Has Consumption Inequality Mirrored Income Inequality?[*]

Mark Aguiar   Mark Bils

December 23, 2013

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.

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 1980-82, excluding health care, equaled 86 percent of that implied by NIPA. By 2008-2010 this ratio fell to only 66 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 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,

---

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

[2] For 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.

[3] We 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.

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.

Our identification assumption is that, once we control for systematic mis-measurement at the good-time and income-time level, the residual measurement error at the household-good-time level is classical. In particular, it is orthogonal to that good's expenditure elasticity conditional on income group. This encompasses a wide range of residual measurement error. Nevertheless, there are scenarios that violate this assumption. For instance, suppose that from 1980 to 2010 high-income households began systematically to under report spending on luxuries, but not necessities, whereas low-income households began under reporting spending on necessities, but not luxuries. Under this scenario, our approach would understate the true relative shift in spending by richer households toward luxuries, thereby understating the rise in consumption inequality. Under the reverse scenario (high-income stop reporting their spending on necessities, low-income stop reporting luxuries), our approach will overstate the rise in consumption inequality. We discuss this identification assumption (and when it may fail) at length at the end of Section 3.1.

To illustrate our approach, 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 25 percent relative to that for food at home between 1980-

82 to 2008-10. Based on our estimated Engel elasticities, this implies an increase in total expenditure of 18 percent (see Figure 3). By contrast, the bottom income quintile reported that entertainment expenditures declined by 40 percent relative to that reported for food at home, suggesting a decline in total expenditure of 29 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 47 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 a little more than 30 percent between 1980 and 2010, roughly 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. This implies that 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 17 percent between 1980 and the early-1990s, by an additional 18 percent through 2007, for a total increase of 35 percent, closely tracking the profile of income inequality. For the Great Recession we estimate a small reduction in consumption inequality of 4 percent.

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 in-

equality significantly in the 1990s. 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. 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.

Browning and Crossley (2009) share our interest in measurement error and also employed an Engel-curve approach. Specifically, Browning and Crossley (2009) argue that multiple noisy measures can dominate a single, relatively accurate measure of household expenditure, building on the insight that the covariance of multiple measures may mitigate measurement error. For noisy measures they suggest using two categories of spending, each with Engel curve elasticities of one, so that the expected covariance of the two measures will be close to the variance of total expenditure. As an alternative, Browning and Crossley suggest employing a luxury and a necessity, rather than two luxuries or two necessities, again so the covariance of the two spending variables will be close to the variance of total expenditure. As an application they employ spending on food, including that at restaurants, as a necessity and entertainment expenditure as a luxury. 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. In particular, our methodology is designed to measure

---

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

consumption inequality, which is not a focus of the Browning and Crossley analysis.

The use of Engel curves to infer total expenditure is often used when only a subset of expenditures is reported. For instance, Blundell et al. (2008) (BPP) use the CE to estimate the demand for food conditional on prices, total nondurable expenditure, and demographics, and then invert this to map the PSID's food expenditure series into an imputed measure of nondurable consumption. In addition to a related methodology, BPP shares our interest in the cross-section of consumption. BPP use income measures from the PSID to argue that the variance of both permanent and transitory income shocks increased in the 1980s. This is consistent with several other studies based on earnings data (for example, Gottschalk and Moffitt (1994, 2009); Heathcote et al. (2010)). They use this finding to reconcile the gap between consumption and income inequality between 1980 and 1992, employing a specification that allows the data to determine the extent of insurance of permanent and transitory income shocks. Their estimates suggest that there is partial insurance for permanent shocks and almost complete insurance of transitory shocks. The estimates suggest somewhat more insurance against permanent income shocks than that implied by the standard incomplete markets permanent income model, particularly for highly educated consumers (see Kaplan and Violante (2010) on this point as well). Our measures of consumption inequality using reported CE data are consistent with BPP's imputed measures. To the extent that reported consumption is systematically mis-measured, our corrected measures of consumption inequality suggest less insurance of income shocks than that implied by reported expenditure. Alternatively, the PSID measures of income may provide an incomplete picture of the increase in permanent income risk. In this regard, several recent studies using administrative data have found a larger role for permanent income risk in explaining the increase in income inequality (for example, Kopczuk et al. (2010); Dahl et al. (2011); DeBacker et al. (2013); Monti and Gathright (2013)). While we do not take a stand on the permanent versus transitory nature of income inequality, we contribute to this literature by providing a methodology that adjusts measured consumption inequality for systematic measurement error, which could be used to shed light on the nature of uninsurable income risk.

Several papers find a smaller rise in consumption inequality than in income in other countries (for example, see the special *Review of Economic Dynamics* issue of January 2010 for studies of inequality in several countries). These studies may appear to contrast with our result that income and consumption inequality mirror each other in the US. However, the studies of other economies are not necessarily inconsistent with our findings, given that there is no a priori reason that the underlying income dynamics are the same in all countries. In particular, the permanent-income paradigm may explain the difference between the US

and Europe. For example, Jappelli and Pistaferri (2010) document that in Italy between 1980 and 2006, transitory idiosyncratic income shocks rather than greater dispersion in the permanent wage structure explains the majority of the rise in income inequality. Similarly, using income data from the British Household Panel Data for 1991 to 2003, Blundell and Etheridge (2009) document a decline in the permanent component of income inequality relative to its transitory component.

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 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 sec-

ond 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. 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 2.2.

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.

When documenting differences across income levels, 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.

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 bin to the mean for the bottom income bin. Keep in mind that the allocation of respondents into the high and low-income groups is based on before-tax income, and so the groups 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

---

[5] Jonathan 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.

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). 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

---

[6] The 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.

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.

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

---

[7] Specifically, 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.

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 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 it 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

---

[8] Specifically, 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-(-.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: $(-0.24)(0.75) = -0.18$.

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

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

3.1   Econometric Approach

To set notation, let $h = 1, ..., H$ index households, the unit of observation in the CE; $i = 1, ..., I$ denote the $I = 5$ income groups; $j = 1, ...J$ index our $J = 20$ goods; and let t index time (year). x_{hjt} denotes reported expenditure on good j at time t by household h. X_{ht} denotes total expenditure at time t by household 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. In particular, let x_{hjt}^* denote the true expenditure, and

$$x_{hjt} = x_{hjt}^* e^{\zeta_{hjt}}. \tag{1}$$

We can decompose ζ_{hit} into three components:

$$\zeta_{hjt} = \psi_t^j + \phi_t^i + v_{hjt}. \tag{2}$$

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-household specific measurement error (e.g., food expenditures of household h 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} across households to be zero for all t. Our identifying assumption is that v_{hjt} is classical measurement error; in particular, it is

14

independent of the characteristics of good j and household 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 estimate a log-linear approximation to the Engel curves. Of course, Engel curves cannot be log-linear globally 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. We estimate a second-order expansion as a robustness check in Section 4. 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 the first-order expansion in true expenditure satisfies:

$$\ln x_{hjt}^* - \ln \bar{x}_{jt}^* = \alpha_{jt}^* + \beta_j \ln X_{ht}^* + \mathbf{\Gamma}_j \mathbf{Z}_h + \varphi_{hjt}, \tag{3}$$

where $\bar{x}_{jt}^*$ is the average expenditure on good j in year t across all households. The term $\mathbf{Z}_h$ is a vector of demographic dummies based on age range (25-37, 38-50, 51-64), number of earners (<2, 2+), and household size (≤ 2, 3-4, 5+). We allow the coefficient vector on demographics $\mathbf{\Gamma}_j$ to vary across goods.[10] The variable α_{jt}^* reflects the expansion point of average total expenditure. Note that first-order good-time specific demand shifters, such as the effect of relative prices, are captured by mean expenditure on each good, a point we discuss in the next paragraph. 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. Note as well that β_j do not have a time subscript, reflecting the assumption that the expenditure elasticity for each good is stable over time. We explore the stability of β_j and robustness to other potential mis-specification issues in Section 4.

