

CEX 7/23/02
BATTISTIN
9:00AM

Errors in Survey Reports of Consumption Expenditures

Erich Battistin*
Institute for Fiscal Studies, London

10th July 2002

Abstract

This paper considers data quality issues to analyze the pattern of consumption inequality in the 1990s exploiting two complementary datasets from the US Consumer Expenditure Survey. 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). Each survey has its own questionnaire and sample. We exploit information from one sample as an instrument for the other to derive a correction for the measurement error affecting observed measures of consumption. We use the implications of our findings as a test for the permanent income hypothesis.

1 Introduction

The aim of this paper is to shed more light on the comparison between recall-based and diary-based data on household consumption using US micro-level data.

There is some evidence that measurement errors in recall consumption data lead to potentially misleading results in analyzing household saving behavior (Battistin, Miniaci and Weber, 2001). Recent studies show how available data can be unsuitable for the analysis of the permanent income (life-cycle) hypotheses and how adjustments provide greater consistency concerning the time series properties of consumption (Wilcox, 1992, Slesnick, 1998, and Rosati, 2001).

Validation data, that is data on the variables of interest collected from an independent assessment of validity study (such as payroll records), are useful - whether available - to infer on the error structure of observed variables (see for example Rodgers, Brown and Duncan, 1993, and Pischke, 1995).

*Preliminary and incomplete. This paper benefited from useful discussions with O. Attanasio, J. Banks, R. Blundell, H. Ichimura, H. Low, E. Rettore, U. Trivellato, G. Weber and from comments by audiences at Padova University, Cemmap (London), ESPE 2002 and at the 10th International Conference on Panel Data, Berlin 2002. Address for correspondence: Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE - UK. E-mail: erich_b@ifs.org.uk.

In what follows we jointly exploit diary and recall data from two independent samples from the US Consumer Expenditure Survey to define a new measure of non-durable consumption.

Recall data are deemed to be reliable for bulky items (major consumer durables: real property, automobiles and major appliances) or for those components either having regular periodic billing or involving major outlays (transports, fuel and rent). Recall data on frequently purchased goods are more likely subject to non-negligible measurement error. For this reason diary surveys are designed to obtain detailed recordings of expenditures on small, frequently purchased items, which are normally difficult to recall.

It turns out that neither diary nor recall-based data provide a reliable aggregate measure of total non-durable consumption. On the basis of evidence reported in a number of previous studies, we split the set of commodities entering non-durable consumption into two groups indicating which one of the two survey methodologies (diary or recall) leads to more accurate data quality. This identification strategy allows us to define an improved measure of total non-durable consumption considering for each commodity the source presenting higher quality data.

We account for the measurement error on each commodity entering non-durable expenditure exploiting information from one sample as an instrument for the other; we derive a correction for the bias affecting observed measures of consumption and we derive bounds for the within cohort inequality of consumption. We use the implications of our findings as a test for the permanent income hypothesis.

The remaining of this paper is organized as follows. Section 2 describes the two data sources and some data collection issues. Section 3 reviews very briefly the economic set-up we will use as a motivation for our exercise. Section 4 compares the two datasets with respect to household characteristics already found to be relevant for data quality in previous studies of expenditure surveys. Section 5 presents a puzzle implied by the comparison of means and inequality indicators of total non-durable expenditure exploiting the information in the two surveys. Section 6 analyzes such discrepancies looking at the contribution of different non-durable goods. Our restrictions to define an improved measure of expenditure combining recall and diary information on non-durable commodities are discussed in Section 7 and Section 8. Section 9 and Section 10 present the econometric background to derive error adjusted bounds on consumption inequality over the life cycle. Results are presented in Section 11 and Section 12 concludes.

2 Data

The Consumer Expenditure Survey (CEX in the following) is a national survey with two separate components: the Diary, completed by respondents for two consecutive weeks, and the Interview, with four quarterly interviews. Each of the two components has its own questionnaire and sample. It is currently the only micro-level data set reporting comprehensive measures of consumption expenditures for a large cross-section of households in the US.

Each component of the survey addresses an *independent* sample of households. The Interview sample is selected on a rotating panel basis targeted at

5000 units each quarter; each consumer unit is interviewed about own monthly expenditures every three months over four consecutive quarters. It turns out that for each household we observe the monthly time series of purchases on different goods over one year (12 observations overall). After the last interview, the sample unit is dropped from the survey and replaced by a new consumer unit.

Diary data are referred to repeated cross sections of households (around 4500 per year) receiving two weekly diaries during a separate visit by a census interviewer over the two-week period interview. Both for the Interview and the Diary sample a number of questions are asked concerning household characteristics (demographics, work-related variables, education and race) and very detailed income information.

The two surveys are designed to collect different types of goods and services. While some items are collected exclusively in only one survey, there is a set of commodities (basically non-durable goods) for which expenditures are captured by both surveys. However, the comparison is not always straightforward. For example, expenditures incurred by members while away from home overnight or longer (for trips or vacation) are not collected in the Diary survey. Moreover, changes in survey instruments characterize the data over the period covered by this analysis. The Interview survey does not collect expenses for housekeeping supplies, personal care products, and nonprescription drugs, which are meant to contribute about 5 to 15 percent of total expenditures.

In the Interview survey households are retrospectively asked for their *usual* expenditure via two major questions. The first type of question asks for the weekly/monthly purchase directly for each reported expenditure; the exact wording is '*What has been your usual weekly/monthly expense for ... in the last quarter?*'. Amongst non-durable goods households are asked to report their usual weekly expenditure only for tobacco products and for food and non-alcoholic beverages consumed at home. The expenditure on the latter category is obtained as the difference between the usual weekly total expenditure at grocery stores or supermarkets and how much of this amount was for non-food items (specified as '*paper products, detergents, home cleaning supplies, pet foods, and alcoholic beverages*'). Expenditures on alcoholic beverages and food away from home (but not food consumed on vacation) are referred to the 'usual' monthly amount.

The second type of question asks for expenses in the last quarter by a detailed collection of expenditures on a list of separate goods (referred to clothing, food consumed on vacation and entertainments). Recall data are collected by a trained interviewer asking questions and providing examples of items in each category.

In the Diary survey respondents are asked to record their purchases made each day for two consecutive one-week periods. Diary respondents are assisted by cues printed on the diary and - whether needed - by interviewers at pick-up; the daily expense record is designed as a self-reporting, product-oriented diary on which respondents keep track of a detailed description of all expenses for two consecutive weeks.

3 The economic background

The permanent income hypothesis implies that, for any cohort of people born at the same time, inequality in both consumption and income should grow with age (see, for example, Deaton and Paxson, 1994). In what follows, we will investigate such implications using cohort data from both the surveys. The aim of this section is to review the economic theory that motivates our exercise.

The simplest formulation of the conventional model of consumption under uncertainty (Deaton, 1992) assumes that, in each period E , individuals maximize expectation of a time-separable utility function

$$r(o_t) + \sum_{s=t+1}^T \delta_s r(o_s)$$

subject to expectation of an intertemporal budget constraint

$$(o_t - i_t) + \sum_{s=t+1}^T n_s(o_s - i_s) = \omega_t S$$

Throughout this section r will denote an utility function invariant over the life cycle, δ_t the individual's rate of subjective preference, n_t the real interest rate and ω_t the accumulated wealth at time E . The length of the life cycle is u , while o_t and i_t represents log real consumption and income in each period.

The first order conditions to solve this problem imply that marginal utility of consumption obeys the following equation

$$r'(o_t) = \beta_t r'(o_{t-1}) + \nu_t$$

where ν_t is a shock to consumption resulting from new information at time E and $\beta_t = \delta \nu_t$ (see Hall, 1978). Otherwise stated, the last expression implies that only the actual level of consumption is informative to predict future values of consumption. In particular, future consumption is independent of actual income and wealth related variables given actual consumption.

The relationship between the evolution of consumption and the evolution of marginal utility depends on the function r . If $r'(o_t)$ is approximately linear in o_t (that is if each subperiod's utility function is quadratic up to discounting by the rate of time preference δ_t), the intertemporal choice of consumption over the life cycle is given by

$$o_t = \beta_t o_{t-1} + \nu_t S$$

The last expression has some testable implications once we are willing to make assumptions on the evolution of the β_t terms over time. To see that, consider the case where $\beta_t = 1$, as in the original Hall (1978) model. First, under such condition individual's variance must increase over time, since

$$E y n(o_t) = E y n(o_{t-1}) + E y n(\nu_t) S$$

Secondly, since

$$\frac{o_t}{E} = \frac{1}{E} \sum_{s=0}^t \nu_s S \quad (1)$$

individual's consumption is a sample average of independent shocks and hence asymptotically normal by means of the central limit theorem. Note, however, that this result applies when E grows to infinity; since individuals have finite lifespans, the distribution of consumption is expected to be normally distributed only amongst older people (see Blundell and Lewbel, 1999).

When β_t varies over time, the implications of the model might be different. The distribution of individual's consumption at time E is more dispersed than the distribution of consumption at time $E-1$ when β_t is greater than one. If the rate of interest is greater than the rate of time preference (i.e. if incentives to postpone consumption dominate impatience), the distribution of individual's consumption can either concentrate or disperse depending on the variance of v_t .

On the other hand, consumption in older age can be normally distributed only if β_t is always centered around one, to avoid having the sum of shocks in (1) be zero or infinity as E grows to infinity.

