditure, demographic, and income data collected by the Federal Government. The data are collected in two component surveys: the (quarterly) Interview Survey and the Diary Survey. Each year, the CE program releases topcoded microdata from these surveys, which are used by researchers in a variety of areas, including academia, government, market research, and other private industries.

In 2006, the Division of Consumer Expenditure Surveys began conducting a workshop each July for users of the CE microdata. Held in the conference facilities of the Bureau of Labor Statistics (BLS) headquarters in Washington, DC, the workshops have included speakers demonstrating features of the data, as well as reports from researchers who have used the data in their work. The BLS recognizes the workshop, which averages 50 participants, as a "BLS Best Practice." Each year, the format has changed to incorporate suggestions from participants, but the basic framework has remained intact.

For the most recent workshop (July 2009), the program was expanded from 2 days to 3 days. The first day was designed especially for new users, including novices and those who had never used the data. The second day

was designed to feature research from users outside the BLS. The third day was designed particularly for more experienced users. The program was arranged in this way to accommodate as many participants as possible. That is, any attendee could attend 1, 2, or all 3 days of the workshop and benefit from sessions geared toward his or her expertise.

The first day opened with a welcome by leader Bill Passero, followed by an overview of the CE, featuring topics such as how the data are collected and published (Veri Crain). This session was followed by an introduction to the microdata, including an explanation of its features (Passero and Jeff Crilley). In the afternoon, participants received practical "hands-on" training, with expert users from the Division of Consumer Expenditure Surveys staff (Laura Paszkiewicz and Crilley) demonstrating introductory computer programming techniques to participants in a classroom equipped with several computers.¹

The second day featured a full day of presentations from researchers outside the BLS. In addition, complementing speakers who described results from the Interview Survey, Geoffrey Paulin gave a presentation on uses of data from the Diary Survey and Gerald Perrins, of the Philadelphia regional office of the BLS, described how regional offices use such data. In all, nine outside researchers presented their work, demonstrating the wide range and diversity of topics that can be studied with this rich source of data. (See the box on page 49 for the names, affiliations, and topics of the presenters.) The day concluded with a special presentation from Terry Schau and Leslie Brown Joyner, the current and immediate past managing editors of the *Monthly*

Labor Review. Their presentation described the publication process from submission to printing, for authors interested in having their works appear in that journal. Next came a brief talk by Steve Henderson about changes to the microdata files that would occur with the release of the 2008 microdata in October 2009, including a description of specially collected data on the 2008 Economic Stimulus payments, also known as "tax rebates." The second day of talks represented the first time that a full day of the workshop was devoted to research presentations; the talks were added in response to comments from past attendees indicating that research presentations were among their favorite parts of the workshops.

The third day featured advanced topics, including technical details about sampling methods and the construction of sample weights (Catherine Hackett, Division of Price Statistical Methods); imputation and the allocation of microdata (Troy Olson); and the proper use of sample weights in computing population estimates (Paulin). The latter session noted that the proper use of weights requires a special technique to account for sample design effects that, if not employed, results in estimates of variances and regression parameters that are incorrect.² Similar sessions in the afternoon addressed the proper use of the multiply imputed income variables (Paulin); using data from participants in all four published interviews, rather than treating observations from each quarter independently (Passero); and applying sales taxes to expenditure reports (Meaghan Duetsch). These presentations were followed by practical training sessions that, in part, provided programming examples of the con-

Geoffrey D. Paulin is a senior economist in the Publication and Development Section, Branch of Information and Analysis, Office of Prices and Living Conditions, Bureau of Labor Statistics. E-mail: Paulin.Geoffrey@bls.gov

cepts described in the earlier sessions of that day. The day concluded with CE program staff soliciting feedback from the participants.

Following are abstracts of the papers read at the conference, listed in the order in which the papers were presented:

Richard Bavier, "Reconciliation of Income and Consumption Data in Poverty Measurement," *Journal of Policy Analysis and Management*, winter 2008, pp. 40–62. A recent series of papers has renewed interest in the question of whether consumption data are superior to income data for poverty measurement. Although the Census Bureau has provided researchers with an experimental series of variables that can produce a comprehensive income measure, this resource has not been fully exploited in previous analyses. When poverty is measured by a comprehensive income measure, income poverty rates and trends are similar to consumption poverty rates. Arguments that income is measured with more error than consumption at the bottom of the distribution are shown to be based upon inferior income data. (© 2008 by the Association for Public Policy Analysis and Management.)