---

[10]We have explored an extension in which demographic taste-shifters are allowed to vary by income as well as good. Specifically, we interact the demographic dummies Z_h with household log after-tax income. The results are nearly the same. In particular, in our benchmark WLS specification, we estimate inequality has increased by 0.35 between 1980/82 and 2005/07 (Table 3 Column 2). The comparable estimate with demographic*income interaction is 0.32. The estimate for the change during the Great Recession is -0.04 in both specifications.

An important concern is whether shifts in spending over time are driven by changes in relative prices. Note that relative prices do not appear explicitly in (3). This reflects that the first-order price effects are embedded in the good-time intercept α_{jt}^*. More precisely, the first-order effect of changes in prices (the cumulation of own price effects and the effects due to cross-price elasticities) on demand for good j at time t are good-time specific effects, and thus captured by the good-time intercept α_{jt}^*. Our specification therefore accommodates changes in demand over time that are driven by shifts in relative prices. A distinct but related question is whether the expenditure elasticity β_j depends on relative prices. Such an interaction is not addressed by the good-time specific intercept. However, to the extent that movements in relative prices over time lead to movements in expenditure elasticities, this issue falls under the question of the stability of expenditure elasticities over time. We discuss this possibility in detail in Section 4. That section also discusses complications due to relative price effects that may arise in a quadratic specification, as noted by Banks et al. (1997).

In terms of observables, equation (3) can be re-written

$$\ln x_{hjt} - \ln \bar{x}_{hjt} = \alpha_{jt} + \beta_j \ln X_{ht} + \mathbf{\Gamma}_j \mathbf{Z}_h + u_{hjt}, \tag{4}$$

where the residual term includes income-specific systematic measurement error ϕ_t^i as well as idiosyncratic taste shocks φ_{hjt} and mis-measurement v_{hjt}:

$$u_{hjt} = \phi_t^i + v_{hjt} + \varphi_{hjt}. \tag{5}$$

Note that the good-time specific measurement error ψ_t^j is differenced out by including mean observed expenditure on the left hand side, leaving $\alpha_{jt} = \alpha_{jt}^* + \beta_j \left(\ln X_{ht}^* - \ln X_{ht} \right)$.

We estimate expenditure elasticities β_j 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.

There are a number of issues that arise in estimating (4). There are cases in which household expenditure on a particular good may be zero, making the log specification inappropriate. In our estimation, we replace $\ln x_{hjt} - \ln \bar{x}_{jt}$ with the percentage deviation from average expenditure on that good in that year: $\tilde{x}_{hjt} \equiv \frac{x_{hjt} - \bar{x}_{jt}}{\bar{x}_{jt}}$. These are equivalent representations in a first-order expansion around average expenditure, but raise the concern that

households with large deviations may influence the estimation in one or the other specification.[11] We have verified that the analysis does not depend on whether we use log total expenditure as the independent variable or the percent deviation from that year's average; we report results using log total expenditure for ease of discussion. We defer discussion of higher order terms for total expenditure until Section 4.

A second 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 report results using two alternative approaches to instrumenting. The first exploits the fact that total expenditure reflects permanent income and will thus be correlated with current income. Specifically, we instrument total expenditure with dummies for the household's income group as well as the continuous variable log after-tax income. The second approach exploits the fact that households in the CE report total expenditure in separate interviews for each of four quarters. This allows us to divide each households spending into that over its first two quarters versus its final two. We then estimate the Engel elasticities from (4) based on the expenditures from the final two quarters, instrumenting for household total expenditure with its total expenditure over the first two quarters. This second approach exploits that total expenditure is a natural proxy for permanent income. As we shall see in the next sub-sections, the two approaches yield nearly identical results.

These IV specifications are designed to address classical measurement that is uncorrelated with income or lagged consumption. As modeled above, there may be systematic measurement error that is common across households within an income group or common to a household over time. 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

---

[11]In a previous version, we averaged expenditure within income-demographic cells and then explored log expenditure on each good across cells. The results are comparable and reported in Aguiar and Bils (2011).

estimate of how consumption inequality evolved over the years of the survey. We first adjust expenditure for demographics. Specifically, let

$$\hat{x}_{ijt} \equiv \tilde{x}_{hjt} - \hat{\Gamma}_j \mathbf{Z}_h,$$

where $\hat{\Gamma}_j$ is the estimate of Γ_j from (4) . Using (3), we have

$$
\begin{aligned}
\hat{x}_{hjt} &= \alpha_{jt} + \phi_t^i + \beta_j \ln X_{ht}^* + \varphi_{hjt} + v_{hjt} \\
&= \alpha_{jt} + \phi_t^i + \beta_j \ln X_{it}^* + \beta_j \left(\ln X_{ht}^* - lnX_{it}^* \right) + \varphi_{hjt} + v_{hjt} \\
&= \alpha_{jt} + \phi_t^i + \beta_j \ln X_{it}^* + \varepsilon_{hjt},
\end{aligned}
\tag{6}
$$

where the middle line has substituted in the average log expenditure for income group i, which is the focus of our analysis. The residual term is $\varepsilon_{hjt} = \beta_j \left(\ln X_{ht}^* - lnX_{it}^* \right) + \varphi_{hjt} + v_{hjt}$.[12]

To implement (6), we regress $\hat{x}_{hjt}$ on a vector of good-time dummies (whose coefficients correspond to α_{jt}), a vector of income-time dummies $D_{i,t}$ (whose coefficients correspond to ϕ_t^i), and the interaction of income-time dummies and the first-stage estimates $\hat{\beta}_j$. The coefficient on this last interaction term will be the respective estimates of $\ln X_{it}^*$ for each income group. To address the issue of normalization, we estimate expenditure relative to the lowest income group $(i = 1)$. That is, we have a consistent estimate of consumption inequality: $\delta_{it} = \ln X_{it}^* - \ln X_{1t}^*$. 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.[13]

Our key identifying assumption is that idiosyncratic measurement errors and preference shocks are not systematically related to the expenditure elasticities across goods. More exactly, we require that v_{ht}, the idiosyncratic component of the mis-measurement of good j for household h in income group i at time t, and the corresponding idiosyncratic taste shock φ_{hjt}, both be orthogonal to the expenditure elasticity β_j conditional on income group. This implies that ε_{hjt} is independent of $\beta_j * D_{i,t}$. 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}.[14] 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 total expenditure of group i increases relative to that of group

---

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

[13] Specifically, we draw with replacement from the micro data for all years and re-estimate both stages.

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

i', that increase will fall disproportionately on luxuries.

Before proceeding, it is useful to discuss the strengths and limitations of our identification procedure. The second stage uses our first-stage estimates $\hat{\beta}_j$ as generated regressors. We require that these estimates (interacted with income-group dummies D_{it}) be orthogonal to ε_{hjt} in the second stage. Note that we never use the same time period for the first and second stages to avoid correlated sampling error arising from our generated regressors. Aside from the generated regressors, we still have the question of whether the residual is correlated with the true β_j. The presence of $\beta_j (\ln X_{ht}^* - \ln X_{it}^*)$ in ε_{hjt} is not an issue, as this will be orthogonal to our regressor by definition.[15] The important identification issues arise due to the presence of the good-time-household mis-measurement, ν_{hjt}, and whether this mis-measurement is orthogonal to $\beta_j * D_{it}$.

