

NBER WORKING PAPER SERIES

WHAT REALLY HAPPENED TO
CONSUMPTION INEQUALITY IN THE US?

Orazio Attanasio
Erich Battistin
Hidehiko Ichimura

Working Paper 10338
<http://www.nber.org/papers/w10338>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2004

First draft 20th August 2003. The paper benefited from useful discussion with Richard Blundell, David Card, Angus Deaton, David Johnson, Tom MaCurdy, Luigi Pistaferri and from comments by audiences at the conference “Hard-to-measure goods and services: essays in honor of Zvi Griliches”, Washington, DC September 2003, and seminars at Bureau of Labor Statistics, Stanford and Berkeley. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

©2004 by Orazio Attanasio, Erich Battistin, and Hidehiko Ichimura. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What Really Happened to Consumption Inequality in the US?
Orazio Attanasio, Erich Battistin, and Hidehiko Ichimura
NBER Working Paper No. 10338
March 2004
JEL No. E21, E26

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 collected 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. We find that consumption inequality, as measured by the standard deviation of log non-durable consumption, has increased by roughly 5% points during the 1990s.

Orazio Attanasio
Department of Economics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM
and NBER
o.attanasio@ucl.ac.uk

Erich Battistin
Institute for Fiscal Studies
erich_b@ifs.org.uk

Hidehiko Ichimura
University College London and Cemmap
h.ichimura@ucl.ac.uk

Contents

1	INTRODUCTION	3
2	THE CEX SURVEYS	9
2.1	Diary and Interview Samples	10
2.2	The information collected in the CEX	11
2.3	The selected sample	14
3	EVIDENCE ON CONSUMPTION AND WAGES	17
3.1	Consumption means	19
3.2	Wage inequality in the CEX and in the CPS	21
3.3	Consumption inequality	23
4	COMBINING INFORMATION FROM INTERVIEW AND DIARY SAM- PLES	27
5	RESULTS	33
6	CONCLUSIONS	36

List of Figures

1	Standard deviation of log per capita monthly expenditure	7
2	Non-durable expenditures in 2000 dollars - Consumer Expenditure Survey (CEX) and Personal Consumption Expenditures (PCE)	18
3	Mean of log monthly expenditure on non-durable goods (2001 dollars)	18
4	Mean of monthly expenditure on non-durable goods (2001 dollars)	19
5	Standard deviation of log wages in the CPS	22
6	Squared coefficient of variation for D goods	26
7	Squared coefficient of variation for R goods	26
8	Inequality growth using observed covariances	33
9	Inequality growth using combined information	34

List of Tables

1	Definitions of expenditure categories	12
2	Commodity split	14
3	Sample sizes	15
4	Percentage of zero expenditures	24
5	Regressions using observed covariances	33
6	Inequality growth over time	34

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, many 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), documenting 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 CEX until 2001 that are roughly consistent with those reported by Slesnick (2001). 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 CEX 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 Cochrane (1991), Mace (1991) and Townsend (1994). 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 are related to the nature of the individual level data currently available on consumption expenditure. The CEX is a relatively small survey, collected mainly to compute weights of the CPI, rather than studying consumption inequality. Moreover, the survey is affected by other problems. There is now substantial evidence that by aggregating CEX data it 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 partly be 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 35%) 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 Sample, IS in the following) and one based on weekly diaries (Diary Sample, DS). The 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. This data structure does not constitute an important problem if one is only interested in means, but it creates a problem if one needs information on total consumption expenditure *for a given household*, as it is the case if one wants to study consumption inequality across households. Of course, in the absence of reliable data on *total* consumption for a given household that would allow to study inequality at the individual level, one could focus on differences in mean consumption across well defined groups of households. But such an approach would miss an important dimension of inequality, that is *within* group inequality.

One of the most puzzling results that arises from the analysis of CEX consumption inequality data is that the evolution over time of consumption inequality as measured in the IS and DS is very different (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 (diary information is only after 1986). 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

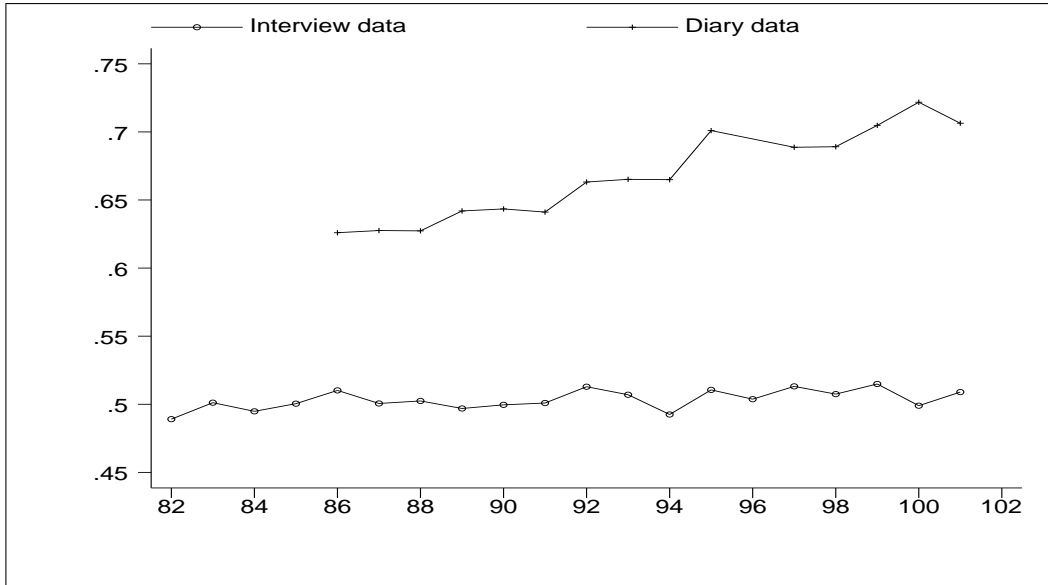


Figure 1: Standard deviation of log per capita monthly expenditure

in the variance *across education groups* (see Attanasio and Davis, 1996 and Attanasio, 2003), a constant overall inequality also constitutes indirect evidence of a decline in inequality within groups.¹ Regressing the two lines on a constant and a linear time trend, one obtains (after 1986) a trend coefficient of 0.04 for the IS and of 0.64 for the DS with a t statistic of 1.17 and 12.66, respectively. The correlation between the residuals of the two regressions is 0.26.

This evidence is particularly puzzling because the differences in mean non-durable consumption between the two surveys is relatively stable over time (as we show in Section 3).² Moreover, in many other dimensions the CEX offers a picture of inequality that is remarkably consistent with that obtained from other (extensively explored) datasets. In what follows we provide some evidence in this respect.

The aim of this paper is twofold. First, we provide additional information on the dynamics of inequality of consumption and its components in the two CEX surveys and relate it to the

¹This follows since $Var(y) = Var(E[y|z]) + E[Var(y|z)]$, where groups are denoted by z .

²Attanasio (2003), however, shows that if one conditions on cohorts defined from education and year of birth, 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.

dynamics of wage inequality, as measured both in the CEX and the Current Population Survey. Second, we use the information that some items are better measured in the DS and others in the IS and some assumptions on the nature of measurement error in the two surveys, to construct an *unified* picture of the dynamics of consumption inequality in the US between 1986 and 2001 (and particularly for the 1990s).