Arpita Biswas, Ph.D. candidate, Clemson University, "Effect of Income Taxes on Charitable Giving," dissertation (in progress). This paper analyzes how lowering tax rates on income affects charitable giving decisions of individuals across various socioeconomic groups. The paper focuses on a particular government intervention—the Economic Growth and Tax Reconciliation Relief Act of 2001—to answer three key questions: (1) What

is the elasticity of charitable giving with respect to income and price? (2) As lower taxes increase the marginal cost of giving, what is the effect on charitable giving of the exogenous increase in income due to lower taxes? (3) With regard to the latter question, does the income effect dominate the substitution effect? The CE data set from 1997 to 2006 is used both to derive empirical results and to investigate how the results vary across income and age groups.

Raymond Ring, Professor of Economics, School of Business, University of South Dakota, "Reconciling Census Households with Consumer Expenditure Survey Consumer Units." The CE uses consumer units to categorize and analyze expenditure data, whereas the Census Bureau uses households. The definitional differences between these sample universes are slight, but significant. This paper emphasizes those differences in order to identify situations in which the results of data analysis vary with the definition used.

John McCollough, assistant professor of business, The Pennsylvania State University-Lehigh Valley, "Consumer Discount Rates and the Decision to Repair or Replace a Durable Product: A Sustainable Consumption Issue," *Journal of Economic Issues*, March 2010, pp. 183–204. This paper attempts to answer the question, "Given two consumers with the same income level and the same socioeconomic background, why does one consumer choose to replace an older, malfunctioning product while the other chooses to have the product repaired for further reuse?" To help address this question, an enhanced replacement model is pre-

sented and then empirically tested. The replacement model has been around for some time, helping consumers and firms make capital budgeting decisions. The replacement model was chosen because it highlights the role of individual discount rates and consumers' time preferences in the decision process. The analysis can be extrapolated to the macrolevel in which, for economies that are similar in that they have comparable levels of gross domestic product per capita and prices, the economy with a higher societal discount rate might be more strongly characterized as a "throwaway society" than the one with a lower societal discount rate.

Megumi Omori, assistant professor of sociology, Department of Sociology, Social Work, and Criminal Justice, Bloomsburg University, "Household Expenditure on Children: Differences in Resource Allocation by Household Type." Children in single-parent families are disadvantaged in many ways, compared with their counterparts in two-parent families. Differences in children's well-being between two-parent and single-parent families are often attributed to differences in resources, such as parental income and time spent with children. Also, the economic disadvantage of single families is clearly shown by statistics: in 2006, the median income for married families was \$69,716, whereas that for single-father families was \$47,078 and for single-mother families was \$31,818 dollars. Although it is well established that income is a strong indicator of children's well-being, little focus has been paid to possible differences in the *allocation* of economic resources.

Accordingly, this paper uses the 2005 CE Interview Survey to explore household expenditures on children, with a special focus on household types. Specifically, the study asks, “Are there any differences in spending patterns for children in different types of households within the same income group, and if so, where is the difference?”

Lisa Kolovich, Ph.D. candidate, University of Maryland, “Home Bias in International Trade: Who Has a Taste for Discrimination?” dissertation (in progress). This paper investigates the role that consumer preferences play in “home bias,” the preference by consumers for products produced in their own country over otherwise identical imports. Data from the U.S. automobile market for the years 2000 to 2004 and from the CE are used to estimate whether there is a home bias for automobiles produced by Ford, General Motors, or Chrysler in the light of Becker’s 1971 theory of discrimination. The results show that, after accounting for a wide array of amenities and for performance, quality, and reliability characteristics of automobiles, there is indeed a home bias for American-badged (that is, Ford, General Motors, and Chrysler) automobiles. In addition, preliminary results show that certain demographic groups of consumers—for instance, older consumers, less educated consumers, and individuals residing in the Midwest—appear to be more likely to purchase American-badged automobiles.

