Has Consumption Inequality Mirrored Income Inequality?*

Preliminary

Mark Aguiar       Mark Bils

December 2, 2009

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 two alternative measures of consumption expenditure, using data from the Consumer Expenditure Survey (CE). We first use reports of active savings and after tax income to construct the measure of consumption implied by the budget constraint. We find that the consumption inequality implied by savings behavior tracks income inequality closely between 1980 and 2007. Second, we use a demand system to correct for systematic measurement error in the CE’s expenditure data. Specifically, we consider trends in the relative expenditure of high income and low income households for different goods with different income 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. This second exercise also indicates that consumption inequality has closely tracked income inequality over the period 1980-2007. Both of our measures show a significantly greater increase in consumption inequality than what is obtained from the CE’s total household expenditure data directly.

*Both authors: University of Rochester and NBER. Email: mark@markaguiar.com and mark.bils@gmail.com. We thank Yu Liu for outstanding research assistance.
1 Introduction

We revisit the issue of whether the increase in income inequality over the last 30 years has translated into a quantitatively similar increase in consumption inequality. Contrary to several influential studies discussed below, we find that consumption inequality has closely tracked income inequality over the period 1980-2007. Like most of the previous literature that argues the opposite, we base our conclusions on the Consumer Expenditure Survey’s (CE) interview survey. However, we focus on two new measures of consumption that under our stated assumptions adjust for the systematic measurement error in the CE. The first measure is the CE’s data on savings, from which we calculate consumption via the budget constraint. The second is a demand system, from which we estimate relative consumption growth using relative expenditures on luxuries and necessities. Both measures show a substantial increase in consumption inequality, similar in magnitude to the increase in income inequality. The increase is particularly large for the period 1980-1995, consistent with the view that changes in income inequality in this period reflected changes to permanent income.

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 similar to Krueger and Perri’s, we show that relative income inequality increased by 35 percent (.35 log points) between 1980 and 2007, where our conservative measure of income inequality is the ratio of those in the 80-95th percentiles to those in the 5-20th percentiles.\(^2\) The corresponding increase in consumption inequality for the same two groups is 20 percent. The gap is particularly large for the most recent 10 years, for which income inequality increased 15 percent and consumption inequality only 7 percent.

We reassess these facts using two alternative measures. Our first exercise is simply budget constraint accounting. The mirror image of the differential trends between income and consumption inequality is a growing gap in savings favoring high income households. (Our benchmark measure of income is total household income after taxes and transfers, which is designed to capture adjustments due to government programs and financial income.) Based on reported consumption expenditures, the high income group increased their savings rate from 27 percent to 39 percent between 1980 and 2007, while the low income group went from a savings rate of -27 percent to -24 percent. The

---

\(^1\)For other contributions to this literature, see Blundell and Preston (1998), Blundell et al. (2008), and Heathcote et al. (2009).

\(^2\)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.
implied savings rates using CE income and consumption are implausible. For the overall mean, the implied savings rate in the CE increases from 9 percent in 1980 to over 20 percent in 2007. This contrasts with savings out of disposable income reported in the national income and product accounts (NIPA), which falls from 10 percent to 2 percent, as well as is inconsistent with other micro data sets (see Bosworth and Anders, 2008 and Bosworth and Smart, 2009). This discrepancy is in line with the well documented decline in aggregate consumption reported in the CE relative to NIPA.

In addition to expenditures and income, the CE asks detailed questions on savings flows directly. These questions include net payments of loans, changes in deposit balances, purchases of stocks, etc. The average reported savings rate in the CE declines over time, consistent with NIPA but in contrast to the savings rate implied by the CE’s consumption data. Calculating implied expenditure as income minus savings, we obtain an increase in relative consumption of 30 percent, close to the relative change in income of 35 percent. The CE’s savings measures are noisy (particularly regarding new mortgages), and so we view them primarily as a consistency check on the reported consumption data, and only secondarily as an independent measure of consumption itself.

Our preferred measure of consumption inequality uses the CE’s expenditure data, but allows for systematic measurement error. Our modeling of measurement error is fairly general. In particular, we allow for time-dependent multiplicative measurement error that is good specific as well as income-group specific. The former allows for the mis-measurement of particular goods to vary over time, such as the possibility that the under-reporting of luxuries has increased relative to the under-reporting of necessities. The latter allows for the measurement to be income-group specific, such as the possibility that the under-reporting of expenditure of high income households across all goods has increased relative to the under-reporting by low income households. This modeling of measurement captures systematic mis-measurement that is correlated with the characteristics of the good and the income-characteristics of the households. We also allow for mis-measurement at the level of good-income group interaction (clothing of the rich versus clothing of the poor), but restrict this joint mis-measurement to be independent of the characteristics of the goods (in particular, the good’s income elasticity).