We now discuss some plausible scenarios and evaluate whether they violate our identification assumption. One scenario is that shifts in expenditure on good j are the result of relative price movements. As discussed previously, such changes are accommodated by the good-time specific intercept α_{jt}.[16] Specifically, this intercept captures the first-order price elasticities that are common across households at a point in time. A particular concern would be if relative price changes induced shifts in spending between luxuries and necessities that differed across high- and low-income households. For example, suppose that the relative prices of luxuries increased and high-income households exhibited an inelastic price response while low-income consumers exhibited an elastic response. The price relative changes would then cause a shift in spending on luxuries for richer households, relative to poorer, beyond that created by their relative changes in total expenditures. Our specification does not control for this heterogeneity in price elasticities, and our good-time intercept will only pick up the average price effect. Nevertheless, we can say something about the likelihood of this scenario. The starting point for this possible concern is that relative price movements in our sample period are correlated with expenditure elasticities, our second-stage regressors. We do not see this in the data. Specifically, we have constructed price indices for our twenty goods and find that the correlation between the change in price between 1980 and 2010 with the good's respective estimated price elasticity is small and not significantly different from zero. In particular, weighting categories by their average spending in NIPA, the correlation is -0.13 with a p-value of 0.58.

---

[15]That is, $\mathbb{E}[\beta_j D_{it} * \beta_j (\ln X_{ht}^* - \ln X_{it}^*) | \beta_j, D_{it}] = \beta_j^2 D_{it} (\mathbb{E}[\ln X_{ht}^* | D_{it}] - \ln X_{it}^*) = 0$, where the last equality follows from the definition of $\ln X_{it}^*$ as average expenditure within income group i.

[16]More precisely, this rests on the assumption that all households at a point in time face a common set of prices. We maintain this standard assumption, but acknowledge the caveat that prices may vary across households due to the ability to search. See Aguiar and Hurst (2007) for an empirical exploration of this phenomenon.

One potential source of measurement error is that household j experienced a rapid change in expenditure on good j, but reports some smoothed average of expenditure over a longer time frame than the current or previous quarter. If this change is idiosyncratic to household h, this does not violate our orthogonality condition. More to the point, suppose household h had an increase in permanent income and increased expenditure on all goods, with the increase governed by the expenditure elasticity β_j. If the household reports a smoothed expenditure number, it will under-report expenditure on all goods, but more so for luxuries. Nevertheless, as long as the change in permanent income is idiosyncratic, it will average out within an income group, and therefore not violate orthogonality with $\beta_j * D_{it}$.

However, suppose that all households in income group i experienced permanent income growth, and mis-reported expenditure on good j by averaging over several periods. For example, suppose high-income household experience rapid income growth relative to poor households, which in turn induces rapid growth in expenditure. This growth will be biased towards luxuries by definition. If households smooth their responses over time, the high-income households will under-report expenditure on all goods, but more so for luxuries. This will lead to a violation of our identification. More precisely, the under-reporting that is biased towards luxuries of the rich will lead us to under-estimate inequality at a point in time. This is an intuitive source of bias. If households are averaging their reported expenditure on all goods over a longer time frame, we will estimate a level of consumption inequality that holds on average over the time frame. If inequality is increasing and households are reporting long-run averages, we will understate true inequality at a point in time (and vice versa if inequality is declining).

These examples provide a sense of when our identification holds and when it fails. It also gives a sense of possible bias; namely, to the extent that mis-measurement leads the high-income households to under-report luxuries relative to necessities (and this mis-measurement is greater for high-income than low-income households), our second stage will underestimate true inequality. If the reverse is true (that is, the rich under-report necessities relative to luxuries or the poor over-report necessities relative to luxuries), we will overstate inequality.

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 first column of elasticities uses log income and dummies for income group to instrument for total expenditure, while the second column of elasticities

uses the initial two interviews' total expenditure to instrument for the final two quarters' total expenditure. 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 correlation coefficient between the two sets of elasticities is 0.93, and each column has a similar amount of dispersion (standard deviation of 0.50 and 0.47, respectively) indicating consistency across specifications.

Both specifications indicate that tobacco has a negative elasticity, while domestic services, education, and entertainment are relative luxuries. Consistent with other studies, food at home has a fairly low expenditure elasticity (0.37 and 0.47), while food away from home has a high elasticity (1.33 and 1.32). Housing services, our largest expenditure category, has an expenditure elasticity of 0.92 and 0.93 in the two respective specifications.

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. For context, the ratio of mean entertainment expenditure to mean food at home expenditure rises slightly, from 0.15 to 0.16, over this period.

The relative shift in expenditure towards a luxury for the high-income households implies a sharp increase in total expenditure inequality. For a sense of how these shifts are informative by total expenditure inequality, consider the increase in the high-income line in Figure 3. On average, high-income expenditure on entertainment increased 48 percent faster than mean expenditure on entertainment, but increased only 4 percent faster than the mean for food at home. Note that comparison across goods within an income class addresses income-group specific multiplicative measurement error (ϕ_t^i), while comparison to mean expenditure on each good in each year addresses good-specific shifters due to price effects and good-specific mis-measurement ($\alpha_{j,t}$). Given the respective expenditure elasticities of 1.74 and 0.37 (Table 2) and temporarily ignoring demographic shifts, a simple calculation suggests an increase of log expenditure for high income households relative to the mean of $(48 - 4)/(1.74 - 0.37) = 32$ points. For low-income households, their expenditure on entertainment fell 16 percent relative to the mean, while their expenditure on food at home increased 4 percent relative to the mean. This suggests a decrease in relative total expen-

diture of 15 points. On net, relative expenditure on these two goods for these two income groups suggests an increase in total expenditure inequality of 47 log points.

This estimate is a noisy measure given the presence of idiosyncratic shocks at the income-good level. A more precise estimate can be obtained using all goods. Figure 4 provides a sense of how the identification scheme works. The figure is a scatter plot of relative consumption growth on each good versus the respective expenditure elasticities. Specifically, consider food at home (the point labelled "foodhome"). The horizontal coordinate is 0.37, the estimated expenditure elasticity for food at home from Table 2. Controlling for demographics, high-income households spent 37 percent more on food at home than low-income households in 1980-82, and 43 percent more in 2008-2010. This relative shift of 0.06 is depicted on the vertical axis of Figure 4. Similarly, the point labelled "ent" for entertainment refers to an estimated elasticity of 1.74 and a relative growth across time and income groups of 62 percent. A fitted line between only these two points would have slope of 0.42, which is the demographically adjusted counter-part of the 0.47 derived in the previous paragraph. Using all 20 goods, the fitted line that is depicted has a slope of 0.425. This suggests that an increase in relative total expenditure of 42.5 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 and the first set of elasticity estimates from Table 2. The first row reports the estimated log inequality in the pooled period 1980–1982, which is 0.86. 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 0.04 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 48 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 6 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 42.5 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 are excessively driving the results. Column (2) 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 down weight 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.[17] 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 17 percent for the early period, 35 percent through 2007, and a decline over the last 5 years of 4 percent.

Column (3) performs the same WLS regression but excludes categories that contain durables.[18] 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 20 percent increase in inequality in the first decade of the sample, and 42 percent through 2007.

The final column of Table 3 uses the alternative elasticities from Table 2 to estimate the WLS specification of Column (2). The alternative elasticities yield slightly lower initial inequality (0.82 versus Column (2)'s 0.90), and a slightly greater increase in inequality over time. Specifically, an increase of 0.21 in the first decade and 0.44 through 2007, compared

---

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

[18] Specifically, from the goods listed in Table 2, we exclude vehicles, appliances, furniture, and entertainment equipment.

to 0.17 and 0.35 in Column (2). The fact that the second stage yields similar estimates is not a surprise given the high correlation of the elasticities in Table 2.