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. 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 CEX and its components. We describe in detail the nature of the information on consumption available in the DS and the IS, as well as the sample we select for the rest of our analysis. In Section 3 we discuss the evolution of average consumption over time in the two surveys. We also report some evidence on comparing the pattern of wage inequality in the CEX and in the Current Population Survey that, complementing that in Attanasio (2003), shows that the two surveys tell similar stories in this dimension. Finally, in this section, we also discuss the puzzle presented in Figure 1. In particular, we consider and dismiss a few simple explanations for the divergence in inequality between 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 the results we obtain using such an approach. Section 6 concludes.

2 THE CEX SURVEYS

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 US. 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.

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

In the CEX, 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 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). Starting in 1999, the sample size has increased by about 40%. 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 Universal Classification Codes (UCC). 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: 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 cate-

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
<i>Housekeeping Services (DS only)</i>
<i>Personal Care (DS only)</i>
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

gories included in our definition of non-durable consumption are listed in Table 1. We consider nine expenditure categories, corresponding to roughly 280 and 400 UCC for IS and DS data, respectively.³

Whilst the bulk of the questionnaires and survey methodology were remarkably stable over time, some minor changes did occur. New diaries with more cues were introduced in the DS after 1991; for the IS, the food question changed in 1982 and 1987. Some UCC have changed, mainly reflecting the diffusion of new goods,⁴ but our aggregates are not affected substantially

³As 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.

⁴For instance, the UCC corresponding to expenditure on personal computers was divided into two UCC relating to software and hardware.

by these changes. Battistin (2003) maps UCC into the nine categories in Table 1 accounting for these changes.

Both surveys are almost exhaustive. This implies that, for most items, we have a measure both for the households in the DS and for those in the IS. The only exception for the definition of consumption considered in our analysis is given by some small items (mainly housekeeping services and personal care - these categories are reported in italic in Table 1 in the “Non-Durable Goods and Services” category) for which information is collected in the DS, but not in the IS. We discuss how we tackle this problem below.

We exclude from our definition of consumption expenditures on durables, health, education as well as mortgage and rent payments. The main reason to exclude expenditure on durables is that it is not directly linked to consumption. One would like to measure the services provided by the existing stock of durables, rather than the increase in the stock of durables. Similarly, education and health expenditure obviously have an important investment component. 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 the consumption of housing services. Whilst all the exercises for which we report results use the expenditure on non-durable and services as our definition of consumption, we also performed some experiments with total consumption, as defined by the BLS. The results we obtained in these cases were substantially similar to those we report.

To estimate average consumption, the BLS follows the standard international procedure of exploiting information from recall questions for more durable and less frequently purchased items bought in the quarter prior to the interview. Diary-based records of purchases carried out

Table 2: Commodity split

Commodities better measured in the Diary Survey: D goods
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
Commodities better measured in the Interview Survey: R goods
Housing and Public Services
Tobacco and Smoking Accessories
Clothing, Footwear and Services
Heating Fuel, Light and Power
Transportation (including gasoline)

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. In what follow, we label the commodities that the BLS thinks are better measured in the DS as D goods and services, and those that are better measured in the IS as R goods and services. In Table 2 we list which categories belong to the D and R groups according to the BLS.⁵

2.3 The selected sample

In this analysis, we focus on households headed by individuals aged at least 25 and no more than 60 and *not* self-employed. The family head is conventionally fixed to be the male in all husband/wife families (representing the 56% and 53% of the whole sample for IS and DS data, respectively). Battistin (2003) presents a detailed description of less important selection criteria used to derive the working sample considered in the analysis.

Although the two surveys are designed to be representative of the same population, there are

⁵It 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.

Table 3: 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

significant differences between the two samples along several dimensions and with a different pattern over time (even using the population weights provided by the BLS). To control for observed compositional 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).

The specification adopted (see Table 1 in Battistin, 2003) includes education, race, age and work-related information of the head, as well as information on household composition (proportion of children and members within certain age bands) and family income. These variables have proved relevant to data quality in previous analysis of CEX data (see Tucker, 1992). We use the same procedure here. However, results obtained using BLS population weights or propensity score weights are basically identical. Table 3 reports, for each year, the size of the sample we end up with for the two surveys. As it 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. The increase in sample size in 1999 we mentioned above is evident in this table.

Monthly expenditure in the DS is defined as $26/12 = 2.16$ times the expenditure observed over two weeks, assuming equally complete reporting. Family consumption is adjusted using the OECD adult equivalence scale, which gives each adult beyond the first a weight of 0.7 and each child (under 18) a weight of 0.5. While such an adult equivalence scale is clearly arbitrary, the results we obtain are only minimally affected by the consideration of alternative scales.⁶ Real

⁶In previous versions we have reported results based on different adult equivalence scales. The choice of the scale, changes the results very little.

expenditures are obtained using the Current Price Index published by the BLS. Consumption data are also corrected for seasonality, using a simple seasonality model estimated on the whole dataset.

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 checked whether the exclusion of households with incomplete income responses makes any difference to our main results.

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 Current Population Survey. 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.

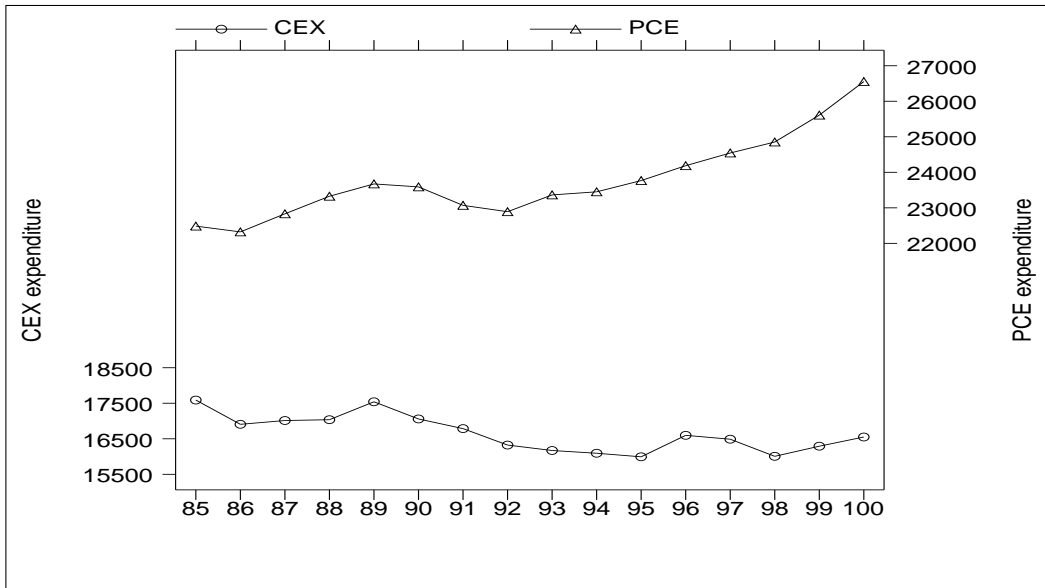


Figure 2: Non-durable expenditures in 2000 dollars - Consumer Expenditure Survey (CEX) and Personal Consumption Expenditures (PCE)