Our estimation procedure consists of two steps. First, we estimate good-specific income elasticities using a simple log-linear demand system. To do this, we use the 1972-73 CE, separating our first stage sample from the post-1980 period of focus. In the second stage, we consider the difference in expenditure growth across goods and across income groups. To see how this approaches addresses mis-measurement, take expenditures on food at home versus nondurable entertainment as an example. The relative expenditure
on food at home across income groups remained essentially constant between 1980 and 2007. Given a non-zero estimated income elasticity of 0.49 for food at home, this suggests zero change in relative total expenditures. While comparing the same good across income groups controls for (multiplicative) mis-measurement of food in each period, it does not control for the possible mis-measurement correlated with income. For this, we can add a second good, nondurable entertainment. Over the same period, the high income-low income ratio of expenditure on nondurable entertainment increased by 0.8 log points. Given an estimated income elasticity of 1.94 for entertainment, this implies a change in relative expenditure of 41 percent. Again, this controls for good-specific measurement error, but not mis-measurement correlated with income. However, any mis-measurement that is specific to income groups, but that is uniform across goods, can be eliminated by differencing across goods. That is, the difference in relative expenditure growth rates will equal the difference in income elasticities times the change in total expenditure inequality (plus an idiosyncratic error term). Solving this equation, the relative growth in these two goods implies a change in consumption inequality of 55 percent. Our procedure is thus a difference-in-difference estimate, where one difference eliminates good-specific mis-measurement and the second difference eliminates income group-specific mis-measurement.

While food and entertainment are interesting due to their extreme income elasticities, the CE data contains expenditure on many goods. We therefore implement this procedure using all goods in a regression framework. Our estimates suggest that consumption inequality increased by 33 percent between 1980 and 2007, approximately the same as the change in income inequality, and slightly larger than that obtained from the budget constraint accounting. We find this estimate is stable across different subsets of goods, different weighting schemes across goods, and alternative first-stage income elasticity estimates.

We also consider trends in inequality in different sub-periods. We find that income inequality increased by 20 percent between 1980 and the mid-1990s, and then by an additional 15 percent between 1995 and 2007. The inequality in reported CE expenditure increased by 12 percent in the first sub-period, and then by 7 percent in the latter half of the sample. Reported consumption inequality does not keep pace with income inequality in either sub-period. Using our demand system estimates, we find that consumption inequality increased by 27 percent between 1980 and the mid-1990s, and then by additional 6 percent through 2007, for a total increase of 33 percent. These estimates more closely track the profile of income inequality, with a larger increase in the 1980s, and a smaller but still significant increase thereafter. In this regard, our estimates support the interpre-
tation that changes in inequality in the 1980s reflected shifts in permanent income, while the change in recent years may be weighted toward transitory changes (see Attanasio and Davis, 1996, Blundell et al., 2008). Our results suggest that this conclusion has actually been under-stated using reported CE consumption inequality.

We are not the first to reassess trends in consumption inequality, particularly with a focus on mis-measurement of CE interview expenditures. Attanasio et al. (2005) uses the diary component of the CE to correct for mis-measurement in the interview survey, and document a large increase in consumption inequality. Their analysis does not extend back to the 1980s due to data limitations. Our analysis uses interview survey data, but brings in data on savings as well as the differential trends across goods and income groups to address mis-measurement. Our paper is complementary to an independent effort by Parker et al. (2009). These authors correct for good-specific mis-measurement by comparing mean CE expenditures good-by-good to those reported in NIPA. They confirm that the CE’s under-reporting (relative to NIPA) has become increasingly pronounced over the sample period, and show that this under-reporting varies across goods. They then correct the CE expenditure by inflating each good by the appropriate factor to ensure the CE aggregates conform to those of NIPA. They show that while this correction affects the means, it does not substantially affect conclusions about consumption inequality post-1980. In our approach, the good-specific adjustments inherent in the NIPA adjustments are handled instead by differencing, although we do make use of the NIPA correction terms as a robustness check to our implied savings measures and the first-stage estimates of income elasticities.

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) and Blundell and Preston (1998). For trends in inequality for a number of countries and time periods, see the papers collected in Krueger et al. (2010). There is also a large literature on consumption versus income inequality over the life cycle, starting with Deaton and Paxson (1994).³ These papers often use the CE for consumption data, and are therefore subject to the measurement error problems addressed

³See also, Storesletten et al. (2004), Heathcote et al. (2005), Guvenen (2007), Huggett et al. (2009), and Aguiar and Hurst (2009).
in this paper. We leave the question of whether our approach has implications for trends in life cycle inequality to future research.

The remainder of the paper is organized as follows. Section 2 describes the data set; section 3 analyzes the CE’s savings data; section 4 performs our main demand-system analysis; and section 5 concludes.

2 Data

In this section we describe our data set, leaving to the data appendix (under construction) more detailed information on the surveys and variable construction. Our data is from the Consumer Expenditure Survey’s interview sample. This is a well known consumption survey that has been conducted continuously since 1980. We also use the earlier waves of 1972 and 1973 for estimation of the demand system. (The survey was not conducted between 1973 and 1980.)

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. An initial interview collects information about household characteristics as well as other baseline information. Each household is then re-interviewed once a quarter for up to four consecutive quarters, with each interview recording expenditures on detailed categories over the preceding three months. The final interview also updates income and demographic information for the preceding 12 months. For income and demographics, we use the responses from the last interview. Income, expenditure, and savings variables are all recorded at the household level. Demographics such as age, sex, and education, are those reported by the “reference person,” who is identified by the response to the question who owns or rents the house. We define income and expenditures at the household level, rather than creating adult equivalence scales. 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.

