What really happened to consumption inequality in the US?*

Orazio Attanasio
University College London, Institute for Fiscal Studies and NBER

Erich Battistin
Institute for Fiscal Studies

Hidehiko Ichimura
University College London and Cemmap

22nd August 2003

Abstract

This paper considers data quality issues for the analysis of consumption inequality exploiting two complementary datasets from the Consumer Expenditure Survey for the United States. The Interview sample follows survey households over four calendar quarters and consists of retrospectively asked information about monthly expenditures on durable and non-durable goods. The Diary sample interviews household for two consecutive weeks and includes detailed information about frequently purchased items (food, personal cares and household supplies). Most reliable information from each sample is exploited to derive a correction for the measurement error affecting observed measures of consumption inequality in the two surveys. Our findings paint rising consumption inequality over time both within cohorts and for the all population. This result strongly contradicts the inequality pattern observed using the Interview sample.

Keywords: Consumption Inequality; Data Collection Methods; Survey Errors

JEL Classification: C13, C42, D12, D91

*This version 20th August 2003. Preliminary and incomplete. Paper for presentation at "Hard-to-measure goods and services: essays in honor of Zvi Griliches", Washington, DC September 2003.

1 INTRODUCTION

The dynamics of inequality over the 1980s and 1990s has received an enormous amount of attention and a voluminous literature studies it. However, most of the existing studies consider either inequality in wages (hourly earnings) or incomes. These studies have documented a very large increase during the 1980s, especially during the first half of that decade followed by some more moderate increases during the last part of the 1980s and the 1990s. Several dimensions of the evolution of inequality have been extensively studied. In particular, several researchers have tried to decompose the observed increase in inequality into increases in inequality between well defined groups (for instance, based on educational attainment) and within groups. Others have focused instead on the decomposition of the increase in inequality between increases in the variance of permanent components of wages and earnings and transitory components.

Very few studies have considered the evolution of inequality in consumption. This is partly due to the paucity of data sources containing individual level consumption data. One of the first papers to use the Consumer Expenditure Survey in the US (CEX in the following) to study the evolution of consumption inequality is Cutler and Katz (1991) that documented an increase in consumption inequality that substantially paralleled the increase in wage and income inequality. Slesnick (1993), on the other hand, analyzes the evolution of poverty in the US and stresses that the picture that emerges when one uses consumption instead of income to measure poverty, is very different, both in terms of levels and of dynamics. Attanasio and Davis (1996) focus on differences across education and year of birth cohorts and report that, coherently with the Cutler and Katz (1991) evidence, especially at relatively low frequencies, relative wage changes are pretty much reflected in relative consumption changes. Slesnick (2001), instead, claims that the evolution of consumption inequality is in sharp contrast to that of income inequality: "...the widely reported U-turn in inequality in the United States is an artifact of the inappropriate use of family income as a measure of welfare. When well-being is defined to be a function of per equivalent consumption, inequality either decreased over the sample or remained essentially unchanged depending on the choice of equivalence scale" (p.154). More recently, Krueger and Perri (2003) discuss results based on the analysis of the Consumer Expenditure Survey until 2001 that are roughly consistent with those reported by Slesnick in his book. In particular, Krueger and Perri (2003) stress that after a modest increase during the first part of the 1980s, consumption inequality is substantially flat. Attanasio (2003) and Battistin (2003), on the other hand, present evidence, based on both the Interview and the Diary segments of the Consumer Expenditure Survey that seems to contradict such a view. Blundell, Pistaferri and Preston (2002) use the PSID until 1992 and show that the inequality of food consumption is increasing in that dataset.

A fair conclusion that can be drawn from the few studies above is that the evidence on the evolution of consumption inequality in the US is far from clearcut and that there is not much agreement in the literature. This state of affairs is particularly unsatisfying because measures of consumption inequality and their evolution can be particularly useful and informative. As Blundell and Preston (1998) stress, under certain condition, consumption comparison can be more informative about welfare differences than income comparisons. Well-being is determined by consumption rather than income. Consumption changes will take into account any mechanism that individual households have to buffer income shocks (either because they are transitory or because they are somehow insured). Deaton and Paxson (1994), spell out some of the implications of the life cycle model for the evolution of the cross sectional variance of consumption inequality. Blundell and Preston (1998) show how to use information on the evolution of income and consumption inequality and some of the insights from the permanent income model to decompose changes in income variances in changes in the variances of transitory and permanent components. An approach complementary to Blundell and Preston (1998) is that of Attanasio and Davis (1996) who frame their evidence in terms of a test of consumption insurance, along the lines proposed by Townsend (1994), Mace (1991) and Cochrane (1991). Essentially, what Attanasio and Davis (1996) label 'uninsured relative wage changes' is closely related to Blundell and Preston 'permanent' shocks, which cannot be self-insured within a life cycle model.

The current lack of consensus and even the small number of studies that have analyzed in detail consumption inequality is in part motivated by the problems of the CEX. The CEX is in fact a relatively small survey mainly collected to compute weights of the CPI, rather then study consumption inequality. There is now substantial evidence that by aggregating CEX data is not easy to obtain figures corresponding closely to figures from National Income and Product Accounts (NIPA) Personal Consumption Expenditure (PCE) data, for many commodities (see McCarthy *et al.*, 2002). While the differences between CEX and NIPA-PCE data can be partly explained by definitional and coverage differences and it does not necessarily arise from problems with the CEX (see, for instance, Slesnick 1992), the amount by which the CEX underestimates national aggregates is massive (around 65%) and compares badly with other surveys such as the Family Expenditure Survey for the United Kingdom. Moreover, the relationship between the aggregated CEX and NIPA-PCE data has worsened considerably during the second part of the 1990s (see for example Battistin, 2003).

The main goal of the CEX (that is the computation of CPI weights) is reflected in the existence of *two* completely separate surveys, one based on retrospective interviews (Interview

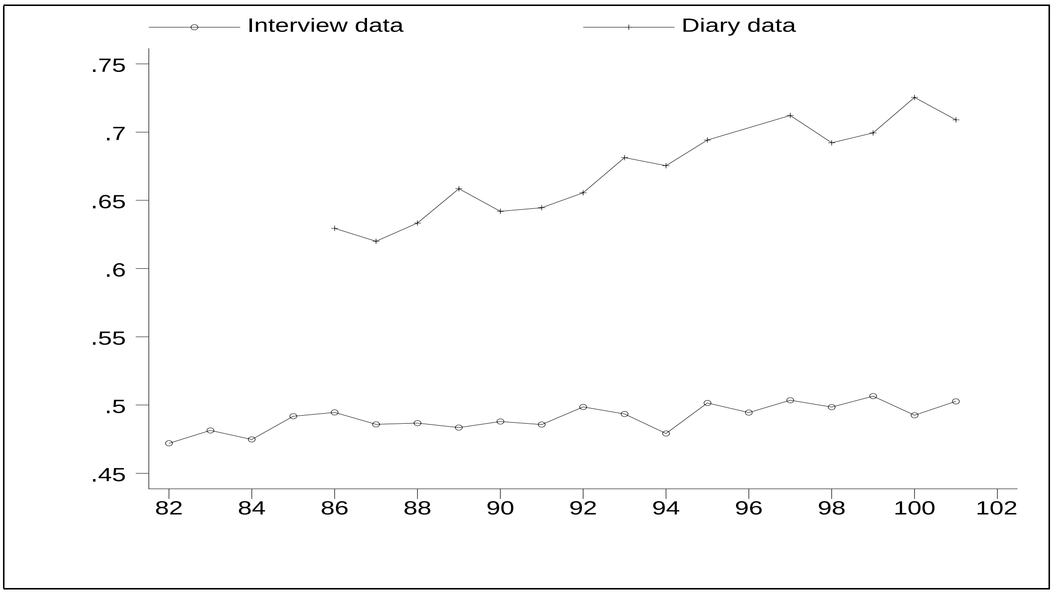

Figure 1: Standard deviation of log per capita monthly expenditure

Sample, IS in the following) and one based on weekly diaries (Diary Sample, DS). The (reasonable) idea is that some expenditure items (such as large, infrequent items) are better measured by retrospective interviews, while others (such as frequently purchased and small items) are better measured by diaries. Indeed, until 1986, the DS only collected information on frequently purchased items. Since 1986, both surveys are in principle exhaustive, but it is quite clear from the BLS literature and from informal communications, that some items are reliably measured in the IS and others in the DS. Of course this creates a problem if one needs information on total consumption expenditure *for a given household*, as it is the case when one wants to study consumption inequality (unless one focuses on inequality as differences in mean consumption across well defined groups of households).