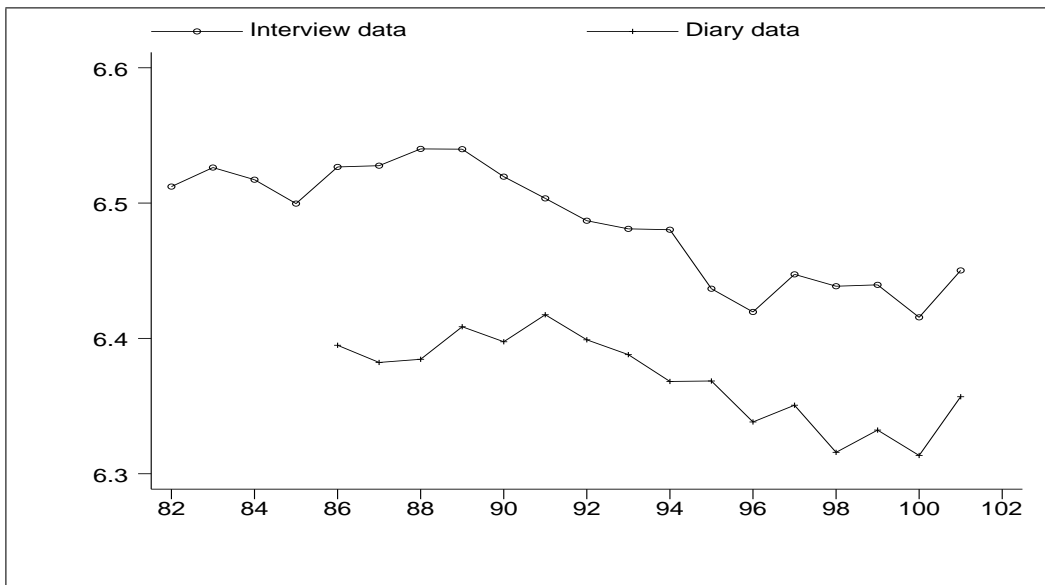


Figure 3: Mean of log monthly expenditure on non-durable goods (2001 dollars)

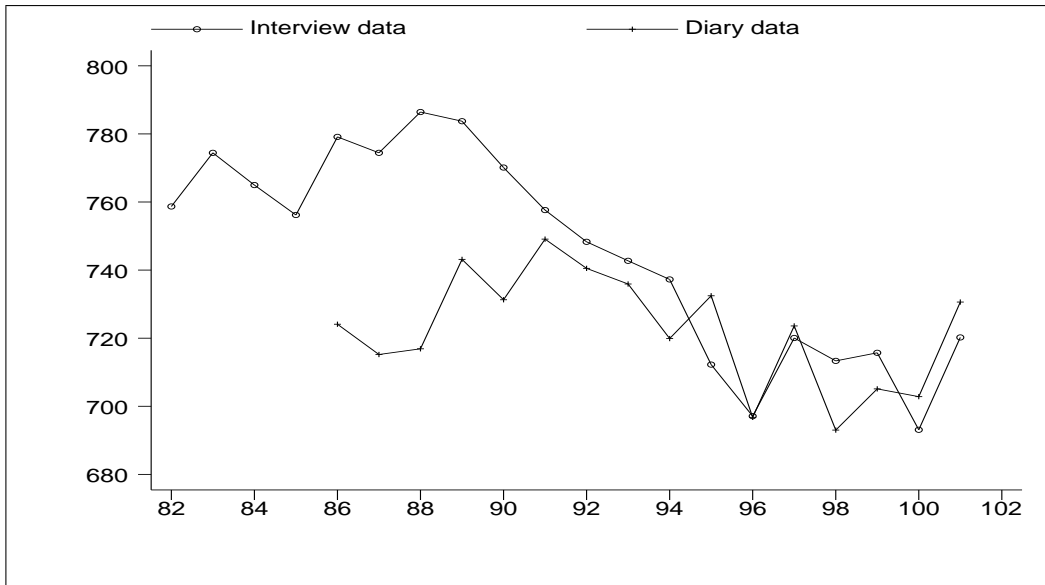


Figure 4: Mean of monthly expenditure on non-durable goods (2001 dollars)

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.⁷ The CEX aggregates are computed using the population weights provided by the BLS and using published information from both the DS and the IS. 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

⁷We are grateful to David Johnson at the BLS for making this graph available to us. Non-durables includes food at home, food away, alcohol, apparel and services, maintenance and repairs, utilities, household operations, housekeeping supplies, gasoline and motor fuel, vehicle maintenance, vehicle rental and other, public transportation, fees and admissions, other entertainment supplies, personal care, tobacco. 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.

consumption, as documented in Banks and Johnson (1998). Second, while the divergence between CEX and PCE is roughly constant in the first part 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 interesting to compare the estimates of aggregate non-durable consumption that emerge from the two surveys. In Figure 3, we plot the time series of average log non-durable consumption computed in the two datasets. While average consumption is consistently higher in the IS than in the DS, we also notice that the relative difference between the two surveys is roughly constant over time and in particular after 1991.⁸

In Figure 4, we plot the time series of the average level (rather than log) of monthly non-durable consumption from the two data sets. The two series are very different in the first part of the sample, but they converge remarkably starting in 1991. Of course Figure 1, where we plotted the standard deviation of log consumption, and Figures 3 and 4, that plot the average level and log consumption, are not independent: under the assumption of log-normality, one would be able to derive any one from the other two figures. We can summarize the evidence from these pictures by saying that, especially from 1991, the average log and level consumption estimated from the two surveys move very closely to each other, while the standard deviation diverges considerably.

⁸This 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.

3.2 Wage inequality in the CEX and in the CPS

There is a widespread perception that income figures in the CEX are not particularly reliable, especially relative to other more established and better explored data sets, such as the Current Population Survey (CPS). Before starting with the analysis of consumption inequality, it is therefore worth reporting some information on how the CEX performs in measuring the level and inequality of wages (hourly earnings) over time relative to the CPS, which has been widely used in the study of wage inequality.

From one of the supplementary CEX files it is possible to obtain measures of earnings and hours worked for each household member. We compare the figures we obtain using these measures to analogous measures obtained from rotating CPS files.⁹ Using this information, we compute the average and the standard deviation of log male hourly earnings in both datasets for the 20 years from 1982 to 2001. We perform the exercise both for the whole sample and for cohort/educational attainment groups.

Top-coding levels are different in the two datasets and have 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. This procedure assumes that the log-normal distribution fits reasonably well the distribution of wages even for the top few percents that are top-coded. This type of correction is particularly important for comparisons over time as top-coding levels change.

We start by computing the mean and standard deviation of log hourly male earnings in the IS and in the CPS for all males living in urban areas aged between 25 and 60. If we correlate

⁹The CPS is not exempted from problems. There are several changes, both in terms of definitions of various variables (such as relation to the household head) and in top-coding levels. We have used the NBER extracts and suggestions to correct for changes in definitions and variable labels over time. We discuss what we do for top-coding below. We thank David Card for providing us with the data for 2000 and 2001 and for some useful suggestions.