The second-stage estimation uses all five income categories; thus it yields 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 estimates also suggest that the vast majority of the increase in consumption inequality occurred between the high and middle groups. In particular, the WLS estimates indicate the 31 point increase in the 90-10 consumption ratio can be attributed to an increase of .29 in the 90-50 ratio and .02 in the 50-10.

Table 4 reports the estimates for income-specific measurement error, ϕ_t^i. In particular, it reports 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 small and positive, implying that the reported consumption inequality is slightly 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. 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.

Before exploring the robustness exercises of the next section, we revisit one of the motivations for this study: the growing discrepancy between aggregate expenditure in the CE and that reported in the national income accounts. In particular, the ratio of reported CE expenditure to that reported in NIPA fell from 0.86 to 0.66 between 1980/82 and 2008/10.[19] That is, real NIPA consumption expenditure increased by 37 percent over this period, while that in the CE increased only 10 percent, generating a difference of 27 log points. This raises the question of whether correcting for systematic measurement reduces or eliminates this increased discrepancy.

---

[19]For this calculation, we omitted healthcare expenses from both the CE and NIPA, as medical expenses represents a major difference in coverage between the two measures.

Our methodology yields a corrected measure of consumption inequality, but does not speak directly to mean expenditure. In particular, our methodology omits one income group and estimates relative expenditure for the remaining income groups. To move from relative expenditure to aggregate we need to take a stand on expenditure for the omitted group. For illustrative purposes, we report some simple calculations along these lines. First, suppose expenditure is correctly reported for the lowest-income group, which in our sample experienced growth of 7 percent. If this is the case, then implied mean expenditure across all households in our sample increased by 20 percent between 1980/82 and 2008/10, or double that estimated directly from reported expenditure.[20] Thus correcting for systematic measurement error can reduce the discrepancy by roughly 40 percent (10/27). Alternatively, we can ask what expenditure growth of low-income households is required to close the entire gap of 27 points. The answer to this question is 24 percent. That is, the implied mis-measurement of low-income household expenditure growth is 17 percent (24 minus the observed 7 percent), while the aggregate mis-measurement relative to NIPA is 27 percent. While these calculations require taking a stand on true expenditure for one income group, they imply that correcting for changes in systematic measurement error may plausibly bridge a significant part of the growing CE/NIPA expenditure gap.

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 specification assumes that 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 consider:

$$\ln x_{hjt}^* - \ln \bar{x}_{hjt}^* = \alpha_{jt}^* + \beta_{1,j} \ln X_{ht}^* + \beta_{2,j} \left(\ln X_{ht}^*\right)^2 + \mathbf{\Gamma}_j \mathbf{Z}_h + \varphi_{hjt}. \tag{2$'$}$$

---

[20]This number is obtained using WLS estimates of $\delta_{i,t}$, the log-difference between expenditure of income group i and the lowest-income group at time t. These estimates imply that the average expenditure increased by 13 log points more than the omitted low-income group between 1980/82 and 2008/10.

As in the benchmark model (3), 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 incorporating the second-order term $(\ln X^*_{ht})^2$. A few issues arise in the quadratic specification. Recall that in our first stage, we instrument for total expenditure to address the classical errors in variables problem. The addition of the squared term implies that 2SLS is no longer appropriate. Hausman et al. (1995) address exactly this issue in the context of Engel curves, and we follow their methodology. Specifically, we divide a household's four interviews into two measures of total expenditure (as in the second specification reported in Table 2). Hausman et al. show how repeated measurements of total expenditure for each household can be exploited to correct for measurement error in our quadratic first-stage.

A second concern was identified by Banks et al. (1997) in their important analysis of quadratic Engel curves. In particular, Banks et al. argue that $\beta_{2,j}$ will in general depend on relative prices. Recall from the benchmark specification that we raised the question of whether the first-order term $\beta_{1,j}$ depends on relative prices. We explicitly deal with this possibility in the next subsection. Banks et al. show that if the linear term is price invariant then the second-order term will not be, except in knife-edge cases. To proceed in such an environment, let $b(\boldsymbol{p})$ be an aggregate price index constructed from the price vector $\boldsymbol{p} = (p_1, ..., p_J)$. Then the coefficient on $(\ln X^*_{ht})^2$ can be generalized to $b(\boldsymbol{p})\beta_{2,j}$. As in Banks et al., we consider a common price index $b(\boldsymbol{p})$ for all goods j, but one that will vary over time due to price movements.[21] If we normalize $b(\boldsymbol{p}) = 1$ at 1994/96 prices, then the first stage estimation on 1994-96 data yields the $\beta_{2,j}$ implied by ($2'$) evaluated at 1994/96 prices. However, the second stage must accommodate changes in the price index $b(\boldsymbol{p})$ over time.

Turning to the second stage, we have:

$$\hat{x}_{hjt} = \alpha_{jt} + \beta_{1,j}\ln X^*_{it} + \beta_{2,j}b(\boldsymbol{p}_t)\left(\ln X^*_{it}\right)^2 + \phi^i_t + \varepsilon_{hjt}, \tag{$5'$}$$

where as before aˆreflects that we have adjusted $\tilde{x}_{hjt}$ for demographics using the estimates of $\boldsymbol{\Gamma}_j$ from the first stage. Note that we explicitly include $b(\boldsymbol{p}_t)$ in the second-stage expression. We maintain the benchmark assumptions regarding systematic measurement, and thus α_{jt} and ϕ^i_t will capture good-specific and income-specific mis-measurement at time t. The good-time specific intercept also captures the first-order effects on demand stemming from relative price movements. As in the benchmark, our assumptions regarding the nature of the good-

---

[21]Banks et al. (1997) discuss general nonlinear demand systems as well as estimate a quadratic extension of the AIDS system (QAIDS). While our benchmark differs from the QAIDS system, the Banks et al. critique remains relevant for a broad class of nonlinear demand systems.

income-time specific measurement error contained in ε_{hjt} ensures the residual is uncorrelated with the regressors. Moreover, ε_{hjt} includes the deviation of $\ln X^*_{hjt}$ and $\left(\ln X^*_{hjt}\right)^2$ from the respective means for household h's income group, as was the case in the benchmark (see equation 6).

In the second stage, the first-stage estimates of $\beta_{1,j}$ and $\beta_{2,j}$ are our generated regressors, which we interact with income-group dummy variables. The coefficient on $\hat{\beta}_{1,j}$ interacted with income group i at time t is our estimate of $\ln X^*_{i,t}$. From (5′), we see that the coefficient on $\hat{\beta}_{2,j}$ is an estimate of $b(\boldsymbol{p}_t)\left(\ln X^*_{i,t}\right)^2$, and is therefore not an unbiased estimate of squared log expenditure. Note that we do not need to include an estimate of $b(\boldsymbol{p}_t)$ directly in the second stage, but rather exploit the fact that the coefficient on $\beta_{2,j}$ is allowed to vary freely across each second-stage sample period. This is sufficient to accommodate the instability of $\beta_{2,j}$ over time, generating a consistent estimate of log expenditure from the coefficient on $\beta_{1,j}$. This is sufficient for our purposes, as we are not interested in estimating $b(\boldsymbol{p}_t)$ and $(\ln X^*_{it})^2$ independently.