Janet Wagner, associate professor of business management and marketing, University of Maryland,

“Seasonality in Household Service Expenditures: A Theoretical Framework and an Empirical Analysis” (coauthored with Manoucher Mokhtari). Using CE microdata, this paper presents a theoretical model of the effects of seasonality and household characteristics on service expenditures. Tobit regression is used to test the model, and the paper concludes that seasonality appears to moderate the effect of personal characteristics, characteristics of the consumer unit, and demographics on the consumer unit’s quarterly service expenditures for household operations, entertainment, and food away from home.

Helen Levy, research assistant professor, University of Michigan, “Consequences of SCHIP for Household Well-Being” (coauthored with Lindsey Leininger (University of Wisconsin) and Diane Whitmore Schanzenbach (University of Chicago)). This paper uses data from the CE and the Survey of Income and Program Participation to analyze how the expansions of the State Children’s Health Insurance Program (SCHIP) affected total household consumption as well as detailed categories of household spending (for example, food, housing, and education). By means of an instrumental variables approach that relies on variation across States in the generosity of SCHIP expansions, the effect of expanding coverage on total consumer spending, as well as the distribution of spending across consumption categories, is isolated. Then, by showing what households do with the money they “save” by switching their children from private

to public health insurance, the analysis provides solid data on how SCHIP has improved the material well-being of the low-income families it is intended to assist, including those who previously had been paying for their own coverage.

Paul Wilson, Minnesota Department of Revenue, “Using CES Data to Estimate Consumption Tax Burdens in Minnesota” (coauthored with Phillip Anthony). The Minnesota Department of Revenue is required to complete a biennial study of Minnesota tax burdens, estimating how State and local tax burdens vary across resident households by income and household type. For the latest edition of this study (the 10th), amounts for each kind of tax were estimated for each household that was included in a stratified random sample of 105,000 households. The CE was used to estimate many of the consumption taxes paid, including the general sales tax, motor vehicle sales tax, and State excise taxes on cigarettes, alcohol, and motor fuels. Consumer expenditure estimates also were used to examine the distribution of spending on categories of goods that were not subject to the sales tax, in order to address the impact of proposals to broaden the sales tax base to include currently nontaxable goods and services.

The next workshop will be held July 14–16, 2010. It will be free of charge to all participants, although advanced registration is required. For more information about the 2009 and forthcoming workshops, visit the CE Web site, www.bls.gov/cex, and look for “Annual Workshop” under the left navigation bar entitled

“PUBLIC USE MICRODATA.” For direct access to this information, the link is www.bls.gov/cex/csxannual-workshop.htm. □

Notes

ACKNOWLEDGMENT: The author wishes to acknowledge William Hawk of the Of-

fice of Prices and Living Conditions, Bureau of Labor Statistics, for compiling the list of abstracts of papers presented at the workshop.

¹ Examples included how to merge data from the FMLY and MEMB files. The FMLY file contains information about the consumer unit as a whole, such as region of residence and summary variables for expenditure categories, including total expenditures, housing, and apparel. (For the definition of a consumer unit, see “About the CE data,” in the appendix.)

The MEMB files contain information about each member of the consumer unit, such as the member’s age, ethnicity, and educational attainment. Data on expenditures are collected for the consumer unit as a whole; therefore, expenditure data on specific members are not available, except in single-member consumer units.

² The CE sample design is pseudorandom. However, a proper use of weights requires the method of balanced repeated replication.

Speakers at the workshop

BLS Staff of Division of Consumer Expenditure Surveys:

- Crain, Veri, economist, Branch of Information and Analysis; day 1
- Crilley, Jeffrey, economist, Branch of Information and Analysis; day 1
- Duetsch, Meaghan, supervisory economist, Chief, Phase 1/Phase 2 Section, Branch of Production and Control; day 3
- Henderson, Steve, supervisory economist, Chief, Branch of Information and Analysis; all days
- Olson, Troy, supervisory economist, Chief, Phase 3 Section, Branch of Production and Control; day 3
- Passero, Bill, senior economist, Branch of Information and Analysis; all days
- Paszkiwicz, Laura, senior economist, Branch of Information and Analysis; day 1
- Paulin, Geoffrey, senior economist, Branch of Information and Analysis; days 2 and 3

Other BLS speakers:

- Hackett, Catherine, mathematical statistician, Division of Price Statistical Methods; day 3
- Joyner, Leslie Brown, supervisory economist, Office of Publications and Special Studies, and Branch Chief, Editorial Services A; day 2
- Perrins, Gerald, supervisory economist, Office of Field Operations, Philadelphia regional office, Division of Economic Analysis and Information; day 2
- Schau, Terry, supervisory economist, Office of Publications and Special Studies, and Branch Chief, Editorial Services 1; day 2

Speakers from outside BLS:

- Richard Bavier, “Reconciliation of Income and Consumption Data in Poverty Measurement,” *Journal of Policy Anal-*

ysis and Management, winter 2008, pp. 40–62.