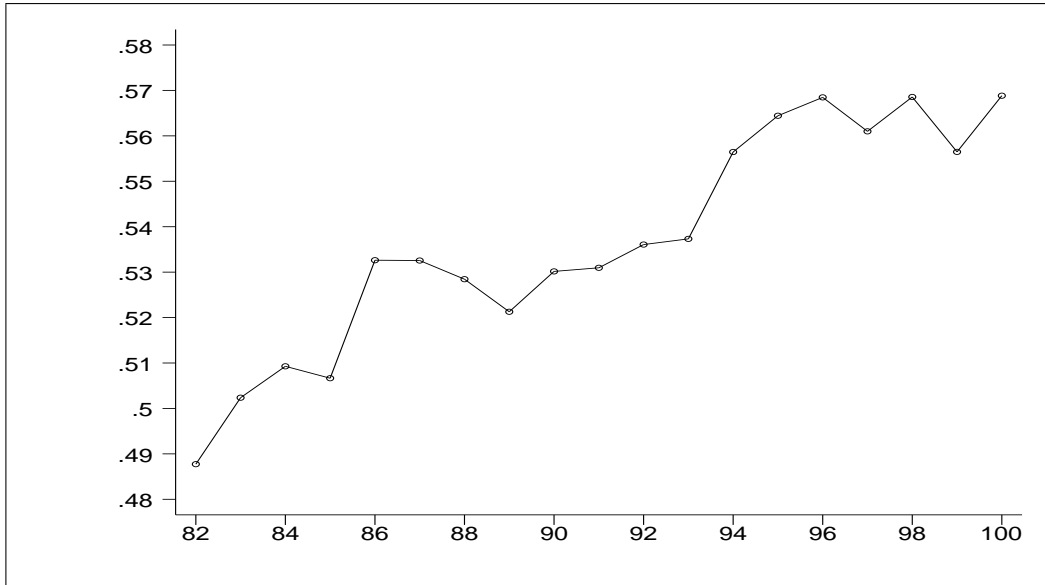


Figure 5: Standard deviation of log wages in the CPS

the averages, we obtain a correlation coefficient of 0.62. Regressing the IS average on the CPS average, we obtain a coefficient of 0.83 (s.e.= 0.26), a constant not statistically different from zero and an R-squared of 38%.

Correlating the standard deviation of log wages we obtain a coefficient of 0.5. A regression of the IS standard deviation on the CPS standard deviation yields a coefficient of 0.32 (s.e. 0.13), a constant of 0.16 (se 0.05) and an R-squared of 26%. Relating changes in mean log wages in the two datasets yields a correlation coefficient of 0.44, while changes in the standard deviation in the two datasets yield a correlation coefficient of 30%. We interpret this evidence as saying that the CEX and the CPS depict similar pictures in terms of the dynamics of inequality in wages.

To conclude this section, in Figure 5 we plot the standard deviation of log wages in the CPS. This figure shows a consistent increase throughout the sample period. Between 1990 and 2000, this measure of wage inequality increases by about 0.04. This will be a useful point of reference when assessing the size of the increase in consumption inequality.

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 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: (i) changes in questionnaires and survey methodology; (ii) changes in the frequency of purchases of commodities; (iii) changes in the willingness to answer surveys; and (iv) 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 consumed at home, changed in 1982 and 1987 (see Battistin, 2003). The first change, 1982, is outside our interval and the second change, 1988, precedes the point in which the two measures of inequality start to diverge (1991). 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

Table 4: 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	0.08	0.09	0.10	0.12	0.15
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	0.17	0.18	0.22	0.23	0.26
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

months for the IS), the DS would result in the same mean (or a stable difference over time) but increasingly larger variances over time in the DS because of the increased number of zeros. Table 4 (see Battistin, 2003) shows that the number of non-zero expenditures for non-durable items varies over time for some groups (particularly, it decreases for ‘alcohol’, ‘tobacco’ and ‘clothing’) but with the same pattern across samples. If anything, the number of zeros seems to increase more in the IS than in the DS (see for instance food and alcohol). We can therefore confidently dismiss such an explanation.

- *Over time people have become less willing to answer accurately or answer at all.* Obviously it is difficult to judge the importance of measurement error over time. However, one can check whether attrition rates and/or 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. Survey response rates did not change much over time, remaining between 80% and 90%. McCarthy *et al.* (2002) and Slesnick (2001) report evidence in this respect. The same is also true for the percentage of incomplete income responses. It is therefore unlikely that this could explain the different trend in inequality painted by the two CEX surveys.

- *Survey non-response.* It is widely believed that households who refuse to take part into surveys are over-represented by very rich and very poor. If so, we would possibly underestimate inequality because of under-representation of CUs in the upper and lower tail of the distribution of consumption. It might be that a larger number of wealthier individual are 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.
- *Changes in sample compositions.* As we have already mentioned the composition of the IS and DS is different, even after using the BLS weights. For instance, the DS is better educated than the IS. While these differences change marginally over time, we control for them by using propensity score weights and show that these differences cannot explain the different dynamics of inequality.

Having discarded some simple explanations of the puzzling patterns observed in Figure 1, we now look at the dynamics of inequality in the two subsets of goods we have defined as R goods and D goods in Table 2. This is possible because both sets of goods are observed in both surveys. In Figure 6, we plot the square of the coefficient of variation of ‘ D goods’ as measured

