

This From Our (Microdata) Files: Research and Instruction Using the Consumer Expenditure Survey, from Novice, Intermediate, and Expert Users

Geoffrey Paulin, Bureau of Labor Statistics¹

Tian Luo, Bureau of Labor Statistics²

Bill Passero, Bureau of Labor Statistics³

Session overview

The Consumer Expenditure Survey (CE) is a unique and rich source of data of interest to researchers, educators, advocates, policymakers, and others in a variety of fields. In addition to expenditures, it provides the most detailed information on income, assets and liabilities, and demographics collected directly from households by the U.S. government. This session was designed both to provide summaries of current research using these data, and also insight into their use.

In addition to viewing standard presentations of research in progress, attendees received materials to take with them. These included a complimentary CD-ROM containing a sampling of the full set of the 2010 Consumer Expenditure Survey microdata; a computer program showing the basics of microdata use was provided; and information about the annual CE microdata users' workshop held in Washington, DC each July since 2006. In the workshop, attendees have an opportunity to learn details about using the data, and to receive "hands on" training in this endeavor.

The standard presentations, for which abstracts follow, represented CE microdata users at all levels of experience—novice, intermediate, and expert. However, they were also designed to provide more than standard recitation of procedures and results. For example, one of these was delivered by an economist from the Bureau of Labor Statistics (BLS) Western Information Office in San Francisco (Tian Luo). On behalf of staff from BLS regional offices across the country, he briefly described the role of his co-workers in providing local area data to researchers who are interested in current economic conditions facing consumers, such as price changes or employment/unemployment rates. Another presenter (Bill Passero) focused on data quality, comparing income data from CE and the Current Population Survey. The final presenter (Geoffrey Paulin) described a new project being undertaken by the CE program: the imputation of Federal and State income tax data. Reasons for the project, proposed methodology, implications for data users, and an invitation to participate in the review process were presented.

The Data

Each presentation described results from the Interview component of the Consumer Expenditure Survey. The Interview Survey is collected in five consecutive quarterly visits to households. Each eligible household contains at least one consumer unit (similar to a family), from which a respondent is selected. In each case, respondents are asked to recall expenditures made by their consumer unit for a variety of goods and services over the last three months. Using a separate sample, the Diary Survey collects information from participants for two consecutive weeks. On each day of the collection period, respondents document expenditures for the consumer unit. The data from both surveys are integrated and published in standard tables, which are available on the Internet. In addition, CD-ROMs containing microdata files from each source are made available for purchase each year. The 2010 microdata files for the Interview and Diary Surveys were made available in September 2011, the earliest release in the program's history.

Materials Distributed

As noted, attendees received folders with numerous materials related to the CE program. Some of these

¹ Senior Economist, Consumer Expenditure Survey Program, 2 Massachusetts Avenue NE #3985, Washington, DC. Phone: (202) 691-5132. Email: Paulin.Geoffrey@bls.gov.

² Regional Economist, Western BLS Information Office (San Francisco), Division of Economic Analysis and Information, 90 7th Street, Suite 14-100, San Francisco, CA. Phone: (415) 625-2286. Email: Luo.Tian@bls.gov.

³ Senior Economist, Consumer Expenditure Survey Program, 2 Massachusetts Avenue NE #3985, Washington, DC. Phone: (202) 691-5126. Email: Passero.Bill@bls.gov.

items are available online, and the rest is available on request. The items include: a BLS News Release entitled “Consumer Expenditures – 2010” (<http://www.bls.gov/news.release/pdf/cesan.pdf>); an article from the BLS “Focus on Prices and Spending” series entitled, “How do U.S. expenditures compare with those of other countries?” (http://www.bls.gov/opub/focus/volume2_number16/cex_2_16.htm); two articles from the Monthly Labor Review (“Household food expenditure patterns: a cluster analysis,” <http://www.bls.gov/opub/mlr/2007/04/art3full.pdf>; and “Household expenditures on children, 2007-08,” <http://www.bls.gov/opub/mlr/2010/09/art1full.pdf>); a printout of website describing BLS onsite researcher program, for researchers interested in accessing confidential data files (<http://www.bls.gov/bls/blsresda.htm>); an announcement of the 2012 Consumer Expenditure Survey Symposium and Workshop (details available at <http://www.bls.gov/cex/csxannualworkshop.htm>); a document providing links to CE and information; a call for presenters at the 2012 CE microdata users’ workshop (deadline: April 13, 2012; <http://www.bls.gov/cex/csx2012workcallpresenters.pdf>); a published report on the 2010 workshop (<http://www.bls.gov/opub/mlr/2011/04/art11full.pdf>); a program from the 2011 CE Microdata Users’ workshop; and a SAS program, with output, using sample-data CD-ROM. The SAS program showed how to obtain information on items such as total outlays, total income, housing characteristics (tenure, and numbers of rooms, bedrooms, bathrooms, and halfbaths), examples of asset (amount in checking, brokerage, or similar account) and liability (type of debt owed—credit card, installment, health care provider, etc.), and assorted demographic characteristics. In addition, the CD-ROM included a document summarizing asset and liability data available in the public use microdata files, and several additional articles using CE data that were not printed for inclusion in the folder.

