Working Papers

FEDERAL RESERVE BANK OF BOSTON

No. 10-12

Imputing Household Spending in the Panel Study of Income Dynamics: A Comparison of Approaches

Daniel Cooper

Abstract:
One of the drawbacks of using household surveys to investigate macroeconomic issues has been a lack of a dataset that contains both adequate household expenditure data and comprehensive household wealth and income data. This paper compares alternative methods of imputing household expenditures in the Panel Study of Income Dynamics (PSID)—that of Blundell et al. (2006) and Cooper (2009). It also analyzes the additional expenditure questions included in the PSID starting in 1999 and expanded in 2005. The paper finds that the Blundell et al. (2006) method works well for imputing households' nondurable expenditures between 1980 and 2007. The results further show that the imputation method in Cooper (2009) dominates that of Blundell et al. (2006) for generating data on households' total expenditures. The decision of which imputation approach to use or whether to use the actual PSID expenditure data from 1999 to 2007 will depend on the user's research question(s) and analysis goals.

JEL Classification: E20

Daniel Cooper is an economist in the research department of the Federal Reserve Bank of Boston. His e-mail address is daniel.cooper@bos.frb.org.

This paper, which may be revised, is available on the web site of the Federal Reserve Bank of Boston at http://www.bos.frb.org/economic/wp/index.htm.

Thank you to Madeleine Weingarten for her research assistance. I could not have completed this paper without her help. Thank you as well to Kathy Bradbury and Bob Triest for helpful comments. All errors are my own.

This version: December 2010

1 Introduction

Economists have often been limited in household surveys by a lack of data on household expenditures. Until 1999 the Panel Study of Income Dynamics (PSID) gathered data primarily on households' food expenditures, although it also gathered detailed information on household wealth, income, and other demographics. In contrast, the Consumer Expenditure Survey (CEX), has very detailed data on household expenditures but limited data on income and wealth. Other household surveys such as the Current Population Survey (CPS) and/or the Survey of Income and Program Participation (SIPP) have little if any information on household expenditures.

Performing microeconomic analysis of macroeconomic issues often requires a comprehensive measure of household expenditures as well as detailed wealth and income data. Household-level data allow researchers to investigate heterogeneity in household behavior—something that cannot be addressed with aggregate analysis. Investigating and/or controlling for household heterogeneity is particularly important when analyzing issues such as the recent housing market and financial crises.

The PSID appeals to researchers because, unlike most household level datasets, it has a long panel dimension, which enables the researcher to control for household- specific effects and changes in household behavior over time. In addition, the PSID is nationally representative in the cross-section. Until recently, a major drawback of the dataset was a lack of detailed household expenditure data. Until 1999, the only consistent measure of household spending in the PSID was households' food expenditures, which did not provide a comprehensive picture of households' spending decisions. Questions were added to the survey beginning in 1999 that provide a broader picture of household expenditures.

A couple of approaches have been proposed for circumventing the dearth of expenditure data in the PSID. Skinner (1987) imputed nondurable consumption in the PSID, based on the observed relationship between nondurable consumption, food consumption, and a group of demographic variables that are common in both the PSID and CEX. Blundell, Pistaferri, and Preston (2006) expand on Skinner's approach and estimate food demand relationships in the CEX, which they then invert to get nondurable consumption in the PSID. The contribution of their approach is using an instrumental variable approach to deal with potential bias in the imputation process. In addition, Cooper (2009) uses an in-sample method to impute households' nonhousing expenditures in the PSID, based on households' budget constraint and the available income and saving data.

This paper compares the different techniques for imputing household expenditures in the PSID. In particular, it extends the approach in Blundell, Pistaferri, and Preston (2006) (BPP) through 2007 and compares the data from BPP's out-of-sample imputation method

and Cooper's in-sample approach to aggregate benchmarks. The paper also looks at how well BPP's imputation method captures the actual spending data reported in the PSID from 1999 onward. This analysis includes extending the work in Charles et al. (2007) to provide a mapping between the disaggregated CEX expenditure categories and the additional PSID spending questions added in 2005.

The results show that BPP's out-of-sample approach does a good job of imputing households' nondurable expenditures in the PSID. The imputed data line up well with the actual CEX data, but tend to be somewhat lower than the equivalent data from the National Income and Product Accounts (NIPA). The divergence between the micro data and the aggregate data becomes worse when one imputes a broader basket of expenditures than BPP's nondurable expenditure measure. In particular, total per capita imputed household expenditures and the actual CEX data are substantially lower than per capita total personal consumption expenditures (PCE) in the NIPA. This finding is consistent with recent work by Sabelhaus (2010) and others that shows that the CEX data under-report aggregate household spending.

In comparison, the in-sample imputation approach, based on households' budget constraints, does a much better job of capturing total household expenditures in the PSID. As predicted, these data lie somewhere between total PCE and total PCE excluding housing, and follow the general trend observed in the NIPA data. This budget constraint based approach clearly dominates BPP's imputation approach when a researcher is interested in examining households' total expenditures in the PSID. This method also is preferable to using households' actual expenditure data recorded in the PSID from 1999 onward in terms of measuring households' composite consumption. The actual PSID data, however, are reasonable and worth using when a researcher is interested in households' more disaggregated spending behavior. The actual PSID data are also preferable to using BPP's technique to impute a comparable basket of goods.

The remainder of this paper proceeds as follows. Section 2 discusses the data in more detail. Section 3 discusses the different consumption imputation techniques. Section 4 describes the sample, and Section 5 discusses the results from the different imputation approaches. Section 6 concludes.

2 Data

2.1 Panel Study of Income Dynamics

The PSID is a nationally representative longitudinal study of households and their offspring that began in 1968. The original sample included roughly 4,000 households and currently

the PSID surveys nearly 8000 U.S. households. The data were collected annually between 1968 and 1997 and every other year since 1997. The most recent data are for 2007.

The PSID collects detailed data on households' economic variables, health status, and social behavior. The survey also contains data on household income, financial wealth, housing wealth, and saving. The nonhousing wealth data come from "wealth supplements" in 1984, 1989, 1994, and 1999 onwards. These wealth supplements contain data on the following household financial wealth categories: other real estate, businesses or farms, cash, stocks, bonds, vehicles, noncollateralized debt, and IRA/401k accounts. The wealth supplements also collect information on so called "active saving," which measures households' saving out of their current income and excludes capital gains. These saving data are discussed in more detail in the appendix.

A drawback of the PSID is that food consumption is the only expenditure category consistently available since the survey began. Food expenditures provide a narrow view of household spending, and do not capture all the variation in household consumption behavior, since they are largely nondiscretionary. Starting in 1999, however, the PSID added additional questions about spending to obtain a more comprehensive measure of household consumption. These spending questions were further expanded in 2005.

The spending categories added in 1999 include households' health care expenses, mortgage or rent payments, housing insurance costs, home transportation expenses, child care expenses, schooling costs, recurring automobile costs, and utilities.[1] Transportation expenses include public transit, cabs, and other costs of getting from one place to another. Automobile costs include monthly loan or lease payments, vehicle maintenance costs, insurance costs, and down payment outlays on newly acquired vehicles. Charles et al. (2007) provide a detailed discussion regarding these additional spending data, and they find that the data line up well with the corresponding data from the CEX. In 2005 the PSID began asking households to report their spending on home maintenance and repairs, home furnishings, recreation expenditures, clothing, and vacations.[2] Section 2.3 discusses these data in more detail and compares them with the corresponding CEX data.

2.2 Consumer Expenditure Survey

The Bureau of Labor Statistics (BLS) uses the CEX to construct expenditure weights for the consumer price index, and thus the data provide very detailed information on the spending patterns of U.S. consumers. There are actually two distinct surveys: a "Diary" component

[1] Utility costs were previously collected in the survey but only sporadically.

[2] The home maintenance and repair data are different from the home improvement (investment) data available in the wealth supplements. In particular, the home maintenance data include ordinary repairs and general upkeep but not addition or remodeling costs.

that surveys consumers' daily spending habits over the course of two weeks, and an "Interview" survey that asks respondents to report their spending habits for the past three months. In the interview survey, households are followed for up to four consecutive quarters. Since the interview survey collects household spending data for a longer horizon than the diary survey, it is the part of the CEX that is used in this paper and others.

The CEX is considered by many to have the best available data on household-level expenditures. Winter (2004) shows that more detailed consumption questions yield more reliable consumption data. The main drawback of the CEX is that it is a short panel and has limited data on household wealth. There has also been discussion recently about how the CEX underestimates aggregate consumption relative to the NIPA data (see for example Gardner, McClelland, and Passero (2009) and Sabelhaus (2010)). This can be seen even with the per capita spending analysis in this paper. The current consensus is that the CEX underestimates aggregate consumption because it fails to capture the spending of households at the upper end of the wealth distribution. This survey issue is not a huge concern for the cross-sectional analysis in this paper, but could be important for some exercises, such as estimating Engle curves. There is a project underway at the Bureau of Labor Statistics, that is tasked with engaging practitioners and researchers to consider different options for improving the CEX survery. The paper by Sabelhaus (2010) is part of this effort.