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. Columns (1) and (2) estimate the second stage via OLS and WLS, respectively, where the latter weights goods by their share in NIPA PCE. Columns (3) and (4) impose the nonlinear restriction that the coefficient on $\beta_{2,j}$ is the square of the coefficient on $\beta_{1,j}$. This restriction is only valid if $b(\boldsymbol{p}_t) = 1$ for the respective second-stage sample period. We report these restricted estimates for reference, as they shed light on the sensitivity of our estimates to price-induced instability in $\beta_{2,j}$. Column (3) implements nonlinear least squares, while Column (4) implements weighted nonlinear least squares. The standard errors in all specifications are bootstrapped.

The different columns of Table 5 indicate that the results are similar across specifications, with the WLS 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 0.80 and 0.77 in the two WLS specifications of Table 5, compared to 0.82 in Column (4) of Table 3. (Note that Column (4) in Table 3 is the appropriate counterpart as that specification also used lagged expenditure to address measurement error.) Similarly the change in inequality is also similar with the nonlinear specification. Table 5 Column (2) reports an increase of 30 log points between 1980-82 and the early 1990s, and of 30 log points through 2007. (The corresponding benchmark estimates is 21 for the 1990s and 44 for 2007.) The change during the Great Recession is

essentially zero in the nonlinear specifications, compared to -0.05 in the linear model. We view this robustness 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

A 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. 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, under the WLS specification, yield an estimated level of log inequality of 0.70 in 1980-82. This contrasts with the estimate of 0.85 obtained 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 0.36 using the 2005-07 elasticities. This is actually greater than the 0.31 point estimate reported in Table 3. We have also estimated the demand elasticities using the 1983-2007 sample, that is, just trimming the beginning and end reference periods. These elasticities generate a change in consumption inequality of 0.34 log points between 1980-82 and 2008-10.

5   Conclusion

The results presented in this paper suggest that increases in consumption inequality mirror that of income inequality to a much 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-household-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, "Lifecycle Prices and Production," *American Economic Review*, 2007, *97*, 1533–1559.

_ **and** _ , "Deconstructing Life Cycle Expenditure," *working paper*, 2009.

_ **and Mark Bils**, "Has Consumption Inequality Mirrored Income Inequality," *NBER Working Paper 16807*, 2011.

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.

Banks, James, Richard Blundell, and Arthur Lewbel, "Quadratic Engel Curves and Consumer Demand," *The Review of Economics and Statistics*, 1997, *79*, 527–539.

Battistin, Erich, "Errors in Survey Reports of Consumption Expenditures," *Institute for Fiscal Studies WP 03-07*, 2003.

Blundell, Richard and Ben Etheridge, "Consumption, Income, and Earnings Inequality in the UK," *Working Paper, UCL*, 2009.

_ **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," *CRRWP 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 Economi Review*, 1992, *82*, 546–551.

Dahl, Molly, Thomas DeLeire, and Jonathan Schwabish, "Estimates of Year-to-Year Volatility in Earnings and in Household Incomes from Administrative, Survey, and Matched Data," *Journal of Human Resources*, 2011, *46*, 750–774.

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.

DeBacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos, "Rising Inequality: Transitory or Permanent? New Evidence from a Panel of US Tax Returns," *Brookings Papers on Economic Activity*, 2013.

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.

Gottschalk, Peter and Robert Moffitt, "Growth in Earnings Instability in the US Labor Market," *Brookings Papers on Economic Activity*, 1994.

_ **and** _ , "The Rising Instability of US Earnings," *Journal of Economic Perspectives*, 2009, *23*, 3–24.

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 Economi Review*, 2007, *97*, 687–712.

Hausman, Jerry, Whitney Newey, and J.L. Powell, "Nonlinear Errors in Variables: Estimation of some Engel Curves," *Journal of Econometrics*, 1995, *65*, 205–233.

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.

Jappelli, Tullio and Luigi Pistaferri, "Does Consumption Inequality Track Income Inequality in Italy?," *Review of Economic Dynamics*, 2010, *13*, 133–153.

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.

Kaplan, Greg and Giovanni Violante, "How Much Consumption Insurance Beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2010, *2*, 53–87.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song, "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937," *Quarterly Journal of Economics*, 2010, *125*, 91–128.

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.

Monti, Holly and Graton Gathright, "Measuring Earnings Instability Using Survey and Administrative Data," *Working Paper, US Census*, 2013.

Murphy, Keving 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 Ameir 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 the 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 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	(I) Elasticity	(SE)	(II) Elasticity	(SE)
Housing	27.3	0.92	(0.02)	0.93	(0.02)
Food at Home	11.7	0.37	(0.02)	0.47	(0.02)
Vehicle Purchasing, Leasing, Insurance	13.2	1.02	(0.08)	0.72	(0.1)
All Other Transportation	7.4	0.89	(0.03)	0.91	(0.04)
Utilities	5.2	0.47	(0.02)	0.55	(0.02)
Health Expenditures including Insurance	5.0	0.91	(0.06)	1.11	(0.08)
Appliances, Phones, Computers with associated services	4.9	0.87	(0.04)	0.94	(0.05)
Food Away from Home	4.6	1.33	(0.06)	1.32	(0.07)
Entertainment Equipment and Subscription Television	4.1	1.26	(0.07)	1.22	(0.08)
Men's and Women's Clothing	2.6	1.35	(0.05)	1.38	(0.06)
Entertainment Fees, Admissions, Reading	2.2	1.74	(0.06)	1.65	(0.07)
Cash Contributions (Not for Alimony/Support)	2.2	1.81	(0.18)	1.26	(0.12)
Furniture and Fixtures	1.5	1.39	(0.1)	1.55	(0.15)
Education	1.3	1.63	(0.18)	1.88	(0.23)
Shoes and other apparel	1.5	1.09	(0.09)	1.19	(0.11)
Domestic Services and Childcare	1.5	1.60	(0.13)	1.80	(0.13)
Tobacco, other smoking	1.0	-0.26	(0.09)	-0.05	(0.08)
Alcoholic Beverages	1.0	1.14	(0.09)	1.14	(0.08)
Children's Clothing (up to age 15)	1.0	0.67	(0.07)	0.83	(0.09)
Personal Care	1.0	0.96	(0.05)	0.96	(0.05)

Note: The first column presents each good's average share of total expenditure for 1994–1996. The remaining columns report estimates of each good's expenditure elasticity, with associated standard errors in parentheses. Specification (I) sums each household's expenditure (on each good and in total) over all four interviews and instruments log total expenditure with dummy variables indicating the household's income category as well as the continuous variable of log real after-tax income. Specification (II) splits the four interviews into two sub-samples. Each household's expenditure is computed using the sum of the final two interviews. Log total expenditure summed over the first two interviews is used as the instrument. The correlation of the two specifications is 0.93. 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	0.86	0.90	0.86	0.82
	(0.07)	(0.06)	(0.05)	(0.08)
Log Change 1980/82–1991/93	0.27	0.17	0.20	0.21
	(0.08)	(0.06)	(0.06)	(0.07)
Log Change 1980/82–2005/07	0.48	0.35	0.42	0.44
	(0.08)	(0.07)	(0.07)	(0.08)
Log Change 2005/07–2008/10	-0.06	-0.04	-0.04	-0.05
	(0.08)	(0.06)	(0.06)	(0.08)
Categories Included	All	All	Without durables	All
Specification	OLS	WLS	WLS	WLS
First-Stage Instrument	Income	Income	Income	Lagged Expenditure

Note: This table reports the estimated change in income inequality obtained from the second-stage regressions. Columns (1) through (3) use the first-stage estimated expenditure elasticities reported in Column (I) of Table 2 as regressors, while Column (4) uses those reported in Column (II) of Table 2. The estimated parameters in the first row represent log inequality between the high-income and low-income households in 1980-82. The next three rows 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 weighted least squares, using the average NIPA shares for 1980–2010 as weights; the third column implements WLS while omitting all good categories containing durables; and the final column implements WLS with the alternative first-stage elasticities. The standard errors are calculated using a bootstrap with 100 replications.