The most puzzling aspect that arises from the analysis of CEX consumption inequality data is that a time series plot of any simple inequality measures reveals completely different patterns for the IS and the DS (see Attanasio, 2003, and Battistin, 2003). In Figure 1, we plot the standard deviation of log of per-adult equivalent non-durable consumption from 1982 to

2001 for IS and DS data. The figures are based on all households headed by an individual aged 25 to 60. The difference is remarkable: the DS plot shows a substantial increase, amounting to around 10 percentage points between 1986 and 2001. The IS plot, on the other hand, shows a path that is substantially flat. As the IS shows an increase in the variance *across education groups* (see Attanasio, 2003, and the evidence we provide below), the plot also implies that the IS shows a decline in inequality within groups.

This evidence is particularly puzzling because the differences in mean non-durable consumption between the two surveys is relatively stable over time (as shown in Section 3).[1] Moreover, in many other dimensions the CEX offers a picture of inequality that is remarkably consistent with that obtained from other (extensively explored) datasets. For instance, Attanasio (2003) shows that the patterns of hourly earnings inequality (and means) across year of birth cohorts and education groups is remarkably similar to the patterns that emerge from the CPS.

The main aim of this paper is to use the information that some items are better measured in the DS and others in the IS and some assumptions (which are partly testable) on the nature of measurement error in the two surveys, to construct an unified picture of the dynamics of consumption inequality in the US over the last twenty years. The fact that some items are better measured in one survey than in the other implies that consistent means for total (non-durable) consumption can easily be obtained for any consistently defined group of consumers combining the two surveys. However, to get an estimate of inequality (say the standard deviation of log consumption or the coefficient of variation of consumption) one needs to deal with the covariance between different consumption items that are well-measured in different surveys.

[1] Attanasio (2003), however, shows that if one conditions on education and year of birth cohort, the differences in the dynamics of inequality between the two surveys is not as remarkable as in Figure 1. Whether this is genuinely due to conditioning or to the small sample sizes in the DS once one crosses education and year of birth cohort is, however, debatable.

However, one can make use of the measurement-error ridden measure in both surveys and, under some assumptions we discuss in Section 4, obtain point estimates of the growth of the coefficient of variation of non-durable consumption.

The rest of the paper is organized as follows. In Section 2, we describe the Consumer Expenditure Survey and discuss the puzzle presented in Figure 1. In particular, we discuss and discard a few simple explanation for the divergence in inequality between the two surveys observed in Figure 1. We also report some of the figures in Attanasio (2003) which show that the pattern of wage inequality in the CEX is very similar to that observed in the CPS. Section 3 presents a puzzle implied by the comparison of means and inequality indicators of non-durable expenditure exploiting information from the two surveys. In Section 4 we write down the basic relationships and assumptions we use to obtain an estimate of the variance of consumption by combining the two surveys we have. In Section 5 we present our results. Section 6 concludes.

2 THE CEX SURVEYS

As we mentioned above, the CEX consists of two separate surveys, the Interview Survey (IS) and the Diary Survey (DS). In this section, we summarize the main features of these two components. In particular, Section 2.1 describes the IS and the DS questionnaires. Section 2.2 discusses the extent to which the IS and the DS are comparable with respect to sample designs, population coverage and information collected. In the same subsection we also discuss the definition of household total consumption we use in the analysis. Finally, Section 2.3 presents some evidence on the sample we use in this paper. The reader interested in more specific details on the survey methodology in the CEX is referred to Battistin (2003) and Bureau of Labor Statistics (2003).

2.1 Diary and Interview Samples

The CEX is currently the only micro-level data set reporting comprehensive measures of consumption expenditures for a large cross-section of households in the United States. Sample consumer units are households (literally, "all members of a particular housing unit who are related by blood, marriage, adoption, or some other legal arrangement"; see Bureau of Labor Statistics, 2003).

The CEX has a long history: the first survey was collected in 1916-1917. More recently, the CEX was collected in 1960-1961 and 1972-1973. As the main scope of the survey is to compute weights for the CPI, data were collected roughly every ten years. As a consequence, the survey methodology and the questionnaires are not homogeneous across the early surveys and this makes inter-temporal comparisons difficult. However, in 1980 it was decided to collect data on a continuous basis with a methodology that was roughly consistent over time. Since then, and especially after 1982, the instruments changed only marginally and in very few occasions. Therefore, with some important caveats (see the discussion in Battistin, 2003) it is conceivable to use the time series of cross sections since 1982 for inter-temporal comparisons.

The survey consists of two *separate* and *independent* samples of households, each of them with its own questionnaire. The IS is a rotating panel including 5000 units each quarter. The DS consists of repeated cross sections of households (around 4500 per year) interviewed over a two-week period. Response rates for the two components are reasonably good (around 80 percent). More detailed characteristics of the two surveys are discussed extensively in Battistin (2003).

In the IS, households are interviewed about their expenditures every three months over five consecutive quarters. The first interview, however, is a contact interview on which there is no information in the public database. After the last interview households are dropped and replaced by a new unit, so that - by design - 20 percent of the sample is replaced out every quarter. Only one person responds for the whole consumer unit, typically the most knowledgeable of expenditures in the family. The percentage of households completing all five interviews is about 75 percent.

In the DS, consumer units are asked to self-report their daily purchases over two consecutive one-week periods using product-oriented diaries. Each diary is organized by day of purchase and by broad classifications of goods and services. Respondents are assisted by printed cues and - whether it is needed - by interviewers at pick-up. The percentage of households completing both diaries is about 92 percent.

Crucial to our exercise is that the two samples drawn are random and representative of the *same* population. The two survey components are in fact based on a *common* sampling frame: the 1980 Census for those households sampled in the 1980s and the 1990 Census for households sampled in the 1990s. Sample designs differ only in terms of frequency and over sampling of DS households during the peak shopping period of Christmas and New Year holidays.

2.2 The information collected in the CEX

In this paper we use twenty years of data from both surveys of the CEX between 1982 and 2001. From 1980 to 1985 the DS only collected information on frequently purchased items, while it became comprehensive in 1986. Because of this our analysis will focus especially on the 1986-2001 period.

Both the DS and the IS collect detailed information on individual commodities, identified by several hundreds of UCC codes. The information on frequently purchased items, and especially food items, is much more detailed in the DS. In the IS food is made only of two large components:

Table 1: Definitions of expenditure categories

Food and Non-Alcoholic Beverages at Home
Food and Non-Alcoholic Beverages Away from Home
Alcoholic Beverages (at home and away from home)
Non-Durable Goods and Services
Newspapers and Magazines
Non-durable Entertainment Expenses
Housekeeping Services
Personal Care
Housing and Public Services
Home Maintenance Services
Public Utilities
Miscellaneous Home Services
Tobacco and Smoking Accessories
Clothing, Footwear and Services
Clothing, Footwear
Services
Heating Fuel, Light and Power
Transportation (including gasoline)
Fuel for Transportation
Transportation Equipment Maintenance and Repair
Public Transportation
Vehicle Rental and Misc. Transportation Expenses

food at home and food away from home. We perform a first level of aggregation on both surveys. This aggregation is mainly dictated by the categories that form the CPI defined by the BLS. We further aggregate these categories into non-durable consumption and other consumption expenditure.

Throughout the analysis we will be focusing on the expenditure on *non-durable goods and services*. The expenditure categories considered have been defined so that definitions are comparable and consistent over time and across surveys (see Battistin, 2003). Expenditure on non-durables is defined according to the definition in Attanasio and Weber (1995): food and non-alcoholic beverages (both at home and away from home), alcoholic beverages, tobacco and expenditures on other non-durable goods such as heating fuel, public and private transports (including gasoline), services and semi-durables (defined by clothing and footwear).

