

Has Consumption Inequality Mirrored Income Inequality?*

Mark Aguiar

Mark Bilis

May 10, 2012

Abstract

We revisit to what extent the increase in income inequality over the last 30 years has been mirrored by consumption inequality. We do so by constructing an alternative measure of consumption expenditure, using data from the Consumer Expenditure Survey (CE), that employs a demand system to correct for systematic measurement error. Specifically, we consider trends in the relative expenditure of high-income and low-income households for different goods with different expenditure elasticities. Our estimation exploits the difference in the growth rate of luxury consumption inequality versus necessity consumption inequality. This “double-differencing,” which we implement in a regression framework, corrects for mis-measurement that can systematically vary over time by good and income group. Our results show that consumption inequality has tracked income inequality much more closely than estimated by direct responses on expenditures.

*Princeton University and NBER, University of Rochester and NBER. Email: mark@markaguiar.com and mark.bilis@gmail.com. We thank Yu Liu for outstanding research assistance.

1 Introduction

We revisit the issue of whether the increase in income inequality over the last 30 years has translated into a quantitatively similar increase in consumption inequality. Contrary to several influential studies discussed below, we find that consumption inequality has tracked income inequality. Like most of the previous literature that argues the opposite, we base our conclusions on the Consumer Expenditure Survey's (CE) interview survey. But rather than measure consumption inequality directly by summing household expenditures, we base our measure of consumption inequality on how richer versus poorer households allocate spending across goods. In particular, we estimate relative consumption growth across income groups by observing how households in these groups have shifted their expenditures toward luxuries versus necessities over time. We show our approach is robust to systematic trends in measurement error that may bias measures based on summing household spending. We find a substantial increase in consumption inequality, similar in magnitude to the increase in income inequality.

An influential paper by [Krueger and Perri \(2006\)](#), building on related work by [Slesnick \(2001\)](#), uses the CE to argue that consumption inequality has not kept pace with income inequality.¹ In an exercise comparable to Krueger and Perri's, we show that both relative before and after-tax income inequality increased by about 33 percent (.33 log points) between 1980 and 2010, where our conservative measure of income inequality is the ratio of those in the 80-95th percentiles to those in the 5-20th percentiles. Based on relative household expenditures, the corresponding increase in consumption inequality for the same two groups is only 11 percent.²

A concern with the CE evidence is the well-documented decline in aggregate consumption reported in the CE relative to NIPA personal consumption expenditures (e.g., [Garner et al., 2006](#).) Aggregate expenditures reported by CE households for 1980, excluding health care, equaled 87 percent of that implied by NIPA. By 2010 this ratio fell to only 63 percent.³ This does not necessarily imply that the CE fails to capture trends in consumption inequality. If the CE's under-reporting is uniform across income groups, then the mis-measurement will

¹For other contributions to this literature, see [Blundell and Preston \(1998\)](#), [Blundell et al. \(2008\)](#), and [Heathcote et al. \(2010\)](#).

²For the period 1980-2004, [Krueger and Perri \(2006\)](#) report a log change in the 90/10 income ratio of approximately 0.36 for income, and 0.16 for consumption.

³We exclude medical expenses from this calculation as the CE only reports a households' insurance premiums and other out-of-pocket expenditures, omitting health care expenses paid by other parties. If health care expenditure is included, the ratio of CE to NIPA expenditures declines by 26 percentage points from 82 to 56 percent.

not bias ratio-based measures of consumption inequality. However, as we illustrate below, that scenario implies extreme shifts in relative saving rates from 1980 to 2010. In particular, the implied savings rate for low-income households must plummet from -23 to -59 percent of income. We document that the savings rates implied by reported expenditure (i.e., income minus expenditure) are inconsistent with the savings data households directly report in the CE; that is, the budget constraint does not hold. The failure of this consistency check motivates the need for an alternative measure of consumption inequality in the CE.

We measure consumption inequality based on how high- versus low-income households allocate spending toward luxuries versus necessities. Intuitively, if consumption inequality is increasing substantially over time, then higher income households will shift consumption toward luxuries more dramatically than lower income households. The key advantage of this approach is that it does not require that the overall expenditures of households be well measured. Starting from consistent estimates of a demand system (Engel curves), the ratio of spending across any two goods with different expenditure elasticities identifies the household's total expenditure. This estimate is clearly robust to household-specific multiplicative measurement error, since the ratio of expenditures will be unaffected. Inequality in consumption across income groups is then estimated by comparing their respective ratios. This estimate of inequality is robust not only to household-specific measurement errors (e.g., more severe underreporting by richer households), but also to good-specific measurement errors (more severe underreporting for some goods than others). Good-specific measurement errors are eliminated once differences are taken across households.

To illustrate, take expenditures on nondurable entertainment (a luxury) versus food at home (a necessity). The top income quintile in the CE increased reported spending on entertainment by 23 percent relative to that for food at home between 1980 to 2010. Based on our estimated Engel elasticities, this implies an increase in total expenditure of 15 percent (see figure 3). By contrast, the bottom income quintile reported that entertainment expenditures declined by 43 percent relative to that reported for food at home, suggesting a decline in total expenditure of 28 percent. Both these calculations are robust to income-specific measurement error in the CE, even if the error changes over time. But, if the CE captures less of actual entertainment expenditures over time, relative to food at home, then both these growth rates are biased downward. Log differencing the two rates eliminates that bias, implying an increase in inequality of 43 log points.

While food and entertainment are interesting due to their extreme expenditure elasticities, a major advantage of the CE data is that it offers detailed expenditures across nearly all categories of goods. We therefore implement this Engel curve approach using all goods

in a regression framework to exploit this richness of the CE. Our estimates suggest that consumption inequality increased by close to 30 percent between 1980 and 2010, nearly as much as the change in income inequality, and nearly three times that estimated based on directly examining relative household expenditures in the CE. We find this estimate is stable across different subsets of goods, different weighting schemes across goods, and alternative first-stage elasticity estimates. The results imply a substantial trend in income-specific mis-measurement in the CE. Specifically, the estimation implies that relative under-measurement of high-income expenditure is growing over time, with an increase of about 20 log points over the entire sample.

We also consider trends in inequality in different sub-periods. We find that after-tax income inequality increased by 20 percent between 1980 and the early-1990s, by an additional 13 percent between 1993 and 2007, then remained stable through the great recession. The inequality in reported CE expenditures increased by only 11 percent in the first sub-period, by 6 percent from 1993 to 2007, then actually reversed (falling) by 6 percent from 2007 to 2010. So reported consumption inequality fails to keep pace with income inequality in any of the three sub-periods. Using our demand system estimates, we find that consumption inequality increased by more than 20 percent between 1980 and the early-1990s, by an additional 13 percent through 2007, for a total increase of over 30 percent, closely tracking the profile of income inequality. For the great recession we estimate a small reduction in consumption inequality, intermediate to that seen in income inequality (no change) and the larger decline in inequality in reported expenditures.

We are not the first to reassess trends in consumption inequality, particularly with a focus on mis-measurement of CE interview expenditures. [Battistin \(2003\)](#) and [Attanasio et al. \(2007\)](#) use the diary component of the CE to correct for mis-measurement in the interview survey. They estimate that the interview survey underestimates the rise in consumption inequality significantly in the 1990s. Related, [Browning and Crossley \(2009\)](#) argue that multiple noisy measures can dominate a single, relatively accurate measure, building on the insight that the covariance of multiple measures may mitigate measurement error. Our approach shares a similar spirit, but exploits differencing across goods within a demand system rather than extracting a common source of variation from covariances. Our paper is also complementary to [Parker et al. \(2009\)](#), who focus on the gap between CE expenditures and those reported by NIPA to obtain a corrected estimate of consumption inequality. Most recently, [Attanasio et al. \(2012\)](#) document that the substantial increases in consumption inequality we report are consistent with other estimates of consumption inequality, including those derived from expenditures in the Panel Study of Income Dynamics, the CE diary survey,

and reported vehicle expenditures.

There is a large literature concerning consumption inequality that precedes or is not focused on the issues raised by Slesnick and Krueger and Perri. An important paper by [Attanasio and Davis \(1996\)](#) documents that the increase in the college premium observed for wages in the 1980s is mirrored by similar increases in consumption inequality. However, [Attanasio and Davis \(1996\)](#) do not address the relative trends within education groups, which is where [Krueger and Perri \(2006\)](#) show the conflict between income and consumption inequality trends is starkest. Other important papers in this earlier literature include [Cutler and Katz \(1992\)](#), [Johnson and Shipp \(1995\)](#), and [Blundell and Preston \(1998\)](#). [Sabelhaus and Groen \(2000\)](#) also discuss mis-measurement in the context of the relationship of consumption and income. For trends in inequality for a number of countries and time periods, see the papers collected in [Krueger et al. \(2010\)](#). There is also a large literature on consumption versus income inequality over the life cycle, starting with [Deaton and Paxson \(1994\)](#).⁴ These papers often use the CE for consumption data, and are therefore subject to the measurement error problems addressed in this paper. We leave the question of whether our approach has implications for trends in life cycle inequality to future research.

The remainder of the paper is organized as follows. Section 2 describes the data, documents trends in income and expenditure inequality, and analyzes the CE's savings data; section 3 performs our demand-system analysis; section 4 examines robustness to potential mis-specification, especially with respect to our Engel curve estimates; and section 5 concludes.

2 Data Description and Inequality Trends

In this section we describe our data set and document trends in income and consumption inequality. The data appendix contains a more detailed discussion of variable construction and our sample.

2.1 Data

Our data are from the Consumer Expenditure Survey's interview sample. This is a well known consumption survey that has been conducted continuously since 1980. We include

⁴See also, [Storesletten et al. \(2004\)](#), [Heathcote et al. \(2005\)](#), [Guvonen \(2007\)](#), [Huggett et al. \(2009\)](#), and [Aguiar and Hurst \(2009\)](#).

waves starting in 1980 and extending through 2010. The survey is large, consisting of over 5,000 households in most waves. Each household is assigned a “replicate” weight designed to map the CE sample into the national population, which we use in all calculations. Each household is interviewed about their expenditures for up to four consecutive quarters. Each interview records expenditures on detailed categories over the preceding three months. The final interview records information on earnings, income, and taxes from the preceding 12 months, aligning with the period captured for expenditures. Income, expenditure, and savings variables are all recorded at the household level. However, when estimating household demand equations we control for demographic dummy variables that reflect the number of household members, number of household earners, and the reference member’s age.

The CE reports expenditure on hundreds of separate items. We aggregate these into 20 groups, which are listed in table 2. The division of expenditures into groups is governed by several criteria. The first is to respect BLS categorization of similar goods. The second is to define groups broadly enough to ensure consistency across the various waves of the survey. The third is to define groups narrowly enough that they span a wide range of expenditure elasticities. We adhere to the groupings created by the BLS in published statistics with minor exceptions. For instance, we group telephone equipment and services with appliances, computers, and related services rather than with utilities, based on priors regarding expenditure elasticities.

For expenditure on housing services, we use rent paid for renters and self-reported rental equivalence for home owners. For surveys conducted in 1980 and 1981 households were not asked about rental equivalence. We impute the rental equivalence for homeowners in these early waves as discussed in the appendix. For durables other than housing we use direct expenditure, and do not impute service flows. This is motivated by our use of income groups as the unit of analysis (described below), and the assumption that aggregating over many households provides a good proxy for the consumption of durable services at a point in time. We show in section 3 that our estimates are not sensitive to excluding durables. Reported expenditures on food at home are notably lower for the 1982 to 1987 CE waves. This disparity appears to reflect different wording in the questionnaire for those years. We adjust food at home expenditures upward by 11% for these years, with the basis for this correction detailed in the appendix.