4 Descriptive analysis

Clearly household characteristics (occupation and economic activity of the head, household composition, region of residence) affect the share of spending and the quality of reporting own expenditures. Table 1 shows t-statistics from a logistic regression of the binary indicator Interview/Diary household over a set of variables including work-related information and characteristics found to be relevant for data quality in previous analysis of CEX data (Tucker, 1992). Weighted results are presented, using population weights from each survey. The specification adopted includes polynomial terms in the age of the reference person and in the proportion of children and members within certain age bands (these terms are not reported because not statistically significant).

Data exploited in this analysis cover ten years between 1988 and 1998. We excluded from our sample households residing in student housing, households whose head is unemployed and households with single women. The family head is conventionally fixed to be the male in all the H/W families (representing the 56% and 53% of the whole sample for Interview and Diary data, respectively). Furthermore, we considered only households headed by individuals aged at least 23 and no more than 73 and not self-employed. These restrictions leave us with a sample of 145178 and 36095 units, for Interview and Diary data respectively.

Although the two surveys are designed to be representative of the same population, significant differences are found along several dimensions and with a different pattern over time. For example, the amount of weeks worked per year by the reference person is higher in the Diary sample.

5 Evidence on consumption behavior

A standard way to analyze the dynamic properties of consumption with repeated cross-sections is to rely on cohort analysis. In what follows we will group households into cohorts on the basis of the year of birth of the reference person (defining six 10-year bands) and we will produce some descriptive graphs

Table 1: T-statistics from propensity score estimates; dependent variable 1=Interview, 0=Diary

Variable	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Total number of components	-0.88	-0.28	-0.25	-0.67	0.90	0.18	-2.22	-0.63	-0.96	1.85	-1.67
Number of components 18-	-1.46	0.80	-0.55	0.85	-0.39	0.04	-2.21	0.53	1.20	-0.25	0.23
Number of components 64+	0.45	-1.96	1.29	-1.21	-1.90	-1.45	-0.14	-0.16	-0.86	-2.62	-0.41
Proportion of children 0-3	2.33	-1.11	1.02	-1.02	-0.79	0.65	2.41	-0.54	-0.77	-1.55	0.81
Proportion of children 4-7	1.08	-2.36	1.11	-0.97	0.65	0.66	4.62	-0.21	0.16	-1.20	0.65
Proportion of children 8-12	1.97	-1.67	-0.03	-2.23	-0.52	-0.82	2.61	0.33	-0.47	-0.40	0.87
Proportion of children 13-18	2.71	-2.25	0.67	0.13	-1.00	-1.39	2.13	0.83	0.10	-0.34	1.39
Husband and wife (H/W) only	-0.70	-2.63	-0.99	-0.82	-2.12	-1.23	0.21	-0.10	-0.27	-3.83	0.14
H/W, own children only, oldest child 0-5	-1.79	-1.16	-1.21	-0.17	-0.78	-1.41	-1.27	-1.15	-0.91	-1.72	-0.52
H/W, own children only, oldest child 6-17	-0.26	0.02	-0.19	-0.03	-1.27	-0.34	0.33	-1.56	-2.33	-2.59	-0.54
H/W, own children only, oldest child over 18	-0.46	0.53	0.89	-0.13	-1.45	1.00	0.41	0.17	-0.11	-3.18	0.35
All other H/W households	1.24	-0.36	-0.08	-0.03	-0.43	-0.66	2.16	-0.16	-0.00	-2.81	0.58
One parent (male) at least one child 0-18	-0.56	1.60	-1.13	-0.19	1.11	-0.05	-0.38	0.81	-1.64	0.70	-0.33
Single persons	1.71	0.66	1.82	-0.05	-0.11	-1.64	0.41	1.67	-0.12	-1.64	1.29
Dummy for Midwest region	-1.51	0.33	0.55	0.59	0.56	0.72	1.54	1.62	1.13	-0.59	-0.01
Dummy for South region	0.39	-1.00	-0.35	-0.51	0.16	-0.08	1.61	2.10	0.53	-0.18	0.13
Dummy for West region	-0.53	-0.62	-0.58	-0.91	0.20	0.47	1.74	0.70	0.69	-0.38	0.08
Age of the reference person	-1.57	-2.06	-2.75	-1.80	-1.56	-0.86	-1.10	-1.55	-1.85	-1.99	-1.26
Dummy for Black	-1.79	-0.73	2.38	0.22	0.08	2.35	0.04	-0.17	0.29	2.06	0.60
Dummy for American Indian	0.84	-0.10	-0.08	1.14	1.62	0.68	-0.61	-0.40	0.34	1.23	0.65
Dummy for Asian or Pacific Islander	1.01	0.29	-1.36	-1.44	0.79	1.04	2.07	-1.15	-0.07	0.64	0.21
Less than High School	-1.89	-0.35	-0.74	-1.27	-2.59	-0.01	-2.11	-2.26	1.92	0.15	-1.98
High School	-1.93	0.48	0.33	-0.73	-2.62	0.61	-1.84	-3.08	-1.17	-0.37	0.37
Post High School	-2.28	-1.09	-1.46	-1.04	-2.42	-0.72	-2.21	-3.56	-0.83	-0.71	-1.31
Weeks worked per year	-4.60	-4.00	-5.25	-4.79	-5.88	-3.74	-3.13	-5.52	-5.69	-5.67	-4.19

for total non-durable consumption using average cohort techniques¹.

The components of non-durable expenditure we consider are the ones already considered by Attanasio and Weber (1995; see the first column of Table 5 below): 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). Health and education expenditures do not enter total expenditure on non-durables.

We account for the difference in reference period defining monthly expenditures in the Diary sample as $26 \times 12 = 312$ times the expenditure observed over two weeks, assuming equally complete reporting².

Because of the small within-quarter variation in reporting Interview expenditures (less than 2% of the total variation in our sample), we consider only the expenditure figure for the month preceding the interview (thus taking only four observations for each household). It has been found that expenditures for many items are reported more frequently for this month than for earlier months (Silberstein and Jacobs, 1989). This could obviously mean a partial recollection of past events (mainly less important purchases) increasing with longer reference period and/or a telescoping effect for the month nearest to the interview.

Since differences in consumption across the two surveys might reflect differences in the composition of the samples with respect to household characteristics, we re-weight diary households exploiting a weighting scheme based on the regressions in Table 1 (see for example Battistin, Miniaci and Weber, 2001)³. Under the assumption that sampling differences are adequately captured by this weighting scheme, the remaining differences reflect solely the nature of the instrument exploited in each survey (i.e. diary vs recall questions).

Our analysis will closely follow Deaton and Paxon (1994). To summarize the evidence from this section, we find that both the surveys highlight an increasing within cohort dispersion of total disposable income and wages over time. On the other hand, the pattern of consumption inequality turns out very different depending on the source we consider. Part of the current economic literature is working to explain why a rising income inequality in 1990s has not been accompanied by a corresponding rise in consumption inequality. This pattern has already been pointed out by several papers both exploiting US data (Krueger and Perri, 2001) and UK data (Blundell and Preston, 1998, and Attanasio, Berloffa, Blundell and Preston, 2001). On the other hand, only few papers motivate such a pattern looking at the quality of data exploited in the analysis.

5.1 Expenditure levels

Figure 1 reports age profiles of log non-durable consumption using the Consumer Price Index published by the Bureau of Labor Statistics. Each data

¹We will tend to use 'expenditure' and 'consumption' as two synonyms since the distinction is not relevant in this context.

²Attanasio, Battistin and Ichimura (2002) consider alternative definitions of monthly figures to account for the frequency of purchasing pattern over the two-weeks diary.

³Intuitively, the aim is to down-weight (up-weight) those households in the Interview sample exhibiting characteristics in Table 1 over represented (under represented) with respect to the Diary sample. Such a weighting scheme depends on the conditional probability of observing those characteristics in the population represented by the Interview sample (the well known propensity score), that is on the binary regressions reported in the table.

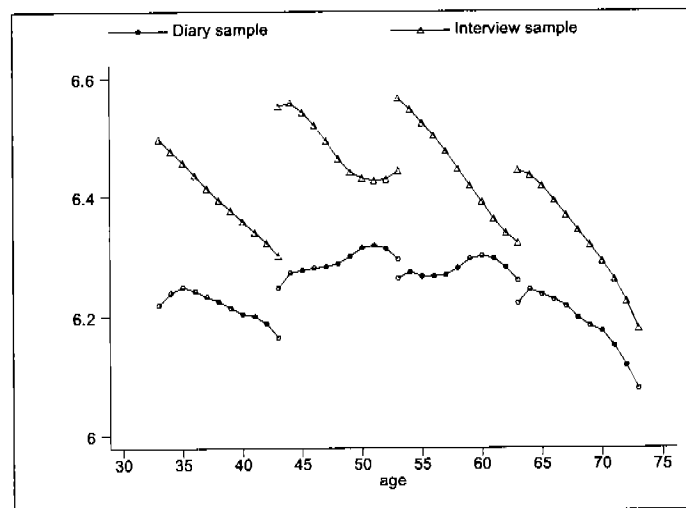


Figure 1: Logs of mean non-durable consumption by cohort

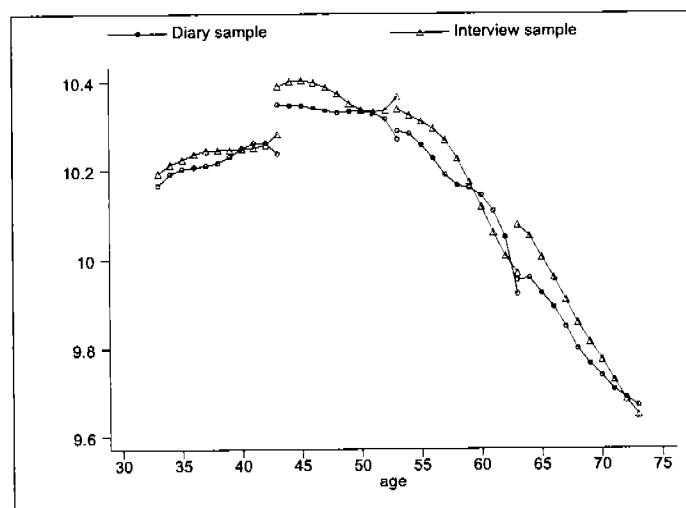


Figure 2: Logs of mean family income by cohort

point represents mean total expenditure on non-durables of the generic cohort in the generic year over the period 1988 – 1998. We present results only for four different cohorts of individuals: those born in 1951-60 (the mid-age of this cohort is 33 in 1988), those born 1941-50 (aged 43 in 1988), those born in 1931-40 (aged 53 in 1988) and those born in 1921-30 (aged 63 in 1988)⁴. Since the weighting scheme adopted reweights Diary households so that the distribution of characteristics in Table 1 is the same across the two samples within each year, the composition of the two samples is comparable across time. Figures reported refer to *per capita* expenditures, exploiting the well established measure of equivalent adults per household - that is the number of adults (older than 18) plus half the number of children (aged up to 18) -, leads to similar results.