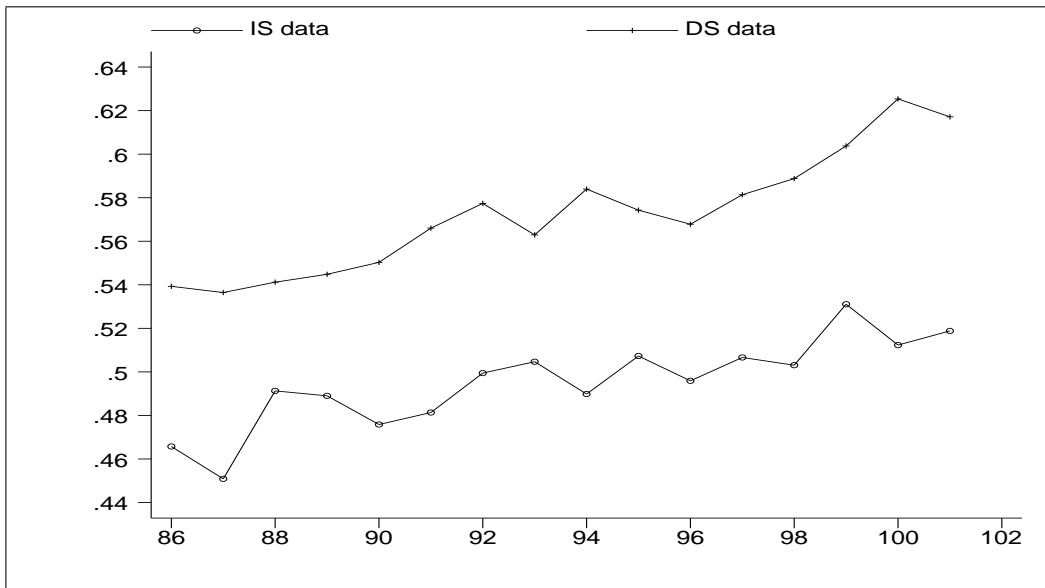


Figure 6: Squared coefficient of variation for D goods

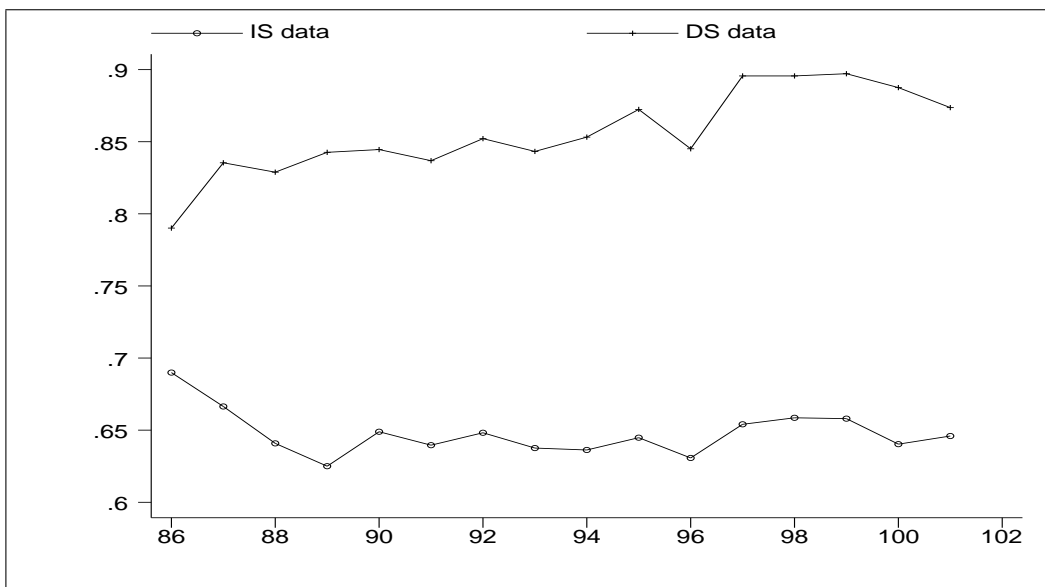


Figure 7: Squared coefficient of variation for R goods

in both surveys, while in Figure 7, we plot the square of the coefficient of variation of ‘ R goods’ as measured in both surveys.

Two features of these pictures are noteworthy. First, the coefficient of variation of ‘ D goods’ seems to be increasing in both surveys. On the other hand, the coefficient of variation of ‘ R goods’ increases slightly in the DS, while it stays constant in the IS after 1990. The drop observed in IS for 86 and 87 may be related to the change in the food question as discussed in Battistin (2003). The divergence in the path of measured coefficients of variation between the two surveys is particularly evident in the first part of the sample, until 1990. Second, for both sets of goods, the coefficient of variation is much larger in the DS than in the IS. This feature might be a consequence of the shorter horizon covered by the DS and the larger number of zeros documented in Table 4.

4 COMBINING INFORMATION FROM INTERVIEW AND DIARY SAMPLES

Rather than pursuing further the attempt to explain the difference between 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 combine information from the IS and the DS, in that some commodities are deemed to be better measured in the IS, while others in the DS.

In Table 2 we listed which categories, according to Bureau of Labor Statistics (2003), are better measure by each survey. Let C^* be total expenditure on *all* non-durable commodities

$$C^* = C_D^* + C_R^*, \tag{1}$$

where C_D^* and C_R^* represent expenditures on items that are better measured in DS and IS data, respectively. Obviously, being both surveys (almost) exhaustive, a measure of ‘*R* goods’ exists also in the DS and a measure of (most) ‘*D* goods’ exists also in the IS.¹⁰

More accurate estimates of average non-durable consumption can easily be obtained by combining information from 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.

In what follows, we will be interested in the squared coefficient of variation

$$CV(C^*)^2 = \frac{Var(C^*)}{E(C^*)^2} = \frac{Var(C_D^*) + Var(C_R^*) + 2Cov(C_D^*, C_R^*)}{[E(C_D^*) + E(C_R^*)]^2}. \quad (2)$$

The reason for this choice is twofold. First, if total consumption is log-normally distributed, the following relationship holds exactly

$$Var(\ln C^*) = \ln \left[\frac{Var(C^*)}{E(C^*)^2} + 1 \right],$$

so that the quantity in (2) is informative on the variance of log consumption. Blundell and Lewbel (1999) provide strong empirical evidence to support the fact that, in a variety of datasets the cross sectional distribution of consumption seems to be very well approximated by a log-normal.¹¹ Second, regardless of the distribution of consumption, the squared coefficient of variation provides a first order approximation to the variance of log-consumption and therefore is of some interest as an index of inequality.

¹⁰The reason for the qualifiers in parentheses is the existence of a small subset of commodities (mainly personal care items) on which there is information in the DS and no information in the IS (as pointed out in Table 1). We discuss the implications of the presence of these goods after we describe our approach.

¹¹It is worth noting that the measure of total consumption considered by Blundell and Lewbel (1999) is derived from the Family Expenditure Survey for the UK, where information on ‘*R*’ and ‘*D*’ goods is collected on the *same* households using the most appropriate methodology as defined in Table 2.

Each household is either observed in the DS or in the IS. However, for each households, in both surveys, we observe expenditure on both D and R commodities. In what follows, we will denote by C^d total non-durable expenditure as measured in the DS and by C^r total non-durable expenditure as measured in the IS.¹² Observed consumption in the two surveys is then given by

$$C^d = C_D^d + C_R^d, \quad (3)$$

$$C^r = C_D^r + C_R^r.$$

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

Assumption 1

$$C_D^d = C_D^*,$$

$$C_R^r = C_R^*.$$

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

$$C^d - C^* = C_R^d - C_R^* = \nu_R^d, \quad (4)$$

$$C^r - C^* = C_D^r - C_D^* = \nu_D^r. \quad (5)$$

If Assumption 1 is satisfied, the mean of IS and DS errors is identified by $E(\nu_D^r) = E(C_D^r) - E(C_D^*)$ and $E(\nu_R^d) = E(C_R^d) - E(C_R^*)$, respectively. By analogy, the mean of total expenditure can be estimated by $E(C^*) = E(C_R^*) + E(C_D^*)$. Figures for IS and DS errors as proportion of

¹²As mentioned above, for the time being we are ignoring the fact that some ‘ D goods’ are not observed in the IS. We tackle this problem below.

total non-durable expenditure are reported in Battistin (2003), where implications on estimated saving rates from the IS are also discussed.¹³

Note that we do not require measurement error to be of the classical type. Indeed, classical measurement error, and in particular the absence of correlation between the measurement error and the true level of the relevant variable and Assumption 1 would imply a larger measured variance of ‘ R goods’ in the DS and of ‘ D goods’ in the IS. This implication is obviously contradicted by Figures 6 and 7.

As we can easily estimate the denominator of (2) 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 (2), it is clear that to estimate the variance of C^* we lack an estimate of the $Cov(C_D^*, C_R^*)$. Equations (4) and (5) together with Assumption 1 imply that observed covariances in the two surveys are informative about such a covariance, since

$$\begin{aligned} Cov(C_D^d, C_R^d) &= Cov(C_D^*, C_R^*) + Cov(C_D^*, \nu_R^d), \\ Cov(C_D^r, C_R^r) &= Cov(C_D^*, C_R^*) + Cov(C_R^*, \nu_D^r). \end{aligned}$$

Clearly, if we assumed that either $Cov(C_D^*, \nu_R^d) = 0$ or $Cov(C_R^*, \nu_D^r) = 0$, 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 $Cov(C_D^*, \nu_R^d) = 0$, we could test whether $Cov(C_R^*, \nu_D^r) = 0$ and viceversa. Clearly each of these alternative two tests are equivalent to testing whether

$$Cov(C_D^*, \nu_R^d) = Cov(C_R^*, \nu_D^r),$$

since this is equivalent to

$$Cov(C_D^r, C_R^r) - Cov(C_D^d, C_R^d) = 0. \tag{6}$$

¹³While 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.