On the income side, we use the CE measures of total household labor earnings, total household income before tax, and total household income after tax. These variables are reported in the last interview and cover the previous 12 months. Before-tax income in the CE includes labor earnings, non-farm or farm business income, social security and retirement

benefits, social security insurance, unemployment benefits, workers' compensation, welfare (including food stamps), financial income, rental income, alimony and child support, and scholarships. Our measure of before-tax income is that reported in the CE, but we add in food as pay and other money receipts (e.g., gambling winnings). For consistency, as we count receipts of alimony and child support as income, we subtract off payments of alimony and child support. Finally, as rental equivalence is a consumption expenditure for home owners, we include rental equivalence minus out-of-pocket housing costs as part of before-tax income as well. Our measure of after-tax income deducts personal taxes from our measure of before-tax income. These taxes are federal income taxes, state and local taxes, and payroll taxes. Note that federal income taxes can be negative, especially as they capture earned income credits. We consider an alternative measure of after-tax income by replacing self-reported federal income taxes with taxes calculated from the NBER's TAXSIM program. We discuss those results as a robustness check in section 2.3.

The CE asks respondents a number of questions on active savings. For example, they record net flows to savings accounts, purchases of assets (including houses and business), payments of mortgages, payments of loans, purchases and sales of vehicles, etc. The detailed components of savings are reported in the data appendix. We use the savings data as a consistency check, via the budget constraint, on reported consumption. We show below that the average saving rate reported in the CE appears broadly consistent with that obtained from the flow of funds or national income accounts, although there are marked differences. In particular, the data on new mortgages in the CE raise the question of whether the CE accurately records the net effect of refinancing on savings. The CE data show sharp up-ticks in new mortgages around 1993 and the early 2000s, consistent with published statistics on refinancing. However, a number of reported new mortgages have no corresponding house purchase or significant pay down of an existing mortgage. The CE data imply an average "cash out" percentage of 73 percent from new mortgages not associated with a house purchase, while studies of refinancing suggest that only roughly 13 percent is taken out as cash, with the balance used to pay off existing mortgages and related costs (see [Greenspan and Kennedy, 2007](#)). For this reason, we construct an alternative measure of household savings that caps the amount of net borrowing (cash out) associated with new mortgages at one third the size of that mortgage. This reduces the average implied cash out ratio of refinanced mortgages to 14 percent, close to the number reported by [Greenspan and Kennedy \(2007\)](#).

Income, saving, and household total expenditures are expressed in constant 1983 dollars using the CPI-U. Note that we use the aggregate CPI to deflate total expenditures, and

do not deflate separately by expenditure category. This keeps all elements of the budget constraint in the same units. All results based on individual expenditure categories are expressed for one set of households relative to others (e.g., high versus low income) at a point in time, so price deflation is not an issue.

CE survey waves from 1981 through 1983 include only urban households, and so for consistency we restrict our analysis to urban residents. Our analysis employs the following further restrictions on the CE urban samples. We restrict households to those with reference persons between the ages of 25 and 64. We only use households who participate in all four interviews, as our income measure and most savings questions are only asked in the final interview. We restrict the sample to those which the CE labels as “complete income reporters,” which corresponds to households with at least one non-zero response to any of the income and benefits questions. We eliminate households that report extremely large expenditure shares on our smaller categories. Finally, to eliminate outliers and mitigate any time-varying impact of top-coding, we exclude households in the top and bottom five percent of the before-tax income distribution. (The extent of top coding dictates the five percent trimming.) We are left with 62,734 households for 1980-2010. The data appendix details how many households are eliminated at each step.

From this sample, we divide households into 5 bins based on before-tax income, with the respective bins containing the 5-20, 20-40, 40-60, 60-80, and 80-95 percentile groups, respectively. For each income group in each year, we average expenditure, income, and savings variables across the member households. Our primary measure of inequality is the ratio of the mean of the top income group to the mean of the bottom income group. When estimating the expenditure elasticities, reported in table 2 below, we control for demographics. To do this, we further divide each income group into 18 demographic cells, based on age range (25-37, 38-50, 51-64), number of earners (<2, 2+), and household size (≤ 2 , 3-4, 5+). The analysis is done by averaging within each cell (using CE household weights) and then weighting the cell by the sum of the underlying household sampling weights. The use of cells has a number of advantages, including mitigating household-specific idiosyncratic measurement error by averaging; ensuring non-zero expenditure when taking logs; and providing consistent groups for comparison of income, expenditure, and savings.

2.2 Trends in Income and Consumption Inequality

In this subsection, we review the trends in income and consumption inequality using our CE sample. We then discuss the CE savings rates and check the consistency of expenditure,

saving, and income inequality from the perspective of the budget constraint.

We begin with labor earnings. The top line in figure 1 depicts the trend in labor earnings inequality. As discussed in section 2, inequality is the ratio of the mean for the top income cells to the mean for the bottom income cells. Keep in mind that the allocation of respondents into the high and low-income groups is based on before-tax income, and so the cells are the same for all lines in figure 1.

There is substantial year-to-year movement, reflecting in large part sampling error, so we average over multiple years in table 1. In particular, we look at four three-year periods: 1980-82, 1991-93, 2005-07, and 2008-10. The fifth column reports the change over the sample period before the “Great Recession” by log differencing the first and third columns. The final column reports the log change between 2005-07 and 2008-10. We break out the recent recession given that inequality behaves somewhat differently during this period and has already attracted some academic interest.⁵ We also break the sample at 1993 to highlight the sharp rise in inequality during the first decade or so of our sample. While that break captures the sharp early rise in inequality, it leaves aside the middle period 1994-96 employed for the Engel curves in the two-step estimation discussed in the next section.

For the 1980-82 period, average household labor earnings in 1983 dollars was \$44,995 for our top income group and \$7,002 for our bottom income group, for a ratio of 6.41. Labor earnings for the top income group grew by 30 percent (in log points) through 2007, while labor earnings for the low income grew by 10 percent, resulting in a ratio of 7.88 in 2005-07. This implies an increase in earnings inequality of 21 log points. The increase in inequality in the first decade of our sample (from 1980-82 to the 1991-93 period) is even larger at 28 percent. But this is largely driven by years 1992-93 which, from figure 1 appear as outliers for earnings. For 2007-10, earnings inequality expanded by 9 log points.

The next line in figure 1 is for before-tax income which, recall, includes transfers. Inequality in this broader measure of income is lower at each point in time, but also shows a steady increase over time. In particular, this ratio increases from 4.75 in 1980-82 to 6.40 in 2005-07 (third row of table 1), for an increase of 30 percent over this period. Inequality in total household income, after deducting taxes, grew by slightly more than in before-tax income, with an increase of 33 percent over the 1980–2007 sample period (Row 3 of table 1).

⁵Jonathan Heathcote, Fabrizio Perri, and Gianluca Violante (*VOX EU*, 2010) examine the CE data through 2008, Ivaylo Petev, Luigi Pistaferri, and Itay Saporta Eksten (In *Analysis of the Great Recession*, D. Grusky, B. Western, and C. Wimer, eds., forthcoming) through 2009. Each find a considerable fall in inequality with the recession, where inequality is measured by relative expenditures at the 90th versus 10th percentile of consumption expenditures. Each find the fall in inequality coincides with a large drop in expenditure at the 90th percentile.

Income inequality was roughly flat during the Great Recession, with increases of only 2 and 1 log points respectively in before and after-tax income between 2005-07 and 2008-10.

As a robustness check on the CE measure of after-tax income, we computed federal income taxes using the NBER's TAXSIM program, and used this in place of the CE's self-reported income tax to calculate after-tax income for the 1980–2010 period. This alternative measure of after-tax income inequality increased from a ratio of 3.79 for 1980-82 to a ratio of 5.01 for both 2005-2007 as well as 2008-10. That equals a log change of 28 points. This exercise suggests that respondents in the CE are under-reporting the progressivity of federal income taxes relative to TAXSIM, and this gap is increasing modestly over time. Nevertheless, the differences do not substantially change the conclusion that income inequality increased significantly over this period, on the order of 30 percent.⁶

Figure 1 also depicts consumption inequality between the top income group and the bottom income group based on reported expenditures. The increase is much less than that of earnings or income before the recent recession, the feature highlighted in [Krueger and Perri \(2006\)](#). In table 1, we see that consumption inequality increased by only 17 percent over the pre-Great Recession period. Consumption inequality fell during the Great Recession, with a decline of 6 log points between the 2005-07 and 2008-10 surveys. So for the full sample inequality in reported expenditures increased by only 11 percent, or about a third of that seen in income.

We have also computed inequality relative to the middle-income group, which represents the 40th to 60th percentiles. For simplicity, we will refer to this as the 50th percentile. The 32 percent increase in before-tax income inequality reported in table 1 can be broken into an increase of 21 percent for the 90-50 ratio, and 11 percent for the 50-10 ratio. Similarly, the 34 percent increase in after-tax income inequality is composed of a 21 percent increase for the 90-50 ratio and 13 percent increase for the 50-10 ratio. For consumption, the 11 percent increase is skewed entirely to the top, with a 13 percent increase in the 90-50 ratio and a 1 percent decrease in the 50-10 ratio. That is, there is actually no reported increase in consumption inequality in the bottom half of the sample.

⁶The rise in income inequality we observe in the CE is broadly consistent with patterns in other data. [Meyer and Sullivan \(2009\)](#) measure income inequality using income information in the Current Population Surveys (CPS). There are differences in methodology from our approach; for instance, their statistics adjust for family size using equivalence scales. Nevertheless, they show for 1980-2007 an increase in the 90-10 differential in after-tax income of 27 percent. [Heathcote et al. \(2010\)](#) also examine after-tax income based on CPS data, but report a larger increase in the 90-10 differential for 1980-2005 of a little over 50 percent.

2.3 Saving Rates

We now turn to implied and observed saving rates, beginning with mean saving rates. Figure 2 depicts the personal saving rate reported in the flow of funds accounts.⁷ There is a clear downward trend in this series, starting from 12.2 percent for 1980-82 and falling to 1.7 percent for 2005-07, and then recovering slightly during the recent recession. This downward trend in the personal saving rate is well known, and is similar to that implied by the national income accounts.

The implied savings rate in the CE data can be computed as one minus the ratio of mean consumption expenditures to mean after-tax income. This series is also depicted in figure 2. The implied saving rate has a dramatically different trend, increasing from 13 percent for 1980-82 to 23 percent for 2005-07, and then continuing upward to 25 percent for 2008-10. This systematic increase in implied savings is at odds with the flow of funds or national income accounts, and is the counterpart to the previously discussed increasing gap between CE and NIPA expenditure.

Figure 2 also reports the saving rate constructed from the CE's savings data. The series labeled "unadjusted" is the sample mean of reported savings divided by mean after-tax income for each year. The mean savings rate falls from 3 percent in 1980 to -12 percent at the end of the sample. This decline is the opposite of the increase implied by consumption data, revealing an inconsistency between the CE's consumption, income, and savings data that is increasing over time. As mentioned in section 2, there is a measurement issue concerning new mortgages, which underlies the large decline generally, and the sharp swings around 1993 and 2003 in particular. As described in section 2, we construct an alternative savings series designed to address the mis-reporting of new mortgages. This series is the "adjusted" series in figure 2. With adjustment, the series more closely tracks the flow of funds savings and eliminates part of the sharp downward spikes in savings in the mid-1990s and 2000s.