The profile obtained exploiting Diary data is always below the profile for Interview data for all the cohorts over the entire life-cycle. In the absence of time effects, vertical distances between broken lines within each survey can be interpreted as cohort effects; these differences don't have the same sign across the two surveys and their absolute value looks smaller exploiting Diary data. Consumption drops after retirement as already found in previous studies on US and UK data (see for example Banks, Blundell and Tanner, 1998). Non-durable consumption is likely to fall as cohorts age because of reduced family size, reduction in work related expenditures (transports and meals, amongst others), investments on durable goods, drops in the cost of leisure.

Figure 2 presents the same statistics referred to available family income, defined as the monthly average income obtained from the amount reported by each household for the whole year. It turns out that Diary data lead to considerably higher values of saving rates over the life cycle.

5.2 Within cohort consumption inequality

In this section we discuss observed differences across the two samples with respect to the dispersion of non-durable expenditure by cohort. We will consider lifetime profiles of the variances of log consumption since the observed pattern of such an indicator discriminates between alternative theoretical models of consumer behavior.

As discussed in Section 3, according to Permanent Income Hypothesis (see Deaton, 1992) the variance of consumption within a fixed-membership group of people should increase over time, since the process describing the intertemporal choice of consumption is distributed as a random walk (provided that innovations to consumption are not perfectly correlated among people within the group). Increasing consumption inequality is financed by an increasing dispersion in total income, defined as the sum of earnings and asset income⁵.

Figure 3 shows the pattern of inequality both for Diary and Interview data. We find inequality (defined as the variance of log monthly non-durable expenditures) to be higher for Diary as compared to Interview data. This may be

⁴Note that - here and in the following graphs - each Interview household appears four times with consumption referred to different months over its one-year interview (the attrition problem is not particularly strong in our data).

⁵However, it is not necessary that there be increasing dispersion in both of these components of total income, and rising consumption inequality will be observed even if the cross sectional distribution of earnings is constant. See for example Attanasio, Berloffa, Blundell and Preston (2002).

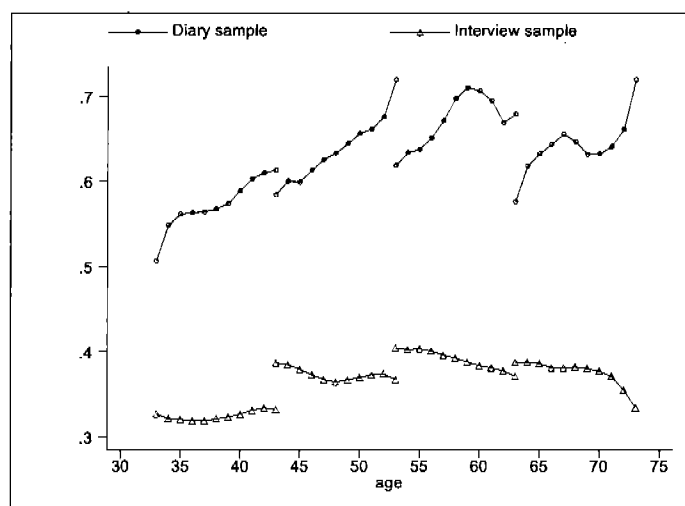


Figure 3: Variance of log non-durable consumption by cohort

due to respondent issues, but is definitely related to the time periods of the two surveys being different: the shorter time period results in data with greater volatility (the Diary reference period is one week). Figure 4 reports the same inequality measure referred to total disposable income. Figure 5 reports hourly wage inequality for the same cohorts.

The information contained in each sample leads to contradictory results with respect to the consumption inequality pattern: Diary inequality increases over time while the recall counterpart gets pretty flat after age 50⁶. Inequality seems to be most pronounced exploiting Diary data for all groups. The bump-shaped pattern peaking before retirement could be explained by the effect of an increasing leisure time due to the retirement age.

Moreover, as pointed out by Deaton (1992), the theoretical behavior of inequality within cohorts depends on people's attitude toward risk and on available mechanisms for sharing risks between people and periods. The Permanent Income Hypothesis assumes that people have certainty equivalent preferences, allows them to lend and borrow as much as they want and permits no direct sharing of risk between people. Changing any of these assumption will generally affect the way in which risk is filtered into consumption inequality un thus its evolution over the life cycle.

Consumption inequality might not have increased over time in spite of the rise in income inequality because spending captures permanent income and income got more volatile. Blundell and Preston (1998) provide some evidence for this using UK data for repeated cross-sections of household over the period 1968-1992 showing a strong growth in transitory income inequality and small variation for the variance of permanent income shocks.

⁶We checked the robustness of this analysis exploiting different measures of inequality selected from the Generalized Entropy family (see for example Shorrocks, 1983) coming out with a picture consistent with the one presented here. Results available on request.

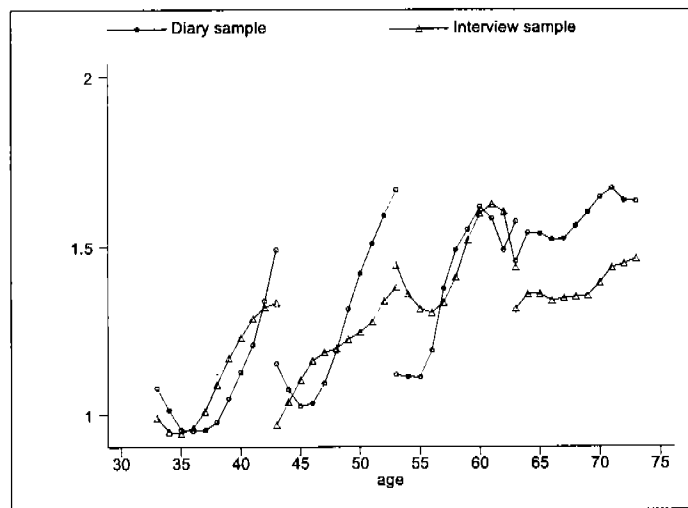


Figure 4: Variance of log family income by cohort

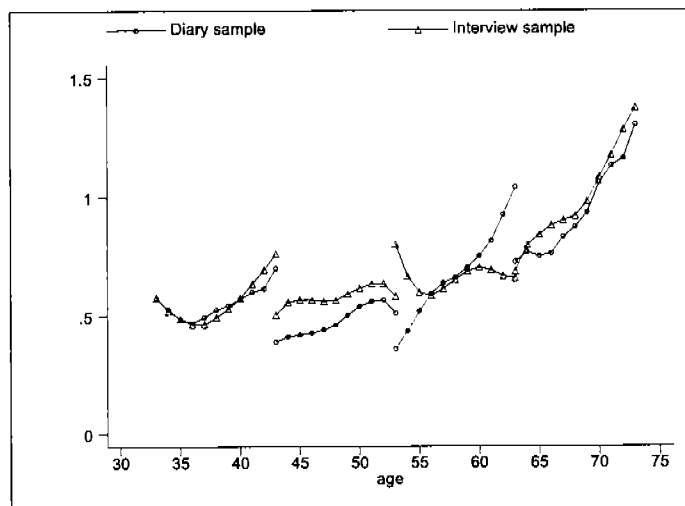


Figure 5: Variance of log wages by cohort

Another reasonable explanation which has not been deeply investigated so far is the measurement error variation in reporting own income and expenditures over time. The reported spending could be less lumpy because of different methods of payment introduced over time (e.g. direct debit) and income might be more error affected because of the increased self-employment (inducing better quality in consumption data and worse quality in income data).

5.3 Aggregate consumption inequality

To conclude this section, we present results on aggregate inequality, without controlling for cohort effects. Figure 6 and Figure 7 report the evolution of inequality over time for Diary and Interview data, respectively. These figures also report the bootstrap 5% critical region resulting from the null hypothesis of constant inequality over the time period 1988-98.

The same figures for income inequality (not reported in this paper) show an increasing pattern both for Diary and Interview data after 1990.

6 Inequality decomposition by factors

In this section we seek to describe how marginal changes in expenditures for specific commodities can affect the inequality of total expenditure.

Table 2 presents reporting rates for Diary and Interview data, that is the proportion of non-zero expenditures for a specific commodity⁷. There is a sizeable proportion of households who have zero recorded expenditure in the Diary sample because of the shorter reference period (two weeks instead of one month). For goods not frequently purchased the expenditure distribution has a spike at zero corresponding to non-consumers. The frequency of purchasing on food at home, clothing and (particularly) heating fuel is lower in the Diary sample; the higher reporting rate for food away from home (which by definition does not include expenditures on vacation) could suggest that many purchases are missed in the Interview sample. The overall pattern remains the same uniformly over time, across samples and for each commodity; this we take as an evidence that changes in consumption habits are well reflected in both the samples⁸.