2.3 PSID Expenditure Categories Post 2003

As noted earlier, Charles et al. (2007) performs a cross-sectional comparison of the additional PSID expenditure data from 1999 to 2003 with the relevant data in the CEX. The authors also create a mapping between the PSID expenditure categories and disaggregated CEX expenditure (UCC) codes (p. 34). Appendix Table A.1 shows a similar mapping of the additional PSID expenditure data added in 2005.[3] The mapping between the CEX and PSID categories for the post 2005 data is relatively straightforward. The one difficult category is "trips and vacations." Households in the PSID are asked to report the amount they spent on trips and vacations including expenditures on transportation. Separately, they are asked to report their transportation expenditures, as they have been since 1999, and there is no mention of including or excluding out-of-town trips. The CEX is careful to distinguish transportation costs "on trips" from other transportation spending. Charles et al. (2007) includes these "on trips" expenditure categories in the mapping for the various transportation categories from 1999 onward. Most fall under "other transportation." The mapping in Table 1 assumes that the "on trip" expenditures are included in the PSID transportation category and not in the vacation category, to be consistent with the mapping in Charles et al. (2007)

[3]The table also includes an updated mapping of the "utility" category, since some utility-related UCC codes were dropped in the CEX in 2005 Q2 and others were added to take their place.

and to avoid double counting.

The PSID acknowledges that the questions as currently worded could result in households including their "on trip" expenditures in both the transportation and vacation categories in 2005 and 2007. The PSID administration plans to correct the wording of the questions in future survey waves. Double counting the "on trip" data in the CEX does not substantially change the imputation results in attempts to replicate the reported PSID expenditure data post 2003.

Figure 1 depicts mean expenditures by age group for the spending categories added to the PSID in 2005 (red line) and compares them with the relevant mean expenditures by group in the CEX data (blue line). Reported housing maintenance expenditures are much higher in the PSID than the CEX. One reason for this could be that households in the PSID confuse home repair expenditures with spending on additions or remodeling (home improvement). The home improvement spending question is part of the wealth module, not the consumption module, so some households may lump the two categories together when answering the home repair question. In contrast, the CEX has very detailed categories for capturing households' home maintenance expenditures.

Households' home furnishing consumption is relatively similar in the two datasets, while spending on clothing and on vacations and recreation are somewhat higher in the PSID, on average, than in the CEX. The vacation and recreation data are combined, since the definitions of vacation and recreation expenditures in the two datasets contain potential overlaps. The CEX vacation data in Figure 1 include households' transportation expenditures "on trips." The CEX data are much lower than the PSID data without these items, suggesting that PSID households include travel expenses in their reported vacation expenditures. In addition, both the recreation/vacation and the clothing data diverge more in the datasets for older households. Such households tend to be wealthier, so perhaps a part of this difference results from the PSID doing a somewhat better job of capturing the expenditures of wealthier households.

3 PSID Consumption Imputation

3.1 Relating Food Expenditures to Nondurable Expenditures

One option to obtain a broader household expenditure measure in the PSID is to impute consumption, based on the observed relationship across household groups between food spending and the broader expenditure measure in another dataset such as the CEX. This approach was first used by Skinner (1987) and is often referred to as "matching" based on observed information. In particular, Skinner regressed total consumption on food consumption, utility

costs, and a series of demographic variables that exist in both the PSID and CEX. He then applied his estimated CEX relationships to the actual PSID food, utility, and demographic data to impute total household expenditures.

Blundell, Pistaferri, and Preston (2006) (BPP) also impute spending in the PSID using parameters estimated with the CEX. There are two key differences, however, between the BPP approach and that of Skinner. BPP estimate a food demand equation that they then invert to get a measure of households' nondurable spending. In addition, they allow households' budget elasticity to shift with observable characteristics, and they estimate the food demand relationship using an instrumental variable approach to reduce bias caused by measurement error in consumption. In particular, BPP use cohort-education-year specific averages of a husband and wife's log hourly wages and as instruments for log nondurable expenditures. BPP argue that their technique yields imputed nondurable consumption data that line up very well with trends in the CEX consumption distribution.[4]

BPP's food demand setup (based on their equation 1) is as follows:

$$ln(f_{i,x}) = D_{i,x}^{'}\beta + \gamma ln(c_{i,x}) + e_{i,x} \ , \tag{1}$$

where f is food expenditures (available in both surveys); D is a vector of variables containing prices and a set of household demographics (available in both datasets); c is total non-durable expenditures (available in the CEX); and e captures unobserved heterogeneity in the demand for food as well as any measurement error in food expenditures. BPP assume that food is a normal good ($\gamma > 0$). The subscript x identifies data from the CEX, while the subscript p signifies data from the PSID.

Nondurable consumption in the PSID ($c_{i,p}$) is calculated based on rearranging equation (1):

$$\hat{c}_{i,p} = exp\left(\frac{ln(f_{i,p}) - D_{i,p}^{'}\hat{\beta}}{\hat{\gamma}}\right) \ , \tag{2}$$

where $\hat{\beta}$ and $\hat{\gamma}$ are the coefficients obtained from estimating equation 1 with the CEX data. BPP's measure of nondurable expenditures includes food (home and away), alcohol, tobacco, utilities, transportation (including gasoline), personal care, clothing, and housing rents.

3.2 Using Income and Saving Data

Imputing total (or composite) consumption using households' income and saving data is based on re-writing their per period budget constraint. In particular,

[4]This paper uses the same IV approach.

$$s_t = y_t - c_t \tag{3}$$

$$\Rightarrow c_t = y_t - s_t \,, \tag{4}$$

where s_t is household saving, y_t is disposable household income, and c_t is household consumption.

Ziliak (1998) used this approach to impute "composite" consumption in the PSID from 1976 to 1986. Because of data limitations he constructed a measure of households' saving, using changes in their financial wealth.[5] His measure of saving therefore included capital gains (or passive saving), which biases downward the estimates of consumption since rising or falling asset prices do not directly affect households' cash flows.

Cooper (2009) uses the same general approach as Ziliak (1998), but uses a measure of household saving that *excludes* capital gains. In particular, he makes use of data in the PSID starting in 1989 that capture households' additions and subtractions from various asset categories since the previous wealth supplement. For example, households report the amount they contribute to 401k or IRA savings plans as well as the amount they withdraw from such plans. These data capture more accurately households' saving out of current income, which is the appropriate measure to use in equation (4). The appendix discusses the active saving data in more detail.

Given the timing of the wealth supplements, active saving data in the PSID are available for the following years: 1984 to 1989, 1989 to 1994, 1994 to 1999, 1999 to 2001, 2001 to 2003, 2003 to 2005, and 2005 to 2007. Given the timing of the data, a household's composite consumption is calculated over the same periods as follows:

$$c^i_{t,t-1} = (y^i_{t,t-1} - T^i_{t,t-1}) - s^i_{t,t-1} \,, \tag{5}$$

where $c^i_{t,t-1}$ is a household's consumption excluding housing between period $t-1$ and t, $s^i_{t,t-1}$ is household saving over the same period, $y^i_{t,t-1}$ is household income (excluding rental income), and $T^i_{t,t-1}$ is the household's federal and state income tax burden. Households' income tax burdens are estimated using the NBER's TAXSIM module, which takes a variety of inputs and returns an estimate of each household's federal and state taxes.[6]

The type of household expenditures measured by equation 5 is somewhere in between total consumption and nonhousing consumption as measured by the national accounts.[7] On

[5] Ziliak imputes households' financial wealth holdings using the approach in Zeldes (1989), since such data are unavailable in the PSID prior to 1984.

[6] The TAXSIM module applies stylized, but reasonably accurate, algorithms to reflect the personal income tax codes at the federal level and for each state.

[7] Technically, $c^i_{t,t-1}$ is not consumption but rather a measure of household expenditures, since it does

a cash-flow basis $c^i_{t,t-1}$ implicitly includes households' mortgage and/or rental payments. Rental payments are counted as consumption in the National Income and Product Accounts (NIPA) while mortgage payments are not. Instead the NIPA attempt to calculate owners' equivalent rents—or the implicit cost of owning a house including mortgage payments, maintenance, depreciation, and other factors. Owner-occupied rent costs rise as house prices increase, while nominal mortgage payments remained fixed. As a result, the consumption measure includes some of households' housing expenditures as measured by the NIPA but not all.[8]

An advantage of this saving and income imputation approach is that it only uses only in-sample information from the PSID, rather than inferring households' spending based on observed consumption relationships in external datasets such as the CEX. A drawback to both approaches is that they likely lead to measurement error. The imputed consumption data generated by the saving and income approach, however, are at least internally consistent with the reported saving and income data in the PSID.

3.2.1 Timing

The timing in equation 5 is somewhat unusual, and it is worth clarifying in greater detail. For example, consider the 1999 survey. That year, households reported their income for the previous year (1998) as well as their current stock of wealth. Since the PSID interviews usually occur in the first quarter of the survey year, I assume that households' wealth information is roughly equivalent to their end-of-year wealth in the previous year (1998). Households are asked to report their active saving between wealth supplements, so these data for 1999 roughly cover year-end 1993/beginning 1994 to year-end 1998. The consumption data estimated by equation (5) for this period cover the beginning of 1994 through the end of 1998. Income and tax data from 1994, 1995, 1996, 1997, and 1998 are used to keep all the timing consistent.[9]

The data timing for the other years' wealth follows the same pattern. For instance, consumption calculated based on the 2003 survey covers 2001 and 2002 and uses income and taxes from those years in the estimation. In addition, the 2003 survey records households' 2002 income, as always, but also includes "off-year" income data from 2001.[10] This off-year income question was not asked in 2005 or 2007, so there are no reported income data for

not include the service flow from household durables. This paper refers to this measure as consumption or expenditures interchangeably, however, for ease of discussion.

[8]Unfortunately, there is not enough information in the PSID to capture housing expenditures exactly as they are in the NIPA.

[9]The PCE price deflators used to convert estimated nominal expenditures into real expenditures are also timed accordingly.

[10]Households are asked for their overall family income two years prior to the survey year.