Table 4: Change in Relative Income-Specific Measurement Error

	(1)	(2)	(3)	(4)
Relative Mis-Measurement 1980-82				
High-Income – Low-Income	0.09	0.04	0.05	0.10
	(0.06)	(0.05)	(0.04)	(0.07)
Change 1980/82–1991/93	-0.13	-0.04	-0.04	-0.08
	(0.06)	(0.05)	(0.05)	(0.06)
Change 1980/82–2005/07	-0.30	-0.16	-0.17	-0.25
	(0.07)	(0.05)	(0.05)	(0.07)
Change 2005/07–2008/10	0.04	0.02	0.02	0.03
	(0.06)	(0.05)	(0.05)	(0.06)
Categories Included	All	All	Without durables	All
Specification	OLS	WLS	WLS	WLS
First-Stage Instrument	Income	Income	Income	Lagged Expenditure

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 100 replications.

Table 5: Trends in Consumption Inequality: Nonlinear Engel Curves

	(1)	(2)	(3)	(4)
Log Inequality 1980-1982	0.98	0.80	0.82	0.77
	(0.16)	(0.15)	(0.08)	(0.09)
Log Change 1980/82–1991/93	0.47	0.30	0.27	0.22
	(0.17)	(0.12)	(0.09)	(0.09)
Log Change 1980/82–2005/07	0.42	0.30	0.38	0.32
	(0.18)	(0.12)	(0.09)	(0.08)
Log Change 2005/07–2008/10	$< |0.01|$	0.05	-0.01	$< |0.01|$
	(0.13)	(0.09)	(0.06)	(0.05)
Specification	Unrestricted OLS	Unrestricted WLS	Restricted NLS	Restricted NWLS