Possible conjectures on the sources of reported zero expenditure for a certain good in one or both components include (i) under reporting the expenditure in an acknowledged purchase, (ii) non-identified item non-response (i.e. not reporting a purchase that was made) and/or (iii) variation in preferences across the sample. It is actually clear that demographic characteristics may have different impact on the share of spending at different expenditure levels. Households may simply not consume some commodities (due, amongst others, to price and income variation); for example, it is reasonable that families with more children allocate more expenditure to - say - food and a lower fraction to alcohol. Because of the weighting scheme adopted, differences in the two data sources across time in Table 2 presumably reflect only two of the three sources of zero

⁷By definition, those households presenting null expenditure on total food (both at home and away from home) are dropped from the analysis (less than 1% in each sample).

⁸Excluding goods with higher proportion of zeros (tobacco, alcohol and heating fuel) from non-durable expenditure confirms the different pattern of aggregate Diary and Interview inequality presented in Section 5.

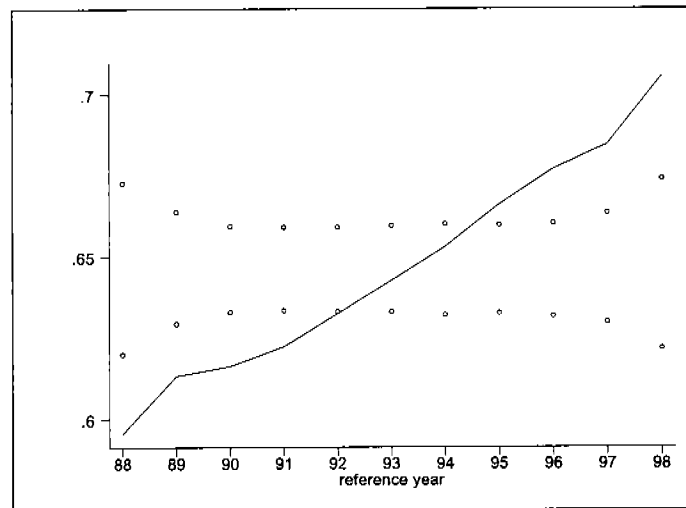


Figure 6: Variance of log non-durable consumption over time in Diary

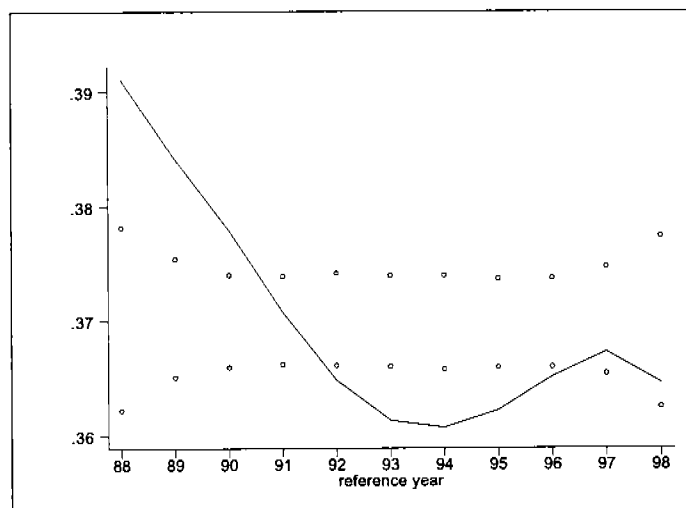


Figure 7: Variance of log non-durable consumption over time in Interview

Table 2: Percentage of consumers: $\Pr(X > 0 | \text{Interview})$ and $\Pr(X > 0 | \text{Diary})$

Interview	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	0.993	0.994	0.993	0.993	0.994	0.994	0.993	0.995	0.995	0.994	0.993
Food and non-alcoholic beverages away	0.836	0.844	0.836	0.826	0.828	0.832	0.832	0.829	0.826	0.840	0.840
Clothing and footwear	0.652	0.654	0.640	0.641	0.633	0.631	0.622	0.612	0.593	0.596	0.561
Tobacco and smoking accessories	0.388	0.375	0.364	0.339	0.338	0.325	0.319	0.325	0.303	0.297	0.285
Alcoholic beverages (at home and away)	0.508	0.495	0.476	0.460	0.464	0.459	0.460	0.446	0.438	0.436	0.431
Transport services (including gasoline)	0.935	0.942	0.939	0.937	0.933	0.940	0.937	0.936	0.932	0.937	0.932
Heating fuel, light and power	0.871	0.879	0.875	0.887	0.888	0.895	0.896	0.901	0.888	0.888	0.891
Housing (rent and services)	0.990	0.991	0.991	0.992	0.992	0.993	0.993	0.991	0.989	0.989	0.989
Personal care, entertainments and services	0.955	0.958	0.954	0.950	0.952	0.947	0.920	0.919	0.910	0.912	0.914
Diary	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	0.977	0.977	0.986	0.984	0.985	0.983	0.978	0.982	0.975	0.977	0.973
Food and non-alcoholic beverages away	0.908	0.904	0.907	0.890	0.887	0.869	0.871	0.862	0.876	0.884	0.904
Clothing and footwear	0.633	0.636	0.619	0.666	0.643	0.630	0.613	0.597	0.598	0.599	0.587
Tobacco and smoking accessories	0.403	0.372	0.377	0.349	0.336	0.309	0.328	0.307	0.316	0.296	0.287
Alcoholic beverages (at home and away)	0.531	0.526	0.517	0.484	0.498	0.477	0.481	0.451	0.443	0.419	0.451
Transport services (including gasoline)	0.932	0.939	0.936	0.917	0.908	0.887	0.909	0.912	0.911	0.914	0.908
Heating fuel, light and power	0.450	0.429	0.459	0.563	0.469	0.434	0.457	0.454	0.446	0.455	0.434
Housing (rent and services)	0.897	0.899	0.888	0.916	0.921	0.886	0.889	0.890	0.884	0.884	0.880
Personal care, entertainments and services	0.909	0.911	0.912	0.905	0.895	0.866	0.872	0.862	0.867	0.860	0.860

expenditures (infrequent purchasing and misreporting) because the variation of preferences (as a function of observable characteristics) is likely to affect in the same way the two samples⁹.

Results in Section 5 rise the problem of identifying the contribution in overall inequality attributable to each commodity defining non-durable consumption. The problem is related to an unique decomposition rule as suggested by Shorrocks (1982), since the inequality contribution assigned to each source can vary arbitrarily depending on the choice of decomposition rule. Particularly important for our purposes is the ability to meaningfully decompose the index into inequality between and within different commodities. The decomposition must be consistent, in the sense that commodities' contribution should add up to the overall amount of inequality.

Table 3 and Table 4 report the percentage contribution of each commodity to total inequality using the 'natural' decomposition rule $e_{yn}(R) = \sum_j p_j \text{tf}(C_j; R)$, both for Diary and Interview data and separately for . The contribution of each commodity C_j to total inequality is then expressed as the slope coefficient of the Engel-type regression of C_j on non-durable expenditure R . Alternative procedures based on decompositions of the Gini coefficient (see for example Garner, 1993) lead to the same result¹⁰.

7 Modelling inaccuracies

The goal of this section is to derive error-adjusted measurements of the inequality indices presented in Section 5. Exploiting jointly the two surveys we (i) discuss on the quality characterizing each commodity entering total non-durable expenditure and (ii) define an improved measure of overall consumption.

A first possible explanation for the evidence we found in Section 5 is due to definitional and collection methodology differences between the two surveys. While the Diary survey collects detailed disaggregated data and then sums these up to get total spending, the Interview survey asks a global retrospective question about total spending. Differences in levels and variances for the considered inequality indices might be determined by different monthly expenditure estimates on each category as a result of this aggregation. Additional effects might depend on periodic changes in the survey instruments over the years: one example of this is that a new diary form with more categories and expanded use of cues for respondents has been introduced in the Diary survey since 1991.

Changes in data quality over time, which may be positive when related to respondent's learning curve or negative when stemming from declining interest as time increases, can affect both the surveys. With regard to the Interview sample, it might be the case that people report similar values of consumption over the four waves conditioning on what reported in the first interview, implying

⁹Attanasio, Battistin and Ichimura (2002) discuss the frequency of purchasing problem looking at Diary data. In particular, they show that the number of days over the two-week diary when expenditures occur gets more concentrated over time. Changes in buying habits over time might be related to the evidence from Section 5 and to the figures of the following two tables.

¹⁰However, under suitable constraints, it can be proved that there is an unique decomposition rule for any inequality measure for which the proportion of inequality attributed to each commodity is the proportion obtained in the natural decomposition rule of the variance (Shorrocks, 1982).