2003 or 2005. These data are imputed by growing out households' reported income in 2002 and 2004, respectively, using aggregate growth rates.

4 Sample Selection

4.1 BPP Imputation Approach

This paper follows BPP's procedure for selecting the CEX and PSID samples. This procedure is outlined in their Tables 1 and 2. The paper first replicates BPP's procedure over their sample horizon (1980–1992) and then extends their procedure through 2007.[11] The sample selection criteria, which are the same for both time horizons except for a few noted differences, are discussed below.

4.1.1 PSID Sample

PSID interviews prior to 1980 are dropped, since the CEX data begin in 1980. Households with a major change in family composition or a change in marital status are also eliminated along with female-headed households and households with missing education or location data.[12] Following BPP, households with heads born prior to 1920 or after 1959 are also dropped. Households with extreme income growth values are also eliminated—in particular those with family income growth above 500 percent, below -80 percent, and those with reported income below $100. Households with income that is less than their reported food expenditures are also removed from the sample along with households in the original low income subsample (also known as the SEO), which made up 39 percent of the original 1968 sample. Households that are part of the representative sample of Latino households added in 1990 are also excluded. Finally, the sample is restricted to households with heads between the ages of 30 and 65.

The main difference between the PSID sample in this paper and BPP's sample is the measure used to determine a household head and wife's educational attainment. In particular, BPP uses "grades of school finished" to compute education levels. This education variable is available for only some waves of the PSID survey; other waves have only categorical variables for a head and wife's education. This lack of data continuity is especially an issue when extending BPP's imputation method through 2007. This paper uses an education series that is consistently coded over time and that attempts to eliminate misreported

[11] 2007 is currently the last year of available PSID data.

[12] The PSID contains a variable on family composition change (V17710 in 1992), which identifies whether there is a change in the household head or wife. Households with such familial changes are the ones eliminated. In addition, the composition change variable is absent in the 1994 to 1999 waves, so the composition change restriction is not made in those years for the full sample results.

changes in individuals' educational status in the PSID. For example, an individual's education level cannot decline between consecutive survey waves. The sample statistics for these constructed education data line up well with the summary statistics for BPP's education data. Overall, the 1980–1992 PSID sample in this paper is somewhat larger than the one in BPP. This could be because some households in this paper's sample were dropped when the alternative education data were constructed.

There are two sample selection changes when the imputation technique is extended through 2007. First, the allowed range for a household head's year of birth shifts from 1920-to-1959 to 1920-to-1978. This allows younger households to enter the sample in the later years of the imputation. In addition, the household composition change variable is unavailable from 1994 to 1999. The family composition change restriction is therefore not applied to these years.

4.1.2 CEX Sample

This paper matches BPP's nondurable consumption measure by summing food at home, food away from home, alcoholic beverages, tobacco, personal services, personal care, heating fuel, transportation costs (including gasoline), clothing, and housing rents.[13] Households that have missing food data and/or have zero reported nondurable expenditures are dropped from the CEX sample. This paper also eliminates households that have incomplete income responses, zero before-tax income, reported income below the amount they spent on food, missing region or education records, and/or changes in family composition over the course of their four interviews.[14] Consistent with the PSID sample, households whose head was born before 1920 or after 1959 (after 1978 for the full sample) are dropped as are all households whose head is younger than 30 years or older than 65 years of age. Finally, the sample is restricted to households that were present for all four quarterly interviews, so that their annual expenditures could be calculated properly.

Households begin their quarterly interviews at any month during the year, so it is important to take this timing into account when calculating households' annual expenditures. If a household is interviewed for at least two quarters in a given year t, then the reference year for its consumption is t; otherwise the reference year for its spending is $t-1$. This timing convention is consistent with the approach in BPP. In addition, households' education categories are recoded relative to the BPP sample to make them compatible with the constructed education variable used in the PSID sample. In particular, BPP classifies a household head as a "high school dropout" or "high school graduate." This paper identifies the educational

[13]Personal services include babysitting and housekeeping services. Personal care includes laundry costs, personal care (grooming) appliances, and professional haircut and grooming services. This measure includes only tenant rent and not the CEX variable that represents owner's equivalent rent.

[14]Income data are measured as of the consumer unit's final interview.

achievement of individuals as less than high school, high school graduate, or at least some college.

Overall, the CEX sample in this paper has roughly 2000 fewer observations than BPP's sample, despite using very similar sample selection criteria. The years with substantially fewer observations than the rest are 1980 and 1985. The reason for the relatively small number of observations in 1980 is unclear. The reduced number of observations in 1985 is likely a result of the CEX sample redesign in that year. In particular, the CEX documentation cautions against linking these data to subsequent years' data because of the methodology change. As a result, this paper drops a number of households that started their interviews in late 1985 that may not have been dropped in the BPP sample.[15] When the imputation is extended through 2007, a similar change in methodology in 1995 results in fewer observations for 1995 than for other years.

4.1.3 Summary Statistics

Table 1 shows summary statistics of this paper's estimates for selected years of relevant variables that are common to the PSID and CEX, as well as the corresponding reported summary statistics from BPP's paper (their Table 3) for comparison purposes. Overall the data are relatively similar both between the PSID and CEX and between the samples in this paper and those in BPP's original paper. The regional distribution of households and the average age of households are similar between the between the CEX and PSID and also between the sample in this paper and in BPP. The education data are also comparable, especially considering the different approach to defining educational achievement in this paper versus BPP's approach. Annual food expenditures on average are lower in the PSID than in the CEX in 1980, but this relationship flips starting in 1986, and food spending in the PSID becomes somewhat higher than in the CEX. This pattern is consistent with BPP's findings in 1992, but not 1986.

The food expenditure means in this paper are slightly lower than in BPP. The average household family size and the number of children are also lower in the samples in this paper than in BPP, but the relative magnitudes of the PSID versus CEX data follow the same general pattern as in BPP's data. The differences between the sample means in this paper and those in BPP's original paper are likely due to the fact that the samples in the two papers are not exactly the same for the reasons discussed above. Overall, the summary statistics seem reasonable and confirm that the samples used for the imputation procedures are similar.

Table 2 reports summary statistics for the extended CEX and PSID samples for selected years through 2007. Again the PSID and CEX data line up reasonably well and similar

[15] BPP do not mention anything about the change in sample methodology.

patterns arise as with the earlier years shown in Table 1. In particular, average family size in the CEX continues to be generally higher than in the PSID, as is the percentage of households headed by a white person. The education data also continue to line up well, even with the recoding relative to BPP's appraoch. In addition, mean food expenditures in the PSID are generally somewhat higher than in the CEX, continuing the trend observed starting in 1986. The remaining statistics confirm that the two samples have a similar distribution of households.

4.2 Saving and Income Approach

The sample selection criteria for the saving and income approach follow that in Cooper (2009), and are unrestrictive relative to the criteria for the BPP approach. Nonhousing expenditures are calculated for everyone with available income and active saving data that are not topcoded. Since the income data are available at a higher frequency than the saving data, the yearly household income data are summed over the relevant years between wealth surveys. In addition, given the different time horizons between wealth supplements, the calculated consumption data are averaged to get yearly equivalent numbers.

Households that have negative imputed consumption are eliminated from the subsequent analysis. In addition, the analysis is run a second time with the sample restricted to those households that are in the PSID sample for the BPP imputation. This restriction tends to result in slightly higher imputed mean spending across households. Overall, the saving and income approach is somewhat less restrictive in the PSID than in the BPP approach, because in the former one does not have to worry about getting the CEX and PSID samples to line up as closely as possible.

5 Results

This section analyzes the different consumption imputation approaches. The average imputed expenditure data are compared with the relevant aggregate NIPA data on a per capita basis. The PSID and CEX data are divided by family size and then averaged across households, while the aggregate data are normalized by the estimated annual U.S. population. Except as noted, the data are deflated by the aggregate PCE deflator and are reported in terms of 2000 dollars. The results could have incorporated expenditure specific deflators; however, it seemed best to use the same deflator for all series for ease of comparison.

5.1 BPP Imputation Approach

Table 3 shows this paper's estimates of BPP's food demand equation through 1992. The elasticity of food expenditures with respect to nondurable consumption is about 1.02, compared with 0.85 in BPP's paper (their Table V). In other words, estimated food consumption increases by about 1 percent for every 1 percent increase in nondurable expenditures. The estimated year-consumption effects have the same signs as in BPP's results, but are quite a bit larger. The rest of the coefficient estimates often have similar signs as their counterparts in BPP's Table V, but the magnitudes are different. This divergence in the food demand estimates between the two papers is not surprising, given that the CEX sample size differs by about 2000 households. The results in this paper, however, seem reasonable. For example, households with more children have a more elastic food consumption response to changes in nondurables than households with fewer or no children. As nondurable expenditures increase, nondiscretionary expenditures take up a larger budget share for households who have more mouths to feed.

Figure 2 shows real per capita imputed nondurable expenditures in the PSID, using the estimates shown in Table 3. The imputed data are adjusted to account for differences in mean food consumption in the PSID versus the CEX, as in BPP.[16] Food data were not collected in the PSID in 1988 and 1989, so the imputation cannot be done for those years. Overall, the imputed nondurable PSID data correspond well with the equivalent CEX data.

Figure 3 plots the results for the extended imputation sample.[17] Once again, the CEX data and the imputed PSID data line up well, especially through the mid-1990s. Starting in 1994, however, imputed nondurable expenditures are a bit higher than the CEX data even after adjusting for differences in average food expenditures in the two datasets. This pattern of the imputed data being higher than the actual CEX data in the more recent years can also be seen in the results discussed below. BPP may have tailored their approach to matching the data through the mid-1990s; however, there is nothing obvious in their procedure that should not apply to the whole sample. As an alternative explanation, the average family size in the CEX is a bit higher than in the PSID later in the sample, which should mechanically should make per capita nondurable expenditures in the CEX lower than the imputed PSID data, all else being equal.