On the income side, we use the CE measures of total household labor earnings (before tax), as well as total income after tax. These variables are reported in the last interview and cover the previous 12 months. After-tax income in the CE includes labor earnings, non-farm business income, social security and retirement benefits, social security insurance, unemployment benefits, workers compensation, welfare, financial income, rental income, alimony and child support, scholarships or stipends, and food stamps, minus taxes. Our measure of after-tax income is that reported in the CE, but we add in food as
pay, other money receipts (lump sum alimony and child care payments, proceeds from the sale of personal items, etc.). For consistency, as we count receipts of alimony and child support as income, we subtract off payments of alimony and child support. Similarly, as benefits are part of income, we subtract social security taxes from income. Finally, as rental equivalence is a consumption expenditure for home owners, but not out of pocket housing expenses (for example, costs stemming from capital improvements and repairs), we include rental equivalence minus out of pocket housing upkeep costs as part of after tax income as well.

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. The CE records the total outstanding credit balances in the 2nd and 5th interviews, which are 9 months apart. From this, we calculate net payments of credit by taking the difference and scaling up to an annual measure by 4/3. The other net worth items are reported as flows and do not require differencing across interviews.

While the CE contains fairly rich data on savings, it is designed to measure consumption and not savings. We use the savings data primarily as a consistency check, via the budget constraint, on reported consumption. As we show in section 3, the average saving rate reported in the CE appear to be broadly consistent with those obtained from the national income accounts. However, 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 implies sharp up-ticks in new mortgages around 1993 and in the 2000s, which is consistent with published statistics on refinancing. However, there are a number of reported new mortgages without a corresponding purchase of a house or a significant paying down of an existing mortgage. New mortgages for households who do not purchase a home are on average nearly 14 times the reported reductions to existing mortgages. In particular, the CE data imply an average “cash out” percentage of 73 percent from new mortgages not associated with a house purchase. This high rate is not supported by other studies of refinancing, which suggest that roughly 13 percent of the new mortgage is taken out as cash and the remainder is used to pay off existing mortgages and related costs (see Greenspan and Kennedy, 2007).

These questionable transactions, while not many in number, nevertheless affect the mean savings rate due to their size. To address this potential measurement error, we identify questionable new mortgages as those that are greater than 1.5 times the sum of the purchase price of a new house plus any lump sum payments or reductions to existing mortgages.
mortgages. Only 7 percent of the sample has a questionable new mortgage, but roughly three quarters of the new mortgages fall into this category. For these mortgages, we top code the amount of the mortgage as the sum of the full amount of any house purchase plus the payment on existing mortgages plus one third the reported mortgage amount. This implies that at most one third of the new mortgage amount is taken out as cash. This reduces the average implied “cash out” ratio of refinanced mortgages to 14 percent, consistent with the number reported by Greenspan and Kennedy (2007). In what follows, we typically present two savings series, the raw series using reported mortgages (labeled “unadjusted”) and the alternative series which uses the adjusted mortgages (labeled “adjusted”). As documented in section 3, it turns out that the adjustment affects mean savings rates, but does not have a significant impact on implied consumption inequality.

The CE reports expenditure on hundreds of separate items. We aggregate these into 20 categories, which are listed in table 2. The mapping of the individual goods into groups was governed by several criteria. The first was to group similar goods together. The second was consistency across the various waves of the survey; to the extent possible, each group has the same components in every year. The third was to ensure our groups spanned a wide range of income elasticities.\footnote{We adhere to the groupings created by the BLS in published statistics with a couple of exceptions. We have grouped telephone equipment and services with appliances, computers, and related services rather than with utilities, based on our prior regarding income elasticities. We combine expenditure on reading materials with other nondurable entertainment expenditures because alone it represents a trivial expenditure share (about 0.2%).}

For expenditure on housing services, we use rent paid for renters and self-reported rental equivalence for home owners. For interviews conducted between 1980 through the second quarter of 1982, households were not asked about rental equivalence. We impute the rental equivalence for homeowners in these early waves based on non-housing expenditures as well as demographics. In particular, we use the 1983 sample 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 durables other than housing we use direct expenditure, and do not impute service flows. This is motivated by our use of income groups as the unit of analysis (described below), and the assumption that aggregating over many households provides a good proxy for the consumption of durable services at a point in time. We show in section 4 that our estimates are not sensitive to the exclusion of housing and other 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. To adjust for this drop, we increase food at home expenditure by 11% for these years. 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 years 1982 to 1987. This adjustment is similar to that employed by Krueger and Perri (2006).

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.

In the early waves the survey only included urban households, and so for consistency we restrict our analysis to urban residents for the entire sample period. (We will explore robustness in a subsequent draft.) Our analysis employs the following further restrictions on the CE urban samples, both for the 1980-2007 and 1972-1973 samples. First, we restrict households to those with reference persons between the ages of 25 and 64. Second, 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 aggregate expenditures for each household across the four interviews, so each household appears once in the sample. We assign households to years based on the month of the first interview, with households starting the survey in the fourth calendar quarter assigned to the next year. Third, we restrict the sample to those with “full income reports”, which is defined by the CE as recording at least one non-zero response to any of the income and benefits questions. Fourth, we eliminate households that report unusually large expenditures on our smaller categories. In particular, we exclude any household that records spending more than half of after tax income on any category, with the exception of housing, food, and vehicle purchases. 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 after tax income distribution. The data appendix details how many households are eliminated at each step. We are left with 9,537 households for the 1972-73 sample, and 49,540 households for the period 1980-2007.