Equation (6) is a *necessary but not sufficient* condition for our identification assumption. If $Cov(C_D^*, \nu_R^d) = 0$, we can identify $Cov(C_D^*, C_R^*)$ from the DS, while if $Cov(C_R^*, \nu_D^r) = 0$ we can identify from the IS. Unfortunately, the restriction in equation (6) is rejected in our data (see Battistin, 2003).

However, if we are only interested in the changes of the variance (or of the coefficient of variation) of consumption, we can use a weaker and less unappealing identification assumption.

Since

$$\Delta Var(\ln C^*) \simeq \Delta \frac{Var(C_R^*)}{E(C^*)^2} + \Delta \frac{Var(C_D^*)}{E(C^*)^2} + 2\Delta \frac{Cov(C_D^*, C_R^*)}{E(C^*)^2}, \quad (7)$$

it is easy to show that the last term on the right hand side of this expression can be identified if we assume either that $\Delta \frac{Cov(C_D^*, \nu_R^d)}{E(C^*)^2} = 0$ or $\Delta \frac{Cov(C_R^*, \nu_D^r)}{E(C^*)^2} = 0$. Once again, assuming the former we can test the latter, and viceversa. Once again testing either of these assumptions is equivalent to testing the hypothesis that

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

which in turn is equivalent to the following hypothesis in terms of observables

$$\Delta \frac{Cov(C_D^d, C_R^d)}{E(C^*)^2} - \Delta \frac{Cov(C_D^r, C_R^r)}{E(C^*)^2} = 0. \quad (8)$$

Such an hypothesis is a *necessary but not sufficient* condition for the point identification of $\Delta Var(\ln C^*)$. A sufficient condition for the point identification of $\Delta Var(\ln C^*)$ is the following

Assumption 2

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

A non-rejection of the relationship in (8) does *not* guarantee that Assumption 2 is correct: it is possible that the covariances of measurement error with the true value change in both

surveys *in the same way*, while we need that such covariances are constant over time. It should be stressed however, that the assumption of a homoscedastic measurement error is implicitly made by all studies that analyze changes in inequality over time.

As we mentioned above, for some commodities, mainly personal care items, the only available information is that in the DS, as the IS does not ask the relevant retrospective questions. This slightly complicates the approach sketched above. Equation (1) now becomes

$$C^* = C_{Dr}^* + C_{Dp}^* + C_R^*, \quad (9)$$

where C_{Dr}^* are ‘ D goods’ also available in the IS and C_{Dp}^* are ‘ D goods’ available only in the DS. Equation (3) now becomes

$$C^d = C_{Dr}^d + C_{Dp}^d + C_R^d.$$

Equation (9) implies that one cannot obtain a complete estimate of average total non durable consumption from the IS. Moreover, when one wants to compute the variance, one gets

$$Var(C^*) = Var(C_{Dr}^* + C_{Dp}^*) + Var(C_R^*) + 2[Cov(C_{Dr}^*, C_R^*) + Cov(C_{Dp}^*, C_R^*)]. \quad (10)$$

The first of the two covariances in the last expression can only be estimated, under the assumption that $Cov(C_{Dp}^* + C_{Dr}^*, \nu_R^d) = 0$. As before, if we are only interested in the changes in this covariances, we can use the less restrictive assumption that $\Delta Cov(C_{Dp}^* + C_{Dr}^*, \nu_R^d) = 0$. However, we will not be able to test the hypothesis that $\Delta Cov(C_{Dp}^* + C_{Dr}^*, \nu_R^d) = \Delta Cov(C_R^*, \nu_D^r)$.

In other words, if all the commodities in the DS were also observable in the IS, we could use either the covariance between D and R goods estimated in the DS or that estimated in the IS to compute the (changes in the) variance of total consumption. Now, at least for $Cov(C_{Dp}^*, C_R^*)$ we are forced to use the DS. For $Cov(C_{Dr}^*, C_R^*)$, however, we can use, as before both surveys.

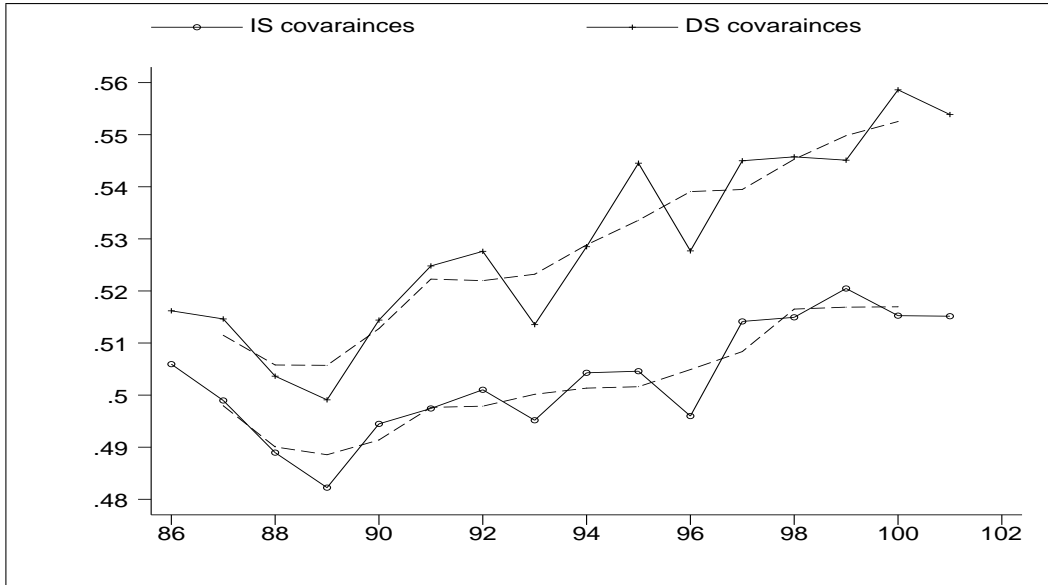


Figure 8: Inequality growth using observed covariances

Table 5: Regressions using observed covariances

	whole sample		1990s	
	Coef.	Std. Err.	Coef.	Std. Err.
time	0.1715	0.0291	0.0870	0.0666
cons	0.0112	0.0028	0.0191	0.0066
R^2	71.15%		17.57%	
obs	16		10	

5 RESULTS

In Figure 8, we report two estimates of the evolution of consumption inequality. Both are derived using equation (10). However, the first is obtained using $Cov(C_{Dr}^r, C_R^r)$ for $Cov(C_{Dr}^*, C_R^*)$, while the second uses $Cov(C_{Dr}^d, C_R^d)$. That is, the first measures the variance using

$$Var(C^*) = Var(C_{Dr}^d + C_{Dp}^d) + Var(C_R^r) + 2[Cov(C_{Dr}^d, C_R^d) + Cov(C_{Dp}^d, C_R^d)],$$

while the second uses

$$Var(C^*) = Var(C_{Dr}^r + C_{Dp}^r) + Var(C_R^r) + 2[Cov(C_{Dr}^r, C_R^r) + Cov(C_{Dp}^r, C_R^r)].$$