Keeping Up Appearances, Or Just Keeping Afloat? How and Why American Households Overspend

Jeffrey D. Lundy, University of California, San Diego⁴
(Described at session by Geoffrey Paulin)

Why do Americans “overspend” (i.e. consume more money than they gain)? The pressures that drive overspending have been a subject of sociological debate since Veblen published his classic text on “conspicuous consumption” at the turn of the 20th century. In recent times, social theorists have built upon this tradition by offering diverse explanations for overspending; such as an increased drive for social status, a redefinition of “luxury” goods, a reduction in families’ financial stability, and a desire to impress racial or class peers. However, none of these scholars have explored empirical data to test their theories about which Americans are overspending, the causes for overspending, or the relative importance of overspending as an explanation for households’ wealth losses.

This paper develops a model for how overspending contributes to households’ wealth losses. Using the *Consumer Expenditure Survey*, the paper explores how wealth losses are distributed among American households, and the empirical conformity of that distribution with the expectations of past theorists.

The causes of wealth loss are found to be contingent on a household’s place in the financial spectrum. Households lacking major assets are likely to lose wealth due to income-consumption gaps like losing a job or living in poverty; whereas wealthy households are affected by fluctuations in the macroeconomy. Past theories of overspending based on status spending and racially-based spending are found to be insignificant predictors of wealth loss. However, problems in survey design may be to blame for the low explanatory power of some of these factors.

⁴Ph.D. Candidate in Sociology. Email: jlundy@ucsd.edu.

The Day the Roof Caved In: Consumer Expenditures for Housing in Major Metropolitan Areas Before, During, and After the Bursting of the Recent Housing Bubble

Karen Ransom, Bureau of Labor Statistics⁵
Jacqueline Midkiff, Bureau of Labor Statistics⁶
Cheryl Abbot, Bureau of Labor Statistics⁷
(Described at session by Geoffrey Paulin)

Over the last few years, national media have discussed in detail the so-called “bursting of the housing bubble.” According to the stories, years of continually low interest rates combined with the availability of increasingly exotic mortgages led to an unsustainable increase in housing purchases and mortgage refinancing. Eventually, consumers who could not ultimately pay their loans defaulted, and the resulting foreclosures caused a downward spiral of lowered property values, decreased liquidity of both homeowners and financiers, and other problems, which eventually led to a greater financial crisis and deep recession, from which the economy is still in recovery; for example, the national unemployment rate peaked at 9.6 percent for 2010, and remained at or above 9.1 percent for six of the nine first months (January through September) of 2011.

This work, which was recently undertaken, investigates how consumers reacted before, during, and after the crest in property values. Data from the Consumer Expenditure Interview Survey are examined in several years for various major metropolitan areas (e.g., New York City, Chicago, Atlanta, and San Francisco) to see how rates of homeownership (and conversely renting), housing expenditure, mortgage principal payment, and other variables changed during these periods. In addition, changes by other demographics, such as age, are examined where sample size permits.

⁵Regional Economist, Chief, Division of Economic Analysis and Information, Southeast BLS Information Office, Atlanta, GA. Phone: (404) 893-4224. Email: Ransom.Karen@bls.gov.

⁶ Regional Economist, Chief, Division of Economic Analysis and Information, Mountain-Plains BLS Information Office, Kansas City, MO. Phone: (816) 285-7001. Email: Midkiff.Jacqueline@bls.gov.

⁷ Regional Economist, Chief, Division of Economic Analysis and Information, Southwest BLS Information Office, Dallas, TX. Phone: (972) 850-4805. Email: Abbot.Cheryl@bls.gov.

Use of Local Mass Transit: Expenditures in Washington, DC and Other Northeastern U.S. Cities Compared

Kara Markley, Bureau of Labor Statistics⁸
(Described at session by Geoffrey Paulin)

This project is a work in progress that uses 2009 data from the Interview component of the Consumer Expenditure Survey to compare public transportation expenditures, particularly mass transit (subway, bus, etc.), in Washington, D.C. to other large cities in the Mid-Atlantic and Northeast regions. The research will look at the percentage of users reporting mass transit expenditures in each city, as well as their average quarterly expenditures on these services. While New York City easily has the greatest percentage of users and largest expenditure share, the research is expected to show that due to its extensive and expanding public transportation system, relative to other cities, Washington, D.C. is not terribly far behind New York in public transportation expenditures and the percent reporting expenditures (despite it also over-taking all other U.S. cities for hours spent idling in traffic congestion each year)!

This research also will examine the demographics of Washington, D.C. mass transit users, in terms of income status, employment status, age groups, educational status, etc. The research will point to trends in percent reporting expenditures as a proxy for ridership status among these different demographic groups, as well as examining their mass transit expenditures.