The fact that aggregate consumption in the CE is falling relative to NIPA does not necessarily bias measures of inequality. For example, if CE expenditures are under reported by the same multiplicative factor for all income groups, then the ratio of consumption across groups will not be biased. However, such an assumption has somewhat extreme implications for relative saving rates. Suppose we uniformly increase expenditures across groups in 2008-10 to generate a decline of 6 percentage points in the aggregate CE savings rate, which is

⁷Specifically, the saving rate is personal saving without consumer durables divided by disposable income. A similar pattern is obtained using the national income and product accounts, where savings is disposable personal income minus personal outlays.

the decline observed in the flow of funds. This implies that consumption should be adjusted upwards by 24 percent.⁸ Given that $\frac{Savings}{Income} = 1 - \frac{Consumption}{Income}$, this implies each income group's saving rate must be adjusted downward by 24 percent of their respective consumption to income ratio. Because the consumption-income ratio is much higher for low-income groups, it requires an extreme decline in their savings rate. In particular, the implied savings rate for the top income group must decline modestly from 28 percent for 1980-82 to 26 percent for 2008-10, while for the bottom group is must go from -23 all the way down to -59 percent. We would suggest that such a trend decline in savings rate for the bottom group is extreme, especially given that income is defined to include transfers and given that the very lowest income households are trimmed from the sample.

These implied saving trends across income groups are also inconsistent with the CE's (admittedly noisy) micro data on active savings.⁹ In particular, high-income respondents report an adjusted savings rate of 2 percent in 1980-82 and a rate of 1 percent in 2008-10. Low-income respondents report corresponding saving rates of 3 percent and 0 percent, respectively.

As previously emphasized, reported savings is not a focus of the CE, and one may reasonably question conclusions drawn solely from reported savings. Our primary focus is to use the savings data as a consistency check on the CE's consumption data. It turns out that the savings data tell a much different story regarding consumption inequality than do the expenditure data. This inconsistency raises the question of whether the expenditure data are subject to systematic measurement error that biases our estimates of consumption inequality. Addressing this potential measurement error is the focus of the next section.

3 Demand System Estimates of Consumption Inequality

In this section we present our main results. We first discuss how our econometric methodology corrects for several classes of mis-measurement. We then estimate a simple demand

⁸Specifically, let γ denote our adjustment factor, so we increase consumption by a factor of $(1 + \gamma)$ uniformly across households. The adjustment to the saving rate is: $\Delta \frac{S}{Y} = -\gamma \frac{C}{Y}$. To match the 6 point decline in the saving rate observed in the flow of funds, the aggregate CE saving must be adjusted down by $0.12 - (-0.06) = 0.18$ points in 2008-2010. As the ratio of aggregate CE consumption to income in 2008-2010 is 0.75, an adjustment factor of $\gamma = .24$ is required: $(-0.24)(0.75) = -0.18$.

⁹It is also not reflected in other micro-data on savings, as documented by [Bosworth and Anders \(2008\)](#) and [Bosworth and Smart \(2009\)](#).

system which we use to generate our estimates of consumption inequality growth.

3.1 Econometric Approach

To set notation, let the index $h = 1, \dots, H$, represent cells defined by income and demographics as described in section 2. Let $i = 1, \dots, I$ denote the $I = 5$ income groups. With 18 demographic groups for each of the 5 income group, we have $H = 90$. Let $j = 1, \dots, J$ index our $J = 20$ goods; and let t index time (year). With this notation, let x_{hjt} denote reported expenditure on good j at time t by income-demographic group h , where we average over households in each cell using the CE replication weights. Let X_{ht} denote total expenditure at time t by group h ; that is, $X_{ht} = \sum_{j=1}^J x_{hjt}$.

We assume that x_{hjt} is measured with error, with the degree of mis-measurement depending on time, income group, and good. Note that this is the systematic measurement error that survives averaging across households within each income-demographic group. In particular, let x_{hjt}^* denote the true expenditure, and

$$x_{hjt} = x_{hjt}^* e^{\zeta_{hjt}}. \quad (1)$$

Given that x and x^* are non-negative, our specification of measurement error is to this point without loss of generality. We can decompose ζ_{hit} into three components:

$$\zeta_{hit} = \psi_t^j + \phi_t^i + v_{hjt}. \quad (2)$$

Here, ψ_t^j reflects mis-measurement of consumption good j at time t that is common across respondents (e.g., food may be under-reported for all households); ϕ_t^i represents mis-measurement specific to i at time t that is common across goods (e.g., the rich may under-report all expenditures); and v_{hjt} is the residual good-group specific measurement error (e.g., food expenditures of the rich are under-reported). Without loss of generality (given the presence of ψ_t^j and ϕ_t^i), we normalize the mean of v_{hjt} to be zero for all t . Our identifying assumption is that v_{hit} is classical measurement error; in particular, it is independent of the characteristics of good j and group h at each date t . We will be more precise about the independence condition after we discuss our estimation strategy.

Our estimation consists of two steps. First, we estimate the total expenditure elasticities for each good. We assume that Engel curves are log-linear and so expenditure elasticities are constant. Of course, this can only be true locally, unless all elasticities are one. Nevertheless,

it provides a tractable framework to address the mis-measurement of expenditure in the CE. A reasonable benchmark is that respondent's errors (positive or negative) are scaled by their level of expenditures. As we show below, the log-linear specification is particularly well suited to handle such measurement error. A popular alternative local approximation is the Almost Ideal Demand System (AIDS) of [Deaton and Meullbauer \(1980\)](#), which assumes that the share of expenditure on good j is log linear in total expenditure. The AIDS approximation has nice features for tractably testing implications of consumer optimization, but is not well suited to handle good-specific measurement error ψ_t^j in our second stage. Multiplicative measurement error is not differenced out in the AIDS specification.

We assume that true expenditure satisfies:

$$\ln x_{hjt}^* = \alpha_{jt}^* + \beta_j \ln X_{ht}^* + \mathbf{\Gamma}_j \mathbf{Z}_h + \varphi_{hjt}. \quad (3)$$

The term \mathbf{Z}_h is a vector of demographic dummies corresponding to age, number of earners per household, and family size, reflecting the categories used to construct the demographic cells. We allow the coefficient vector on demographics $\mathbf{\Gamma}_j$ to vary across goods. The error term φ_{hjt} represents idiosyncratic relative taste shocks as well as the second-order error from the log-linear approximation, which we assume are independent of total expenditure and independent of expenditure elasticities β_j . We discuss mis-specification further in section 4 in which we test the robustness of the log-linear specification.

We estimate expenditure elasticities using the 1994-96 Consumer Expenditure Survey. These three waves represent the mid-point of our sample. In previous work, we have used the 1972-73 CE survey as the basis for estimating expenditure elasticities. It turns out our second-stage estimates are relatively stable with respect to the first-stage time period, a point we discuss in detail in the robustness section. Specifically, we estimate expenditure elasticities using observed expenditures:

$$\ln x_{hjt} = \alpha_{jt} + \beta_j \ln X_{ht} + \mathbf{\Gamma}_j \mathbf{Z}_h + u_{hjt}, \quad t = 1994, 95, 96, \quad (4)$$

where $\alpha_{jt} \equiv \alpha_{jt}^* + \psi_t^j$ subsumes the good-time specific measurement error into the intercept, and

$$u_{hjt} = \phi_t^i + v_{hjt} + \varphi_{hjt}. \quad (5)$$

A concern with estimating a demand system like (4) is that mis-measurement of individual goods is cumulated into total expenditure, inducing correlation between the measurement error captured in the residual and observed total expenditure. A standard technique is to

instrument total expenditure with income and other proxies for total expenditure. We are already using income-category averages, which eliminates measurement error uncorrelated with income. Nevertheless, as modeled above, there may be measurement error that is common across households within an income group. That is, the fact that the 1994-96 CE may contain systematic measurement will lead to biased estimates of the expenditure elasticities. In particular, if consumption inequality is understated in 1994-96, the expenditure elasticities will be biased away from one. When we invert the demand system, as described below, this will lead to understatement of consumption inequality in other years as well. A bias in the opposite direction will be in effect if inequality is overstated in the 1994-96 surveys. For this reason, our ultimate estimates of inequality must be interpreted as conditional on the level of inequality observed in the first-stage surveys. In the robustness section, we discuss how the results vary when we use alternative years for the first stage.

The second stage of our estimation is to invert the demand system (3) to recover an estimate of consumption inequality in other years of the survey. We first adjust expenditure for demographics and pool by income group. Specifically, let

$$\ln \hat{x}_{ijt} \equiv \frac{I}{H} \sum_{h \in i} \omega_{ht} \left(\ln x_{hjt} - \hat{\Gamma}_j \mathbf{Z}_h \right),$$

where ω_{ht} is the normalized sum of the CE sample weights for demographic group h in year t , and $\hat{\Gamma}_j$ is the estimate of Γ_j from (4). That is, $\ln \hat{x}_{ijt}$ is the average expenditure of income group i in year t on good j adjusted for demographics. Using (3), we have

$$\ln \hat{x}_{ijt} = \alpha_{jt} + \beta_j \ln X_{it}^* + \phi_t^i + \varepsilon_{ijt}, \quad (6)$$

where $\varepsilon_{ijt} = \varphi_{ijt} + v_{ijt} = \frac{I}{H} \sum_{h \in i} \omega_{ht} (\varphi_{hjt} + v_{hjt})$.¹⁰ Our key identifying assumption regarding measurement error is that v_{hit} is orthogonal to β_j . That is, the idiosyncratic component of the mismeasurement of good j for group h at time t does not vary systematically with the expenditure elasticity. This, plus our assumptions regarding the taste shock φ_{ijt} , imply that ε_{ijt} is independent of β_j . Therefore, we can obtain a consistent estimate of $\ln X_{it}^*$, up to a normalization, by least squares. We only have identification up to a normalization given the presence of α_{jt} .¹¹ Note that changes in systematic measurement error over time are captured by good-time and income group-time dummies. Identification comes from the fact that if the income of group i increases relative to that of group i' , it will increase its relative expenditure, but the increase will fall disproportionately on luxuries.

¹⁰The residual term will also contain estimation error related to $\hat{\Gamma}_j$, which we suppress in the notation.

¹¹That is, the mean of $\ln X_{it}^*$ is not identified as $\alpha_{jt} + \beta_j \ln X_{it}^* = \alpha_{jt} - \beta_j \delta + \beta_j (\ln X_{it}^* + \delta)$.

To implement (6), we regress $\ln \hat{x}_{ijt}$ on a vector of good-time dummies (whose coefficients correspond to α_{jt}), a vector of income-time dummies (whose coefficients correspond to ϕ_t^i), and the interaction of income-time dummies and β_j . The coefficients of the last group of variables will be the estimate of $\ln X_{it}^*$. To address the issue of normalization, we estimate expenditure relative to the lowest income group. That is, we have a consistent estimate of consumption inequality: $\delta_{it} \equiv \ln X_{it}^* - \ln X_{1t}^*$. To estimate trends over time, we restrict ϕ_t^i and δ_{it} to be constant within each three-year window 1980–1982, 1991–1993, 2005–2007, and 2008–2010, but allow the good-time intercept terms α_{jt} to vary year by year. Our two-step procedure requires adjusting the second stage standard errors, which we do by bootstrapping.¹²

3.2 Results

Table 2 reports the results of our first stage estimates of each good’s total expenditure elasticity. The table also includes the average share of each good out of total expenditure for our 1994–96 CE sample. The standard errors are reported next to each estimate and suggest that our first stage has a fair degree of precision, particularly for the goods with large expenditure shares.