In addition, both the actual CEX data and the imputed data are a good deal below the equivalent NIPA per capita nondurable data.[18] The imputed PSID data are somewhat closer to the NIPA data than are the CEX data in recent years, given that the PSID and CEX data

[16] All of the imputed results are adjusted to account for differences in mean food expenditures between the two datasets.

[17] Table A.2 shows the results from estimating the food demand equation between 1980 and 2007.

[18] The NIPA series is constructed to match BPP's definition of nondurable expenditures, using data from the PCE underlying detail tables (2.4.5U).

diverge. One reason for this could be that the PSID does a better job of capturing actual household size than the CEX, especially recently.

The fact that the data are below the NIPA data overall is consistent with existing work that suggests the CEX data under-report household expenditures relative to the aggregate data (see, for example, Sabelhaus (2010)). Some of these differences are due to the expenditure categories not aligning exactly from a conceptual and/or definitional point of view. In addition, many of the NIPA data are imputed based on past trends and other information, which can yield somewhat different results than households' reporting their actual expenditures over the previous quarter. See Gardner, McClelland, and Passero (2009) for a further discussion of these issues.

Overall, BPP's nondurable imputation approach for the PSID is reasonable and does a very good job of transferring the average level of nondurables from the CEX to the PSID. The actual and imputed data diverge from the NIPA data, but this is somewhat expected, and the imputation approach enables researchers to capture a much broader measure of household expenditures over time than is possible with the food expenditure data currently available in the CEX.

5.2 Further Analysis of BPP's Approach

One way of checking the accuracy of the BPP imputation approach is to alter the food demand equation to substitute a composite expenditure measure that corresponds to the sum of the household expenditure data available in the PSID from 1999 onward for households' nondurable expenditures.[19] The imputed data from this procedure are compared with households' reported expenditures. This exercise is performed twice, once with a composite spending measure equal to the PSID data available consistently from 1999 to 2007, and a second time including the additional data available in 2005 and 2007.

Figures 4 and 5 compare the imputed results to the actual data. The bar charts show average real per capita spending. The average expenditure data recorded in the PSID and CEX for the relevant consumption categories between 1999 and 2007 are very similar in each year. This result is consistent with Charles et al. (2007), who find that the PSID expenditure questions added in 1999 do a good job matching the equivalent CEX expenditure categories. The imputed PSID data, however, are a bit higher than the actual PSID data. This finding is consistent with BPP's approach, which seems to overpredict PSID expenditures relative to what is recorded in the CEX for recent years, although overall the imputed data are reasonable.

[19]The composite measure includes the CEX categories that most closely match the PSID categories. This measure includes data on housing and car costs. Excluding these categories does not qualitatively impact the results.

In comparison, Figure 5 shows that the actual CEX data are a good bit lower than the actual PSID expenditure data available in 2005 and 2007. This may be because the vacation and recreation expenditure questions added to the PSID do not line up as well with the CEX data as some of the other categories. As noted earlier, there is a potential for households in the PSID to report their vacation travel expenses in 2005 and 2007 in both the transportation category and the vacation category.[20] The imputed data, however, are roughly in line with the actual PSID data, given the tendency of BPP's method to over-predict expenditures relative to the CEX.[21]

The results in Figures 4 and 5 suggest that BPP's method does a reasonably good job of capturing reported expenditures in the PSID. BPP's approach appears to be an accurate way to capture actual household expenditures in the PSID. Not surprisingly, however, it seems best to use the actual PSID expenditure data when they are available and are broad enough for the purposes of the intended analysis.

5.3 Saving-Income Approach

Figure 6 plots total real PCE per capita and total real PCE per capita excluding housing from the NIPA accounts. The figure also shows average total household expenditures from the CEX as well as imputed total expenditures in the PSID, using a modified version of BPP's food demand approach.[22] The results provide further evidence that the CEX data under-report household expenditures relative to the NIPA, and confirm that this under-reporting has worsened over time.

In contrast, households' imputed total expenditures based on the saving and income approach are much more in line with the NIPA data. Figure 7 shows these results. The NIPA and CEX data are averaged over the relevant years to make them compatible with averaged PSID data, given the timing of the PSID data discussed in Section 3. On average, total imputed expenditures lie between the two NIPA series, as predicted earlier. Indeed, if a researcher wishes to utilize a measure of total household expenditures in his or her PSID research that is comparable to the NIPA data, then the saving and income approach dominates BPP's method.

The saving and income imputation approach also dominates actual reported expenditures in the PSID from 1999 onward if one wants to capture households' total expenditures. The actual PSID data starting in 1999 are useful if a researcher wants to look at specific components of households' expenditures. In addition, there is potentially less measurement error

[20] Attempts to account for this double counting with the CEX data do not noticeably alter these findings.

[21] Both the 1999-to-2007 data and the 2005-to-2007 data are below the equivalent NIPA data (not shown). This is not surprising, given that the PSID expenditure questions are designed to match the CEX categories that tend to have lower reported expenditures than in the NIPA.

[22] The data include the CEX's measure of imputed housing rent for owner occupants.

in households' reported spending than in their imputed spending. Even though households' imputed spending is based on their budget constraint, it requires comparing two potentially noisy reported measures (saving and income) rather than using one potentially noisy data point.

The imputed data from the saving and income approach drop off a bit in 2005–2006 relative to trend. It is hard to know whether this is a problematic pattern or temporary noise without having the data yet to calculate 2007–2008 consumption. One potential explanation is that no off-year income data are available in the 2005 and 2007 surveys, as discussed earlier. As a result, the 2003–2004 and 2005–2006 consumption measures are estimated using imputed income for 2003 and 2005, respectively. Therefore, these estimates may lack precision compared with the earlier years' estimates when the off-year income data are reported in the survey. Overall, the saving-income (budget constraint) approach for imputing consumption seems to be fairly accurate for capturing total household expenditures in the PSID especially compared to the available alternatives.

6 Conclusion

This paper investigates different approaches for imputing a broader basket of household expenditures in the PSID than food consumption, which was the only consistent measure of household expenditures available in the PSID prior to 1999. The paper analyzed the approach of Blundell, Pistaferri, and Preston (2006) along with that of Cooper (2009). The former authors invert the calculated coefficients from food demand equations estimated using CEX data to obtain nondurable expenditures in the PSID. Cooper's approach is based on households' budget constraint and uses the available saving and income data in the PSID to impute households' total expenditures.

The paper replicates the approach in BPP and extends the imputed expenditure data through 2007. The results show that BPP's approach works well for imputing households' nondurable expenditures. The paper also confirms the validity of BPP's approach by using it to impute data for a basket of goods that are the same as those in the actual expenditure data available in the PSID starting in 1999. Indeed, the imputed and actual expenditure measures are closely aligned. The CEX data and PSID data, however, are a good bit lower than the equivalent per capita NIPA data. This is especially true when applying BPP's technique to impute total household expenditures in the PSID based on the total expenditure data available in the CEX.

The paper demonstrates that the imputation technique in Cooper (2009) does a very good job of replicating total household expenditures in the NIPA. The approach dominates that of BPP for analyzing such expenditures in the PSID. Given the lack of consumption

data in the PSID, however, either imputation technique discussed in this paper is useful for analyzing household spending behavior in the PSID especially prior to 1999.

This paper also shows that none of the imputation techniques is perfect. The perceived accuracy of the imputation approaches depends somewhat on what one believes is the appropriate spending benchmark for comparison purposes. The CEX under-reports expenditures relative to the NIPA, but this underreporting does not mean that the CEX data should be hastily dismissed as a valid benchmark for disaggregated household expenditure measures, especially given the proposed reasons for the CEX's shortcomings. More work needs to be done to improve the accuracy of imputed expenditures in the PSID, but, as this paper demonstrates, the two existing techniques are very reasonable given their goals.

References

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2006. "Imputing Consumption in the PSID Using Food Demand Estimates from the CEX." Mimeo.

Charles, Kerwin Kofi, Sheldon Danziger, Geng Li, and Robert F. Schoeni. 2007. "Studying Consumption with the Panel Study of Income Dynamics: Comparisons with the Consumer Expenditure Survey and an Application to the Intergenerational Transmission of Well-being." Finance and Economics Discussion Series 2007-16. Federal Reserve Board of Governors.

Cooper, Daniel. 2009. "Impending U.S. Spending Bust? The Role of Housing Wealth as Borrowing Collateral." Public Policy Discussion Paper Series 09-9. Federal Reserve Bank of Boston.

Gardner, Thesia I., Robert McClelland, and William Passero. 2009. "Strengths and Weaknesses of the Consumer Expenditure Survey from a BLS Perspective." Paper presented at the NBER Summer Research Institute Conference on Research in Income and Wealth joint with Aggregate Implications of Microeconomic Consumption Behavior Workshop. Cambridge, MA: National Bureau of Economic Research. Available at http://www.nber.org/confer/2009/SI2009/PRCR/Garner_McClelland_Passero.pdf.

Sabelhaus, John. 2010. "The Joint Distribution of Consumption and Income." Mimeo University of Maryland.

Skinner, Jonathan. 1987. "A Superior Measure of Consumption from the Panel Study of Income Dynamics." *Economic Letters* 23: 213–16.

Winter, Joachim. 2004. "Response Bias in Survey-Based Measures of Household Consumption." *Economics Bulletin* 3: 1–12.