From this sample, we construct our units of observation, which are a panel of income-demographic cells. Specifically, in each year we divide the sample into 6 equally weighted bins based on after tax income. As we have excluded the top and bottom five percent, each bin corresponds to 15 percent of the sample before trimming of the tails. Each income
group is further divided into 18 demographic cells, based on age range (25-37, 38-50, 51-64), number of earners (<2, 2+), and household size (≤2, 3-4, 5+). For each cell and each year, we average across expenditure, income, and savings variables. These cell averages are the unit of analysis and our primary measure of inequality is the ratio of the mean of the top income group to the mean of the bottom income group.

3 Budget Constraint Accounting

In this section, we review the trends in income and consumption inequality using our CE sample. We then discuss the CE savings rates, and introduce our first alternative measure of consumption based on the budget constraint.

3.1 Trends in Income and Consumption Inequality

We begin with labor earnings. The top line in figure 1 depicts the trend in labor earnings inequality. As discussed in section 2, inequality is the ratio of the mean for the top income cells to the mean for the bottom income cells. Keep in mind that the allocation of respondents into the high and low income groups is based on after tax income, and so the cells are the same for all lines in figure 1. There is substantial year-to-year movement, reflecting in large part sampling error, so we average over multiple years in table 1. In particular, we look at the three multi-year periods 1980-82, 1992-1995, and 2005-2007. For the 1980-1982 period, average household labor earnings in 1983 dollars was $43,716 for our top income group and $8,791 for our bottom income group, for a ratio of 4.97. Labor earnings for the top income group grew by 33 percent through 2007, while labor earnings for the low income respondents fell by 2 percent in real terms, resulting in a ratio of 7.10 in 2005-2007. This implies an increase in earnings inequality of 36 percent over the full period, almost all of which occurred in the first half of the sample period.

Inequality in total household income, after taxes and transfers, grew by nearly as much as earnings (Row 2 of table 1). However, after tax income displays a more steady trend over time, with 20 of the overall 35 percent increase in inequality occurring before the mid-1990s.

Figure 1 also depicts consumption inequality between the top income group and the bottom income group. The increase is much less than that of earnings or after tax income, the feature highlighted in Krueger and Perri (2006). In table 1, we see that consumption inequality increased by only 19 percent over the full period, with 12 percent of that change occurring in the first half of the sample. Adjusting the consumption expenditures
to ensure that good-by-good CE aggregates track those from the national income accounts does not change this conclusion. In figure 1, the two measures of consumption inequality are nearly identical.

3.2 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 national income and product accounts, where savings is disposable personal income minus personal outlays. There is a clear downward trend in this series, starting at 9.8 percent in 1980 and falling to 1.7 percent in 2007. This downward trend in the personal saving rate is well known, and is similar to that implied by the flow of funds data.

The implied savings rate in the CE data can be computed as one minus the mean ratio of 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 9 percent in 1980 to over 20 percent in 2007. This sharp increase in implied savings is at odds with the 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 -20 percent at the end of the sample. This decline is nearly a mirror image of the increase implied by consumption data, implying an inconsistency between the CE’s consumption, income, and savings data that is increasing over time. The decline, while of the right sign, is much larger than that obtained from NIPA. 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, which more closely tracks NIPA savings and eliminates part of the sharp downward spikes in savings in the mid-1990s and 2000s.

While mean savings rates are a useful check on the data, we are primarily interested in relative consumption. Using the budget constraint, we can use reported income and savings at the household level to construct an alternative measure of consumption. Specifically, we define implied consumption as after tax income minus reported savings, which is denoted $Y - S$ in figure 1 and table 1. Starting with table 1, the unadjusted savings data implies an increase in consumption inequality of 20 percent between 1980 and 1995,
and 29 percent for the full sample. The adjusted savings implies nearly identical trends in consumption inequality. Both series track income inequality exactly in the first half of the sample, and very closely over the full sample. Moreover, the trends in consumption inequality implied by the budget constraint are markedly different from that obtained from reported consumption data.

Figure 1 depicts the ratio of high income to low income consumption implied by reported savings for each year. Given that the unadjusted and adjusted savings series yield nearly identical trends in inequality, we only plot the adjusted series to avoid clutter. The savings-implied consumption inequality tracks income inequality quite closely, and is different in both level and trend from the reported consumption data.

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

4 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.

4.1 Econometric Approach

To set notation, let the index \( i = 1, \ldots, I \), represent cells defined by income and demographics, as described in section 2. Our 20 goods are indexed by \( j \), and time is indexed by \( t \). Let \( x_{ijt} \) denote reported expenditure on good \( j \) at time \( t \) by group \( i \). Let \( X_{it} \) denote total expenditure at time \( t \) by group \( i \); that is, \( X_{it} = \sum_{j=1}^{J} x_{ijt} \).

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

\[
 x_{ijt} = x^*_{ijt} e^{\phi_i^j + \psi_i^j + v_{ijt}}. 
\]
Here, $\phi_j^t$ 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); $\phi_i^t$ represents mis-measurement specific to $i$ at time $t$ that is common across goods (e.g., the rich may under-report all expenditures); and $v_{ijt}$ is good-group specific measurement error (e.g., food expenditures of the rich are under-reported). We assume that $v_{ijt}$ is classical measurement error; in particular, it is independent of the characteristics of good $j$ and group $i$ at each date $t$. Without loss of generality (given the presence of $\phi_j^t$ and $\phi_i^t$), we normalize the mean of $v_{ijt}$ to be zero for all $t$.