The estimated elasticities range from -0.40 for tobacco to greater than 2.0 for education and domestic services. Consistent with other studies, food at home has a fairly low expenditure elasticity (0.36), while food away from home has a high elasticity (1.41). Vehicle purchases is also a large category with a fairly high income elasticity. Housing services, our largest expenditure category, has an expenditure elasticity of 0.92.

To provide a sense of how these expenditure elasticities are informative about relative consumption inequality, we first consider two goods – food at home and non-durable entertainment. These goods have reasonably large shares and very different expenditure elasticities. We plot the relative expenditure (entertainment over food at home) for the high- and low-income households in figure 3. High-income households display a shift in expenditure from food to entertainment over the sample period. Specifically, the ratio increases from 0.21 in 1980–82 to 0.27 for 2008–10. Conversely, low-income households display a shift away from nondurable entertainment, with their ratio falling from 0.09 to 0.06. In log differences, the ratio for the rich increased 23.4 log points, while that of the low-income group fell 43.4

¹²Specifically, we draw with replacement from the micro data for all years and re-estimate both stages. In a previous version, we adjusted the standard errors following [Murphy and Topel \(1985\)](#). Neither methodology implies a substantial adjustment to the standard errors.

points.

The relative shift in expenditure towards a luxury for the high-income households suggests a sharp increase in total expenditure inequality. A simple calculation based on this shift provides insight into how our methodology quantifies this increase. Let $j = \{e, f\}$ denote our two goods, entertainment and food at home. Recall from (3) that

$$\ln x_{ie} - \ln x_{if} = \alpha_e - \alpha_f + (\beta_e - \beta_f) \ln X_i + (\varepsilon_{ie} - \varepsilon_{if}),$$

where we have suppressed the time subscripts and are abstracting from demographic controls for this exercise. Note that any income-group specific multiplicative measurement error (ϕ_t^i) is eliminated by comparing across goods within an income group. This can be inverted to yield

$$\ln X_i = \frac{\ln x_{ie} - \ln x_{if}}{\beta_e - \beta_f} + \frac{\alpha_e - \alpha_f + \varepsilon_{ie} - \varepsilon_{if}}{\beta_e - \beta_f}. \quad (7)$$

The estimates from our first stage imply that the denominator on the right hand side is $1.93 - 0.36 = 1.57$. If we were to assume no good-specific measurement error or relative price shifts (i.e., $\alpha_j = 0$), we can use the change in relative expenditure on entertainment and food to estimate the change in total expenditure. In particular, the 23 point increase for high-income households suggests an increase in log total expenditure of 14.6 log points, while the decline for low-income households implies a decline of 27.6 points in total expenditures.

However, both of these estimates are prone to good-specific measurement error (ψ_t^j) as well as responses to changes in prices, both of which are captured by α_{jt} . For example, if the CE is increasingly under-reporting entertainment expenditure relative to food over time, the log changes will both be biased downwards. However, proportional mis-measurement is eliminated by comparing the difference across income groups. That is, α_j is invariant across income groups. Thus the implied change in inequality of 42 log points is cleaned of both types of systematic measurement error. This estimate is a noisy estimate given the presence of the idiosyncratic shocks ε_{ij} . However, we can address this residual mis-measurement by using many goods in a regression framework.

To obtain a better sense of the identification strategy using all goods, figure 4 is a scatter plot of relative consumption growth versus expenditure elasticity. Specifically, the vertical axis depicts the change between 1980/82 and 2008/10 of log relative expenditure across income groups for each good, adjusted for demographic controls. The horizontal axis depicts the expenditure elasticities reported in table 2. The upward sloping pattern indicates that the high-income respondents increased their expenditure relatively more on luxury goods.

The fitted slope is 0.35. This suggests that an increase in relative total expenditure of 35 log points is consistent with the relative shifts across luxuries and necessities over this period.

More formally, table 3 reports our second-stage regression estimates of the log change in consumption inequality from (6). We focus on the change in consumption inequality between the highest income and lowest income groups relative to 1980-82, and discuss other inter-group comparisons below. The first row of table 3 reports the estimated inequality in the pooled base period 1980–1982. This is the estimate of $\ln X_5^* - \ln X_1^*$ for the first three years of our sample. The row labeled “Log Change 1980/82–1991/93” is the estimated change in inequality between 1980-82 and 1991-93. Similarly, the next row corresponds to the estimated change in consumption inequality between 1980-82 and 2005-07. The final row of estimates reports the change in inequality during the Great Recession based on the change between 2005/07 and 2008/10.

Column (1) reports the second-stage estimates using ordinary least squares. The first row reports the estimated log inequality in the pooled period 1980–1982, which is 0.84. For comparison, table 1 reports a log ratio for reported expenditures of $\ln(2.47)=0.90$ for 1980-82, which differs from our second-stage point estimate for that period by 0.06 points. This implies that the level of consumption inequality estimated with our two-step procedure is similar to that obtained from reported expenditure for the beginning of our sample. This similarity, however, does not persist over time. The next two rows of estimates in column (1) report that the estimated change in consumption inequality is 27 percent for the early period and 40 percent through 2007. These numbers are similar in magnitude (or larger) than those for after-tax income reported in table 1, and differ from changes in reported consumption inequality. The final row of column (1) reports a decline in consumption inequality of 5 points, which is similar to that reported for reported consumption in table 1, suggesting that the recent decline in consumption inequality is reflected in the shifting of relative consumption baskets. The estimated increase in consumption inequality for the entire period 1980–2010 is 35 log points, which is the slope in figure 4.

One issue with OLS is that it weights all goods equally in the second stage. This raises the question of whether goods with small shares or greater heteroscedasticity are driving the results. Column (2) estimates the second stage using two-step feasible generalized least squares. Specifically, we allow heteroscedasticity across goods to capture that the size of taste shocks or idiosyncratic measurement error may differ across goods. To estimate good-specific residual variances, we use residuals from the OLS specification of column (1). We use these to weight the final estimation. GLS implies an initial log inequality of 0.85, and an increase in inequality in the first decade of the sample of 24 percent, an increase through

2007 of 0.32 percent and a decline over the last 5 years of 4 percent.

Column (3) performs the same GLS regression but excludes categories that contain durables.¹³ Non-durable consumption avoids the issue of imputed service flow that complicates measures of durable consumption. But, because we maintain the same first stage, these estimates are still of total consumption inequality, not just non-durable consumption inequality. We find that the estimated increase in inequality is stable to this alternative sample. Specifically, we find a 24 percent increase in inequality in the first decade of the sample, and 34 percent through 2007.

The final column of table 3 implements weighted least squares, where the weights reflect the share of each good in personal consumption expenditures (PCE) from the national income accounts. Specifically, we calculate the share of each good out of total PCE for each year, then average the shares over the sample period 1980-2010 and use these shares to weight the goods in the second stage regression. For health expenditures we downweight its share for each year to a factor equal to the share of private expenditures, out of pocket and private insurance, out of total national health expenditures; this factor averages 49 percent for 1980-2010.¹⁴ The baseline log inequality is slightly lower (0.90) in this specification, and the corresponding increase over time slightly lower as well. Specifically, we estimate a change in consumption inequality of 20 percent for the early period, 29 percent through 2007, and 26 percent through 2010.

The second-stage estimation uses all five income categories, and therefore produces an estimate of inequality across any two income groups. As discussed at the end of section 2.2, between 1980 and 2010 the 90-50 after-tax income ratio increased by 21 points, while the 50-10 ratio increased by 13. By sharp contrast, reported total expenditures indicate that the 90-50 consumption ratio increased by 13 points, while the 50-10 ratio actually declined by a point. Our two-stage OLS estimates suggest that two-thirds of the increase in consumption inequality occurred between the high and middle groups, with the estimated 90-50 ratio increasing by 23 points and the 50-10 ratio increasing by 11 points. Our preferred GLS estimates show an equal split, with both the 90-50 and 50-10 ratios increasing by 14 points. These estimates indicate a sharp under-reporting of consumption inequality in the bottom of the distribution over our sample period.

Table 4 report the estimates for income-specific measurement error, ϕ_t^i . In particular,

¹³Specifically, from the goods listed in table 2, we exclude vehicles, appliances, furniture, and entertainment equipment.

¹⁴The data source is Centers for Medicare & Medicaid Services, Office of the Actuary, National Health Statistics Group

we report the difference between the highest income group and the lowest income group: $\phi^5 - \phi^1$. These are the estimates of the income-specific intercepts. The rows and columns are arranged in the same manner as in table 3. The first row point estimates suggest that (relative) income-specific mis-measurement is positive, implying that the reported consumption inequality is overstated in the first years of the CE. Specifically, a positive estimate implies a combination of the high-income respondents over-reporting expenditure and the low-income respondents under reporting expenditure. Over time, the estimated relative mis-measurement falls, implying a trend toward under-reporting by high-income households relative to low-income households. The estimates suggest that the CE is increasingly “missing” expenditure by the high-income households (relative to low-income households), generating an understatement of the true increase in consumption inequality. In the recent period (since 2005), mis-measurement seems to be stable, with the point estimates in the last row of table 4 close to zero.

4 Robustness

The two key assumptions of our Engel curve approach are (i) the demand system is log-linear, and (ii) the income elasticities are stable over time. In this section we explore the sensitivity of our results to relaxing these assumptions.

4.1 Non-linear Engel Curves

Recall that our benchmark estimation assumes the log expenditure on good j is linear in log total household expenditure, conditional on demographics (equation 3). We can relax that assumption by allowing for higher order terms. Specifically, for demographic cell h and good j in time t , we assume:

$$\ln x_{hjt}^* = \alpha_{jt}^* + \beta_{1,j} \ln X_{ht}^* + \beta_{2,j} (\ln X_{ht}^*)^2 + \mathbf{\Gamma}_j \mathbf{Z}_h + \varphi_{hjt}. \quad (2')$$

As in the benchmark model (2), α_{jt}^* and \mathbf{Z}_h represent a good-time specific intercept and a vector of demographic dummies, respectively, and x_{hjt}^* and X_{ht}^* represent cell h expenditure on good j and total expenditure in year t , respectively. Specification (2') extends the benchmark by allowing an impact from the second-order term $(\ln X_{ht}^*)^2$.

As in the benchmark, we estimate this demand system using the 1994-96 CE sample.

Maintaining the same assumptions about measurement error, our second stage is now:

$$\ln \hat{x}_{ijt} = \alpha_{jt} + \beta_{1,j} \ln X_{ht}^* + \beta_{2,j} (\ln X_{ht}^*)^2 + \phi_t^i + \varepsilon_{ijt}, \quad (5')$$

where as before a $\hat{}$ reflects that we have adjusted x_{ijt} for demographics using the estimates of Γ_j from the first stage. Recall that in the second stage $\beta_{1,j}$ and $\beta_{2,j}$ are our generated regressors, whose associated coefficients yield estimates of log expenditure and expenditure squared. Note that this specification nests our benchmark specification (which restricts $\beta_{2,j} = 0, \forall j$). As in the benchmark, our assumptions regarding the nature of the good-income-time specific measurement error contained in ε_{ijt} ensures the residual is uncorrelated with the regressors.