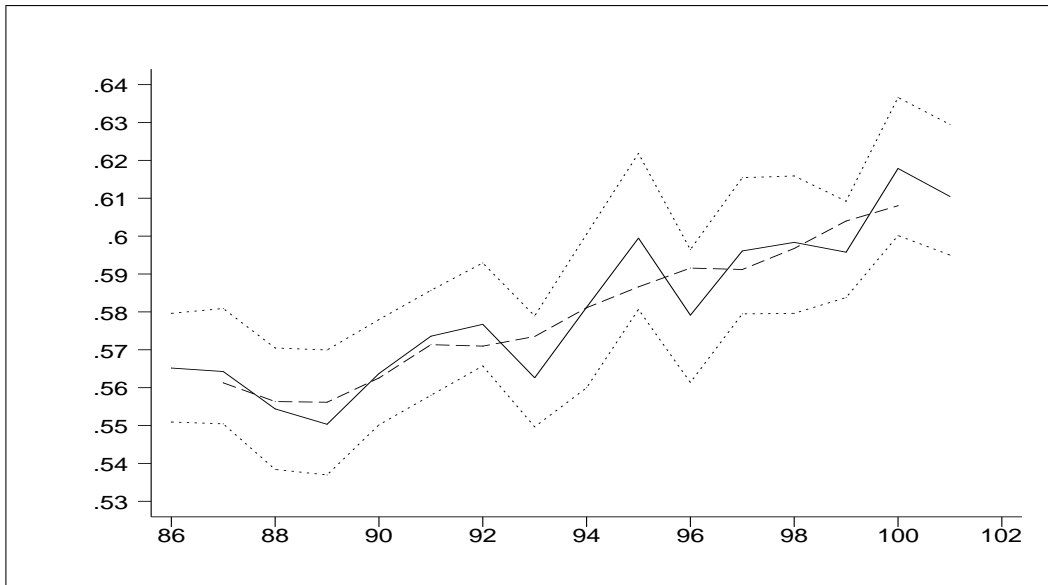


Figure 9: Inequality growth using combined information

time	Interview	Diary	Estimated
$\Delta 1995 - 90$	1.10	5.75	3.58
$\Delta 2000 - 95$	-1.16	2.09	1.84

In addition to the raw estimate, we also plot a smoothed version of each measure obtained by a third order moving average.

The necessary condition in equation (8) is equivalent to the hypotheses that the distance between the two lines in equation (7) is constant. In Table 5 we test such a hypothesis by regressing the difference between the two lines on a constant and a linear trend and test for the coefficient of the linear trend to be zero. Unfortunately, we can reject the null if we use the whole sample (see the left hand side panel). However, if we restrict the sample to the 1990s (see the right hand side panel), we do not reject the null. The changes in the DS questionnaire in the early 1990s, might constitute a justification for restricting the sample to the 1990s.

A possible interpretation of Figure 8 is that regardless of whether we use the DS or the IS based measure of $Cov(C_{Dr}^*, C_R^*)$, the evidence indicates a slight decline of inequality in the

late 1980s followed by a sustained increase during the 1990s. The magnitude of the increase, however, is debatable. If we use estimates of $Cov(C_{Dr}^*, C_R^*)$ based on the DS, we get a larger increase than if we use estimates from the IS.¹⁴

If we decide that Assumption 2 is a valid one, the two lines in Figure 7 give us two measures of the changes in consumption inequality. An issue is therefore how to combine this information to obtain a single efficient measure. Figure 9 presents our estimates of the evolution of consumption inequality (from an arbitrary starting point) obtained from the entire CEX sample, as it results from combining *efficiently* the two lines in Figure 8. In the Figure we also plot confidence intervals (dotted lines).

Denoting with $\hat{\theta}_R$ and $\hat{\theta}_D$ the estimate of (2) obtained using IS and DS covariances respectively, we obtain the estimates in Figure 9 as

$$\alpha\hat{\theta}_R + (1 - \alpha)\hat{\theta}_D,$$

where the value of α is chosen so to minimize the variance of the estimator. The first order condition gives

$$\alpha^* = \frac{Cov(\hat{\theta}_D, \hat{\theta}_D - \hat{\theta}_R)}{Var(\hat{\theta}_D - \hat{\theta}_R)},$$

which corresponds to the coefficient of an OLS regression of $\hat{\theta}_D$ on $\hat{\theta}_D - \hat{\theta}_R$. We estimate the optimal value of α for each time period t between 1986 and 2001 via re-sampling methods using 1000 pseudo-samples from the original dataset.

To improve the readability of figure 9, Table 6 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

¹⁴By using stronger assumptions, one can also relax Assumption 1 of no measurement error in some components of the survey. In the Appendix we spell out which assumptions are necessary to achieve identification in this case. The results we present below would be unaffected by this slightly different approach.

the level so to place the initial level of inequality between the initial levels of inequality in the DS and IS in 1986.

Perhaps not surprisingly, the estimate we obtain for the changes in inequality are in between the 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 rises by about 5.4 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 and points out to an order of magnitude similar to the increase in the inequality in male wages mentioned in Section 3.

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 patterns. We 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 a *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 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 5.4

percentage points, is economically significant.

Our results are reasonably robust. Moreover, of the two assumptions we have used we can conceivably relax the one that states there is no measurement error in some commodities in the Diary Survey and in other commodities in the Interview Survey. By assuming that stationarity of the measurement error processes we would obtain essentially the same results we have presented in this paper.

Assumption 2 is *crucial* to identify changes over time in consumption inequality but the assumption does not help identification of its level. Without that assumption, one can try to put bounds on 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 Assumption 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] Campbell, J.Y. (1993), “Intertemporal Asset Pricing without Consumption Data”, *American Economic Review*, Vol. 83, No. 3, 487-512
- [12] Cochrane, J.H. (1991), “A simple test for consumption insurance,” *Journal of Political Economy*, 99, 957-976
- [13] Cutler, D. and Katz, L. (1991), “Macroeconomic Performance and the Disadvantaged,” *Brookings Papers on Economic Activity*, 2, 1-74
- [14] Deaton, A. and C. Paxson (1994), “Intertemporal Choices and Inequality,” *Journal of Political Economy*, 102, 437-68
- [15] Hirano, K. Imbens, G. and Ridder, G. (2003), “Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score,” *Econometrica*, Vol. 71, No. 4,
- [16] 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”
- [17] Mace, B.J. (1991), “Full insurance in the presence of aggregate uncertainty,” *Journal of Political Economy*, 102, 384-394

- [18] 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
- [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
- [24] Tucker, C. (1992), “The Estimation of Instrument Effects on Data Quality in the Consumer Expenditure Survey,” *Journal of Official Statistics*, 8, 41-61

Appendix

In what follows, we discuss a set of assumptions on the measurement error sufficient to identify

$$\Delta \frac{Var(C^*)}{E(C^*)^2} \tag{11}$$

from observed DS and IS data. Let

$$C_D^d = C_D^* + \nu_D^d,$$

$$C_R^r = C_R^* + \nu_R^r.$$

be the DS and IS measurements of expenditures on D and R items, respectively. Throughout the paper, we maintained the assumption that $\nu_D^d = 0$ and $\nu_R^r = 0$, that is to say DS and IS figures are not affected by measurement error for those items each survey is explicitly targeted to. In what follows, we will weaken this assumption and provide a set of regularity conditions on ν_D^d and ν_R^r such that identification of (11) is secured.