- Arpita Biswas, Ph.D. candidate, Clemson University, “Effect of Income Taxes on Charitable Giving,” Ph.D. dissertation (in progress).
- Lisa Kolovich, Ph.D. candidate, University of Maryland, “Home Bias in International Trade: Who Has a Taste for Discrimination?” Ph. D. dissertation (in progress).
- Helen Levy, research assistant professor, University of Michigan, “Consequences of SCHIP for Household Well-Being,” coauthored with Lindsey Leininger (University of Wisconsin) and Diane Whitmore Schanzenbach (University of Chicago).
- John McCollough, assistant professor of business, The Pennsylvania State University, Lehigh Valley, “Consumer Discount Rates and the Decision to Repair or Replace a Durable Product: A Sustainable Consumption Issue,” *Journal of Economic Issues*, March 2010.
- Megumi Omori, assistant professor of sociology, Department of Sociology, Social Work, and Criminal Justice, Bloomsburg University, “Household Expenditure on Children: Differences in Resource Allocation by Household Type.”
- Raymond Ring, professor of economics, School of Business, University of South Dakota, “Reconciling Census Households with Consumer Expenditure Survey Consumer Units.”
- Janet Wagner, associate professor of business management and marketing, University of Maryland, “Seasonality in Household Service Expenditures: A Theoretical Framework and an Empirical Analysis,” coauthored with Manoucher Mokhtari.
- Paul Wilson, Minnesota Department of Revenue, “Using CES Data to Estimate Consumption Tax Burdens in Minnesota,” coauthored with Phillip Anthony.

APPENDIX: About the CE data

Consumer unit. The basic unit of analysis in the Consumer Expenditure Survey (CE) is the consumer unit. In general, a consumer unit consists of (1) all members of a particular household who are related by blood, marriage, adoption, or some other legal arrangement; (2) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in permanent living quarters in a hotel or motel, but who is financially independent; or (3) two or more persons living together who use their incomes to make joint expenditure decisions. Financial independence is determined by spending behavior with regard to the three major expense categories: housing, food, and other living expenses. To be considered financially independent, the respondent must provide at least two of these expenditure categories, either entirely or in part.

Collection and methodology. Since 1980, the Interview and Diary Surveys have been collected on an ongoing basis. The Inter-

view Survey is designed to collect expenditures for big-ticket (for example, major appliances, and cars and trucks) and recurring (for instance, payments for rent, mortgage, and insurance) items. Data on some expenditures, such as food at home, are collected globally.¹ In addition to data on expenditures, demographics, and income, information about assets and liabilities is collected. In this survey, participants are visited once every 3 months for five consecutive quarters. Data from the first interview are collected only for bounding purposes and are not published.² Since April 2006, about 7,000 consumer units have participated each quarter.

In the Diary Survey, participants record expenditures daily for two consecutive weeks. The survey is designed to collect expenditures for small-ticket and frequently purchased items, such as detailed types of food (white bread; ground beef; butter; lettuce). Since April 2006, about 7,000 consumer units have participated annually. Because they complete a separate diary each week, approximately 14,000 diaries are collected each year.

Notes to the appendix

¹ That is, the respondent is asked to provide an estimate of the consumer unit's total expenditure for these items, rather than collecting detailed information on the items composing food expenditures.

² A bounding interview collects information to alert the interviewer to probe in cases where the purchase of a big-ticket or an infrequently purchased item reported in one interview is reported, perhaps inadvertently, in

the next interview. For example, if, in both the first and second interviews, the respondent reports that he or she purchased a refrigerator, the interviewer can ask followup questions during the second interview to ascertain whether the refrigerator that was purchased was the one reported in the first interview. The same process is followed in the second through fifth interviews when similar cases occur. That is, the second interview provides bounding information for the third interview, and so forth.