We estimate (5') for 1980-82, 1991-1993, 2005-2007, and 2008-2010 surveys. In table 5, we report the estimated level of inequality in 1980-82, and then the log change over time, following the layout of our benchmark table 3. The fact that the coefficient on $\beta_{2,j}$ is the square of the coefficient on $\beta_{1,j}$ is a nonlinear restriction on the parameters. We impose this restriction and estimate the model nonlinearly. Column (1) utilizes nonlinear least squares. Column (2) implements nonlinear generalized least squares.¹⁵ As in the benchmark, weighting by the residual variance down weights the noisier (often smaller) categories. Column (3) reports the estimates from an unrestricted second stage, in which we implement linear GLS without imposing the relationship between log expenditure and its square. For the unrestricted specification, we report the level and change in inequality implied by the coefficient on the first-order term, $\beta_{1,j}$. The standard errors in all specifications are bootstrapped.

The different columns of table 5 indicate that the results are similar across specifications, with the GLS specifications producing slightly smaller changes inequality in line with the linear estimates of table 3. More importantly, comparing table 5 to our benchmark table 3 suggests that nonlinearities do not have a significant impact on the estimated inequality. Specifically, the baseline inequality for the 1980-82 period is 0.84 in column (2) of table 5, and 0.85 in the corresponding column of table 3. The change in inequality is also similar with the nonlinear specification. Table 5 column (2) reports an increase of 27 log points between 1980-82 and the early 1990s, and of 30 log points through 2007. The corresponding benchmark estimates are 24 and 32, respectively. The change during the Great Recession is -0.02 in the nonlinear specification, compared to -0.04 in the linear case. We view this ro-

¹⁵Specifically, we allow for good-specific heteroscedasticity, and weight the observations by the inverse variance estimated using residuals from nonlinear least squares, making column (2) the nonlinear counterpart to column (2) of table 3.

bustness exercise as suggesting that nonlinear terms do not change our conclusions regarding systematic measurement error biasing down reported consumption inequality.

4.2 Stability of Expenditure Elasticities

In addition to a correctly specified demand system, another key assumption in our methodology is that expenditure elasticities are stable over time. The danger in this regard is that expenditure elasticities may depend on relative prices or other attributes of the goods that may have changed over time (any such changes that affect the intercept of the Engel curve are accounted for by the good-time dummy variables). This concern may be relevant given the 30 year span between the 1980 and 2010 surveys, although we have designed broad good categories with this issue in mind. A second concern involves the baseline expenditure inequality in 1994-96. Our benchmark results suggest that inequality in recent surveys is understated. An alternative interpretation is that inequality is correctly measured in the recent surveys, but overstated in the 1994-96 surveys, generating systematically biased first-stage elasticities.

To explore these issues, we have re-estimated expenditure elasticities at different points in the sample. Figure 5 depicts a scatter plot of the 1994-96 elasticities (horizontal axis) against those estimated from the 1980-82 surveys (panel a) and from the 2008-10 surveys (panel b). The ordinal ranking of goods is extremely stable. The correlation of our benchmark elasticities with these alternative elasticities is 0.89 with the 1980-82 elasticities and 0.96 with the 2008-10 elasticities. There is a slight fanning out of the elasticities over time. In particular, the standard deviation of the elasticities across goods is 0.59 in 1980-82, 0.69 in the benchmark, and 0.73 in 2008-10. This difference is consistent with the benchmark results showing mis-measured inequality in the later sample that worsens with time. If total consumption inequality is systematically under-reported in the later samples, the estimated expenditure elasticities will tend to be increasingly biased away from one. Of course, it could also be consistent with the hypothesis that inequality is correctly measured in the later sample, but *under*-reported in 1994-96 (and earlier), biasing our benchmark elasticities towards one.

We can explore this issue further by performing our second-stage estimation using elasticities estimated from the later sample. In doing so, we avoid using the same sample in the first and second stages. Doing otherwise would undermine the orthogonality assumption necessary for the second stage. In particular, if the residual error terms are the same in the first and second stages, our second stage regressors (the estimated elasticities with the

inherited sampling error) will not in general be orthogonal to the error term. This was one motivation for our use of the 1994-96 CE. To circumvent this problem in our robustness exercise, we use the elasticities estimated in 2005-2007 to compute the change in inequality between 1980-82 and 2008-2010.

These alternative elasticities and the GLS specification yield an estimated level of log inequality of 0.68 in 1980-82 and 1.08 in 2008-10 . This contrasts with the 0.85 and 1.13 estimates using the 1994-96 elasticities. As expected, the more dispersed elasticities from 2005-07 generate lower levels of estimated inequality. Nevertheless, the change in log consumption inequality between 1980 and 2010 is estimated to be 0.40 using the 2005-07 elasticities. This is actually greater than the 0.28 point estimate reported in table 3. We have also estimated elasticities using the 1983-2007 sample. These elasticities generate a change in consumption inequality of 0.31 log points between 1980-82 and 2008-10.

5 Conclusion

The results presented in this paper suggest that increases in consumption inequality may mirror that of income inequality to a greater extent than implied by reported total expenditure. The basis of this reinterpretation is the reported shift of high-income households' consumption toward luxuries and away from necessities relative to the consumption baskets of low-income households. The Engel curve approach allows us to use the detailed expenditure reports on different classes of goods to correct for systematic measurement error. Our modeling of measurement error is broad in that we allow biases to vary across good-year and income class-year, as well as allowing for classical (non-systematic) mis-measurement at the level of good-income-year interaction. The attraction of the CE is that it is a comprehensive survey of expenditure across many goods, and this richness can be exploited using a simple demand system. The approach requires assumptions, including that our demand system is correctly specified and that the expenditure elasticities are stable across periods. We have explored the validity of these assumptions in section 4 and found the results are robust to alternative specifications. Our interpretation of the data provides a parsimonious explanation of the inconsistency between reported expenditure inequality, reported savings and income inequality, and the fact that the high-income households report a substantial shift in expenditure towards luxuries relative to low-income households.

References

- Aguiar, Mark and Erik Hurst**, “Deconstructing Life Cycle Expenditure,” *working paper*, 2009.
- Attanasio, Orazio and Steven Davis**, “Relative Wage Movements and the Distribution of Consumption,” *Journal of Political Economy*, 1996, *104*, 1227–1262.
- , **Erich Battistin**, and **Hidehiko Ichimura**, “What Really Happened to Consumption Inequality in the US?,” in “Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches,” Chicago: Chicago University Press, 2007.
- , **Erik Hurst**, and **Luigi Pistaferri**, “The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980-2010,” *NBER Working Paper 17982*, 2012.
- Battistin, Erich**, “Errors in Survey Reports of Consumption Expenditures,” *Institute for Fiscal Studies WP 03-07*, 2003.
- Blundell, Richard and Ian Preston**, “Consumption Inequality and Income Uncertainty,” *Quarterly Journal of Economics*, 1998, *113*, 603–640.
- , **Luigi Pistaferri**, and **Ian Preston**, “Consumption Inequality and Partial Insurance,” *American Economic Review*, 2008, *98*, 1887–1921.
- Bosworth, Barry and Rosanna Smart**, “Evaluating Micro-Survey Estimates of Wealth and Saving,” *CRR WP 2009-4*, 2009.
- and **Sarah Anders**, “Saving and Wealth Accumulation in the PSID, 1984-2005,” *CR-RWP 2008-2*, 2008.
- Browning, Martin and Thomas Crossley**, “Are Two Cheap Noisy Measures Better than One Expensive, Accurate One?,” *American Economic Review*, 2009, *99*, 99–103.
- Cutler, David and Lawrence Katz**, “Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980s,” *American Economic Review*, 1992, *82*, 546–551.
- Deaton, Angus and Christina Paxson**, “Intertemporal Choice and Inequality,” *Journal of Political Economy*, 1994, *102*, 437–467.
- and **John Meullbauer**, “An Almost Ideal Demand System,” *American Economic Review*, 1980, *70*, 312–326.
- Garner, Thesai, George Janini, William Passero, Laura Paszkiewicz, and Mark Vendemia**, “The Consumer Expenditure Survey: A Comparison with Personal Expenditures,” *Bureau of Labor Statistics*, 2006.
- Greenspan, Alan and James Kennedy**, “Sources and Uses of Equity Extracted from Homes,” *Federal Reserve Board Finance and Economics Discussion Series 2007-20*, 2007.

- Guvenen, Fatih**, “Learning your Earning: Are Income Shocks Really Very Persistent,” *American Economic Review*, 2007, *97*, 687–712.
- Heathcote, Jonathan, Fabrizio Perri, and Gianluca Violante**, “Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006,” *Review of Economic Dynamics*, 2010, *13*, 15–51.
- , **Kjetil Storesletten, and Giovanni Violante**, “Two Views of Inequality over the Life Cycle,” *Journal of the European Economic Association*, 2005, *3*, 765–775.
- Huggett, Mark, Gustavo Ventura, and Amir Yaron**, “Sources of Lifetime Inequality,” *working paper*, 2009.
- Johnson, David and Stephanie Shipp**, “Trends in Inequality Using Consumer Expenditures: 1960 to 1993.,” *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1995.
- Krueger, Dirk and Fabrizio Perri**, “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *The Review of Economic Studies*, 2006, *73*, 163–193.
- , – , **Luigi Pistaferri, and Gianluca Violante**, “Cross-Sectional Facts for Macroeconomists,” *Review of Economic Dynamics*, 2010, *13*.
- Meyer, Bruce and James Sullivan**, “Five Decades of Consumption and Income Poverty,” *NBER Working Paper 14827*, 2009.
- Murphy, Kevin and Robert Topel**, “Estimation and Inference in Two-Step Econometric Models,” *Journal of Business and Economic Statistics*, 1985, *3*, 370–379.
- Parker, Jonathan, Annette Vissing-Jorgensen, and Nicholas Ziebarth**, “Inequality in Expenditure in the Twentieth Century,” *working paper*, 2009.
- Sabelhaus, John and Jeffrey Groen**, “Can Permanent-Income Theory Explain Cross-Sectional Consumption Patterns,” *The Review of Economics and Statistics*, 2000, *82*, 431–438.
- Slesnick, Daniel**, *Consumption and Social Welfare: Living Standard and their Distribution in the United States*, Cambridge: Cambridge University Press, 2001.
- Storesletten, Kjetil, Chris Telmer, and Amir Yaron**, “Consumption and Risk Sharing over the Life Cycle,” *Journal of Monetary Economics*, 2004, *51*, 609–633.

Data Appendix

In this appendix we describe construction of the variables in our data set and the impact of sample restrictions. All data are available from the authors' web page.

Construction of variables from CE

The income variables we examine are total household labor earnings, total household income before tax, and total household income after tax. These variables are principally based on responses in the last quarterly interview that cover income from the previous 12 months. Household labor income sums all household member earnings, before deductions, over the past 12 months. The before-tax income in the CE (FINCBTAX) includes labor earnings, business (including farm), and professional income, interest, dividend, rental, and royalty income, income from social security and railroad retirement benefits, income from pensions and annuities, scholarships or stipends, workers' compensation and veterans' benefits, and alimony and child support received. It also includes the following transfer payments: public assistance (welfare) payments including those related to job training, food stamps, supplemental security income, and unemployment benefits.