⁸Regional Economist, Chief, Division of Economic Analysis and Information, Mid-Atlantic BLS Information Office, Dallas, TX. Phone: (215) 861-5603. Email: Markley.Kara@bls.gov.

Expenditures for Higher Education by Race

Tian Luo

Human capital theory suggests that greater investment in education will yield individual productivity improvements. These productivity improvements may result in an increase in discounted lifetime earnings returns. Ideally these investments will exceed the present value costs of delaying earnings and higher education purchases. Labor force data support the conclusion that higher education pays off through a general improvement in earnings and lower unemployment. While various researchers have correlated higher educational attainment with higher earning power, there is little information as to the nexus between educational attainment choice and earnings effects associated with different ethnicities or race.

Using BLS Consumer Expenditure (CE) microdata, we investigate the decision to and the amount to invest in higher education and discern any differences by racial and ethnic groups. We show how socioeconomic factors such as parents' education, income, assets, and occupation, may affect the investment decision and discuss any differences by racial and ethnic groups. This analysis will provide measures in an area where primarily anecdotal, subjective, or partial evidence exists.

Preliminary analysis shows that socioeconomic differences explain many of the expenditure differences among racial and ethnic groups. We found that in general, Blacks and White Hispanics, have a lower probability of any higher educational expenditures, and those who do, tend to spend much less. The opposite results were found for Asians. However, when controlling for socioeconomic differences, these differences converged for White Hispanics as well as partially for Blacks, while Asians had significantly higher probability of expenditure and higher level of expenditures, relative to White non-Hispanics. Results also showed that Asians tend to have a much higher probability of attending school full time while the opposite is observed for Blacks and White Hispanics.

Comparison of Post-Imputation Income Estimates from the Consumer Expenditure Survey and Current Population Survey

Bill Passero

The Consumer Expenditure Survey (CE) collects data both on the buying habits of American consumers and extensive demographic and socioeconomic data, including income data. CE staff assesses the quality of the survey's data by comparing CE data to similar data from other sources. Income data from 1980, the year the CE became a continuous survey, until 2004, could not be effectively compared to other sources of income data as the CE had no procedures to impute income values for sample units whose respondents acknowledged receiving sources of income, but did not provide values.

However, the release of 2004 data marked the introduction of imputation for missing income responses. With a number of years of imputed income data now available, it is possible to evaluate the impact and efficacy of imputation by comparing pre- and post-imputation estimates of CE income with estimates from the Current Population Survey (CPS), a large-scale household survey that has employed imputation in producing its income estimates for many years. This paper analyzes changes in the ratio of post-imputation CE income estimates to income estimates from the 2004-2010 Current Population Survey. The paper also gauges the impact of imputation on CE income estimates, comparing income reports prior to imputation in 2002 and 2003 with post-imputation data from 2004 to 2010.

Imputing Income Taxes in the Consumer Expenditure Survey: The Problem, Proposed Solution, and Call for Review and Comment

Geoffrey Paulin

In theory, there are three uses for money: to spend (consume), to save, and to pay taxes. Given this, the importance of accurately measuring after-tax income is of paramount importance for researchers and policy makers.

For some time, the staff of the Consumer Expenditure Survey program has observed a serious underestimation of taxes paid by consumer units within the survey. In 2010, the most recent year for which data are available, the average consumer unit reported almost \$62,500 in income, but only about \$1,100 in Federal taxes. For single-member consumer units, most of which comprise at most one tax unit, average reported income was

nearly \$33,000, with average Federal taxes reported to be \$730. In part, this is due to a low percent reporting: Less than half (49 percent) of all consumer units report an amount for Federal taxes paid, and while a substantially smaller share (almost 40 percent) of single-member consumer units do so. For those who do not provide a value, \$0 is averaged in with those who do provide values, which lowers the mean considerably.

The problem became even more noticeable when income imputation was introduced with the publication of data collected in 2004. In part, this is because imputation increased the size of the sample used in the publication of income data. That is, prior to 2004, only data from “complete income reporters” (essentially, reporters of values for at least one major source of income) was used to compute published income data, and corresponding tax data; therefore, if complete income reporters also were more likely than incomplete reporters to provide tax information, then using imputed income data from incomplete reporters and unimputed tax data from incomplete reporters in published tables would exacerbate the problem. This hypothesis is supported by the fact that 56 percent of complete income reporters provided information on Federal taxes in 2003, while only 50 percent of all consumer units provided this information in 2004, which, as noted, was the first year in which imputed data became available. Therefore, it is reasonable to investigate possibilities for imputing income tax data.

After much research and consideration, the Consumer Expenditure Survey program is embarking on a project to impute taxes, including Federal and State income taxes, using TAXSIM, a tax simulator produced by the National Bureau of Economic Research. In this presentation, attendees will learn about the software and imputation process; see how test results compare to reported results; understand the implications of tax imputation for their research; and be invited to participate in an evaluation of the imputation process and product.