Our estimation consists of two steps. First, we estimate the income elasticities for each good. We assume that Engel curves are log-linear and so income elasticities are constant. Of course, this can only be true locally, unless all elasticities are one. Nevertheless, it provides a tractable framework to address the mis-measurement of expenditure in the CE. We revisit our key results using a popular alternative demand system in the robustness section [to be added]. We assume that true expenditure satisfies:

$$\ln x_{ij}^* = \alpha_{jt} + \beta_j \ln X_{it}^* + \Gamma_j Z_{i0} + \varphi_{ijt}. \tag{2}$$

The term $Z_{i0}$ is a vector of demographic dummies corresponding to age, number of earners per household, and family size, reflecting the categories used to construct the income-demographic cells. We allow the coefficient vector on demographics $\Gamma_j$ to vary across goods. The error term $\varphi_{ijt}$ represents idiosyncratic relative taste shocks which we assume are independent of total expenditure and independent of income elasticities $\beta_j$.

We estimate income elasticities using the 1972-73 Consumer Expenditure Survey. Using the early sample allows us to separate estimation of the income elasticities from the estimation of the implied growth in total consumption expenditures post-1980. Specifically, let $t = 0$ denote observations from 1972 and 1973. We estimate income elasticities using observed expenditures:

$$\ln x_{ij0} = \alpha_{j0} + \beta_j \ln X_{i0} + \Gamma_j Z_{i0} + u_{ij0}. \tag{3}$$

where

$$u_{ij0} = \phi_j^0 + \phi_i^0 + v_{ij0} + \varphi_{ij0}. \tag{4}$$

We pool the years 1972 and 1973, but allow for a different good-specific intercept for each year.

A concern with estimating a demand system like (3) 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. However, we are already using income-category averages, which eliminates measurement error uncorrelated with income. Nevertheless, as modeled above, there may be measurement error that is common across households within an income group. This issue is mitigated by the fact that, at least using NIPA expenditure as our metric, measurement error is less of an issue in the 1972-73 survey than in the later waves. (See Meyer and Sullivan, 2009.) Moreover, we can use NIPA to adjust expenditure for each good \( j \). Specifically, we inflate or deflate expenditure on good \( j \) with a good specific constant so that aggregate expenditure for each good \( j \) in the 1972-73 CE sample equals the corresponding NIPA expenditure for 1972-73. If this adjustment is not sufficient, our income elasticities may be biased. To the extent our income elasticities are subject to attenuation bias, we will over-predict expenditure differences across income groups. In other words, when we invert the demand system in the second stage, we will predict expenditure inclusive of the measurement error characteristic of the 1972-73 survey.\(^5\)

The second stage of our estimation is to invert the demand system (2) to recover an estimate of consumption growth post-1980. Specifically, let \( \gamma_{it} \equiv \Delta \ln X_{it}^* = \ln X_{it}^* - \ln X_{it1}^* \) denote the true consumption growth for income group \( i \) between \( t_1 \) and \( t \), where we take \( t_1 \) to be the pooled base period 1980-1982. (Here we abuse notation by letting \( i \) represent the six income groups, rather than the full 108 income group-demographic cells above, and so \( \gamma_i \) represents the average growth in consumption for all demographic cells in an income group.) Then (2) implies that

\[
\Delta \ln x_{ijt} - \Gamma_j \Delta Z_{it} = \delta^j_i + \delta^i_t + \beta_j \gamma_{it} + \varepsilon_{ijt},
\]

(5)

where \( \delta^j_i = \Delta \phi^j_i \), \( \delta^i_t = \Delta \phi^i_t \), and \( \varepsilon_{ijt} = \Delta (\psi_{ijt} + \varphi_{ijt}) \). Given our assumptions, \( \varepsilon_{ijt} \) is independent of \( \beta_j \). Therefore, \( \gamma_{it} \) can be consistently estimated by least squares. Note that the terms \( \delta^j_i \) and \( \delta^i_t \) reflect changes in systematic measurement error over time, which can be captured by good-time and income group-time dummies, respectively. Identification of \( \gamma_{it} \) comes from the fact that if the income of group \( i \) increases relative to that of group \( i' \), it will increase its relative expenditure, but the increase will disproportionately fall on luxuries.

\(^5\)As our unit of observation is an income-demographic cell, we can instrument for expenditure with any variables that predict expenditure across the income-demographic cells, other than the demographic controls. The obvious instrument is a cell’s (after-tax) income, which is equivalently to averaging over all demographic cells within an income group. We have explored this specification and found that it generates essentially the same estimates for the income elasticities.
To implement (5), we replace $\Gamma_j$ and $\beta_j$ with their estimated values from our first stage. We capture $\delta^i_j$ and $\delta^i_i$ with dummy variables. The lowest income group is omitted, so our estimated $\gamma_{i,t}$ represent the growth in total expenditure for group $i$ relative to the lowest income group. We focus on the coefficient for the top income group, but include all six groups in each stage of estimation. Our two-step procedure requires adjusting the second stage standard errors, which we do following Murphy and Topel (1985).