We adjust this measure of before-tax income in the following ways to be consistent with budget accounting. We add in food as pay and other money receipts. The latter includes lump-sum receipts of alimony and child support, lump-sum receipts from estates, selling household items, prizes or gambling winnings, and refunds of insurance payments, property taxes, or employer over withholding on social security taxes. We subtract alimony and child support payments, to be consistent with those receipts being treated as income. We also subtract expenditures that we do not treat as consumption. These include life insurance premiums, occupational expenses, fees for financial services, finance charges, legal fees, funeral expenses, moving expenses, and support for college students. We treat the implicit rental from owner-occupied housing both as a component of expenditures and a part of income. So we add home owner's estimate of rental equivalence to before-tax income. At the same time we subtract expenses of home ownership for mortgage interest, property taxes, expenditures for capital repairs and replacements, home insurance, security systems, pest control, and other maintenance expenses both from income and expenditures.

We subtract personal taxes from our measure of before-tax income to arrive at a measure of after-tax income. These taxes include federal, state and local income taxes. We also subtract the income contributed to social security by all household members during the year, as well as contributions for government or railroad retirement programs. The CE measure of social security contributions is estimated by the BLS. Our measure of after-tax income differs from the CE measure (FINCATAX) due to all the adjustments listed above to before-tax income, and because we subtract contributions to social security, government, and railroad retirement programs. We consider an alternative measure of after-tax income by replacing self-reported federal income taxes with taxes calculated from the NBER's TAXSIM program. We do not adjust for state and local taxes, as we do not know the state of residence

for many households in the CE. We also considered replacing social security taxes with TAXSIM values, but this has little effect on the results. This is not surprising, as the social security contributions in the CE are estimated by the BLS as well.

We aggregate CE expenditure items into 20 groups, as described in the text. Our definitions of expenditures by good closely follow definitions in the CE with a few exceptions, most notably for housing services. As in the CE, for renters we define housing by rent paid. But for home owners we use self-reported rental equivalence rather than out of pocket expenditures. This adjustment was described above in discussing adjustment to income. For the eight quarterly surveys conducted in 1980 and 1981 households were not asked about rental equivalence. We impute the rental equivalence for homeowners in these early waves as follows. We use the two years of surveys conducted in 1982 and 1983 and regress reported rental equivalence on total expenditures minus out of pocket housing expenditure, after-tax income, and a set of dummies for age, marital status, family size, and number of earners. We then fit this regression for the earlier waves that do not report a housing service measure. For vacation homes there is no measure of self-reported rental equivalence before 1999. So, for all years, we measure expenditures on vacation homes, like the CE, based on expenditures for mortgage interest, taxes, and maintenance.

We differ from the CE measure of expenditures on vehicles in that we subtract the value of used vehicles that are sold by a household, even when this is separate from any vehicle purchase. (Both our measure and the CE expenditure on vehicles, by using net payments for vehicle purchases, implicitly deducts the value of vehicles traded in as part of purchases.) We also adjust the reported expenditures on food at home in the CE for the 1982 to 1987 waves. Spending on food at home shows a distinct drop for these waves, apparently reflecting a difference in the questionnaire wording from other waves. To adjust for this drop, we increase food at home expenditure by 11% for these waves. This 11% adjustment is derived from a regression for surveys 1980 to 1989 of log food at home expenditures on log after-tax income, log total expenditure, quadratic time trends, and a zero/one dummy variable that equals one for the waves from 1982 to 1987. This adjustment is similar to that in [Krueger and Perri \(2006\)](#).

Our measure of total expenditure will differ from the BLS measure of total expenditure in the CE (TOTEXP) due to these adjustments. It also differs because we treat a set of expenditures (e.g., alimony payments, life insurance, financial fees, social security contributions) as deductions from income, rather than as consumption expenditures. We also treat payments to private pensions as a component of savings, whereas the CE includes these as part of total expenditure.

The CE asks respondents a number of questions on active savings. The BLS employs these responses to publish statistics on net changes in assets and liabilities (see addenda to Current Expenditure Tables, www.bls.gov/cex/home.htm#tables). In each quarterly interview, households report the net change in savings accounts and purchases and sales of stocks and other financial assets. In addition, households report new loans undertaken, including mortgages and home equity loans, and reports equity payments against mortgages and other loans. Households also report purchases and sales of real assets including houses, busi-

nesses, home improvements, and vehicles. They report the net changes in money borrowed or loaned to other households. The CE records the total outstanding credit balances in the first and fourth interviews covering expenditures, which are 9 months apart. We estimate net payments of credit by subtracting the fourth interview’s value from that in the first, and annualize by multiplying by 4/3. (Because all other responses for savings already reflect changes in assets or changes in liabilities, these do not require differencing across interviews.)

Our measure of net changes in assets and liabilities differs in a couple respects from the CE measures reflected in BLS published statistics. The primary difference is that we add payments into private pensions as a form of savings (not as a component of expenditures). Secondly, we do not include net purchases of vehicles, as we treat these as a component of expenditures.

As discussed in the text, the data on new mortgages in the CE raise the question of whether the CE accurately records the net effect of refinancing on savings. We observe a number of reported new mortgages without a corresponding purchase of a house or a significant paying down of an existing mortgage. The CE data imply an average “cash out” percentage of 73 percent from new mortgages not associated with a house purchase, a rate not supported by studies of refinancing. For instance, [Greenspan and Kennedy, 2007](#), find that 13 percent of the value of new mortgages is taken in the form of cash, not used to pay off existing mortgages or to pay related fees. To address this potential measurement error, we construct an alternative measure of household savings that caps the amount of net borrowing (cash out) associated with new mortgages at one third the size of that mortgage. This reduces the average implied cash out ratio of refinanced mortgages to 14 percent, close to the number reported by [Greenspan and Kennedy \(2007\)](#).

Lastly, we create demographic variables for age of the reference person (identified by who owns or rents the residence), the number of household members, and number of household earners, with all variables based on responses in the households final quarterly interval. These variables are used to divide households within each of five income groups into cells, as described in the text.

The impact of sample restrictions

We impose a set of sample restrictions; the impact of these restrictions is reported [Table A 1](#). We begin with 252,758 households for 1980-2010 surveys. We aggregate expenditures for each household across the four interviews—so each household appears only once in the sample. There is considerable attrition across surveys. The BLS responds to attrition by introducing households with the second, or later, survey instrument, so as to keep a balanced panel across interview quarters. Focusing on households that begin with the first survey instrument reduces the potential sample of households to 186,716.

We make the following restrictions on the sample. The 1981 through 1983 surveys include only urban households. For consistency we restrict the samples to urban households for their entirety. This reduces the sample by 9 percent to 170,319. We restrict households to those

with reference persons between the ages of 25 and 64, reducing the sample for by 28 percent to 122,514. In order to contrast household expenditures with income, it is necessary to have measures of expenditures and income over comparable periods. In turn this requires that households participate in all four interviews in order to be present for the income variables in the final interview. This reduces the sample by 31 percent to 84,850. We require households to be “complete income reporters,” which the BLS defines as respondents with values for some major source of income, such as wages, self-employment income, or Social Security income. (Even complete income reporters might not have provided full accounting for all household members.) This restriction reduces the sample by 14 percent to 72,791. We drop households that report implausibly large spending on smaller goods categories. More exactly, we require that households spend less than half of their after-tax income on any category, unless it is housing, food, or vehicle purchases. This restriction reduces the sample by 4 percent to 69,702. (Of those eliminated, 928 households showed negative or zero after-tax income.) Lastly, in order to eliminate outliers and to mitigate the impact of time-varying top-coding, we exclude households in the top and bottom five percent of the before-tax income distribution. (The fraction of households top coded on income fluctuates from about one to just over four percent across survey waves.) This results in a sample of 62,734 households.

Table 1: Trends in Inequality – Ratio of High-Income to Low-Income Respondents

	1980–82	1991–1993	2005–2007	2008–2010	Log Change 1980/82– 2005/07	Log Change 2005/07– 2008/10
Labor Earnings	6.41	8.47	7.88	8.59	0.21	0.09
Before-Tax Income	4.75	5.80	6.40	6.50	0.30	0.02
After-Tax Income	4.21	5.12	5.87	5.92	0.33	0.01
Consumption Expenditures	2.47	2.77	2.93	2.77	0.17	-0.06

Note: High income refers to respondents who report before-tax household income in the 80th through 95th percentiles. Low income refers to respondents in the 5th through 20th percentiles. The elements of the first three columns are the ratio of the average of high-income respondents to the average for low-income respondents, where the averages are taken over the pooled years indicated at the head of the respective column. The last two columns are the log difference the first and third columns and the third and fourth columns, respectively. All variables are converted into constant dollars before averaging. Definitions of each series and sample construction are given in the data section.

Table 2: Engel Curves from 1994–1996 Expenditure Survey

Good Category	1994-96	
	CE Share	Elasticity (SE)
Housing	27.1	0.92 (0.03)
Food at home	12.2	0.36 (0.03)
Vehicle purchasing, leasing, insurance	12.2	1.39 (0.11)
All other transportation	8.1	0.98 (0.04)
Utilities	5.2	0.45 (0.03)
Health expenditures including insurance	4.9	0.99 (0.09)
Appliances, phones, computers, with services	4.8	0.93 (0.04)
Food away from home	4.7	1.41 (0.08)
Entertainment equipment and subscription television	3.8	1.35 (0.06)
Men's and women's clothing	2.6	1.51 (0.07)
Entertainment fees, admissions, reading	2.1	1.93 (0.08)
Cash contributions, not for alimony/support	2.0	1.87 (0.16)
Furniture and fixtures	1.7	1.44 (0.12)
Education	1.6	2.08 (0.28)
Shoes and other apparel	1.5	1.26 (0.08)
Domestic services and childcare	1.5	2.11 (0.23)
Tobacco, other smoking	1.0	-0.40 (0.12)
Alcoholic beverages	1.0	1.36 (0.11)
Children's clothing (up to age 15)	0.9	0.59 (0.17)
Personal care	0.9	1.06 (0.06)

Note: The first column presents each goods average share of total expenditure for 1994–1996, calculated as the cross-sectional average expenditure on each good for each year divided by average total expenditure for that year and then averaged across years. The remaining columns report estimates of each goods expenditure elasticity, with associated standard errors in parentheses. See text for details of sample construction and regression specification. All specifications include demographic control dummies for age, household size, and number of earners.

Table 3: Trends in Consumption Inequality Based on Relative Expenditure Patterns

	(1)	(2)	(3)	(4)
Log Inequality 1980–1982	0.84 (0.07)	0.85 (0.05)	0.84 (0.06)	0.87 (0.06)
Log Change 1980/82–1991/93	0.27 (0.08)	0.24 (0.06)	0.24 (0.06)	0.20 (0.07)
Log Change 1980/82–2005/07	0.40 (0.07)	0.32 (0.07)	0.34 (0.07)	0.29 (0.06)
Log Change 2005/07–2008/10	-0.05 (0.06)	-0.04 (0.06)	-0.02 (0.06)	-0.03 (0.05)
Categories Included	All	All	Without durables	All
Specification	OLS	GLS	GLS	WLS (NIPA Shares)

Note: This table reports the estimated change in income inequality obtained from the second-stage regressions. The estimated parameters in the first row represent log inequality between the high-income and low-income households in 1980–82. The next three columns represent the relative growth in total expenditure for high-income households relative to low-income households for the period specified. See the specification in the text for full details. The first column implements the second stage by OLS; the second column implements the second stage using two-step generalized least squares, where we allow expenditure on each good to have its own residual variance; the third column implements GLS while omitting all good categories containing durables; the last column implements weighted least squares, using the average NIPA for 1980–2010 shares as weights. Standard errors are calculated using a bootstrap with 500 replications.