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 two columns report estimates of the coefficient on $\beta_{1,j}$ from (5′) using OLS and WLS, respectively. The final two columns impose the restriction that the coefficient on $\beta_{2,j}$ is the square of the coefficient on $\beta_{1,j}$ in (5′), and report the estimates, respectively, from nonlinear least squares and weighted nonlinear least squares with weights given by the share in NIPA PCE. Standard errors are calculated using a bootstrap with 100 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 pctile (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
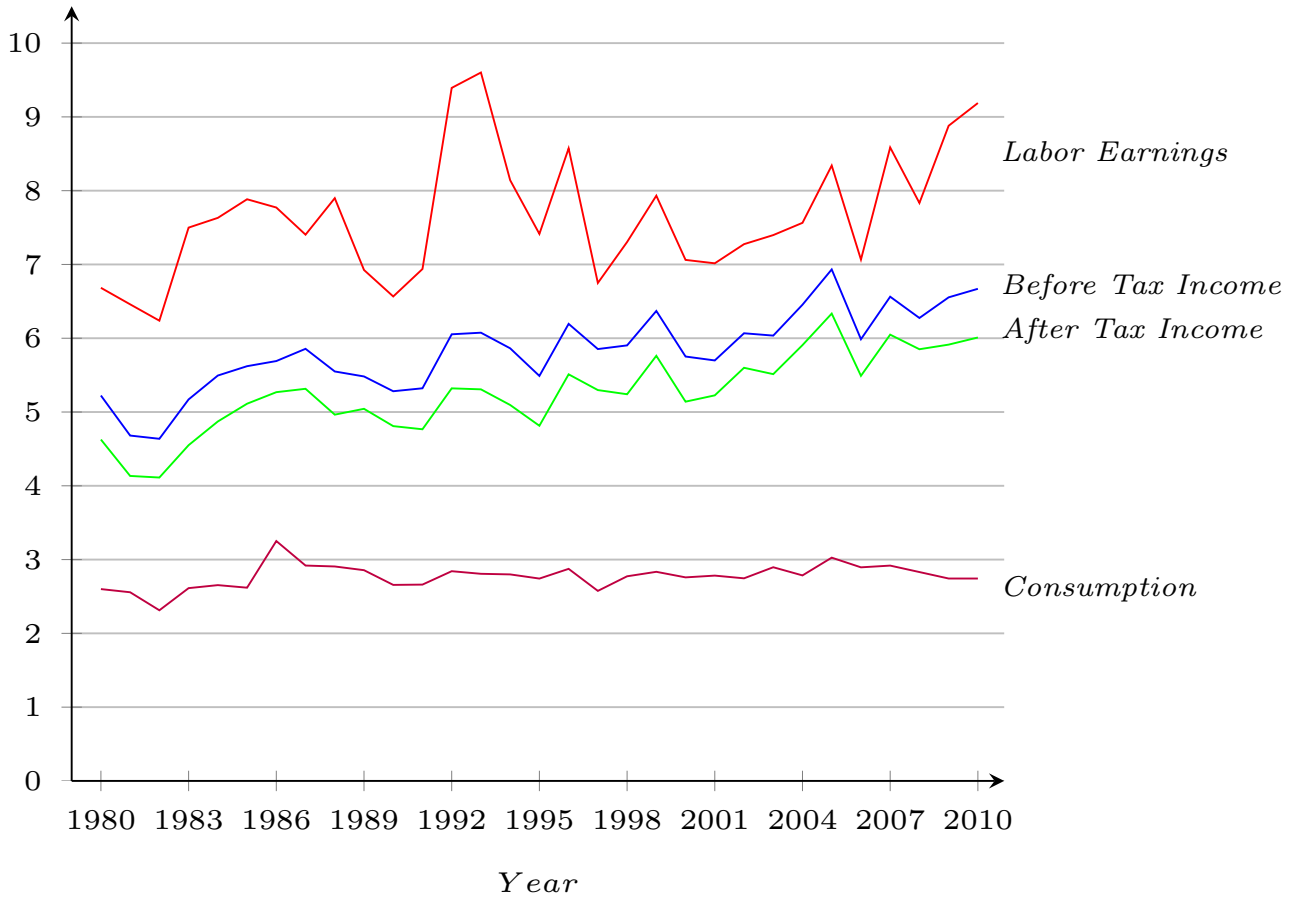
Note: This figure depicts the ratio of high-income to low-income respondents' reported labor earnings, before-tax income, after-tax income, and consumption expenditures. High income refers to respondents who report before tax household income in the 80th through 95th percentiles. Low income refers to 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
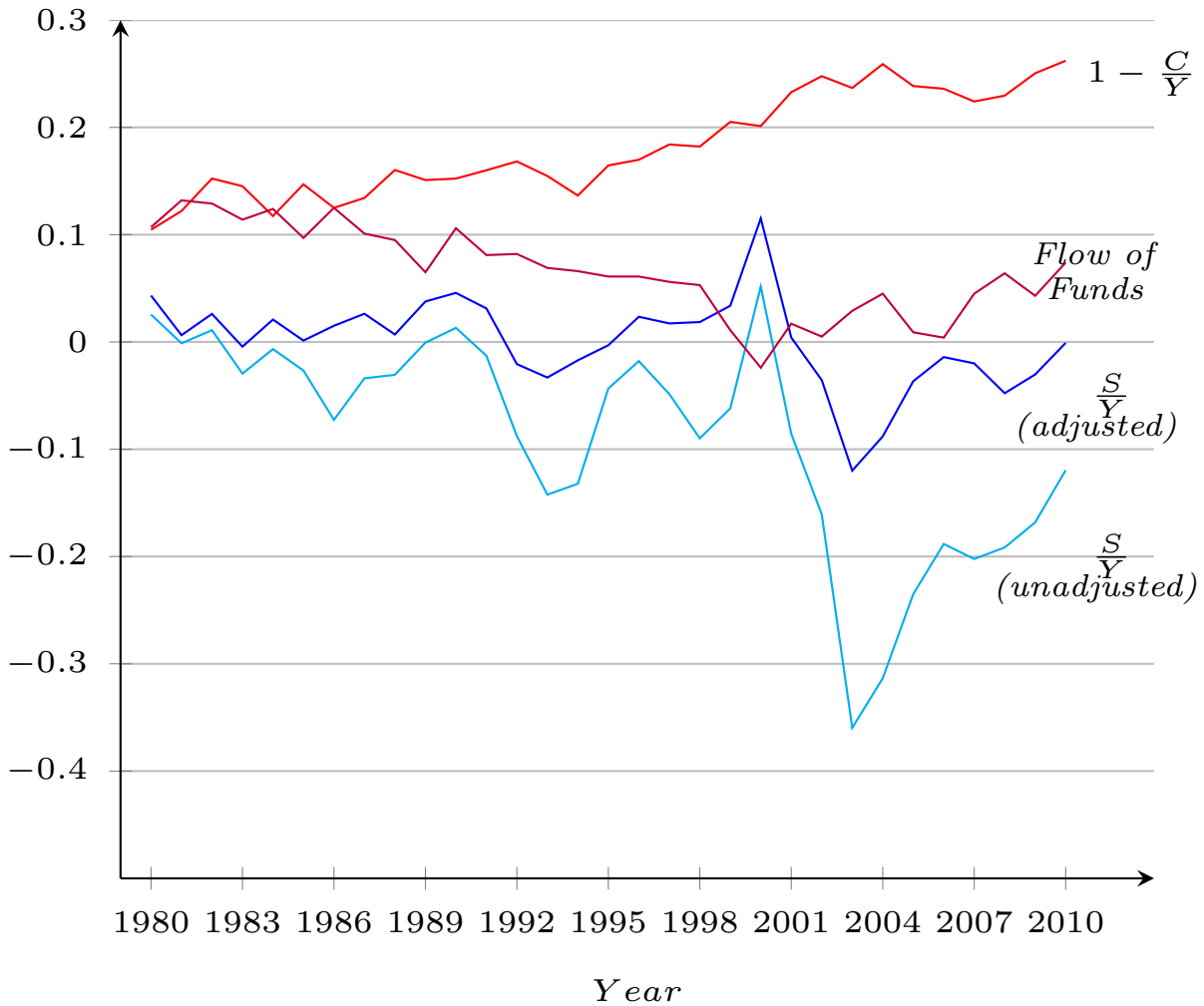
Note: This figure depicts the mean savings rates. The line labeled 1-C/Y refers to implied savings computed as after-tax income minus reported consumption expenditures. The line labeled "Flow of Funds" is the flow of funds aggregate private savings rate out of disposable income. The lines labeled S/Y refer to CE 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
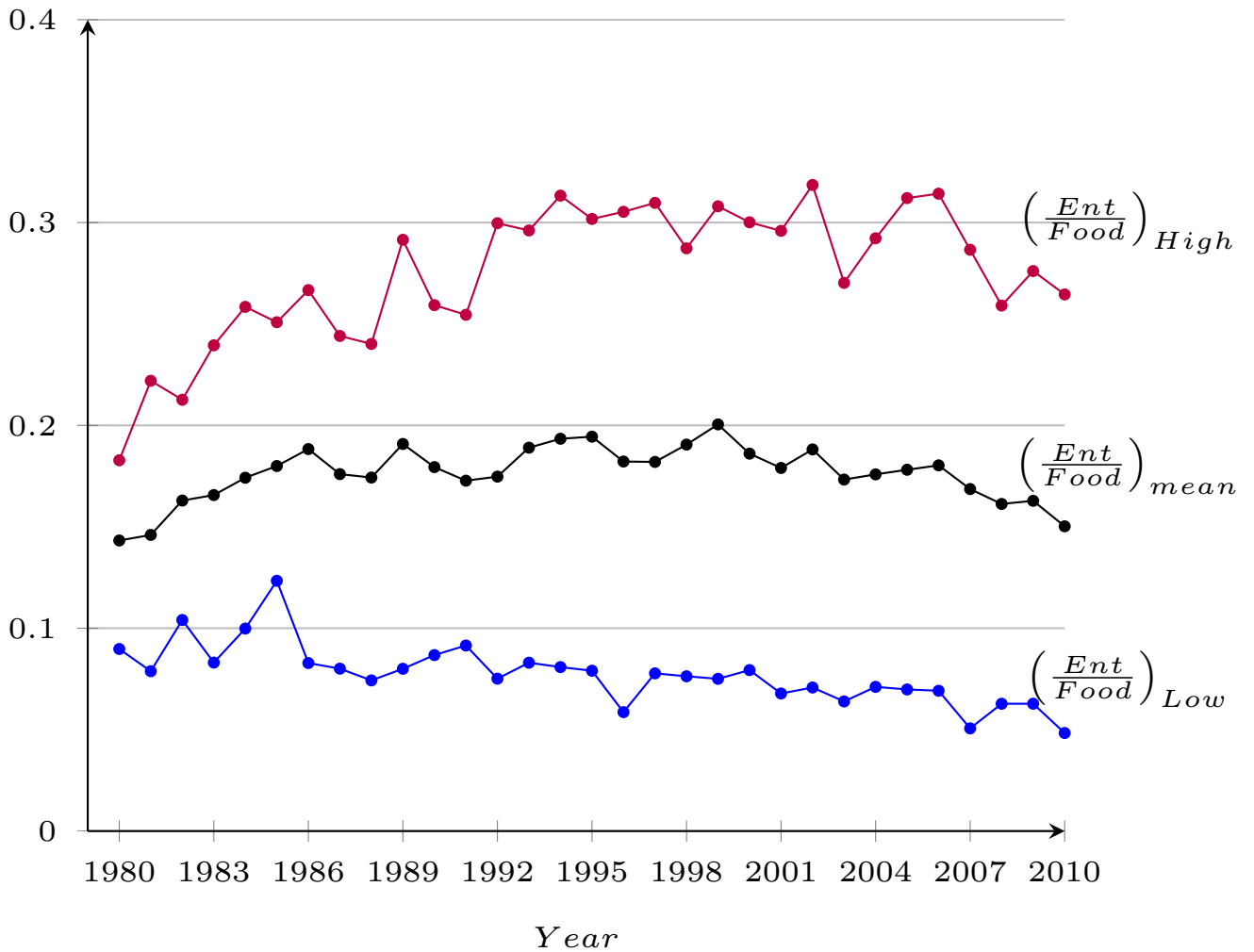

Note: This figure depicts the ratio of spending on nondurable entertainment to food at home for high- and low-income households.