The list of categories included in our definition of non durable consumption are reported in Table 1. Nine expenditure categories are considered, corresponding to roughly 280 and 400 UCC codes for IS and DS data, respectively. [2] We excluded from our definition of consumption expenditures on durables, health, education and mortgage/rent payments are excluded. The main reason to exclude these expenditure items is that they are conceptually very different: for durables we would like to measure the services provided by the existing stock of durables, while education and health expenditure have an important investment component to it. Moreover, in the case of health, the CEX only measures out-of-pocket expenditures. Finally, we excluded rent because we do not have a reliable measure of rental equivalent for home owners and mortgage payments because they are not directly related to housing services. Having said that, we also performed some experiment with total consumption, as defined by the BLS, obtaining results substantially similar to those we report.

The period of time covered by our analysis is characterized by changes in questionnaires and survey methodology in the two surveys. New diaries with more cues were introduced in the DS after 1991; for the IS, except for the food question (that changed in 1982 and 1987), the questionnaire is similar over time. Some UCC codes have changed, but our aggregates are not affected substantially by these changes. Battistin (2003) maps UCC codes into the nine categories in Table 1 accounting for these changes.

Table 2: Sample sizes

Diary sample					
year	**born 1960-69**	**born 1950-59**	**born 1940-49**	**born 1930-39**	**Totals**
1986	257	864	675	419	2,215
1987	383	849	633	466	2,331
1988	345	756	515	374	1,990
1989	422	738	603	412	2,175
1990	497	809	578	459	2,343
1991	574	808	571	396	2,349
1992	603	744	555	352	2,254
1993	624	726	527	370	2,247
1994	560	663	476	295	1,994
1995	542	587	444	265	1,838
1996	688	758	504	328	2,278
1997	722	740	579	411	2,452
1998	674	751	543	347	2,315
1999	985	997	716	481	3,179
2000	1,021	931	722	457	3,131
2001	1,044	950	691	418	3,103

Interview sample					
year	**born 1960-69**	**born 1950-59**	**born 1940-49**	**born 1930-39**	**Totals**
1982		2,881	2,883	1,979	7,743
1983	242	3,226	2,804	1,901	8,173
1984	435	3,059	2,577	1,729	7,800
1985	572	2,765	2,201	1,587	7,125
1986	989	3,498	2,690	2,023	9,200
1987	1,217	3,376	2,609	1,995	9,197
1988	1,436	2,852	2,411	1,611	8,310
1989	1,724	2,768	2,412	1,590	8,494
1990	1,943	2,904	2,309	1,472	8,628
1991	2,030	2,862	2,181	1,568	8,641
1992	2,334	2,869	1,978	1,467	8,648
1993	2,424	2,899	2,159	1,424	8,906
1994	2,380	2,869	2,132	1,384	8,765
1995	2,133	2,526	1,849	1,213	7,721
1996	2,954	3,244	2,380	1,541	10,119
1997	3,088	3,363	2,347	1,585	10,383
1998	3,043	3,221	2,268	1,691	10,223
1999	4,331	4,493	3,147	2,232	14,203
2000	4,393	4,381	3,259	2,216	14,249
2001	4,314	4,099	3,207	1,932	13,552

2.3 The selected sample

In this analysis, we focus on household headed by individuals aged at least 23 and no more than 73 and not self-employed. The family head is conventionally fixed to be the male in all husband/wife families (representing the 56 percent and 53 percent of the whole sample for IS and DS data, respectively). Battistin (2003) presents a detailed description of less important selection criteria used to derived the working sample considered in the analysis.

Although the two surveys are designed to be representative of the same population, significant differences in the two samples are found along several dimensions and with a different pattern over time (even using the population weights provided by the BLS). To control for observed composition differences between the two samples (for instance the DS is slightly more educated than the IS sample), Battistin (2003) weights DS households with the inverse of the probability of being in the IS sample, estimated as a function of characteristics common across the two samples (propensity score weighting; see Battistin *et al.*, 2003, and Hirano *et al.*, 2003). We use the same procedure here. However, results using BLS population weights or these weights are substantially identical.

To give an idea of sample sizes, Table 2 reports, for each year, the size of the sample we end up using from each of the two surveys. As is obvious from the table, sample sizes are not huge, particularly for the DS sample. This represents a real problem if one wants to control for several observable characteristics, such as year of birth and education.

Monthly expenditure in the DS is defined as $26/12 = 2.16$ times the expenditure observed

[2] As we mentioned above, expenditures referring to "Housing and Public Services" and "Non-durable Services" have been introduced in the DS only after 1986, with the exception of very few items for "Home Maintenance Services" and "Non-durable Entertainment Expenses". Similarly, information on "Fuel" and "Transportation" expenses is not available from public tapes between 1982 and 1985. As for IS data, the time series of food at home expenditure presents discontinuities introduced by changes in survey design in 1982 and 1987 (see Battistin, 2003). A detailed description of the items used to define the categories of non-durable consumption can be downloaded at http://www.stat.unipd.it/~erich/papers.html, separately for IS and DS data.

over two weeks, assuming equally complete reporting. Family consumption is adjusted using an equivalence scale which depends on the number of adults and children in various age ranges (see McClements, 1977). Real expenditures are obtained using the Current Price Index published by the BLS.

Both surveys collect information on a very large set of household characteristics (demographics and work-related variables) as well as on income and assets (using a twelve-month recall period). The latter information is subject to top-coding in both components of the CEX and known to be not as reliable as the expenditure information: the amount of incomplete income reporters is about 13 percent in the two surveys and missing values are currently not imputed (see McCarthy *et al.*, 2002). Since the percentage of incomplete income reporters is so high, we included all of them in the final sample. In our robustness analysis we check whether the exclusion of households with incomplete income responses makes any difference to our main results.

Being the CEX a time series of repeated cross sections we will perform most of the analysis using synthetic panel techniques. We will define groups on the basis of year of birth and educational attainment of the household head. We define cohorts as groups of households headed by individuals born in a given time interval. In particular, we consider individuals born in the 1930s, 1940s, 1950s and 1960s. We define three education groups: individuals with less than 12 years of education ("HS dropouts"), with 12 years of education ("HS graduates") and more than 12 years of education ("Other").

3 EVIDENCE ON CONSUMPTION AND WAGES

In this section we present *three* sets of results. First, we compare expenditure means from the two CEX surveys to aggregate values from national accounts data. We find important differences between IS and DS figures and, crucially, in the ratio of CEX to NIPA figures over time. Second we present some data on the evolution of wage inequality exploiting CEX and CPS. This evidence shows that the overall picture painted in the two surveys is essentially the same. Finally, we present some additional information on the evolution of consumption inequality from the two survey components of the CEX.

3.1 Consumption means

In Figure 2, we compare total non durable expenditure in published CEX tables to the figures one obtain for a similar category in the NIPA accounts for PCE.[3] The CEX aggregates are computed using the population weights provided by the BLS. Two elements are worth stressing from this picture. First, even though there are some important definitional differences, discussed extensively by Slesnick (1992, 2001) amongst others, one can not help noticing that the CEX figure massively *understates* the one from PCE data. While this does not necessarily mean that for every single consumption item the PCE provides superior information (see for instance the discussion in McCarthy *et al.*, 2002), this evidence contrasts sharply with similar comparisons for the UK, where aggregating a time series of individual cross sectional data, one obtains close to 95% of non durable consumption, as documented in Banks and Johnson (1998). Second, while the divergence between CEX and PCE is roughly constant in the first of the sample, the difference seems to increase in the second part of the 1990s. This evidence is consistent with that reported by other researchers including Slesnick (1992), Sabelhaus (1996), Slesnick (2001) and several BLS internal publications.

As we are interested in combining the information from the IS and the DS, it might be

[3]We are grateful to David Johnson at the BLS for making this graph available to us. The contents of this figure are comparable to those of Figure 3.2 in Slesnick (2001; page 51), although the latter figure looks at total expenditure on durable and non-durable goods.