Zeldes, Stephen. 1989. "Consumption and Liquidity Constraints: An Empricial Investigation." *Journal of Political Economy* 97(2): 305–346.

Ziliak, James. 1998. "Does the Choice of Consumption Matter: An Application to the Permanent Income Hypothesis." *Journal of Monetary Economics* 41: 201–16.

Table 1

Comparison of Sample Means, PSID and CEX Compared to BPP Sample for Selected Years

Year	1980				1986				1992			
Survey	PSID	BPP-PSID	CEX	BPP-CEX	PSID	BPP-PSID	CEX	BPP-CEX	PSID	BPP-PSID	CEX	BPP-CEX
Age	42	43	43	44	43	44	46	46	45	43	46	47
Family size	3.54	3.61	3.65	3.98	3.30	3.48	3.30	3.60	3.24	3.42	3.22	3.55
# Children	1.33	1.31	1.37	1.49	1.15	1.21	1.12	1.17	1.10	1.14	0.99	1.15
White (%)	91	91	89	89	84	92	86	89	84	90	85	88
HS dropout (%)	21	21	21	20	17	16	20	14	15	13	17	15
HS grad (%)	32	30	32	33	31	32	29	30	31	39	30	30
College (%)	47	49	46	47	52	53	51	56	54	56	53	55
Northeast (%)	20	21	23	20	20	22	21	23	20	22	21	22
Midwest (%)	32	33	28	28	30	30	26	28	30	30	28	29
South (%)	30	31	28	28	31	30	27	27	31	30	26	26
West (%)	16	15	21	24	17	18	26	23	18	18	25	23
Food Expn ($)	4019	4449	4403	4656	4893	5306	4783	6135	6134	6620	5845	6431

Source: Author's estimates and Blundell, Pistaferri, and Preston (2006), Table 3. The first column of data for each year shows summary statistics for the PSID based on estimates in this paper. The second column shows the equivalent summary statistics for the PSID sample in Blundell, Pistaferri, and Preston (2006). The third column shows summary statistics for the CEX based on this paper's sample, and the fourth column shows the CEX summary statistics from the sample in Blundell, Pistaferri, and Preston (2006). Percentages may not add due to rounding.

Table 2

Comparison of Sample Means, PSID and CEX (Selected Years)

Year	1996		1999		2001		2003		2005		2007	
Survey	PSID	CEX	PSID	CEX	PSID	CEX	PSID	CEX	PSID	CEX	PSID	CEX
Age	44	44	45	45	46	45	46	46	46	47	47	47
Family size	3.17	3.26	3.17	3.28	3.11	3.26	3.05	3.20	3.01	3.17	3.02	3.22
# Children	1.07	1.09	1.05	1.06	0.97	1.01	0.92	0.99	0.89	0.97	0.92	1.00
White (%)	77	85	80	88	81	86	79	86	79	85	82	86
HS dropout (%)	13	14	12	12	11	11	10	12	9	11	9	12
HS grad (%)	31	29	30	26	30	27	31	27	31	23	31	23
College (%)	55	57	58	62	58	62	58	61	59	66	60	65
Northeast (%)	19	16	19	18	18	16	18	16	18	18	17	19
Midwest (%)	31	26	31	24	31	24	30	26	30	24	31	23
South (%)	32	33	31	32	31	32	33	34	33	34	32	36
West (%)	17	25	18	26	19	27	18	23	19	24	19	23
Food Expn ($)	6397	6153	7449	6626	7712	6898	7758	6912	8144	7668	8619	8810

Source: Author's estimates.

Table 3

Food Demand Equation Results (1980 to 1992)

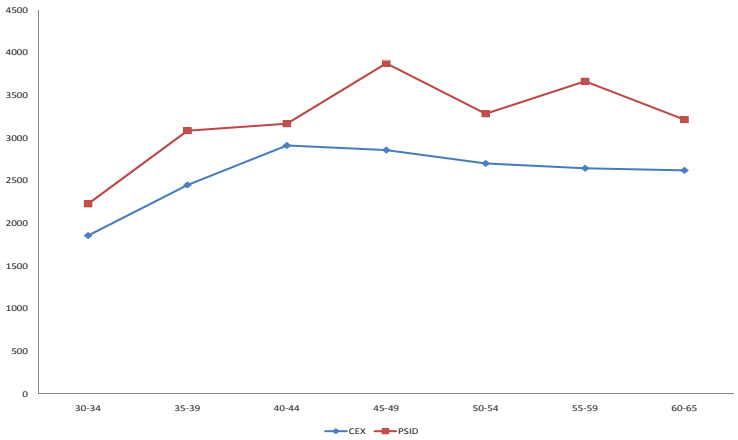
Variable	Estimate	Variable	Estimate
ln c	1.024***	ln p_{food}	7.639
	(0.097)		(10.035)
ln c*1981	0.034	ln p_{fuel}	−0.440
	(0.056)		(6.324)
ln c*1982	0.049	ln $p_{alcohol+tobacco}$	−8.476
	(0.079)		(8.163)
ln c*1983	0.072	ln $p_{transports}$	−2.273
	(0.094)		(11.845)
ln c*1984	0.075	White	0.073***
	(0.110)		(0.009)
ln c*1985	0.092	Family Size	−0.015**
	(0.111)		(0.006)
ln c*1986	0.097	Born 1925-29	0.040*
	(0.090)		(0.016)
ln c*1987	0.094	Born 1930-34	0.059**
	(0.088)		(0.023)
ln c*1988	0.105	Born 1935-39	0.068**
	(0.088)		(0.031)
ln c*1989	0.108	Born 1940-44	0.076*
	(0.098)		(0.039)
ln c*1990	0.116	Born 1945-49	0.060*
	(0.110)		(0.046)
ln c*1991	0.189*	Born 1950-54	0.053
	(0.102)		(0.053)
ln c*1992	0.210*	Born 1955-59	0.041
	(0.110)		(0.062)
ln c*HS Graduate (head)	0.116**	Age	−0.033***
	(0.059)		(0.005)
ln c*Some College (head)	0.100*	Age^2	0.000***
	(0.053)		(0.000)
ln c*One Child	0.191***	Northeast	−0.021**
	(0.042)		(0.008)
ln c*Two Children	0.221***	Midwest	0.005
	(0.054)		(0.008)
ln c*Three or More Children	0.095*	South	−0.020**
	(0.054)		(0.008)
One Child	−1.763***	HS Graduate (head)	−1.088**
	(0.392)		(0.540)
Two Children	−1.981***	Some College (head)	−0.989**
	(0.507)		(0.493)
Three or More Children	−0.753	Constant	15.234
	(0.530)		(12.666)
N	12990		

Source: Author's estimates.

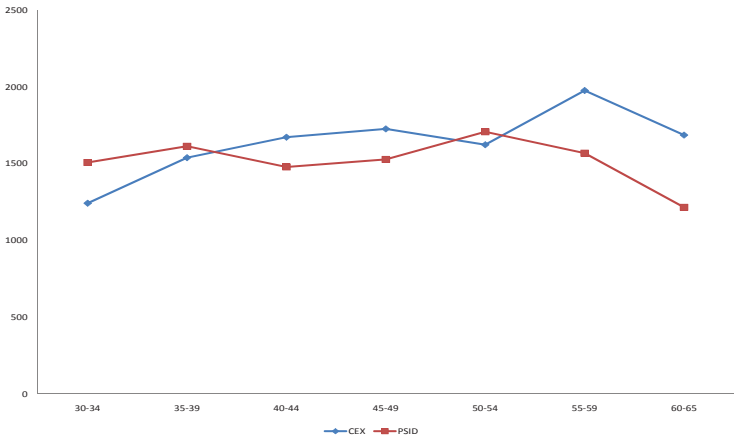
Figure 1

PSID and CEX Expenditures (2000 Dollars)

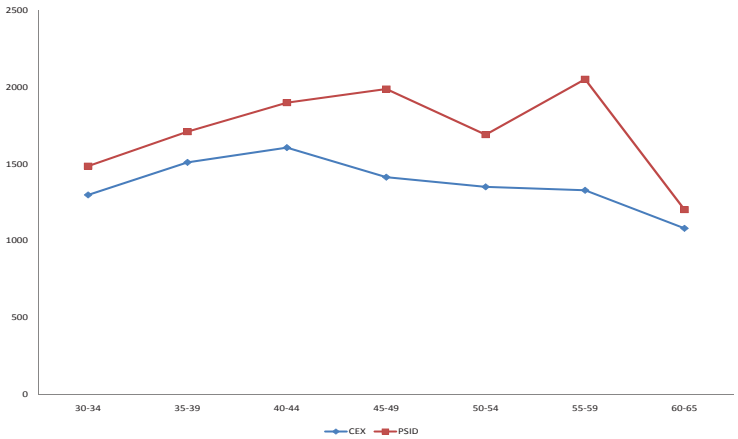
Vacation & Recreation



Household Furnishings



Clothing



Household Repair

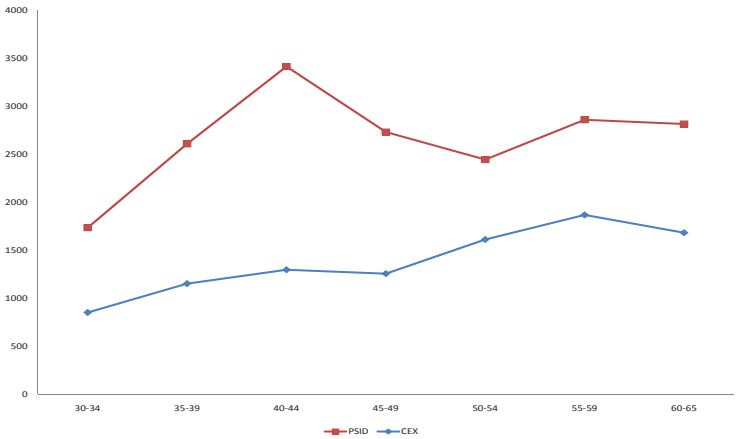


Figure 2

Nondurable Expenditures (1980 to 1992)