Table 3: Commodities contribution as a percentage of total inequality - first two cohorts

Interview	1988-1991	1992-1995	1996-1998	1988-1991	1992-1995	1996-1998
	born in 1951-60			born in 1941-50		
Food and non-alcoholic beverages at home	7.1957	7.4574	5.5964	4.9647	6.3624	5.9633
Food and non-alcoholic beverages away	8.0654	9.3997	9.0776	10.0610	12.4359	18.0766
Clothing and footwear	32.6521	32.7853	22.3858	28.3309	26.6339	24.9554
Tobacco and smoking accessories	0.1088	0.0829	0.1092	0.0146	0.0035	0.0275
Alcoholic beverages (at home and away)	1.3902	1.2824	1.4906	1.0429	1.2395	1.2865
Transport services (including gasoline)	16.0307	16.3818	27.6270	25.9189	16.4867	16.9660
Heating fuel, light and power	1.7474	2.0944	1.4286	1.7171	1.9328	1.7397
Housing (rent and services)	25.7172	22.1616	20.1835	21.4973	24.6232	23.8070
Personal care, entertainments and services	7.0925	8.3546	12.1014	6.4527	10.2822	7.1781
Diary	1988-1991	1992-1995	1996-1998	1988-1991	1992-1995	1996-1998
	born in 1951-60			born in 1941-50		
Food and non-alcoholic beverages at home	7.2666	8.0383	8.2972	5.4152	6.9732	4.6523
Food and non-alcoholic beverages away	6.2144	8.0274	7.1537	5.9134	6.3015	7.2969
Clothing and footwear	16.2819	11.9874	11.6327	11.7353	10.7301	11.9053
Tobacco and smoking accessories	0.2837	0.4270	0.3756	0.1504	0.1291	0.2300
Alcoholic beverages (at home and away)	1.6075	1.4391	1.8149	0.9954	1.1985	0.5860
Transport services (including gasoline)	13.2221	19.9447	20.5062	39.2361	25.8378	48.4306
Heating fuel, light and power	5.0863	5.7421	5.8729	4.5715	4.4891	1.9368
Housing (rent and services)	31.9880	33.5607	32.5094	21.5242	34.6579	17.7348
Personal care, entertainments and services	18.0494	10.8333	11.8374	10.4584	9.6828	7.2271

Table 4: Commodities contribution as a percentage of total inequality - last two cohorts

Interview	born in 1931-40				born in 1921-30			
	1988-1991	1992-1995	1996-1998	1988-1991	1992-1995	1996-1998	1988-1991	1992-1995
Food and non-alcoholic beverages at home	5.4385	6.5723	6.1477	6.5249	7.0083	2.9995		
Food and non-alcoholic beverages away	14.1535	16.4862	15.5665	12.3103	9.2851	7.5462		
Clothing and footwear	31.5390	22.9064	14.6965	22.8693	21.9766	8.3954		
Tobacco and smoking accessories	0.1527	0.2030	0.1336	0.1850	0.2075	0.1582		
Alcoholic beverages (at home and away)	1.2734	1.5880	1.9024	1.2005	1.5376	1.8347		
Transport services (including gasoline)	15.3221	18.9076	28.3261	15.9605	21.3428	42.5272		
Heating fuel, light and power	2.3731	2.2117	1.2486	1.9517	1.8787	1.2126		
Housing (rent and services)	21.9318	22.9141	23.8998	29.6327	28.2458	18.0866		
Personal care, entertainments and services	7.8159	8.2108	8.0787	9.3651	8.5176	17.2397		
Diary	1988-1991	1992-1995	1996-1998	1988-1991	1992-1995	1996-1998		
Food and non-alcoholic beverages at home	7.8211	6.7855	5.2242	3.9726	7.5562	8.2895		
Food and non-alcoholic beverages away	7.9938	6.0270	5.0306	3.8778	5.8114	5.1250		
Clothing and footwear	17.5275	13.3047	9.2236	10.1982	10.4631	11.0031		
Tobacco and smoking accessories	0.4798	0.3517	0.2955	0.2444	0.4728	0.3382		
Alcoholic beverages (at home and away)	2.0322	1.9121	1.1116	1.2452	1.7420	1.7838		
Transport services (including gasoline)	13.9578	22.8222	9.6119	11.4166	14.5685	22.7666		
Heating fuel, light and power	8.4247	4.8559	4.7426	5.6475	5.8729	5.9594		
Housing (rent and services)	28.4459	33.9726	51.8917	53.7371	43.7127	27.6213		
Personal care, entertainments and services	13.3171	9.9683	12.8683	9.6605	9.8004	17.1131		

a flat pattern of the within-household variance over the interview period (one year). Indeed, several papers have shown that the negative effect of poor quality information as the interview-time increases is bigger than the positive effect due to respondent's learning-by-doing process, both for recall and diary data. Turner (1961) and Silberstein and Scott (1991) find that the average of reported food expenditures for diary data decreases across day and week of participation, probably reflecting under-reporting related to a declining interest. Silberstein and Jacobs (1989) find similar results with respect to the time-in-sample (i.e. the number of cycles of participation) for the Interview survey¹¹.

Let n and m be the two potential reported expenditures as the result of being interviewed exploiting a recall or a diary based questionnaire, respectively. Clearly, the difference between these two terms is informative about the effect of reporting expenditures exploiting recall rather than diary based questionnaires. Since the two surveys are referred to separate samples, an identification problem arises from the fact that - by design - we do observe only one of these measurements on each household.

In what follows we proceed approximating the counterfactual measurement at cohort level, that is as if we knew what Interview (Diary) cohorts would have reported had they been interviewed using Diary (Interview) questionnaire. Under this assumption n and m are two measurements of the same latent variable of interest (non-durable expenditure) for each cohort over the time 1988-98¹².

More precisely, reported expenditures might be thought as the sum of two separate components: the first component is due to respondent's failure to correctly report the amount of goods actually purchased (we will refer to this component as the *measurement* or *reporting error* characterizing each survey). The second component is the true (unobserved) individual consumption on that good. The evidence we produced in Section 5 about differences in means and variances across the two surveys suggests that the error processes leading to n and m are likely to be not identically distributed across samples and over time. If the distribution of measurement errors is not stationary over time we cannot separately identify the effect of a real change in the inequality level from the effect induced by variation in the quality of reporting¹³.

A possible approach is to represent the association between n and m by a suitable model that embodies latent expenditures in such a way that n is independent of m given the true value of consumption (the resulting analysis is known as *factor analysis*; see for example Kim and Mueller, 1978).

¹¹In the following we will not consider any bias arising from measurement effects in self vs proxy response questions, interviewers effect on data collection or measurement errors affecting the considered deflator.

¹²Battistin (2002) estimates counterfactual measurements at household level by matching Interview households to a set of 'twin' Diary households presenting similar characteristics.

¹³To give a flavor of such a problem, think about the case of a multiplicative error affecting real expenditures Y whose intensity is given by a parameter σ (Chesher and Schluter, 2001). If we assume independence between Y and the reporting error process, a second-order approximation for the Gini coefficient of the error-contaminated consumption is given by

$$G_Y + \sigma^2 \frac{E[Y^2 f_Y(y)]}{E[Y]},$$

where G_Y is the Gini coefficient associated to Y . It follows that the distance between 'true' and observed Gini coefficients might be different over time because of variations in σ or in the shape of expenditure distribution (indeed, the incidence of the measurement error is not particularly high when Y is heavily right skewed).

If the latent expenditure and the measurement errors affecting each of the two surveys were mutually independent and we observed both n and m on the same household, the joint distribution (n, m) would uniquely determine the density functions (i) of the 'true' unobserved non-durable expenditure and (ii) of the measurement errors in n and m (for example, under the additional assumption of their symmetry about the origin)¹⁴. Based on this identification result, several nonparametric procedures could be implemented to estimate all the densities involved (see, for example, Horowitz and Markatou, 1996, and Li and Vuong, 1998).

While the assumption of a 'classical' measurement error (i.e. zero mean error independent of the true unobserved variable) on total non-durable expenditure can be a reasonable starting point in any model specification exploiting diary data, it has been largely criticized when using recall data. There is a clear evidence that several kinds of non-classical error can affect the precision in reporting work-related variables (Rodgers, Brown and Duncan, 1993, and Torelli and Trivellato, 1993, Manning and Dickens, 2001, Bound, Brown and Mathiowetz, 2001), earnings (Pischke, 1995) and non-durable expenditures. In particular, Battistin, Miniaci and Weber (2001) provide some evidence for the case in which the magnitude of reporting errors is endogenously determined by the real amount of expenditure (allowing higher expenditure levels to increase the probability of large errors), so that the independence assumption is no longer valid.

8 Restrictions

In our analysis we make the following identifying assumptions.

Condition 1 Either diary or recall data identify correctly (i.e. report without any error) true expenditures for all the commodities amongst non-durables.

Condition 2 We know which source (Diary or Interview) provides the actual amount of spending on each commodity.

Indeed since the two survey components of the CEX - the Interview survey and the Diary survey - are targeted to collect different types of expenditures, errors affecting the non-durable measurements (n, m) result from reporting errors affecting those categories 'less reliable' in each survey component. It follows that those categories either having regular periodic billing or involving major outlays easily recalled for a period of three months or longer are better described exploiting the Interview Survey. On the other hand, those categories referred

¹⁴The result follows immediately considering the characteristic functions $\Phi_r(t)$, $\Phi_d(t)$ and $\Phi_{r-d}(t)$ and assuming - for example - a symmetric reporting error in the Diary sample. If u_r and u_d are the reporting errors associated to Interview and Diary data, respectively, it follows that

$$\begin{aligned}\frac{\Phi_r(t)}{\Phi_d(t)} &= \frac{\Phi_{u_r}(t)}{\Phi_{u_d}(t)}, \\ \Phi_{r-d}(t) &= \Phi_{u_r}(t)\Phi_{u_d}(t),\end{aligned}$$

where the last equality is derived exploiting the symmetry of u_d .

to frequently purchased smaller items are presumably more reliable exploiting the Diary Survey.