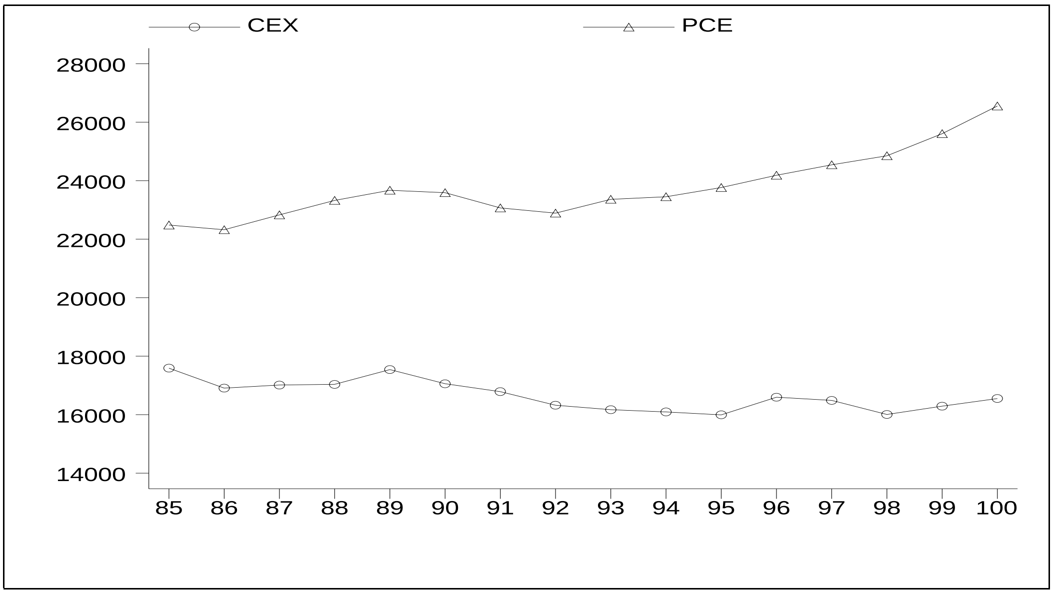

Figure 2: Non-durable expenditures in 2000 dollars - Consumer Expenditure Survey (CEX) and Personal Consumption Expenditures (PCE)

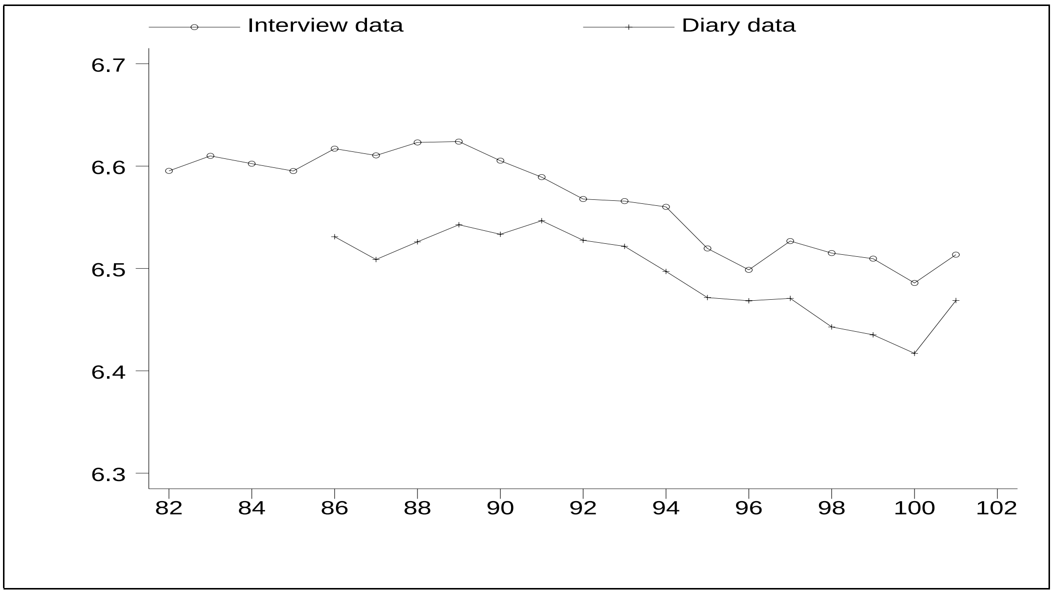

Figure 3: Mean of log monthly expenditure on non-durable goods (2001 dollars)

Table 3: Between-group inequality: growth by education groups

education	Diary		Interview	
	1 year	5 year	1 year	5 year
HS dropouts	-0.004	-0.022	-0.001	-0.011
HS graduates	-0.002	0.009	0.001	0.004
other	0.006	0.012	0.001	0.007

interesting to compare the estimates of aggregate non durable consumption that emerge from the two surveys. In Figure 3, we plot the two time series and notice that non durable consumption is consistently and substantially higher in the IS than in the DS. The other important fact to notice in Figure 3 is that the difference between the two surveys is substantially constant (particularly after 1991). This hypothesis can be tested statistically and can not be rejected at standard significance levels. Battistin (2003) finds that the relationship between mean expenditures in the two surveys varies a great deal considering a similar analysis by expenditure group.

In what follows we will extensively discuss consumption inequality as measured by the coefficient of variation of non-durable consumption (or the standard deviation of log consumption) within cohorts of individuals or for the all population. However, it is also interesting to consider the difference in mean consumption across different, well-identified groups of households as a measure of between-group inequality. Since the literature on wage inequality in the US has paid much attention to changes in the return to education, it is natural to consider groups on the basis of the educational attainment of the household head. As we have already said, *three* groups will be considered: households headed by high-school dropouts, households headed by high school graduates and households headed by individuals with more than high school education.[4]

[4]The analysis of the relative consumption of different education groups is much easier than the analysis of the cross sectional variance of consumption and does not require the methodology we propose below for combining information from the two available surveys. As we deal with means and clearly identified groups of households, we can easily combine the information from the two surveys. Moreover, the results we obtain are very clearcut.

Table 4: CEX-CPS Comparison

correlation of hourly wages				**correlation of residuals**			
	CPS	IS	DS		CPS	IS	DS
CPS	1.0000			CPS	1.0000		
IS	0.9669	1.0000		IS	0.4581	1.0000	
DS	0.9466	0.9301	1.0000	DS	0.4359	0.4139	1.0000

In the two surveys, having divided households according to the education attainment of the household head and year of birth cohort, we compute one-year and five-year changes in the mean of log non durable consumption for each group cell and regress these changes on year and cohort dummies. We then compute, for each of the three education group, the mean of the residuals of these regressions, which we report in Table 3. As is clear from the table, the average rate of growth of consumption over this period was highest for the college graduate and lowest for the high school dropouts.[5] This evidence indicates that the pattern of consumption inequality across education groups is consistent across surveys and is roughly in line with the evidence on wage inequality across education groups (see Attanasio, 2003). This type of evidence is behind the results obtained by Attanasio and Davis (1996).

3.2 Wage inequality in the CEX and in the CPS

Before starting with the analysis of consumption inequality, it is worth reporting some information on how the CEX performs in terms of inequality dynamics of wages (hourly earnings). This part of the analysis is taken from Attanasio (2003). From one of the supplementary files it is possible to obtain measures of earnings and hours worked for each household member. We can compute the analogous figure from March CPS files. Using this information, the average and the standard deviation of log male hourly earnings is computed for our groups in both datasets.

[5]We obtain virtually identical results if we do not condition on cohort dummies in the first step regressions.

Top-coding is different in the two datasets and has changed over time. In both data sets, we compute the mean and standard deviation of each cell by fitting (in each cell) a log normal distribution truncated at the top-coding level. To check whether the information on wages and wage inequality is similar in the CEX and in the CPS, we first divide each sample in cells defined by cohort (year of birth), time and education level. We then compute the correlation coefficient between average log hourly earnings in the CPS, the CEX-IS and the CEX-DS and report them in the first panel of Table 4. As can be seen, the correlation is remarkably high. As it may be argued that a large fraction of this correlation can arise because of difference across education level, we regress, for each survey, the average log wage on time and group dummies (where groups are defined by education and year of birth). In the second panel of the table we report the correlation coefficients among the residuals of these regressions. Once again we obtain remarkably high coefficient, indicating that change in relative wages are similar in the two datasets.