2000 Dollars

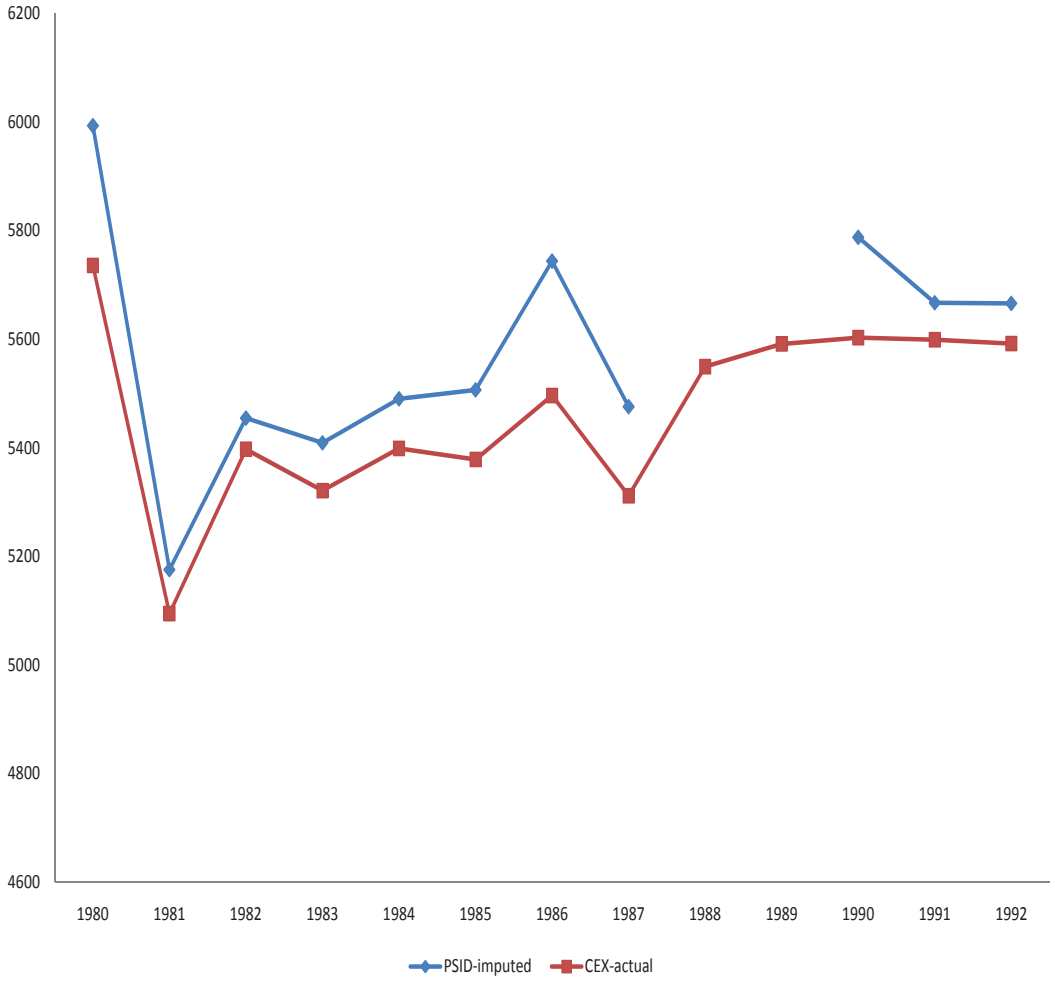


Figure 3

Nondurable Expenditures (1980 to 2007)
2000 Dollars

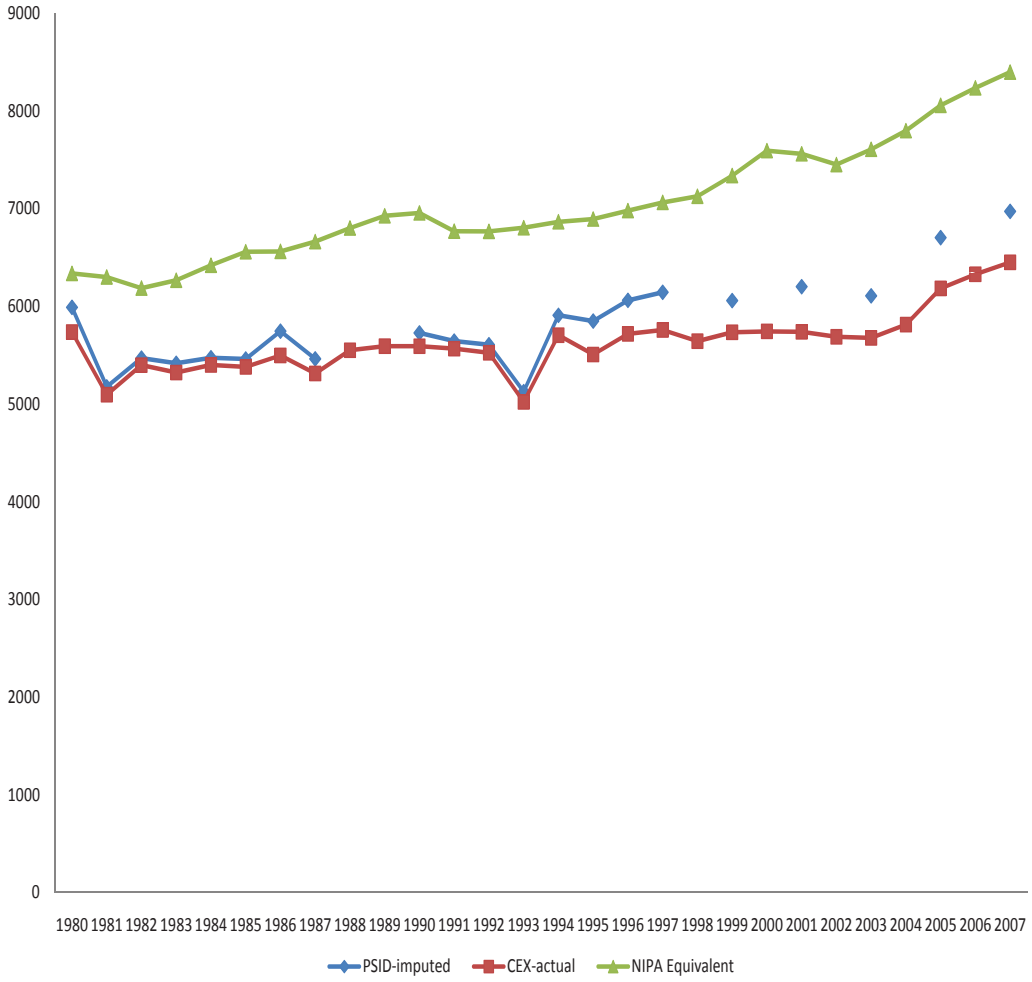


Figure 4

PSID Composite Expenditures (1999 to 2007)