Even if potentially the sign of bias in recall and diary data could be in both directions depending on different commodities (over- or under-reporting of true expenditures, respectively), the available evidence from several countries suggests that under-reporting is more likely to affect the great part of items in expenditure surveys. Complete information on small expenditures is likely to be not always available since the respondent may forget to report less important purchases below a certain amount (see for example Van Praag and Vermeulen, 1993, or Alessie, Gradus and Melemborg, 1990, for a model based approach to this problem).

The magnitude of partial recollection of past events varies for different commodities exploiting recall and diary keeping methods. Those components having regular periodic billing are more likely to be well reported by respondents in the Interview Survey. Indeed, exploiting validation data from the National Income and Product Accounts (NIPA) recall expenditures for *transports, fuel* and *rent* expenditures have been found to be reliable and heavily under-reported by diary data (Gieseman, 1987).

Spending on *alcoholic beverages* and *tobacco* traditionally has been under-reported in household surveys; some authors refer to this evidence as a 'puritan' element in household data. Diary were found to give more reliable information about alcohol consumption than recall data (see Poikolainen and Kakkainen, 1983, and Atkinson, Gomulka and Stern, 1990); comparisons on tobacco expenditures based on mean squared error methods exploiting NIPA data suggest better quality from recall data (Branch and Jayasuriya, 1997).

Clothing is a category which requires fuller investigation. Several studies reveal heterogeneity in results exploiting diary or recall information amongst goods within this category. As expected, recall data seem to be more reliable for costly and salient apparel items (with quite variable results exploiting different methods of source selection), but diaries generally capture more apparel spending (Silberstein and Scott, 1991).

Expenditures on food need particular attention. We take as a meaningful indicator for the quality in reporting Diary *food at home* expenditure the difference between (log-)weekly expenditure derived from diaries (as the sum of detailed food data) and (log-)weekly 'usual' expenditure for food and non-alcoholic beverages at grocery stores given by respondents. This information is collected for each household at the beginning of the (two-week) diary period and its accuracy is therefore not influenced by how respondents learn about own expenditures during the interview (see Figure 8)¹⁵.

It is worth noting that the average of this indicator over households conditional on each month (to control for seasonal effects) is always negative for the period of time covered by this analysis, with values generally decreasing in absolute value over time¹⁶. There is a mild effect of the interview month

¹⁵Note that 'food and non-alcoholic beverages' is the only commodity entering non-durable expenditure for which we observe on the same household both a recall and a diary-based measurement. This would enable us to exploit also the information on the joint distribution (r, d) as described above. This additional information is used by Attanasio, Battistin and Ichimura (2002).

¹⁶This might be related to the introduction of new cues in early 1990's; Tucker (1992) studies the effect of such procedural variations on non-sampling errors.

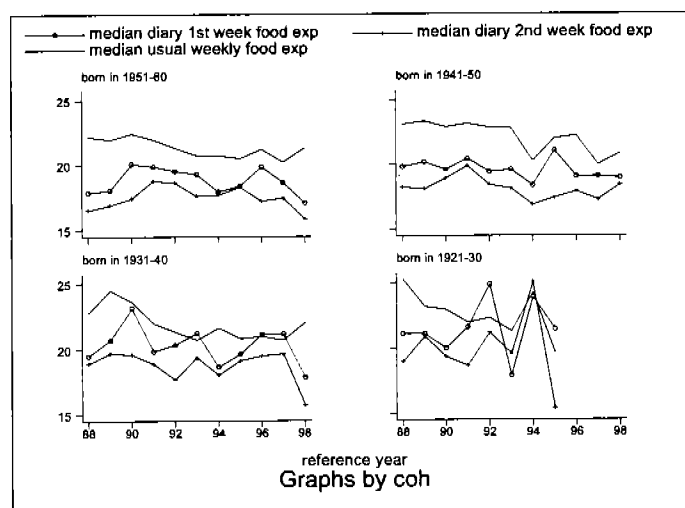


Figure 8: Median of food at home consumption by cohort for Diary households: usual weekly expenditure and reported expenditures for the 1st week and the 2nd week of the diary

on the magnitude of under-reporting, since households interviewed in December usually present values closer to zero. This evidence supports the idea that recall questions overestimate the real spending on food-related items probably because in their reporting households include more goods¹⁷. Indeed the recall question about food expenditure in the Interview survey is derived subtracting to the usual amount spent at the grocery store the usual amount on non-food items.

Finally, we know that even if by definition the Diary Survey does not include purchases during trips or vacation, *food away from home* expenditure is greater than the one reported in the Interview Survey. Because of the nature of the two instruments, it would be reasonable thinking at diaries as the most reliable source to measure food away from home purchases (Stanton and Tucci, 1982). See also Lyberg *et al.* (1997) for a review of data quality in survey measurements.

To summarize, following the literature about official statistics recall data will be considered more reliable to identify expenditures on those components having regular periodic billing or involving major outlays. Diary data will be exploited as the reference source for expenditures on grocery items and personal care, entertainments and other services. Table 5 summarizes the discussion of this section.

¹⁷Battistin, Blow and Smith (2002) find a similar pattern comparing recall and diary information using data from the BHPS, the Family Expenditure Survey and the National Food Survey.

Table 5: Survey selection

Factors in \mathcal{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
personal care, entertainments and other services	Diary

Factors in \mathcal{I}	Survey
clothing and footwear	Interview
tobacco and smoking accessories	Interview
transport services (including gasoline)	Interview
heating fuel, light and power	Interview
housing (rent and services)	Interview

9 Error correction

To summarize the contents of Section 8, we motivate the different pattern of means and inequality indices in the two samples as the aggregate result of inaccuracies affecting both the surveys (Interview and Diary) in reporting spending habits. These inaccuracies are mainly referred to those components of non-durable expenditure each survey is not targeted to: frequently purchased smaller items and services (Interview Survey) and large expenditures occurring on a regular basis such as rent or utilities (Diary Survey). The aggregate effect of these inaccuracies is likely to vary over time, because it depends (i) on significative changes in the structure (i.e. design and collecting strategies) of the two surveys and (ii) on time-in-sample effects.

The aim of this section is to formalize the restrictions presented in Table 5. Let

$$(C_j^d, C_j^r) \quad (2)$$

be the two outcomes that result from being interviewed on commodity C_j via a diary or a recall based questionnaire, respectively, and let C_j^* be the ‘true’ expenditure on the same commodity we don’t observe.

It is worth stressing again that only one of the two variables in (2) is observed for each household since they are potential outcomes resulting from using alternative survey instruments to collect information on household consumption. Measurement errors on each component would be identified if we observed the counterfactual expenditure for each household, that is what the same household would have reported had it participated the other survey.

What we would like to observe is

$$R^* = \sum_j C_j^*$$

that is the sum of ‘true’ expenditures on commodities entering non-durables expenditure R (i.e. those commodities reported in Table 5). What we actually observe are instead two error affected measurements of R^* given by aggregate

expenditures that result from Diary and Interview data

$$\begin{aligned} R^d &= \sum_j C_j^d s \\ R^r &= \sum_j C_j^r S \end{aligned}$$

Note that - by sampling design - (R^r 's R^d) don't refer to the same units. By means of the assumption we discussed in Section 8, the measurement error affecting R^r and R^d depends on the measurement error on different subsets of commodities entering total expenditure (i.e. those commodities either in \mathcal{D} or in \mathcal{I} , respectively). That is, total expenditure from Interview data can be written as

$$\begin{aligned} R^r &= \sum_{j \in \mathcal{I}} C_j^r + \sum_{s \in \mathcal{D}} C_s^r \\ &= \sum_{j \in \mathcal{I}} C_j^* + \sum_{s \in \mathcal{D}} C_s^r s \end{aligned}$$

and the last expression follows since expenditures reported on those commodities in \mathcal{I} are - by assumption - not affected by measurement error. By analogy, it follows that

$$\begin{aligned} R^d &= \sum_{j \in \mathcal{I}} C_j^d + \sum_{s \in \mathcal{D}} C_s^d \\ &= \sum_{j \in \mathcal{I}} C_j^d + \sum_{s \in \mathcal{D}} C_s^* S \end{aligned}$$

Thus, the error components affecting (R^r 's R^d) are the aggregate result of the reporting error on different subsets of commodities, so that we can write

$$R^r = R^* + \overbrace{\sum_{j \in \mathcal{D}} (C_j^r - C_j^*)}^{\text{recall error}} = R^* + r^r s \quad (3)$$

$$R^d = R^* + \underbrace{\sum_{j \in \mathcal{I}} (C_j^d - C_j^*)}_{\text{diary error}} = R^* + r^d s \quad (4)$$

where r^r and r^d represent the measurement error due to recall and diary interviews, respectively. Moreover, note that by definition (r^r 's r^d) are likely to be dependent of R^* since they depend on a common set of commodities.

The following of this section exploits restrictions reported in Table 5 to derive error corrected figures for consumption levels and variances over time. We will focus on means and variances because the identification of higher moments (and percentiles) of the distribution would require modelling the relationship between true expenditures and reporting errors (see Battistin, 2002, or Attanasio, Battistin and Ichimura, 2002).