Solution 1

Under the following

Assumption 3

$$E(\nu_D^d) = 0,$$

$$E(\nu_R^r) = 0,$$

the denominator of the ratio in (11) is identified using the fact that

$$E(C^*) = E(C_D^d) + E(C_R^r).$$

Accordingly, the following equalities are also satisfied

$$\begin{aligned}
Var(C_D^d) &= Var(C_D^*) + E(\nu_D^d)^2 + 2E(C_D^* \nu_D^d), \\
Var(C_R^r) &= Var(C_R^*) + E(\nu_R^r)^2 + 2E(C_R^* \nu_R^r), \\
Cov(C_D^d, C_R^d) &= Cov(C_D^*, C_R^*) + Cov(C_D^*, \nu_R^d) + E(C_R^* \nu_D^d) + E(\nu_D^d \nu_R^d), \\
Cov(C_D^r, C_R^r) &= Cov(C_D^*, C_R^*) + E(C_D^* \nu_R^r) + Cov(C_R^*, \nu_D^r) + E(\nu_D^r \nu_R^r).
\end{aligned}$$

If the following

Assumption 4

$$\begin{aligned}
\nu_D^d &\perp (C_D^*, C_R^*, \nu_R^d), \\
\nu_R^r &\perp (C_D^*, C_R^*, \nu_D^r),
\end{aligned}$$

is also satisfied,¹⁵ it follows that

$$\begin{aligned}
Var(C_D^d) &= Var(C_D^*) + E(\nu_D^d)^2, \\
Var(C_R^r) &= Var(C_R^*) + E(\nu_R^r)^2, \\
Cov(C_D^d, C_R^d) &= Cov(C_D^*, C_R^*) + Cov(C_D^*, \nu_R^d), \\
Cov(C_D^r, C_R^r) &= Cov(C_D^*, C_R^*) + Cov(C_R^*, \nu_D^r).
\end{aligned}$$

Using DS covariances and IS covariances we get

$$\begin{aligned}
\Delta \widehat{Var}(\ln C^*)_{DS} &= \Delta \frac{Var(C_D^d) + Var(C_R^r) + 2Cov(C_D^d, C_R^d)}{[E(C_D^d) + E(C_R^r)]^2} = \Delta \frac{Var(C^*)}{E(C^*)^2} + \Psi_{DS}, \\
\Delta \widehat{Var}(\ln C^*)_{IS} &= \Delta \frac{Var(C_D^d) + Var(C_R^r) + 2Cov(C_D^r, C_R^r)}{[E(C_D^d) + E(C_R^r)]^2} = \Delta \frac{Var(C^*)}{E(C^*)^2} + \Psi_{IS},
\end{aligned}$$

¹⁵Assumption 3 and Assumption 4 together characterize ν_D^d and ν_R^r as classical measurement errors, which may be a problem since we work with expenditure levels.

respectively, where

$$\Psi_{DS} = \Delta \frac{E(\nu_D^d)^2 + E(\nu_R^r)^2 + 2Cov(C_D^*, \nu_R^d)}{E(C^*)^2},$$

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

Identification of (11) follows if¹⁶

Assumption 5

$$\Psi_{DS} = \Psi_{IS} = 0.$$

Solution 2

In what follows, we build on the approximation suggested by Campbell (1993). First note that

$$\ln C^* = \ln(C_D^* + C_R^*) = \ln C_D^* + \ln[1 + e^{\ln C_R^* - \ln C_D^*}],$$

then consider a first order Taylor expansion around x_0 of the last term on the left-hand side of the previous equation

$$\ln(1 + e^x) = \ln(1 + e^{x_0}) + (x - x_0) \frac{e^{x_0}}{1 + e^{x_0}} + o(|x - x_0|).$$

It follows that

$$\ln C^* \simeq [a(x_0) - x_0 b(x_0)] + b(x_0) \ln C_R^* + [1 - b(x_0)] \ln C_D^*,$$

$$a(x_0) = \ln(1 + e^{x_0}),$$

$$b(x_0) = \frac{e^{x_0}}{1 + e^{x_0}},$$

¹⁶Notice that the stationarity of ν_D^d and ν_R^r is *not* sufficient to get

$$\Psi_{DS} = 2\Delta \frac{Cov(C_D^*, \nu_R^d)}{E(C^*)^2},$$

$$\Psi_{IS} = 2\Delta \frac{Cov(C_R^*, \nu_D^r)}{E(C^*)^2}.$$

In other words, even if ν_D^d and ν_R^r were stationary over time, Assumption 2 would not be enough to identify (11).

where $A \simeq B$ indicates $A = B + o(|x - x_0|)$. Accordingly, it also follows that

$$\begin{aligned} \text{Var}(\ln C^*) &\simeq b(x_0)^2 \text{Var}(\ln C_R^*) + [1 - b(x_0)]^2 \text{Var}(\ln C_D^*) + \\ &+ 2b(x_0)[1 - b(x_0)] \text{Cov}(\ln C_R^*, \ln C_D^*). \end{aligned}$$

Now, take

$$x_0 \equiv E(\ln C_R^*) - E(\ln C_D^*)$$

and define ν_D^d and ν_R^r such that

$$\ln C_D^d = \ln C_D^* + \nu_D^d,$$

$$\ln C_R^r = \ln C_R^* + \nu_R^r.$$

If ν_D^d and ν_R^r are classical measurement errors (that is, if Assumption 3 and Assumption 4 are jointly satisfied)¹⁷, then $b(x_0)$ is identified from DS and IS information since

$$x_0 = E(\ln C_R^r) - E(\ln C_D^d).$$

Moreover, the following relationships hold true

$$\text{Var}(\ln C_D^d) = \text{Var}(\ln C_D^*) + E(\nu_D^d)^2,$$

$$\text{Var}(\ln C_R^r) = \text{Var}(\ln C_R^*) + E(\nu_R^r)^2,$$

$$\text{Cov}(\ln C_D^d, \ln C_R^d) = \text{Cov}(\ln C_D^*, \ln C_R^*) + \text{Cov}(\ln C_D^*, \nu_R^d),$$

$$\text{Cov}(\ln C_D^r, \ln C_R^r) = \text{Cov}(\ln C_D^*, \ln C_R^*) + \text{Cov}(\ln C_R^*, \nu_D^r).$$

Using DS covariances and IS covariances we get

$$\Delta \widehat{\text{Var}}(\ln C^*)_{DS} \simeq \Delta \text{Var}(\ln C^*) + \Upsilon_{DS},$$

$$\Delta \widehat{\text{Var}}(\ln C^*)_{IS} \simeq \Delta \text{Var}(\ln C^*) + \Upsilon_{IS},$$

¹⁷By abuse of notation, we continue to write ν_D^d and ν_R^r as in the previous section, although now they refer to errors in the reporting of *log*-expenditure.

respectively, where

$$\Upsilon_{DS} = \Delta b(x_0)^2 E(\nu_R^r{}^2) + \Delta[1 - b(x_0)]^2 E(\nu_D^d{}^2) + 2\Delta b(x_0)[1 - b(x_0)]Cov(\ln C_D^*, \nu_R^d),$$

$$\Upsilon_{IS} = \Delta b(x_0)^2 E(\nu_R^r{}^2) + \Delta[1 - b(x_0)]^2 E(\nu_D^d{}^2) + 2\Delta b(x_0)[1 - b(x_0)]Cov(\ln C_R^*, \nu_D^r).$$

Identification of (11) follows if

Assumption 6

$$\Upsilon_{DS} = \Upsilon_{IS} = 0.$$