4.2 Results

Table 2 reports the results of our first stage estimates of each good’s income elasticity. The table also includes the average share of each good out of total expenditure for our 1972-73 CE sample. The first column of estimates uses the reported expenditure in the CE. The second column of estimates adjusts each good by a constant proportion to bring the aggregate CE expenditure for each good in line with its share in the national income and product accounts. We base this adjustment on conversion factors reported in Meyer and Sullivan (2009). This only affects the estimated elasticities through any affect on the distribution of total household expenditures, our right hand side variable. As can be seen from the table, this adjustment does not have a substantial impact on the estimated income elasticities. The standard errors reported next to each estimate suggest that our first stage has a fair degree of precision, particularly for the goods with large expenditure shares.

The estimated income elasticities range from 0.2 for tobacco to 1.9 for non-durable entertainment. Consistent with other studies, food at home has a fairly low income elasticity (0.5), while food away from home has a high income elasticity (1.4). Vehicle purchases is also a large category with a fairly high income elasticity. Housing services, our largest expenditure category, has an income elasticity of 1, as does out of pocket health expenditures.

To provide a sense of how these income 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 income elasticities. We plot the relative expenditure (high income to low income) for each good in figure 3, along with total expenditures. Food at home shows essentially no change between 1980 and 2007, implying equal growth rates for high income and low income households. Conversely, over this period high income households substantially increased entertainment expenditure relative to low income households. Specifically, the log ratio of expenditure on entertainment increased 0.8 log points. On the one hand, the stable ratio of food at
home expenditure suggests little change in consumption inequality.

On the other hand, the dramatic rise in relative entertainment expenditures suggests the opposite – even with the large income elasticity of 1.94, this change in entertainment expenditure still implies an increase in consumption inequality of 41 percent. However, in the terminology of the previous subsection, both these series are contaminated with systematic measurement error.

To see what we can learn from these two series, recall from equation (5) that $\Delta x_{ijt} = \delta^j_t + \delta^i_t + \beta_j \gamma_{i,t} + \epsilon_{ijt}$, for each good $j$ and income group $i$, where we have omitted the demographic shifters for this exercise. The series depicted in figure 3 are log ratios of high income spending to low income spending, so they are already cleaned of good specific measurement error $\delta^j$. Differencing across the two goods eliminates the income specific measurement error $\delta^i$. An unbiased estimate of the change in consumption inequality is therefore the change in relative entertainment expenditure minus the change in food at home, all divided by the relative income elasticities. We perform this calculation for 2005-2007 versus 1980-82. The change in consumption inequality implied by relative spending on entertainment and food is 0.55 log points, more than double the 0.2 log point change implied by reported expenditure, and even considerably larger than the 35 percent change in income inequality.

While food at home and non-durable entertainment are informative due to their large difference in income elasticities, they represent only two of our expenditure categories. Using all goods reduces the impact of the idiosyncratic error terms and provides more precise estimates. Table 3 reports our second stage regression estimates of the log change in consumption inequality. The estimates correspond to $\gamma_{ij}$ from (5), where all estimates are relative to the lowest income group (the omitted group). We consider the change in inequality between 1980-82 and 1992-95 (first row) and between 1980-82 and 2005-07 (second row). When pooling years, we allow for a different intercept for each year. Standard errors are corrected for the presence of generated regressors following Murphy and Topel (1985).

Column (1) reports the second stage estimates using weighted least squares, where the weights are the expenditure shares reported in table 2. The estimated change in consumption inequality is 27 percent for the early period, and 33 percent for the full sample.\(^6\) These numbers are close to those for after tax income reported in table 1, and differ from

\(^6\)OLS produces essentially identical estimates as the food at home versus entertainment comparison: an inequality increase of 0.39 log points through 1995 and 0.55 through 2007. We do not see it as reasonable, however, to weight equally categories with such large differences in expenditure shares. We prefer the table-reported estimates that weight by expenditure share or correct for the predictably larger errors for smaller categories (GLS).
changes in reported consumption inequality. The change in consumption inequality is most dramatic in the first half of the sample, a point we discuss below.

Column (2) performs the same regression but excludes categories that contain durables. 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 somewhat larger when we exclude durables. Specifically, we find a 42 percent increase in inequality in the first half of the sample, and 53 percent over the full sample.

Column (3) uses all consumption categories, but weights by NIPA expenditure shares rather than reported CE expenditure shares. This corrects for the fact that some categories are particularly under-reported in the CE relative to NIPA, and so their shares will be biased down. Keep in mind that there is no need to adjust the dependent variable itself, as any proportional gap between NIPA and the CE is subsumed in our good-specific measurement error term \( \phi_j \) and so is already differenced out. The alternative weighting raises the estimates a few points from our benchmark specification. Specifically, the estimated change in inequality for the first half of the sample increases from 0.27 to 0.31, and the full sample estimate increases from 0.33 to 0.41. These number suggest an increase in consumption inequality that is actually a little greater than the change in after tax income.

The final column of table 3 estimates the second stage using two-step feasible generalized least squares. Specifically, we allow heteroscedasticity across goods to capture that the size of taste shocks or idiosyncratic measurement error may differ across goods. To estimate good-specific residual variances, we use residuals from the WLS specification of column (1). We use these to weight the final estimation. GLS implies an increase in inequality in the first half of the sample of 33 percent, and a slight decline (2 percent) of consumption inequality post 1995, generating an increase of 31 percent over the entire sample.