2000 Dollars

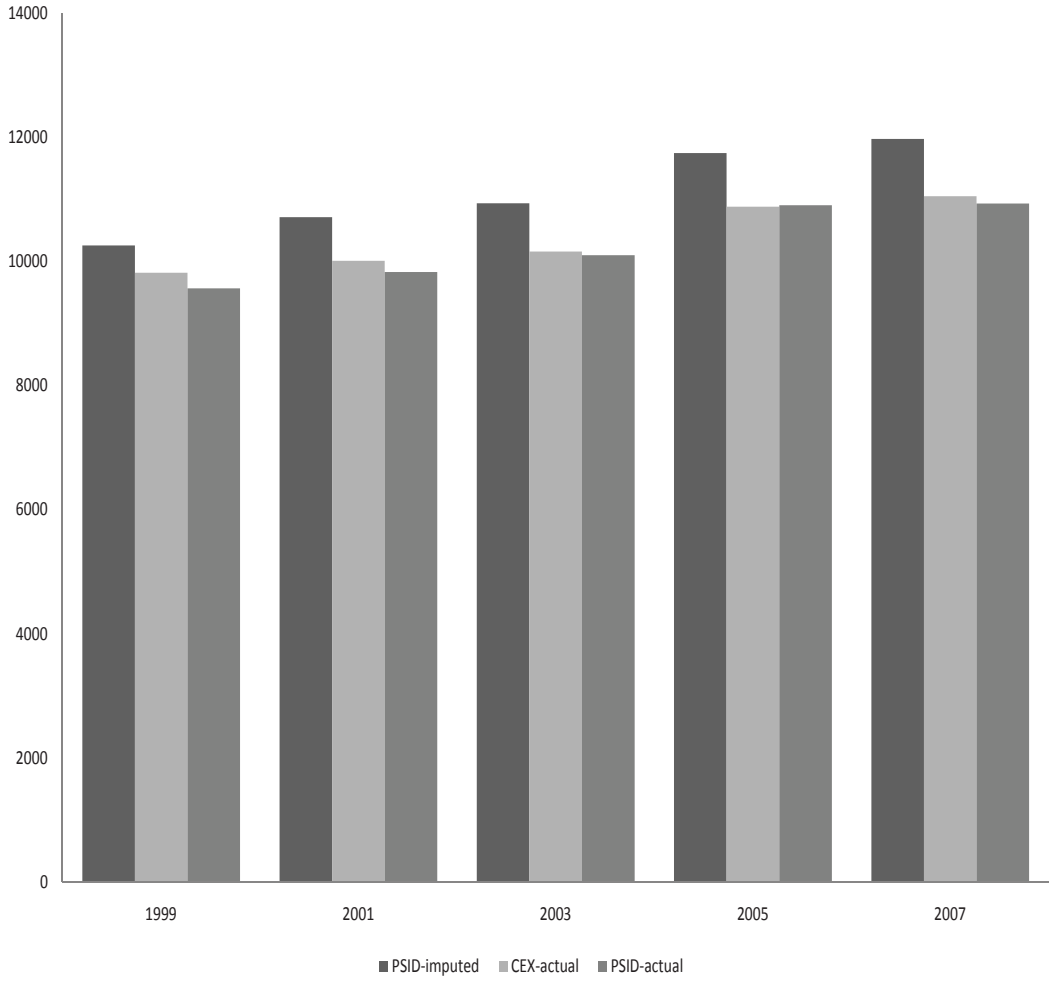


Figure 5

PSID Composite Expenditures Categories (2005 to 2007)
2000 Dollars

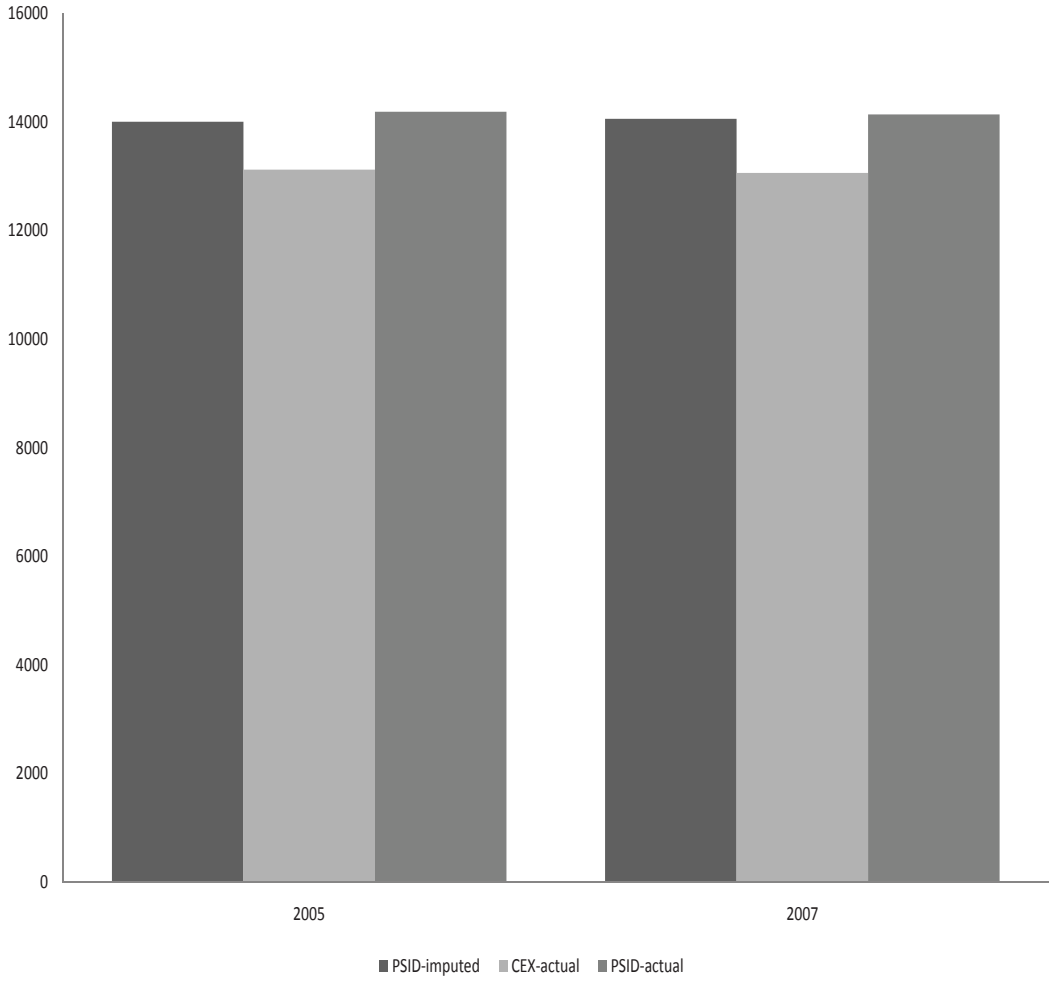


Figure 6

Total Household Expenditures (1980 to 2007)
2000 Dollars

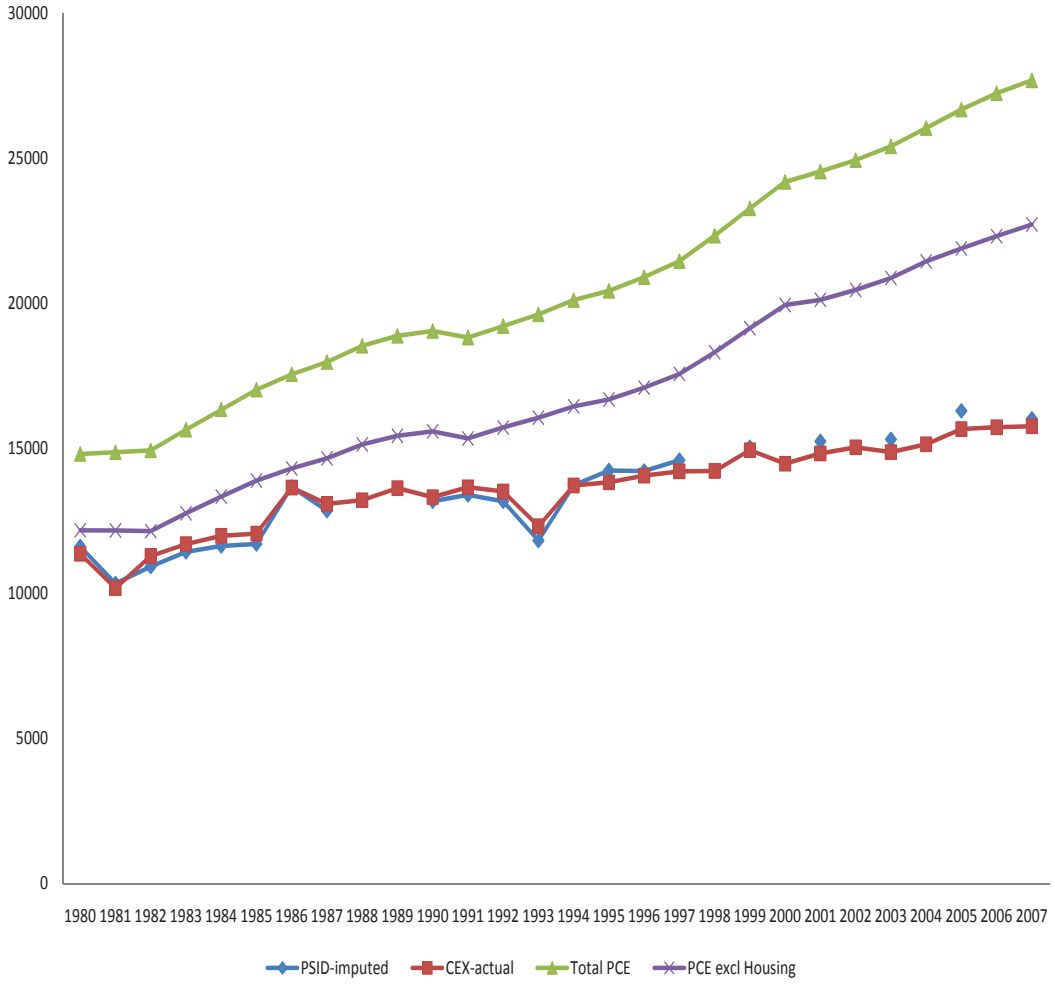
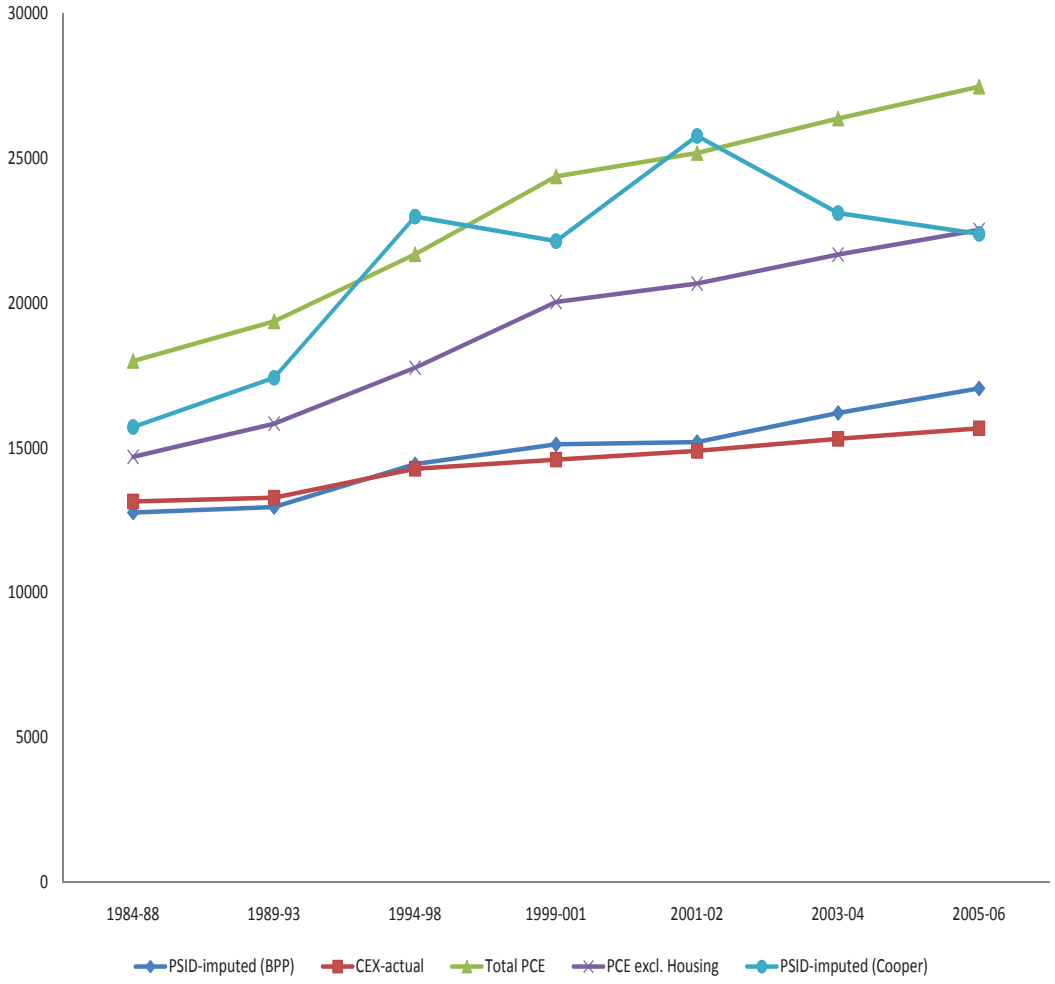