Next we compute the standard deviation of log hourly earnings within each cell in the three datasets. To check that they move together, we regress the standard deviation in the CPS on the standard deviation in the CEX. We obtain a strong positive relation that explains 64% of the movements in the CPS standard deviations.

3.3 Consumption inequality

In Figure 1 we plotted the standard deviation of log non-durable consumption in the IS and the DS. We have already stressed the difference in the time series pattern of the two measures of consumption. Very similar evidence can be obtained considering other measures of inequality. Battistin (2003) for instance, reports evidence on the Gini coefficient, and various measures belonging to the Generalized Entropy Family. This difference is particularly puzzling given

Table 5: Percentage of zero expenditures

Interview	86-89	90-92	93-95	96-98	99-2001
Food and non-alcoholic beverages at home	0.00	0.00	0.00	0.00	0.00
Food and non-alcoholic beverages away	0.10	0.12	0.12	0.13	0.15
Alcoholic beverages (at home and away)	0.37	0.43	0.44	0.49	0.52
Non-durable goods and services	0.03	0.03	0.03	0.05	0.05
Housing and public services	0.03	0.02	0.02	0.02	0.02
Tobacco and smoking accessories	0.56	0.61	0.64	0.66	0.71
Clothing and footwear	0.13	0.15	0.18	0.22	0.28
Heating fuel, light and power	0.11	0.09	0.08	0.07	0.07
Transport (including gasoline)	0.02	0.02	0.02	0.02	0.03

Diary	86-89	90-92	93-95	96-98	99-2001
Food and non-alcoholic beverages at home	0.02	0.01	0.02	0.02	0.02
Food and non-alcoholic beverages away	0.06	0.08	0.11	0.10	0.11
Alcoholic beverages (at home and away)	0.42	0.47	0.50	0.55	0.55
Non-durable goods and services	0.06	0.06	0.08	0.09	0.11
Housing and public services	0.34	0.35	0.36	0.37	0.37
Tobacco and smoking accessories	0.57	0.62	0.67	0.68	0.71
Clothing and footwear	0.26	0.27	0.30	0.32	0.36
Heating fuel, light and power	0.54	0.49	0.54	0.54	0.55
Transport (including gasoline)	0.04	0.06	0.07	0.06	0.07

the substantial stationarity between the difference in mean consumption in the two surveys over time. To make sense of the remarkably different pattern we observe in Figure 1, we start analyzing simple explanations. In particular, we check whether the difference could be explained by changes in questionnaires and survey methodology, by changes in the frequency of purchases of commodities, by changes in the willingness to answer surveys and by changes in the differences in sample compositions.

- Changes in questionnaires and survey methodology. From official BLS documents, analysis of the questionnaires and conversations with BLS staff, we could not identify any substantive change that would explain the observed differences. The only substantive change occurs in the IS for the question for food in, that changes in 1982 and 1987 (see Battistin, 2003), both years substantially outside our interval. Moreover, such an explanation would be difficult to square with the absence of changes in the difference of means.

- Increase in the number of zeros in the DS. A potentially attractive explanation is the following. Over time, people shop less frequently and purchase larger quantities in each shopping trip. As the horizon of the two surveys is different (two weeks for the DS, three months for the IS), the DS would result in the same mean (or the same difference) but larger variances over time because of the increased number of zeros. Table 5 (see Battistin, 2003) shows that the number of non-zero expenditures for non-durable items varies over time for same groups (particularly, it decreases for 'alcohol', 'tobacco' and 'clothing') but with the same pattern across samples.

- Over time people have become less willing to answer accurately or answer at all. One can check whether attrition rates the fraction of incomplete income responses have changed substantially over time and differentially so for the two surveys. Anecdotal evidence shows that wealthier individuals are less willing to answer. An increase in income inequality might result on a larger number of wealthier individual being lost in the IS survey. While such an hypothesis has been suggested (by Sabelhaus, 1996, among others) it is unlikely that it could explain the observed different pattern between IS and DS. We control for differences in sample composition by using our propensity score weights and do not find that these differences (or other differences in sample composition) can explain the different dynamics of inequality in the two samples.

Table 6: Survey selection

Commodities in D	**Survey**
Food and Non-Alcoholic Beverages at Home	Diary
Food and Non-Alcoholic Beverages Away from Home	Diary
Alcoholic Beverages (at home and away from home)	Diary
Non-Durable Goods and Services	Diary

Commodities in R	**Survey**
Housing and Public Services	Interview
Tobacco and Smoking Accessories	Interview
Clothing, Footwear and Services	Interview
Heating Fuel, Light and Power	Interview
Transportation (including gasoline)	Interview

4 OUR APPROACH

Rather than pursuing further the attempt to explain the difference between the inequality measures observed in the IS and DS, in this section we propose a different approach. The main reason for the existence of two different samples is that the BLS believes that different methodologies are more appropriate in measuring different commodities. Indeed, the weights for the CPI, as well as aggregate estimates produced by the BLS, ultimately use both information from the IS and the DS.

To estimate average consumption, the BLS follows the standard international procedure of exploiting information from recall questions for more durable items bought in the quarter prior to the interview. Diary-based records of purchases carried out within a two-week period are used for more frequently purchased items such as food. According to Bureau of Labor Statistics (2003), neither survey is expected to measure accurately all components of consumption.

Since aggregate consumption is obtained from integrated survey data, the question then arises of which survey component provides more accurate information on different items. Let

C^* be total expenditure on *all* non-durable commodities

$$C^* = C_D^* + C_R^*,$$

where C_D^* and C_R^* represent expenditures on items that are better measured in DS and IS data, respectively. In Table 6 we list which categories belong to the D and R groups according to Bureau of Labor Statistics (2003).[6] Obviously, being both surveys exhaustive, a measure of R goods exist also in the DS and a measure of D good exist also in the IS.

More accurate estimates of average non-durable consumption can obviously be easily obtained combining the two surveys. Nevertheless, it is worth noting that straightforward pooling can not be implemented since diary and recall expenditures are *not* observed for the same survey households. If one is interested in the variance of non-durable consumption, the problem is more complicated. The quantity we are interested in is the variance of C^*, where $C^* = C_D^* + C_R^*$ and

$$Var(C^*) = Var(C_D^*) + Var(C_R^*) + 2Cov(C_D^*, C_R^*). \tag{1}$$

Remember that each household is either observed in the DS or in the IS. For each households in both surveys we observe expenditure on both D and R commodities. We denote by C^d total non-durable expenditure measured in the DS and with C^r total non-durable expenditure observed in the IS. Observed consumption in the two surveys is then given by

$$\begin{aligned} C^d &= C_D^d + C_R^d, \\ C^r &= C_D^r + C_R^r. \end{aligned}$$

[6] It is worth noting that, although the level of aggregation considered by the BLS is finer, the classification procedure exploited in what follows broadly reflects the one currently being used in the publication of CEX data. See also the discussion in Battistin (2003), where evidence on the validity of this classification is produced with respect to other expenditure surveys in the world.

As we mentioned above, the BLS thinks that the DS measures accurately commodities in D, while the IS measures well R commodities. We translate this assertion in the following extreme assumption.