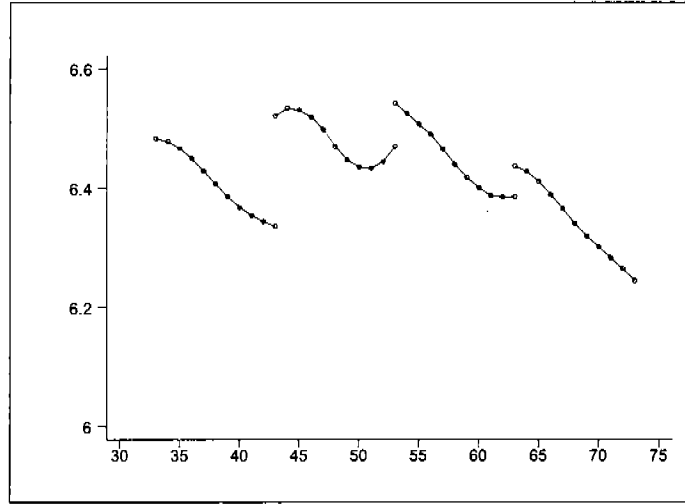


Figure 9: Logs of mean non-durable consumption after correction

9.1 Consumption means

The aim of this section is to derive the analogue of Figure 1 correcting for the reporting errors (r_{sr}^d). Since by definition

$$x(R^*|o) = x\left(\underbrace{\sum_{j \in \mathcal{R}} C_j^*|o}_{\text{recall data}}\right) + x\left(\underbrace{\sum_{j \in \mathcal{D}} C_j^*|o}_{\text{diary data}}\right)$$

we can use either Interview or Diary means according to the identification rule described in the last column of Table 5. Results are reported in Figure 9.

9.2 Consumption inequality

The goal of this section is to derive the analogue of Figure 3 correcting for reporting errors. Throughout our analysis, we will consider figures for the squared coefficient of variation of total expenditure and not for the variance of logs. The reason for this choice will be clear from what follows.

The variance of total expenditure can be expressed as a function of the variance of commodities in \mathcal{D} or in \mathcal{I} and between-group covariances

$$\begin{aligned} e_{yn}(R^*|o) &= e_{yn}\left(\underbrace{\sum_{j \in \mathcal{I}} C_j^*|o}_{\text{recall data}}\right) + e_{yn}\left(\underbrace{\sum_{j \in \mathcal{D}} C_j^*|o}_{\text{diary data}}\right) \\ &+ 2p \text{tf}\left(\underbrace{\sum_{j \in \mathcal{I}} C_j^*s \sum_{j \in \mathcal{D}} C_j^*|o}_{\text{restrictions}}\right) S \end{aligned} \quad (5)$$

The ratio of the previous quantity to the squared mean of total consumption represents a first order approximation for the variance of log consumption¹⁸. The assumptions we made in Table 5 identify only the first two terms in the previous expression, using either Diary or Interview data. On the other hand, the covariance term cannot be identified from the available information since by definition $\mathcal{D} \cap \mathcal{I} = \emptyset$ and we cannot observe true expenditures on \mathcal{D} and \mathcal{I} for the same individuals.

We actually observe two measurements of such covariance in our data. The first one is derived by taking the covariance between commodities in \mathcal{D} and commodities in \mathcal{I} from the Interview sample; the second one is derived from the Diary sample. However, neither of these two covariances identify the true covariance in (5), since the difference between true and observed covariances represents the correlation between commodities in \mathcal{I} (in \mathcal{D}) and recall errors (diary errors, respectively). In fact, using (3) and (4) we can write

$$\begin{aligned} p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^r s \sum_{j \in \mathcal{D}} C_j^r | o \right) &= p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^{*s} \sum_{j \in \mathcal{D}} C_j^r | o \right) \\ &= p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^{*s} \sum_{j \in \mathcal{D}} C_j^{*s} | o \right) + p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^{*s} r^r | o \right) s \end{aligned} \quad (6)$$

and by analogy

$$\begin{aligned} p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^d s \sum_{j \in \mathcal{D}} C_j^d | o \right) &= p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^d s \sum_{j \in \mathcal{D}} C_j^{*s} | o \right) \\ &= p \text{ tf} \left(\sum_{j \in \mathcal{I}} C_j^{*s} s \sum_{j \in \mathcal{D}} C_j^{*s} | o \right) + p \text{ tf} \left(\sum_{j \in \mathcal{D}} C_j^{*s} r^d | o \right) s \end{aligned} \quad (7)$$

Figure 10 reports the difference between observed Interview and Diary covariances (multiplied by a scaling factor). Figure 11 reports the squared coefficient of variation of consumption obtained from (5) when the covariance term is estimated by taking (6) or (7). It turns out that the within cohort inequality is increasing over time with the exception of those individuals born in 1931-40 (i.e. those aged 53 in 1988). Such variance growth is particularly evident for the oldest cohort of individuals.

10 Bounds on inequality

The pattern observed in Figure 11 can be driven by the error components affecting (6) or (7). The strategy we will take to infer the covariance term in (5) consists of three different steps.

¹⁸A Taylor expansion of $\ln Y^*$ in a neighborhood of its mean gives

$$\ln Y^* = \ln E(Y^*) + \frac{Y^* - E(Y^*)}{E(Y^*)} - \frac{[Y^* - E(Y^*)]^2}{2E(Y^*)^2} + \text{error}$$

so that the variance of logs is approximated by taking the squared coefficient of variation

$$\frac{\text{Var}(Y^*|c)}{E(Y^*|c)^2} = CV(Y^*|c)^2 \simeq \text{Var}(\ln Y^*|c).$$

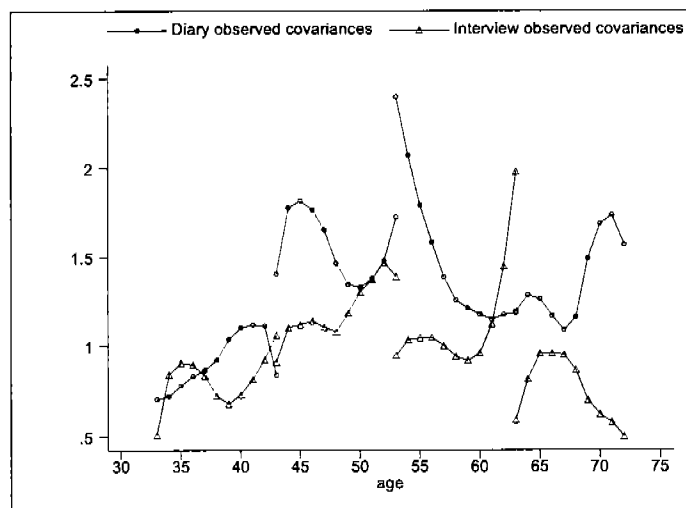


Figure 10: Differences in observed covariances

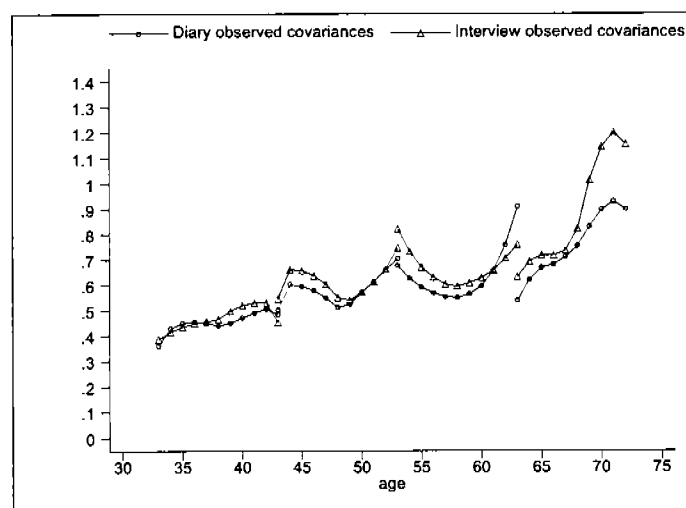


Figure 11: Consumption inequality using observed covariances

- We will be at first non-informative about the magnitude of such covariance and we will bound it using Cauchy-Schwartz bounds. For simplicity, assume that we have expenditures only on two commodities C_j and C_s , where $d \in \mathcal{I}$ and $s \in \mathcal{D}$. The absolute value of their covariance can be bounded by the root of the product of their variances (that turns out identifiable using restrictions in Table 5)

$$|\text{p tf}(C_j^* C_s^* | o)| \leq \sqrt{e \text{ yn}(C_j^* | o) e \text{ yn}(C_s^* | o) S}$$

- The previous bounds can be tightened by a regression adjustment on a set of characteristics (thus considering the residual variation after controlling for \mathbf{z}). The set \mathbf{z} includes family income and a bunch of household characteristics common across the two surveys. We define the following regressions

$$\begin{aligned} C_j^* &= x(C_j^* | \mathbf{z}) + \varepsilon_j^* \\ C_s^* &= x(C_s^* | \mathbf{z}) + \varepsilon_s^* \end{aligned}$$

so that the ε terms are orthogonal to the corresponding regression functions $x(C^* | \mathbf{z})$ that we estimate (semi-)parametrically using Interview or Diary data. We then consider Cauchy-Schwartz bounds for the residuals of each equation

$$|\text{p tf}(\varepsilon_j^*, \varepsilon_s^* | o)| \leq \sqrt{e \text{ yn}(\varepsilon_j^* | o) e \text{ yn}(\varepsilon_s^* | o) S}$$

- We assume that the sign of $\text{p tf}(\varepsilon_j^*, \varepsilon_s^* | o)$ is correctly identified when observed Diary and Interview covariances have the same sign. This corresponds to assuming that the error components affecting (6) or (7) are not that big to reverse the sign of the true covariance. Combining this assumption with the Cauchy-Schwartz bounds derived in the previous point, we obtain one-tail bounds for the covariance, since the lower bound (upper bound) is zero when the sign is positive (negative, respectively).

11 Results

Cauchy-Schwartz bounds, regression adjusted bounds and one-tail bounds on inequality are reported in Figure 12, Figure 13 and Figure 14, respectively¹⁹. Figure 15 combines the within-cohort information of Figure 14 and presents result for consumption inequality over time.