Table 4: Change in Relative Income-Specific Measurement Error

	(1)	(2)	(3)	(4)
Relative Mis-Measurment 1980–1982				
High Income – Low Income	0.13 (0.07)	0.10 (0.05)	0.10 (0.05)	0.08 (0.05)
Change 1980/82–1991/93	-0.17 (0.08)	-0.12 (0.06)	-0.11 (0.06)	-0.10 (0.06)
Change 1980/82–2005/07	-0.35 (0.08)	-0.24 (0.07)	-0.23 (0.07)	-0.16 (0.06)
Change 2005/07–2008/10	0.06 (0.07)	0.03 (0.06)	0.04 (0.05)	0.00 (0.04)
Categories Included	All	All	Without durables	All
Specification	OLS	GLS	GLS	WLS (NIPA Shares)

Note: This table reports the change in the estimated income-specific measurement error for high-income respondents relative to low-income respondents: $\phi^5 - \phi^1$ from equation (6). The specification for each column is the same as in table 3. The first row is the level for the period 1980-82, and the next three rows report the change over the indicated period. Standard errors are calculated using a bootstrap with 500 replications.

Table 5: Trends in Consumption Inequality: Nonlinear Engel Curves

	(1)	(2)	(3)
Log Inequality 1980–1982	0.83 (0.06)	0.84 (0.06)	0.85 (0.06)
Log Change 1980/82–1991/93	0.31 (0.07)	0.27 (0.06)	0.25 (0.07)
Log Change 1980/82–2005/07	0.36 (0.07)	0.30 (0.06)	0.29 (0.07)
Log Change 2005/07–2008/10	-0.05 (0.06)	-0.02 (0.05)	-0.03 (0.06)
Specification	Restricted NLS	Restricted NGLS	Unrestricted GLS

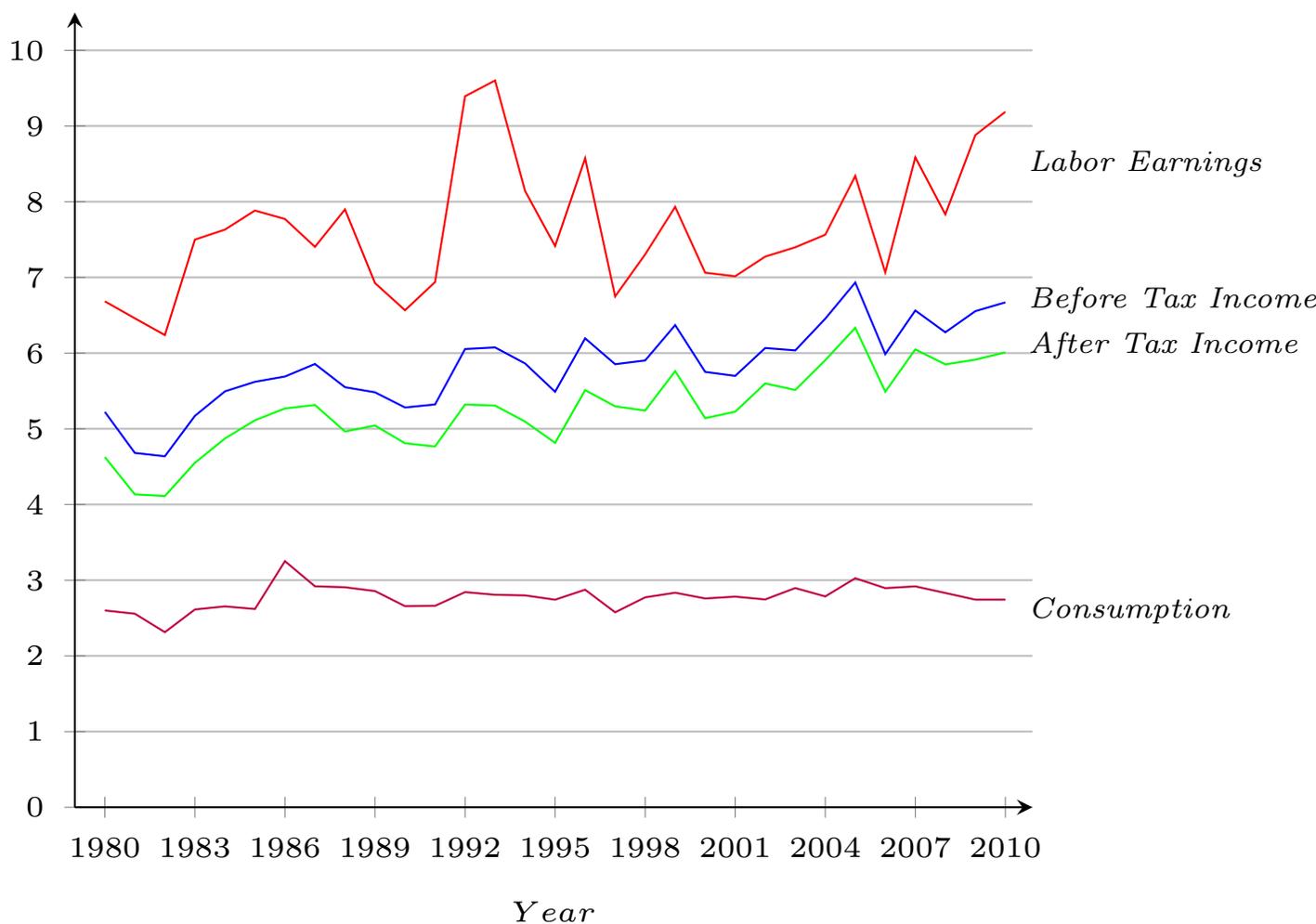
Note: This table depicts the results from our nonlinear specification (5') described in section 4. The rows of the table correspond to the rows in table 3. The first column reports the estimates using nonlinear least squares; the second column implements generalized nonlinear least squares, weighting by the inverse good-year residual variances; the third column does not impose the restriction that the coefficient on $\beta_{2,j}$ is the square of the coefficient on $\beta_{1,j}$ in (5'), and is estimated using generalized least squares. Standard errors are calculated using a bootstrap with 500 replications.

Table A 1: Sample Construction

Total Number of Households	252,758
Households who enter at “first” interview”	186,716
After Sample Restriction:	
Urban	170,319
Ages 25 to 64	122,514
Full-year of Interview Coverage	84,850
Complete Income Reporter	72,791
No Expenditure Outliers	69,702
Truncate Before-Tax Income: 5-95 ptile (Final Sample)	62,734

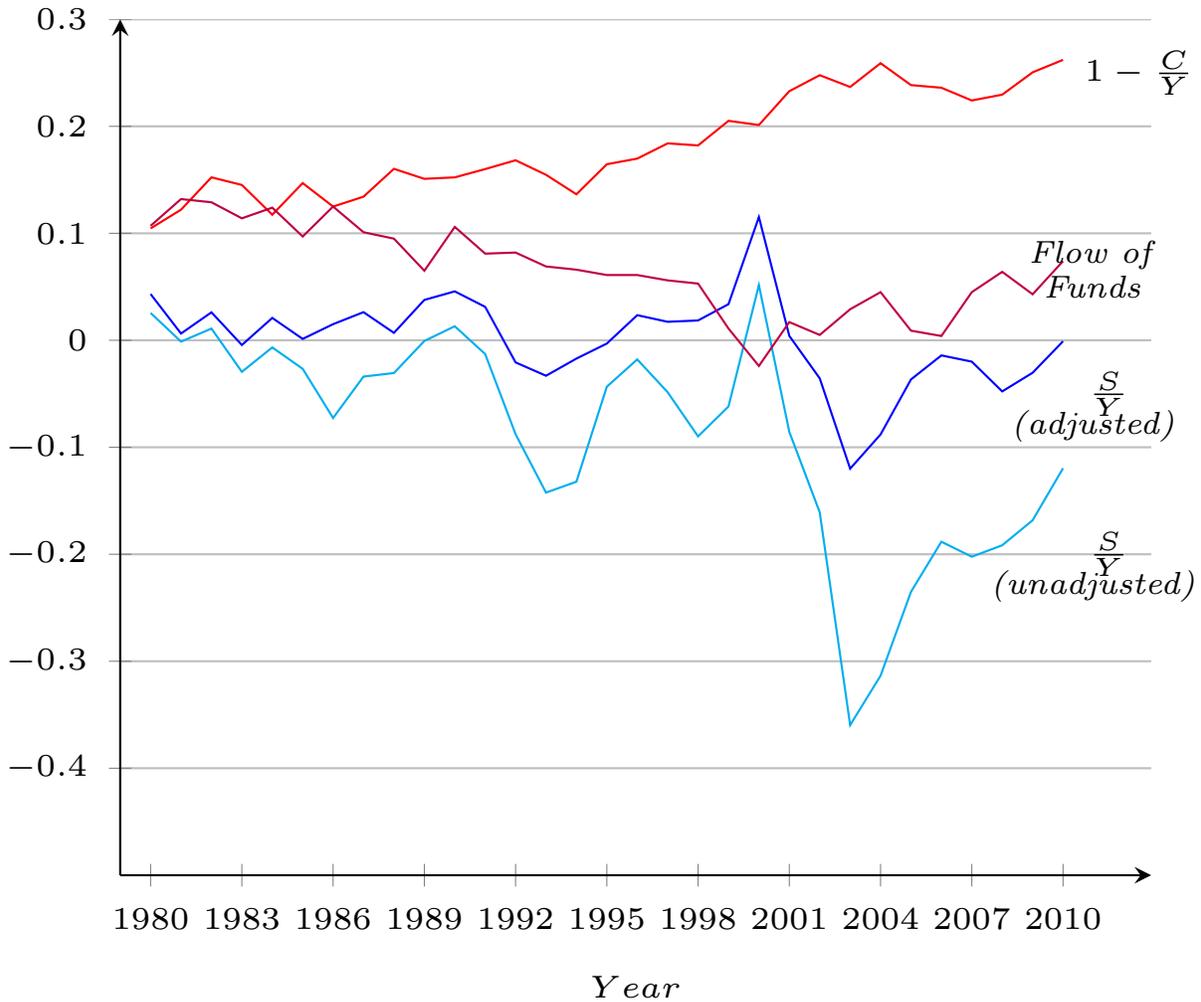
Note: This table reports the sample size after each restriction. The first row reports the original CE sample obtained from the BLS. Each sample restriction is discussed in the data appendix. The final row represents the sample used in the analysis.