Assumption 1

$$\begin{aligned} C_D^d &= C_D^*, \\ C_R^r &= C_R^*. \end{aligned}$$

Accordingly, the reporting error in DS and IS figures comes from C_R and C_D, respectively

$$\begin{aligned} C^d - C^* &= C_R^d - C_R^* = \nu_R^d, \\ C^r - C^* &= C_D^r - C_D^* = \nu_D^r. \end{aligned} \tag{2}$$

If Assumption 1 is satisfied, the mean of IS and DS errors is identified by $E(\nu_D^r) = E(C_D^r) - E(C_D^d)$ and $E(\nu_R^d) = E(C_R^d) - E(C_R^r)$, respectively. By analogy, the mean of total expenditure can be estimated by $E(C^*) = E(C_R^r) + E(C_D^d)$. Figures for IS and DS errors as proportion of total non-durable expenditure are reported in Battistin (2003), where implications on estimated saving rates from the IS are also discussed.[7]

In what follows we will be interested in the variance of log consumption and we will be approximating it by the square of the coefficient of variation

$$Var(lnC^*) \simeq \frac{Var(C^*)}{E(C^*)^2} = CV(C^*)^2.$$

Such a relationship holds exactly if C^* is log-normally distributed. Blundell and Lewbel (1999) provide strong empirical evidence to support the fact that, in a variety of datasets the cross

[7]While the assumption of no measurement error in D commodities in the DS and of no measurement error in R commodities in the IS is extreme, it is made here only for expositional convenience. As we discuss later, it can be slightly relaxed without changing the substance of our argument or our results.

sectional distribution of consumption seems to be very well approximated by a log-normal. As we can easily estimate the denominator of the coefficient of variation by combining the two data sets, we will focus here on the estimation of the cross sectional variance for a given group of individual households. From equation (1), it is clear that to estimate the variance of C^* we lack an estimate of the $Cov(C^*_D, C^*_R)$. From equation (2) and Assumption 1, it is clear that the observed covariances in the two surveys are informative about such a covariance, since

$$\begin{aligned} Cov(C^d_D, C^d_R) &= Cov(C^*_D, C^*_R) + Cov(C^*_D, \nu^d_R), \\ Cov(C^r_D, C^r_R) &= Cov(C^*_D, C^*_R) + Cov(C^*_R, \nu^r_D). \end{aligned}$$

Clearly, if we assumed that either $Cov(C^*_D, \nu^d_R) = 0$ or $Cov(C^*_R, \nu^r_D) = 0$, it would be possible to identify $Cov(C^*_D, C^*_R)$, which is what we are interested in. However, notice that if we assumed that $Cov(C^*_D, C^*_R) = 0$, we could test whether $Cov(C^*_R, \nu^r_D) = 0$ and viceversa. Moreover, we can easily test whether

$$Cov(C^*_D, \nu^d_R) = Cov(C^*_R, \nu^r_D),$$

since this is equivalent to

$$Cov(C^r_D, C^r_R) - Cov(C^d_D, C^d_R) = 0.$$

Figure 4 presents bootstrapped confidence intervals for this difference. We clearly reject the null. We therefore conclude that it would be dubious to base our inference on either of the assumptions mentioned above.

However, if one is interested in the dynamics (rather than the level) of consumption inequality, one could focus on assumptions that allow the identification of such changes. Since

$$\Delta Var(lnC^*) \simeq \Delta \frac{Var(C^*_R)}{E(C^*)^2} + \Delta \frac{Var(C^*_R)}{E(C^*)^2} + 2\Delta \frac{Cov(C^*_D, C^*_R)}{E(C^*)^2}. \tag{3}$$

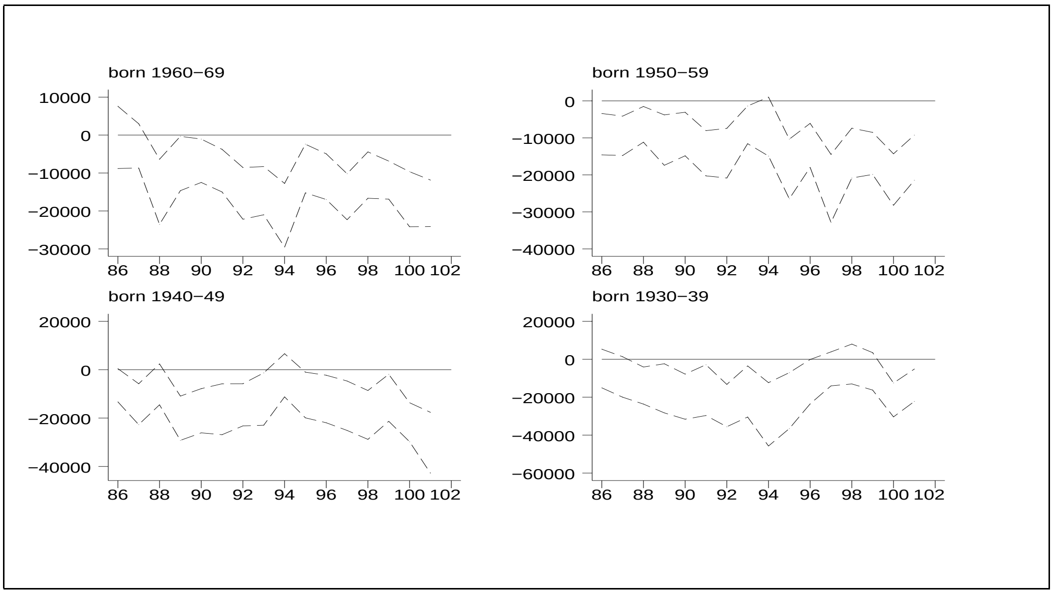

Figure 4: Confidence intervals for $Cov(C_D^d, C_R^d) - Cov(C_D^r, C_R^r)$

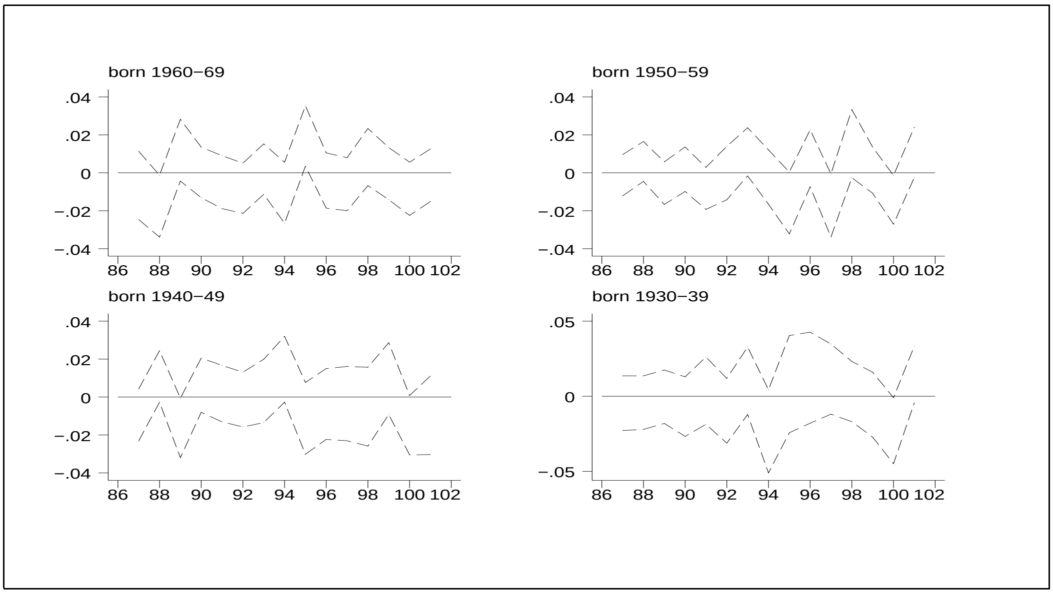

Figure 5: Confidence intervals for $\Delta\frac{Cov(C_D^d, C_R^d)}{E(C^*)^2} - \Delta\frac{Cov(C_D^r, C_R^r)}{E(C^*)^2}$

It is easy to show that the last term on the right hand side of this expression can be identified if we assume

Assumption 2

$$\Delta\frac{Cov(C_D^*,\nu_R^d)}{E(C^*)^2}=\Delta\frac{Cov(C_R^*,\nu_D^r)}{E(C^*)^2}=0.$$