A common feature across all of our specifications is that the majority of the increase in consumption inequality occurs in the first half of the sample. This pattern reflects the trends in income inequality, but is more pronounced. For example, table 1 reports that roughly 55 percent of the total change in income inequality occurs by 1995 (.20/.37). In

\[7\] Specifically, from the goods listed in table 2, we exclude vehicles, appliances, furniture, and entertainment equipment.

\[8\] We also re-estimated employing the alternative set of estimates of income elasticities from the final column of Table 3. The estimates are essentially unaffected, producing increases in income inequality of 0.28 through 1995, and 0.34 through 2007.
column (1), we see that 80 percent of the change (.27/.33) occurs by the mid-1990s. In this regard, our estimates differ most dramatically from reported CE consumption inequality for the first half of the sample. For example, Krueger and Perri (2006) also find fairly stable inequality in recent years, but our estimates suggest that they under-report the increase in the 1980s.

One way to interpret the large change in consumption inequality in the first half of the sample is that the change in income inequality in the 1980s represented permanent income changes, an interpretation echoed in Attanasio and Davis (1996) for between education group inequality. The relatively small change in consumption inequality in the latter half of the sample is consistent with the interpretation that this latter period involved changes in transitory income inequality. Our results confirm the conclusion of previous studies that the increase in income inequality observed in the 1980s had a larger effect on consumption than the changes post-1995, but suggest that this conclusion has actually been under-stated using reported CE consumption.

5 Conclusion

It is well known that the increase in income inequality in the past three decades has not been mirrored in reported consumption inequality. It is equally well known that the micro data used to measure consumption inequality has become increasingly disconnected from other measures of expenditure and savings, such as those derived from the national income and product accounts. This raises the question of whether the relatively small change in consumption inequality is due, at least in part, to systematic trends in measurement error. We have explored this concern by using the CE’s data on reported savings and income to construct an alternative measure of consumption, via the budget constraint. We have also explicitly modeled a fairly general class of measurement error in reported expenditures, and proposed an econometric approach to obtain a consistent estimate of consumption inequality. Both exercises suggest that the increase in consumption inequality has been large and of the same magnitude as the change in income inequality.

Data Appendix

[to be added]
References


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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor Earnings</td>
<td>4.97</td>
<td>7.04</td>
<td>7.10</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>After Tax Income</td>
<td>4.23</td>
<td>5.16</td>
<td>5.99</td>
<td>0.20</td>
<td>0.35</td>
</tr>
<tr>
<td>Consumption Expenditures</td>
<td>2.44</td>
<td>2.76</td>
<td>2.96</td>
<td>0.12</td>
<td>0.19</td>
</tr>
<tr>
<td>Income minus Saving</td>
<td>4.02</td>
<td>4.93</td>
<td>5.35</td>
<td>0.20</td>
<td>0.29</td>
</tr>
<tr>
<td>Income minus Saving (adjusted)</td>
<td>4.01</td>
<td>4.83</td>
<td>5.42</td>
<td>0.19</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Note: High income refers to respondents who report after tax household income in the 80th through 95th percentiles. Low income refers to respondents in the 5th through 25th 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 change in the first two columns and the first and third columns, respectively. All variables are converted into constant dollars before averaging. The row labeled “Income minus Savings” is reported after tax income minus reported savings. The final row, labeled “Income minus Saving (adjusted)” limits the amount of cash taken out of refinanced mortgages, as described in the text. Definitions of each series and sample construction are given in the data section.
Table 2  Engel Curves from 1972/1973 Expenditure Survey