Figure 7

Total Household Expenditures
2000 Dollars



7 Appendix

7.1 Detailed Active Saving Calculations

The PSID tracks households' active saving in seven categories in addition to their 401k/IRA saving mentioned in the text. These other active saving categories include: investment in businesses or farms, checking and saving accounts, bond holdings, stock holdings, housing, other real estate, vehicles, and noncollateralized debt (NCD).[23]

The method for calculating households' active saving in the PSID depends on the asset in question. In particular, active saving for assets with potentially large capital gain components, such as stocks, IRA accounts or annuities, other real estate, and investment in businesses or farms is defined as follows:

$$as^{i,j}_{t-1,t} = I^{i,j}_{t-1,t} - R^{i,j}_{t-1,t} \,, \tag{6}$$

where $as^{i,j}_{t-1,t}$ is active saving for household i in asset j, $I^{i,j}_{t-1,t}$ is the amount invested by household i in asset j between $t-1$ and t, and $R^{i,j}_{t-1,t}$ is the amount withdrawn from asset j by household i over that same period.

For asset categories where capital gains are not a factor, active saving is simply the difference between a household's reported asset value in period t and its value in period $t-1$. These assets include: households' checking and saving account holdings, bond holdings, vehicle values, and noncollateralized debt. In particular,

$$as^{i,j}_{t-1,t} = V^{i,j}_{t} - V^{i,j}_{t-1} \,, \tag{7}$$

where V^{j}_{t} is the value of asset j at time t. The remaining active saving category is housing ($j = h$). The actual calculation of households' saving in housing depends on whether or not a household moves. Households that do not move "save" by paying down their mortgage principal, while households that move may potentially save or dis-save by altering the amount of equity in their homes. In particular,

$$as^{i,h}_{k-1,k} = \begin{cases} D^{i,h}_{k-1} - D^{i,h}_{k} & \text{if move } = 0 \\ E^{i,h}_{k} - E^{i,h}_{k-1} & \text{if move } = 1 \,, \end{cases} \tag{8}$$

where $D^{i,j}_{k}$ is a household's amount of outstanding mortgage debt in period k, $E^{i,j}_{k}$ is the amount of equity a household has in its home at time k, and *move* is an indicator variable that equals 1 if a household moved between $k-1$ and k, and is 0 otherwise. I use k as the time subscript to represent the fact that the time horizon for active saving in housing is different than for the other assets. For years prior to 1999, housing data are available yearly, and the difference between k and $k-1$ represents one year, while $t-1$ to t covers five years. After 1999, the housing and active saving data cover two-year horizons and $t = k$. More formally:

$$as^{i,h}_{t-1,t} = \begin{cases} \sum_{k=t-1}^{t} as^{i,h}_{k,k+1} & t \leq 1999 \\ as^{i,h}_{k,k+1} & t > 1999 \,. \end{cases} \tag{9}$$

[23] Other real estate includes vacation homes, rental properties, and land holdings. NCD includes credit card debt as well as student loans and other unsecured debt.

Yearly active saving in housing prior to 1999 is added together so it covers the same time horizon as the other active saving measures.

Total active saving for a given household is simply the sum of its saving in the individual asset components.

$$as^{i}_{t-1,t} = \sum_{j} as^{i,j}_{t-1,t} \, . \tag{10}$$

Table A.1

UCC Code Mapping for PSID Categories Post-2005

PSID Consumption Category	CE UCC Code
Clothing	360110 - 370902, 370904 - 390322, 390902-430120, 640130
Trips & Vacations	470113, 470212, 520212, 520522, 520532, 520542, 520905-530210, 530312, 530411, 530510, 530901, 610900, 620122, 620212, 620222, 620903, 620909, 620919, 690116, 810400
Other Recreation	310240, 310340-310350, 590111-590410, 600210-610320 620111, 620121, 620211, 620221, 620310, 620330, 620904-620908, 620912, 620921-620930
Household Furnishings & Equipment	220612, 220615, 220616, 230133, 230134, 280110, 280120-310230 310311-310334, 320111-320522, 320633-320904, 340902,340904, 340905,340907,990900 , 230117, 230118, 790611
Home Repair & Maintenance	230112-230115, 230121, 230122, 230123, 230150-230142, 240111-240323, 270211-270214, 270901-270904, 320611-320633, 330511, 340620 340630, 340901, 340903, 340914, 790600, 990930, 990940
Utility (as of 2005 Q2)	250111, 250112, 250113, 250114, 250911-250914, 260111, 260112, 260113, 260114, 260211, 260212, 260213, 260214, 270211, 270212, 270213, 270214, 270310, 270411, 270412, 270413, 270414, 270901, 270902, 270903, 270904

Table A.2

Food Demand Equation Estimates (1980-2007)

Variable	Estimate	Variable	Estimate	Variable	Estimate
ln c	1.134***	ln c*2004	0.081	Age	−0.016***
	(0.048)		(0.068)		(0.003)
ln c*1981	−0.028	ln c*2005	0.063	Age^2	0.000***
	(0.030)		(0.076)		(0.000)
ln c*1982	0.006	ln c*2006	0.080	ln p_{food}	1.837
	(0.031)		(0.083)		(7.937)
ln c*1983	0.038	ln c*2007	0.093	ln p_{fuel}	−6.979**
	(0.038)		(0.091)		(2.799)
ln c*1984	0.018	ln c*2008	0.093	ln $p_{alcohol+tobacco}$	−8.388
	(0.042)		(0.093)		(6.954)
ln c*1985	0.014	ln c*HS Graduate (head)	0.099**	ln $p_{transports}$	14.264**
	(0.045)		(0.048)		(7.256)
ln c*1986	0.093**	ln c*Some College (head)	0.086**	HS Graduate (head)	−0.950**
	(0.041)		(0.040)		(0.450)
ln c*1987	0.055*	White	0.090***	Some College (head)	−0.865**
	(0.030)		(0.006)		(0.372)
ln c*1988	0.046	ln c*One Child	0.096***	Family Size	−0.010**
	(0.031)		(0.027)		(0.005)
ln c*1989	0.020	ln c*Two Children	0.082**	Northeast	−0.020***
	(0.037)		(0.033)		(0.005)
ln c*1990	−0.009	ln c*Three$^+$ Children	0.090**	Midwest	0.029***
	(0.048)		(0.037)		(0.006)
ln c*1991	0.055	One Child	−0.918***	South	0.041***
	(0.045)		(0.260)		(0.006)
ln c*1992	0.061	Two Children	−0.732**	Constant	−5.424
	(0.054)		(0.318)		(5.798)
ln c*1993	0.048	Three$^+$ Children	−0.784**		
	(0.054)		(0.358)		
ln c*1994	0.012	Born 1925-29	0.012		
	(0.054)		(0.015)		
ln c*1995	−0.024	Born 1930-34	0.017		
	(0.060)		(0.017)		
ln c*1996	−0.027	Born 1935-39	0.012		
	(0.062)		(0.019)		
ln c*1997	−0.003	Born 1940-44	0.013		
	(0.058)		(0.022)		
ln c*1998	0.027	Born 1945-49	0.004		
	(0.053)		(0.026)		
ln c*1999	0.014	Born 1950-54	0.011		
	(0.056)		(0.029)		
ln c*2000	−0.007	Born 1955-59	−0.003		
	(0.065)		(0.033)		
ln c*2001	0.063	Born 1960-64	−0.012		
	(0.062)		(0.036)		
ln c*2002	0.063	Born 1965-69	−0.032		
	(0.061)		(0.040)		
ln c*2003	0.083	Born 1970-78	−0.039		
	(0.065)		(0.045)	N	36547

Source: Author's estimates.