This assumption is equivalent to

$$\Delta\frac{Cov(C_D^d,C_R^d)}{E(C^*)^2}-\Delta\frac{Cov(C_D^r,C_R^r)}{E(C^*)^2}=0,$$

which can be easily tested. In Figure 5 we report the bootstrapped confidence intervals for the this difference. The figure clearly indicates that, in this case, we do not reject the null. Under Assumption 1 and Assumption 2, changes in $Var(lnC^*)$ are identified by using observed data (Battistin, 2003, exploits this result to bound the *level* of inequality). In what follows, we use this approach extensively to report our results on the dynamics of consumption inequality.[8]

5 RESULTS

Figure 6 presents estimation results for consumption inequality in the entire sample. In particular, we add to the two lines in Figure 1 our point estimate of the standard deviation of consumption over the period we study. We also draw the bootstrap 5% critical regions for the hypothesis of constant inequality over time for our estimates (dotted lines). To improve the

[8]It is worth noting that

$$\Delta\frac{Cov(C_D^r,C_R^r)}{E(C^*)^2},\qquad\Delta\frac{Cov(C_D^d,C_R^d)}{E(C^*)^2}$$

are both good estimators for the unknown covariance term in (3). The problem then arises of how to combine the two. The estimation problem can be set up as a GMM, where the estimation target is

$$\Delta\frac{Cov(C_D^*,C_R^*)}{E(C^*)^2}$$

which is over-identified. One can estimate this parameter using efficient GMM framework. In what follows, we will use the *average* of the two terms instead.

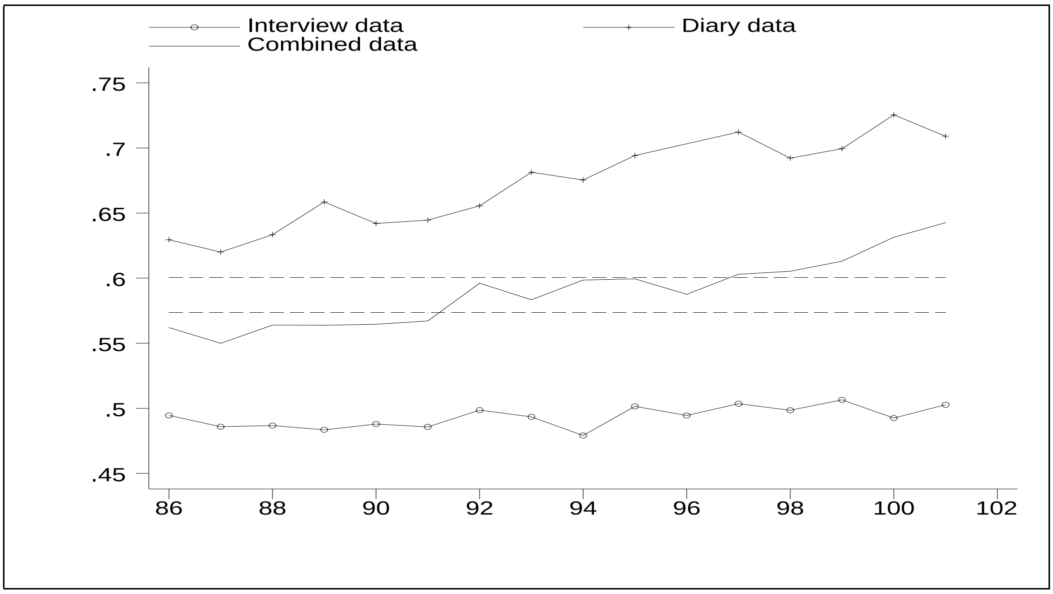

Figure 6: Inequality growth using combined information

Table 7: Inequality growth over time

time	Interview	Diary	Estimated
$\Delta 1990-86$	-0.007	0.013	0.003
$\Delta 1995-90$	0.014	0.052	0.035
$\Delta 2000-95$	-0.009	0.031	0.032

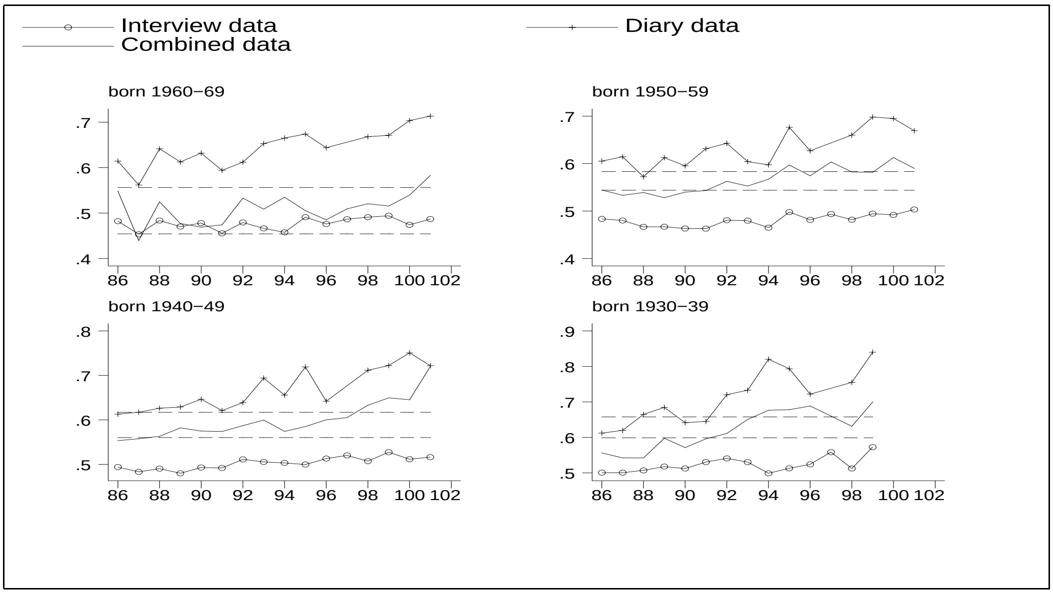

Figure 7: Inequality growth by cohort using combined information

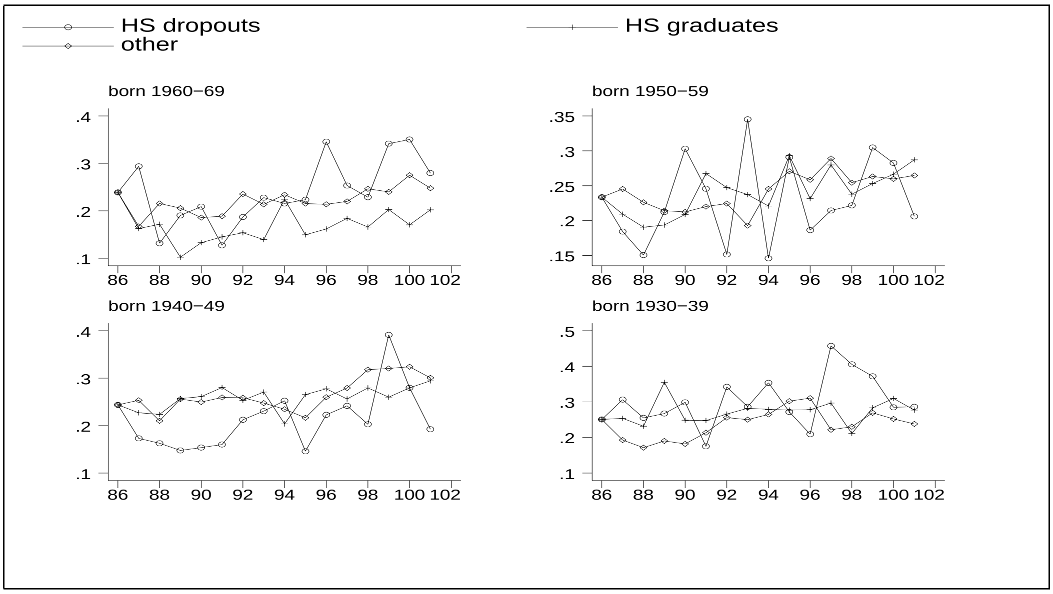

Figure 8: Inequality growth by cohort and education group using combined information

readability of this figure, Table 7 reports observed inequality growth over time from IS and DS data. As we mentioned above, the *level* of inequality is not identified. Strictly speaking the figure is only informative about *changes* in inequality over time. We pin down the level so to place the initial level of inequality between the initial levels of inequality in the DS and IS.[9]

Perhaps not surprisingly, the estimate we obtain for the changes in inequality are in between the two divergent paths for the DS and the IS. However, having gone through the exercise of using optimally the information coming from the IS and the DS, the interesting question is a quantitative one: by how much does inequality increase according to our estimates? The answer is: by a substantial amount. According to our results, inequality raises by about 7 percentage points over the 1990s. This results is economically very different from that reported by Krueger and Perri (2003) and from the change observed in the IS, which is approximately 1 percentage point increase during the same period.