Figure 1: Trends in Inequality



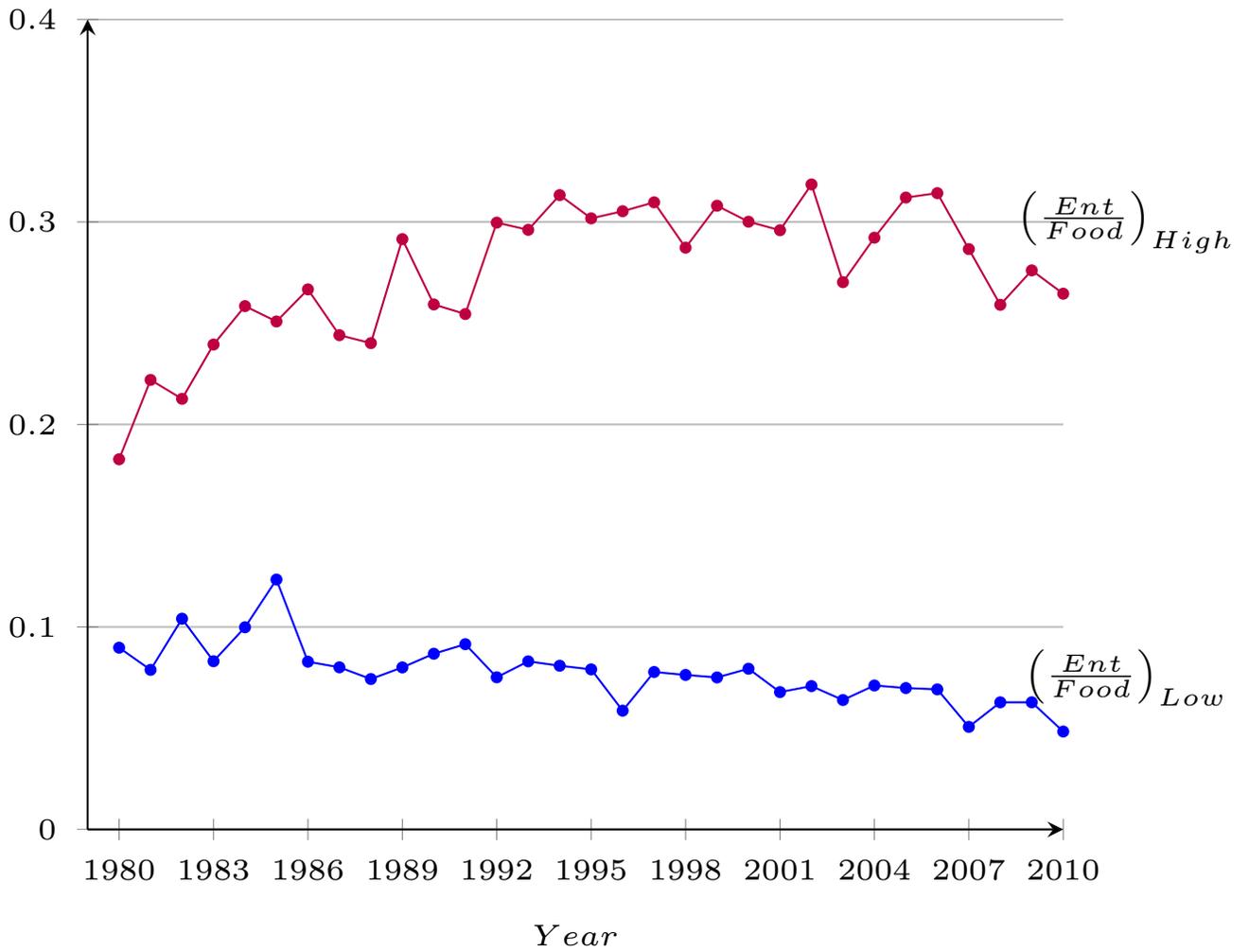
Note: This figure depicts the ratio of high-income to low-income respondents reported labor earnings, before-tax income, after-tax income, and consumption expenditures. High income refers to respondents who report before tax household income in the 80th through 95th percentiles. Low income refers to respondents in the 5th through 20th percentiles. Definitions of each series and sample construction are given in the data section.

Figure 2: Mean Saving Rates



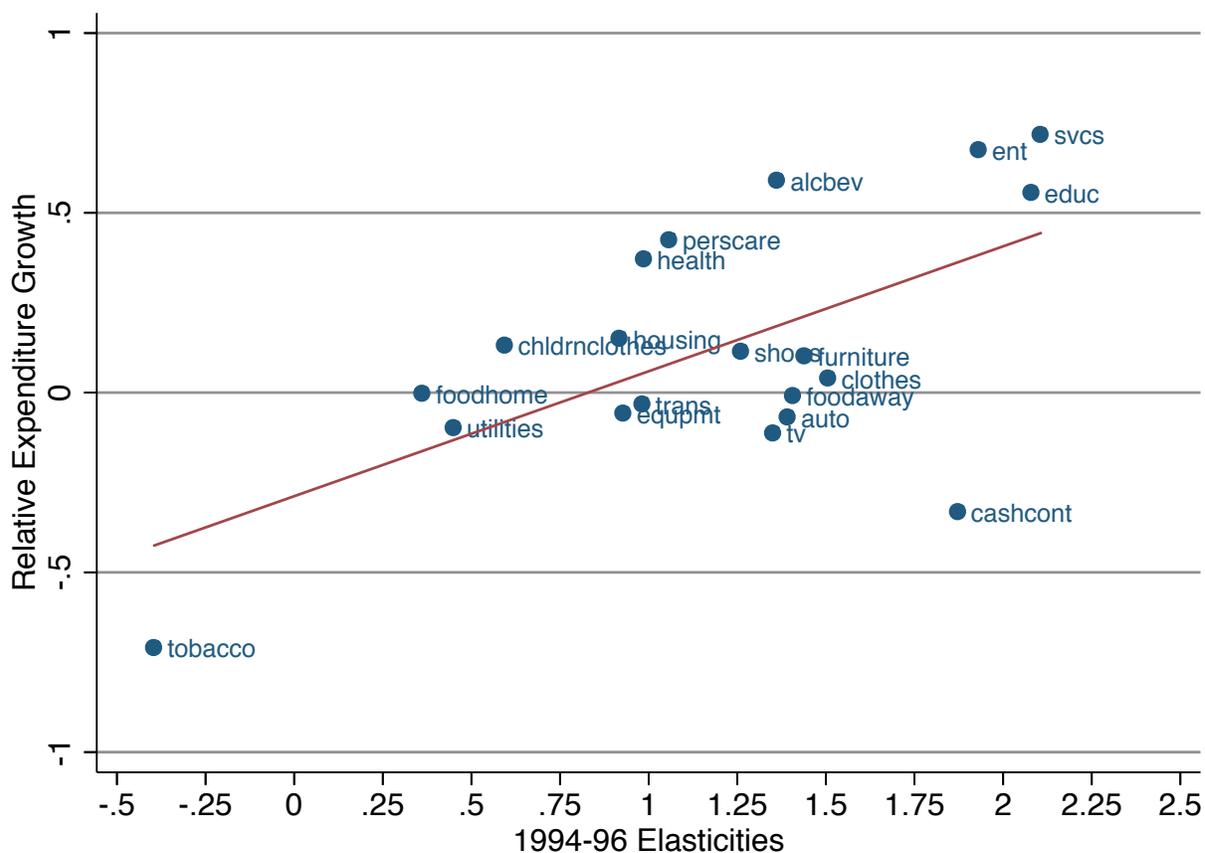
Note: This figure depicts the mean savings rates. The line labeled $1 - C/Y$ refers to implied savings computed as after-tax income minus reported consumption expenditures. The line labeled “Flow of Funds” is the flow of funds aggregate private savings rate out of disposable income. The lines labeled S/Y refer to average reported savings divided by average reported after tax income. Adjusted and unadjusted refer to whether we adjust reported new mortgages, as described in the data section of the text. Definitions of each series and sample construction are given in the data section of the text.

Figure 3: The Ratio of Entertainment to Food Expenditure – High-Income and Low-Income Households



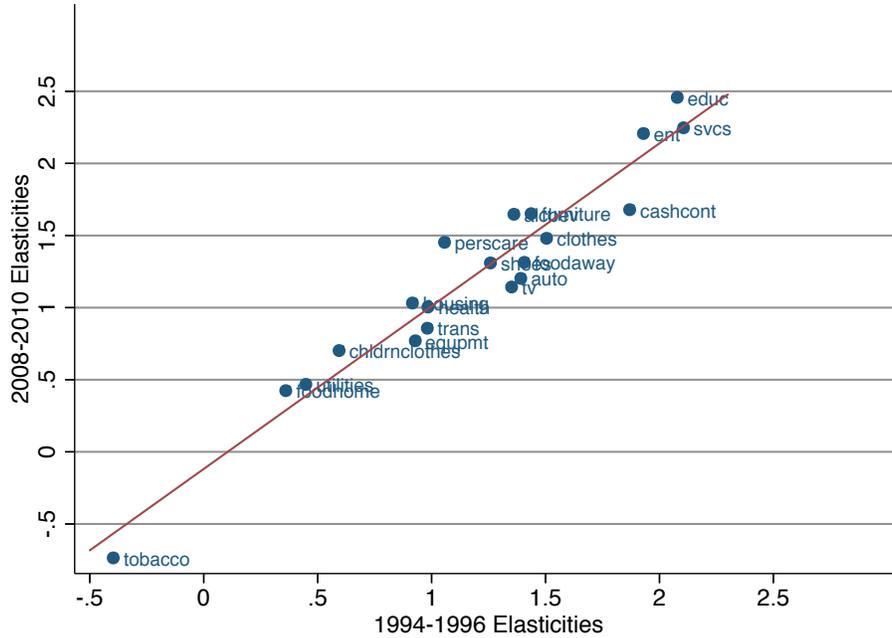
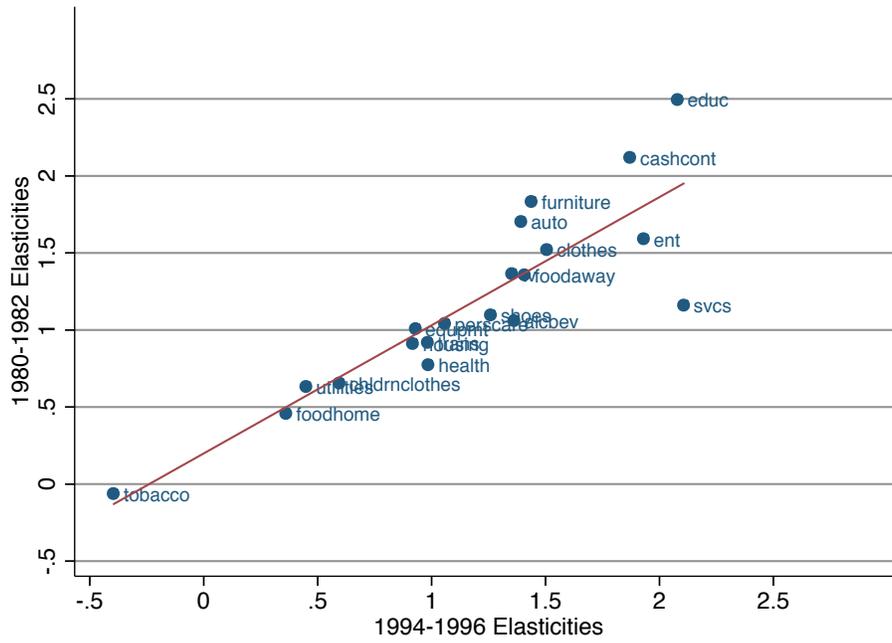
Note: This figure depicts the ratio of spending on nondurable entertainment to food at home for high- and low-income households.

Figure 4: Relative Expenditure Growth for 20 Goods



Note: This figure is a scatter plot of relative (high- versus low-income) expenditure growth over the sample period for each good versus expenditure elasticity. The vertical axis depicts the difference across high-income and low-income households in the log growth in expenditure for each good between 1980/82 and 2008/10. The horizontal axis is each good's estimated expenditure elasticity from table 2. The slope of the scatter plot's regression line is 0.35.

Figure 5: Stability of Expenditure Elasticities over Time



Note: The top panel depicts a scatter plot of the 20 expenditure elasticities estimated using the 1980-82 CE sample versus our benchmark 1994-96 first stage reported in table 2. The fitted regression line has a slope of 0.83 and an R^2 of 0.76. The bottom panel replaces the 1980-82 estimates with estimates using 2008-10 surveys. The fitted line has slope 1.13 and an R^2 of 0.93.