Figure 4: Relative Expenditure Growth for 20 Goods
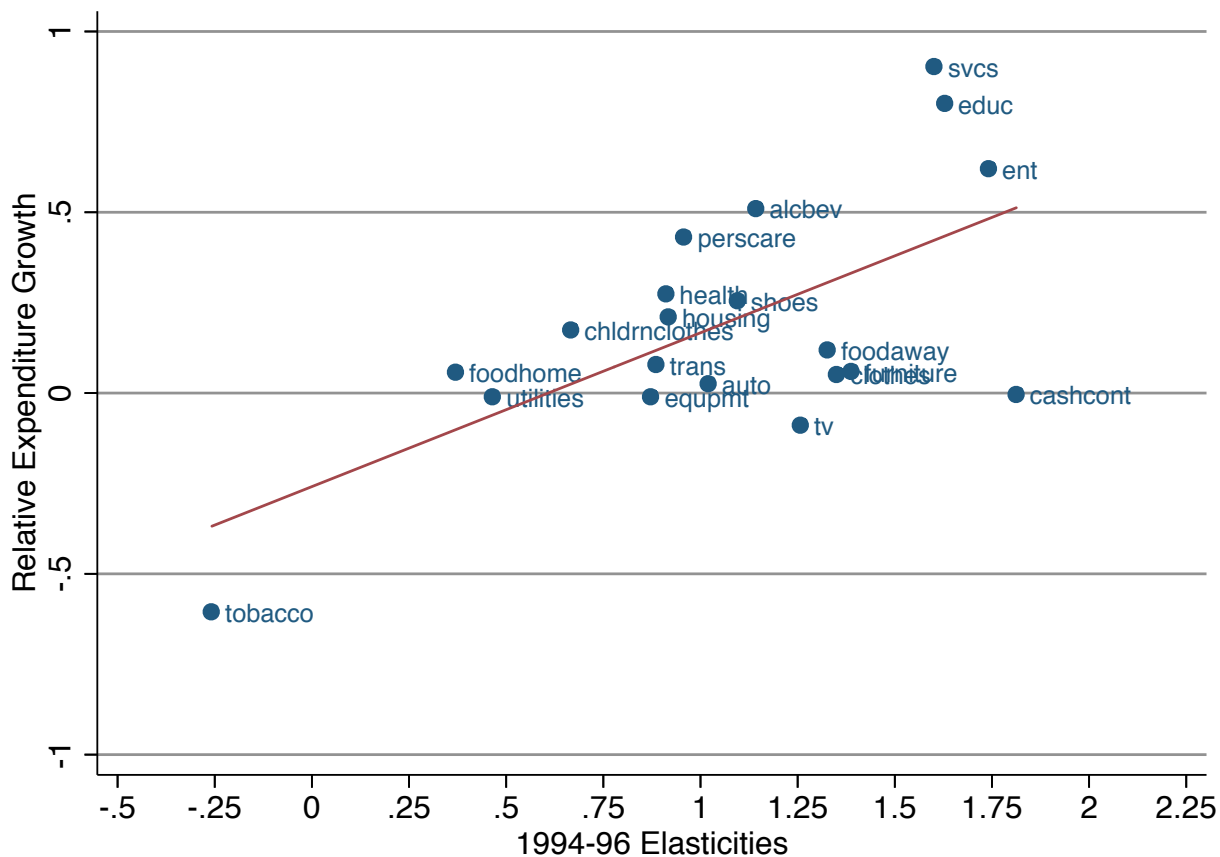
Note: This figure is a scatter plot of relative (high- versus low-income) expenditure growth over the sample period for each good versus expenditure elasticity. The vertical axis depicts the difference across high-income and low-income households in the log growth in expenditure for each good between 1980/82 and 2008/10. The horizontal axis is each good's estimated expenditure elasticity from Table 2 Column (I). The slope of the scatter plot's regression line is 0.425.

Figure 5: Stability of Expenditure Elasticities over Time
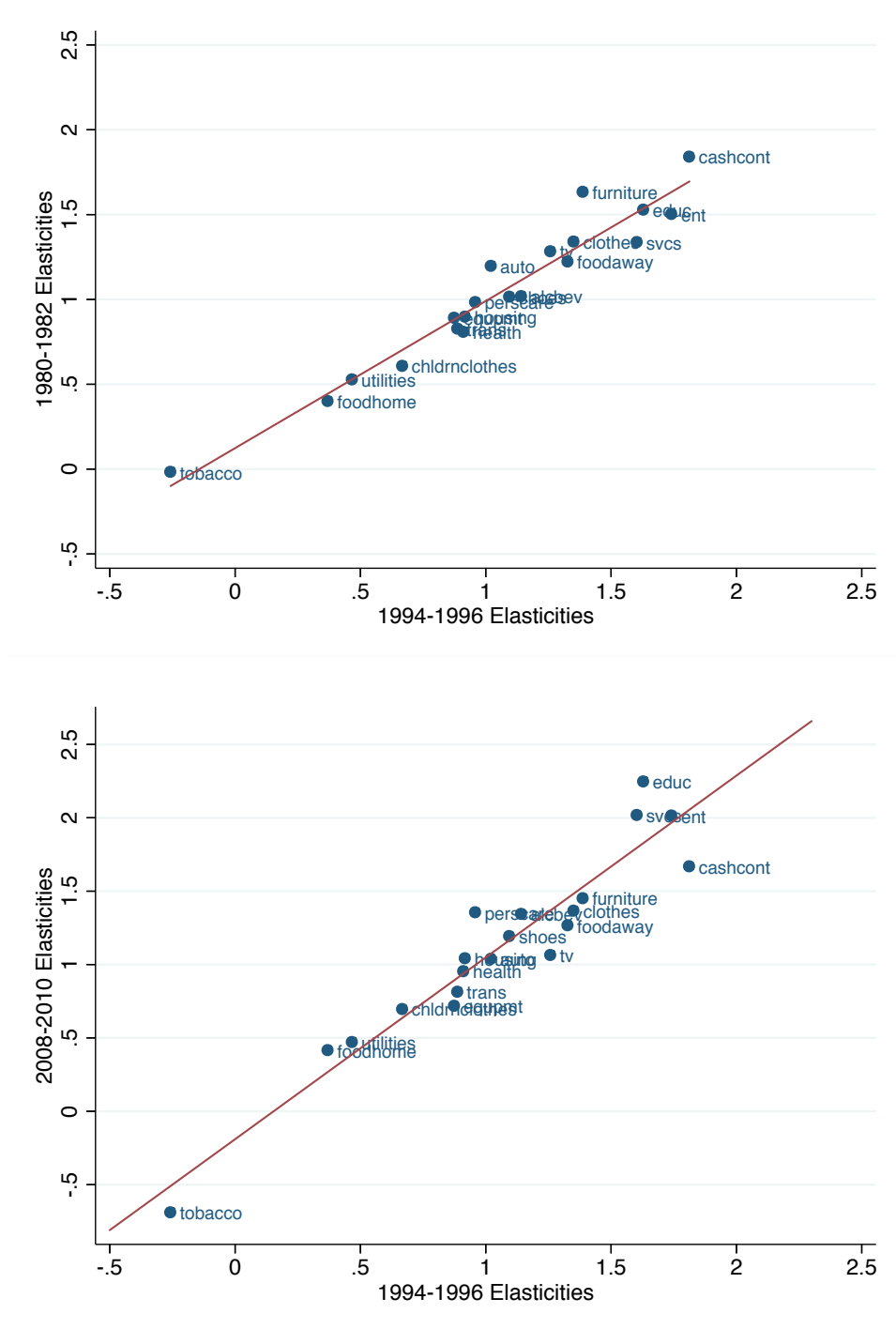
Note: The top panel depicts a scatter plot of the 20 expenditure elasticities estimated using the 1980-82 CE sample versus our benchmark 1994-96 first stage reported in table 2. The fitted regression line has a slope of 0.87 and an R^2 of 0.93. The bottom panel replaces the 1980-82 estimates with estimates using 2008-10 surveys. The fitted line has slope 1.24 and an R^2 of 0.90.