[9] The level in 1986 is taken as the average of observed levels in IS and DS data for that year. Accordingly, dotted lines refer to bootstrap confidence intervals for the median of estimated inequality over the period 1986-2001.

In Figure 6, we compute consumption inequality using all households headed by an individual aged 23-60. As stressed by Deaton and Paxson (1994) and by Blundell and Preston (1998), from an economic point of view, it makes more sense to follow the evolution of consumption inequality of a fixed cohort as it ages. We therefore divide our sample on the basis of the year of birth of the household head and follow four year of birth cohorts as they age. In particular, we consider households headed by individuals born in the 1930s, 1940s, 1950s and 1960s and plot the standard deviation of consumption as estimated following the same procedure we used to obtain Figure 6. Our results indicate an increase in consumption inequality for every cohort except, perhaps the youngest one. This is consistent, as stressed by Deaton and Paxson (1994), with a life cycle model with (uninsured) permanent shocks. Once again, the picture that emerges from the IS is quite different: the standard deviation of log consumption is remarkably flat for all four cohorts considered.

In Table 3, we showed that inequality across education groups has increased during both the 1980s and the 1990s. It is therefore interesting to check what happens to consumption inequality within education groups. Such an exercise can be informative about the nature of the shocks that have affected the households in the sample: permanent shocks (for instances driven by skill biased technical change) which are more likely to be reflected in consumption inequality might be identified more easily in Figure 6 and 7, which considers the population at large, or Table 3, which considers differences among education groups. Instead, the role of temporary shocks (which might not be reflected into consumption inequality) might be more prominent when one considers the evolution of inequality within education groups. Therefore, in Figure 8, we consider the evolution of the coefficient of variation within education groups (only figures for estimated inequality are reported).

Two elements emerge from such a figure (see Attanasio, 2003). First, the difference between IS and DS is much less variable over time than in Figure 6 and 7. Second, and as a consequence, our estimates that combine the information from the IS and the DS, show a modest increase in inequality. This is consistent with the hypothesis mentioned above that shocks within education groups are more easily absorbed and insured, either by self-insurance or by other mechanisms.

6 CONCLUSIONS

This paper starts from the puzzle that, when following the evolution of consumption inequality during the late 1980s and the 1990s in the US in the two available surveys, one obtains very different and contradictory patters. We have tried to use the information that some components of consumption are better measured in the Diary survey while others are better measured in the Interview survey to obtain an *new* view on the pattern of inequality in the US. Obviously, as we do not observe the same households in the two surveys, we can get a point estimate of the cross sectional variance of total non-durable consumption only by making some assumptions on the nature of measurement error.

From our analysis we conclude that consumption inequality has increased substantially more than what is indicated by the path of the standard deviation of log consumption in the Interview Survey (and substantially less than what indicated in the Diary Survey). The increase, of 7 percentage points, is economically significant. We also find that, when we perform the analysis by year of birth cohort, all cohorts but one show a substantial increase in inequality. When we instead split the sample by education groups, we find that the increase is much more modest.

Some of the assumptions we make can be tested. We end up using a set of assumptions that are not rejected by the available data and that allow to identify *changes* in consumption inequality (rather than the level of inequality). Assumption 2 is *crucial* to identify changes over time in consumption inequality but not its level. Interval estimates of the inequality level can be obtained if bounds are defined for the unknown covariance between expenditure on items in R and D. Battistin (2003) shows that Cauchy-Schwartz bounds on the *level* of consumption inequality can be derived using Assumption 1. By looking at these bounds, Battistin (2003) shows that the effect of reporting errors should be *massive* to discard increasing inequality over time, even if Condition 2 were not satisfied. Battistin (2003) performs a sensitivity analysis with respect to the *true* value of the correlation coefficient between items in R and D, ρ^* say. Inequality is found to be statistically increasing over time unless the effect of reporting errors on ρ^* is such that observed correlations in IS and DS data and ρ^* are more than 100% apart (in absolute terms).

References

[1] Attanasio, O.P. (2003), "Consumption Inequality: What we know and what we can learn from it," Lecture to the Annual Meeting of the Society of Economic Dynamics, New York, June 2002

[2] Attanasio, O.P., and Weber, G. (1995), "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103, 95, 1121-1157

[3] Attanasio, O.P., and Davis, S.J. (1996), "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy*, 104, 6, 1227-1262

[4] Banks, J., and Johnson, P. (eds.) (1998), *How Reliable is the Family Expenditure Survey? Trends in Incomes and Expenditures over Time*, London: The Institute for Fiscal Studies

[5] Battistin, E. (2003), "Errors in survey reports of consumption expenditures," Working Paper W03/07, Institute for Fiscal Studies, London. Revised version available at www.ifs.org.uk/staff/erich_b.shtml

[6] Battistin, E., Miniaci, R., and Weber, G. (2003), "What do we learn from recall consumption data?," *Journal of Human Resources*, 38, 2, 354-385

[7] Blundell, R., and Lewbel, A. (1999), "Puzzles of consumption and income distribution explained: Gibrat's law for permanent income," unpublished manuscript, University College London

[8] Blundell, R., and Preston, I. (1998), "Consumption inequality and income uncertainty," *Quarterly Journal of Economics*, 113, 603-640

[9] Blundell, R., Pistaferri, L., and Preston, I. (2002), "Partial insurance, information and consumption dynamics," working paper W02/16, Institute for Fiscal Studies, London

[10] Bureau of Labor Statistics (2003), *Handbook of Methods, Chapter 16, Consumer Expenditures and Income*, available at www.bls.gov

[11] Cochrane, J.H. (1991), "A simple test for consumption insurance," *Journal of Political Economy*, 99, 957-976

[12] Cutler, D. and Katz, L. (1991), "Macroeconomic Performance and the Disadvantaged," *Brookings Papers on Economic Activity*, 2, 1-74

[13] Deaton, A. and C. Paxson (1994), "Intertemporal Choices and Inequality," *Journal of Political Economy*, 102, 437-68

[14] Hirano, K. Imbens, G. and Ridder, G. (2003), "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score," *Econometrica*, Vol. 71, No. 4,

[15] Krueger, D. and Perri, F. (2003), "On the Welfare Consequences of the Increase in Inequality in the United States," paper presented at the "NBER Eighteenth Annual Conference on Macroeconomics"

[16] Mace, B.J. (1991), "Full insurance in the presence of aggregate uncertainty," *Journal of Political Economy*, 102, 384-394

[17] McCarthy, M., Johnson, D., Garner, T., and Passero, B. (2002), "Issues in Construction and Research Use of the Consumer Expenditure Survey," paper presented at the 2002 NBER Summer Institute

[18] McClements, B. (1977), "Equivalence Scales for Children," *Journal of Public Economics*, 8, 191-210

[19] Sabelhaus, J. (1996), "Consumer Expenditure Survey: Family-Level Extracts, 1980:1-1994:1," Congressional Budget Office, available at www.nber.org

[20] Slesnick, D. T. (1992), "Aggregate Consumption and Saving in the United States," *Review of Economics and Statistics*, 74, 585-97

[21] Slesnick, D.T. (1993), "Gaining Ground: Poverty in the Postward United States," *Journal of Political Economy*, 101, 1-38

[22] Slesnick, D.T. (2001), *Consumption and Social Welfare. Living Standards and Their Distribution in the United States*, Cambridge: Cambridge University Press

[23] Townsend, R.M. (1994), "Risk and Insurance in Village India," *Econometrica*, 62, 539-591