<table>
<thead>
<tr>
<th>Good Category</th>
<th>1972/1973</th>
<th>CE Share</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing</td>
<td>21.0</td>
<td>0.99 (.03)</td>
<td>0.95 (.03)</td>
<td></td>
</tr>
<tr>
<td>Food at home</td>
<td>15.5</td>
<td>0.49 (.04)</td>
<td>0.46 (.04)</td>
<td></td>
</tr>
<tr>
<td>Vehicle purchasing, leasing, insurance</td>
<td>10.7</td>
<td>1.47 (.07)</td>
<td>1.41 (.07)</td>
<td></td>
</tr>
<tr>
<td>All other transportation</td>
<td>9.9</td>
<td>0.96 (.03)</td>
<td>0.93 (.03)</td>
<td></td>
</tr>
<tr>
<td>Heath expenditures including insurance</td>
<td>5.2</td>
<td>0.95 (.06)</td>
<td>0.92 (.06)</td>
<td></td>
</tr>
<tr>
<td>Food away from home</td>
<td>5.1</td>
<td>1.44 (.06)</td>
<td>1.41 (.05)</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>4.5</td>
<td>0.72 (.06)</td>
<td>0.67 (.06)</td>
<td></td>
</tr>
<tr>
<td>Appliances, phones, computers, with services</td>
<td>4.5</td>
<td>0.71 (.03)</td>
<td>0.69 (.03)</td>
<td></td>
</tr>
<tr>
<td>Men's and women’s clothing</td>
<td>4.4</td>
<td>1.39 (.05)</td>
<td>1.36 (.04)</td>
<td></td>
</tr>
<tr>
<td>Shoes and other apparel</td>
<td>2.9</td>
<td>0.89 (.05)</td>
<td>0.87 (.05)</td>
<td></td>
</tr>
<tr>
<td>Furniture and fixtures</td>
<td>2.8</td>
<td>1.73 (.08)</td>
<td>1.67 (.08)</td>
<td></td>
</tr>
<tr>
<td>Entertainment equipment and subscription television</td>
<td>2.7</td>
<td>1.21 (.06)</td>
<td>1.18 (.06)</td>
<td></td>
</tr>
<tr>
<td>Entertainment fees, admissions, reading</td>
<td>2.0</td>
<td>1.94 (.06)</td>
<td>1.89 (.06)</td>
<td></td>
</tr>
<tr>
<td>Domestic services and childcare</td>
<td>1.7</td>
<td>1.39 (.19)</td>
<td>1.37 (.18)</td>
<td></td>
</tr>
<tr>
<td>Tobacco, other smoking</td>
<td>1.6</td>
<td>0.24 (.06)</td>
<td>0.22 (.06)</td>
<td></td>
</tr>
<tr>
<td>Children’s clothing (up to age 15)</td>
<td>1.4</td>
<td>0.40 (.10)</td>
<td>0.38 (.09)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>1.3</td>
<td>1.94 (.28)</td>
<td>1.89 (.28)</td>
<td></td>
</tr>
<tr>
<td>Personal care</td>
<td>1.2</td>
<td>1.06 (.07)</td>
<td>1.02 (.07)</td>
<td></td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>1.0</td>
<td>1.41 (.09)</td>
<td>1.41 (.08)</td>
<td></td>
</tr>
<tr>
<td>Cash contributions, not for alimony/support</td>
<td>0.5</td>
<td>1.90 (.16)</td>
<td>1.85 (.15)</td>
<td></td>
</tr>
</tbody>
</table>

Total expenditure adjusted based on NIPA? No No Yes

Demographic control dummies for Age, household size, number of earners Age, household size, number of earners

Note: The first column presents each good’s share of total expenditure. The remaining columns report two alternative estimates of each good’s income elasticity, with associated standard errors in parentheses. Column (1) uses the reported expenditure data, while column (2) adjusts expenditure on each good by a constant multiple to ensure aggregate expenditure equals NIPA expenditure. The data come from the 1972 and 1973 CE surveys. See text for details of sample construction and regression specification.
Table 3  Trends in Consumption Inequality Based on Relative Expenditure Patterns

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Change 1980-1995</td>
<td>0.27</td>
<td>0.42</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>(.060)</td>
<td>(.049)</td>
<td>(.059)</td>
<td>(.045)</td>
</tr>
<tr>
<td>Log Change 1980-2007</td>
<td>0.33</td>
<td>0.53</td>
<td>0.40</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(.078)</td>
<td>(.074)</td>
<td>(.074)</td>
<td>(.051)</td>
</tr>
</tbody>
</table>

Categories included

- All
- Those without durables
- All
- All

Weighting by good

- WLS (CE shares)
- WLS (CE shares)
- WLS (NIPA adjusted)
- GLS

Note: This table reports the estimated change in income inequality obtained from the second stage regressions. The dependent variable in each specification was the growth in consumption for each good by income group, with the first row corresponding to growth between 1980-82 and 1992-1995, and the second row corresponding to growth between 1980-82 and 2005-2007. The independent variables were a full set of income-group and good dummies as well as the estimated income elasticities for each good (from Table 2) interacted with income group. The estimated parameters represent the relative growth in total expenditure for high income households relative to low income households. See the specification in the text for full details. The first column implements the second stage by weighted least squares, using CE consumption shares as weights. The second column omits all good categories containing durables. The third column uses NIPA expenditure shares as alternative weights in WLS. The last column implements the second stage using two-step generalized least squares, where we allow expenditure on each good to have its own residual variance. Standard errors are in parenthesis and are corrected for the presence of a generated regressor using the method of Murphy and Topel (1995).
Figure 1: Trends in Inequality:
Ratio of High Income Households to Low Income Households

Earnings
After Tax Income
Y-S
Consumption (NIPA adjusted)
Consumption


Note: This figure depicts the ratio of high income to low income respondents’ reported earnings, after tax income, income minus savings, NIPA adjusted consumption expenditures, and consumption expenditures. High income refers to respondents who report after tax household income in the 80th through 95th percentiles. Low income refers to respondents in the 5th through 25th percentiles. The income minus saving (Y-S) sample uses the adjusted mortgage series as described in the text. Definitions of each series and sample construction are given in the data section.

Figure 2: Mean Saving Rates

1-C/Y
NIPA Savings
S/Y (adjusted)
S/Y (unadjusted)


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 NIPA savings is the national income accounts aggregate private savings rate out of disposable income. The lines labeled S/Y refer to average reported savings divided by average reported after tax income. Adjusted and unadjusted refer to whether we adjust reported new mortgages, as described in the data section of the text. Definitions of each series and sample construction are given in the data section of the text.
Figure 3: Food and Entertainment-- Ratio of High Income Households to Low Income Households

Note: This figure depicts the relative spending of high and low income households on nondurable entertainment and food at home, as well as total expenditures.