12 Conclusions

In this paper we have discussed how to account for the measurement error affecting both diary-based and recall-based data on non-durable consumption. In

¹⁹The regression functions to adjust for the observable characteristics Z are estimated by a semiparametric specification

$$X^* = \beta Z + \psi(\text{income}) + \varepsilon^*,$$

where ψ is unknown but suitably smooth.

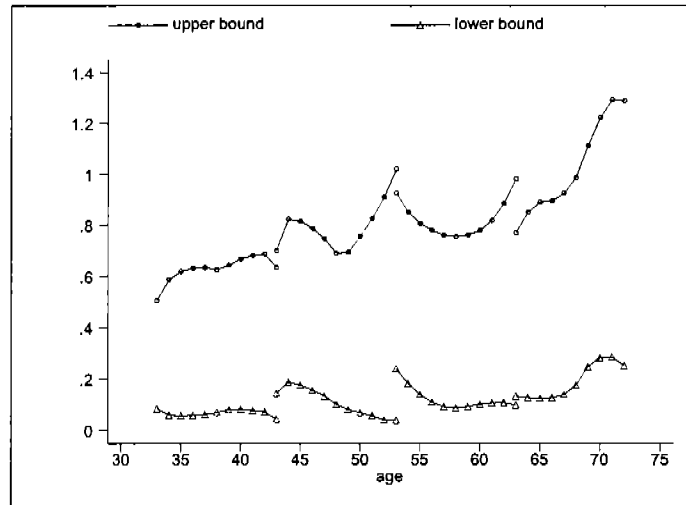


Figure 12: Cauchy-Schwartz bounds on inequality

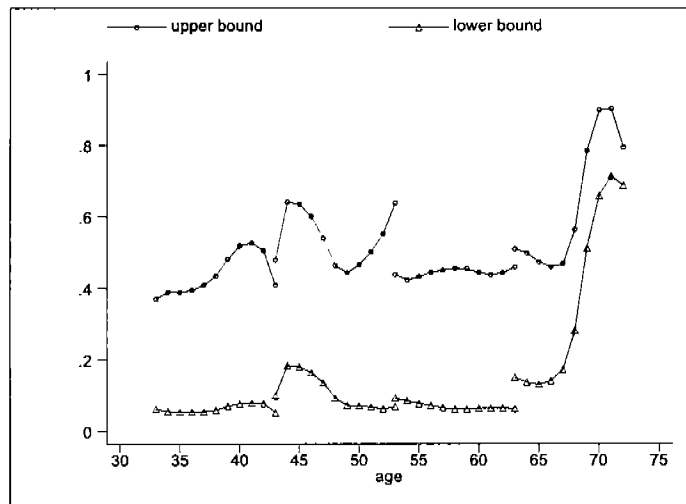


Figure 13: Regression adjusted bounds on inequality

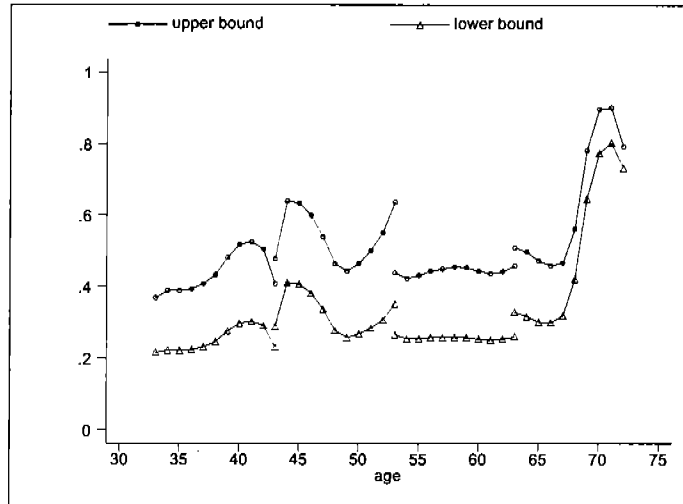


Figure 14: Regression adjusted bounds on inequality assuming known correlation

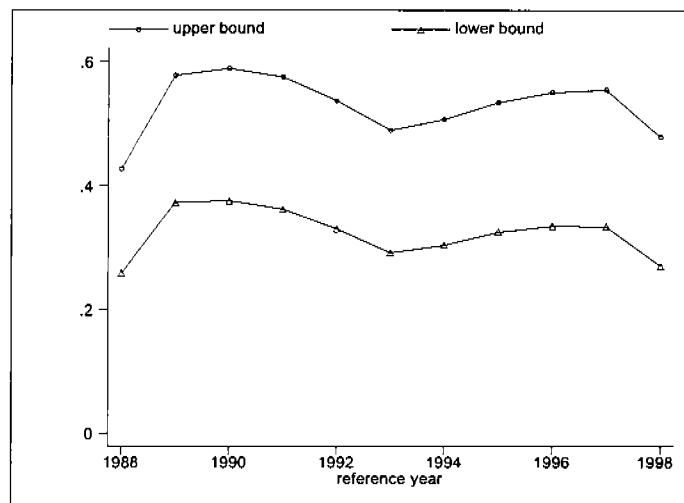


Figure 15: Regression adjusted bounds on inequality assuming known correlation

fact, it is likely that not all the commodities entering non-durable consumption are well reported exploiting only one of these two components. Commodities made of frequently purchased, smaller items are presumably correctly measured using diaries while commodities made of large expenditures or expenditures occurring on a regular basis are presumably more accurate exploiting recall data. It turns out that neither diary nor recall-based data provide a reliable aggregate measure of non-durable consumption.

Integrating different datasets presents the problem of determining the appropriate survey component from which to select the expenditure items. On the basis on evidence reported in a number of previous studies, we split the set of commodities entering non-durable consumption into two groups indicating which one of the two survey methodologies (diary or recall) leads to more accurate data quality.

If data from both the survey sources were available on the same unit, we would be able to define a new measure of non-durable expenditure at micro-level just considering the more reliable source for each commodity. If the samples are referred to different units as for the CEX case, we can provide improved aggregate measures of consumption looking at its mean and - under suitable conditions on the covariance between commodities - its variance following macro-units (i.e. cohorts of households presenting the same observable characteristics).

Our procedure allows us (i) to define an improved measure for mean and variance of non-durable expenditure over the 1990s and (ii) to characterize the measurement error affecting the commodities whose quality is doubtful according to other studies in the literature. We have presented the implications of our findings for the estimation of inequality indices. We show that - exploiting jointly diary and recall data to account the measurement problem - we come out with an inequality pattern not completely against the permanent income hypothesis.

Further considerations

- In this paper we consider the estimation of non-durable expenditures at household level. Estimating the true expenditure level on frequently purchased items for the recall sample (and on bulky items in the diary sample) can be seen as the problem of inferring counterfactuals: what is the counterfactual diary (recall) expenditure measure for recall (diary) respondents? A possible solution requires using information on common observable characteristics at household level to predict recall expenditures in the diary sample and *viceversa*. Battistin (2002) exploits Diary information to approximate - even if the samples are different - counterfactual measurements in the Interview survey coming out with pictures similar to those presented in this paper.

- Can we characterize the measurement error affecting non-durable commodities? If we observed both recall and diary outcomes for commodity C on the same household, according to the assumptions we made the error on that commodity would be (non-parametrically) identified. In fact, if we write

$$\xi = \text{---}^*s$$

we could identify the distribution of ξ by taking the difference between observed diary and recall outcomes depending on the rule discussed in

Table 5. In our case, we can easily recover the first moment of ξ by taking

$$x(\xi) = x(\cdot) - x(\cdot^*)S$$

The variance of the error (and by analogy its higher moments) is not identifiable, since it depends on the correlation between \cdot and \cdot^*

$$e_{yn}(\xi) = e_{yn}(\cdot) + e_{yn}(\cdot^*) - 2p_{tf}(s \cdot^*)S \quad (8)$$

Again, we could bound (8) as we did in Section 10; such bounds can be tightened by imposing the restriction that $p_{tf}(s \cdot^*) \geq 0$, that is imposing that the correlation between what individuals report and their real expenditure is non-negative (or, eventually, not below a given threshold).

References

- [1] Alessie, R. Gradus, R. and Melemborg, B. (1990), *The problem of not observing small expenditures in a consumer expenditure survey*, Journal of Applied Econometrics, Vol. 5, 151-166
- [2] Atkinson, A. B. Gomulka, J. and Stern, N.H. (1990), *Spending on Alcohol: Evidence from the Family Expenditure Survey 1970-1983*, The Economic Journal, Vol.100, No.402, 808-827
- [3] Attanasio, O.P. Battistin, E. and Ichimura, H. (2002), unpublished manuscript, Institute for Fiscal Studies, London
- [4] Attanasio, O.P. Berloffa, G. Blundell, R. and Preston, I. (2001), *The growth in income inequality and income volatility: incomes, earnings and consumption*, unpublished manuscript, University College London
- [5] Attanasio, O.P. and Weber, G. (1995), *Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey*, Journal of Political Economy, Vol. 103, n. 95, 1121-1157
- [6] Banks, J. Blundell, R. and Tanner, S. (1998), *Is There a Retirement-Savings Puzzle?*, American Economic Review, Vol. 88, n.4, 769-788
- [7] Battistin, E. (2002), Ph.d. thesis in Statistics, University of Padova.
- [8] Battistin, E. Blow, L. and Smith, Z. (2002), unpublished manuscript, Institute for Fiscal Studies, London
- [9] Battistin, E. Miniaci, R. and Weber, G. (2001), *What do we learn from recall consumption data?*, forthcoming Journal of Human Resources
- [10] Blundell, R. and Lewbel, A. (1999), *Puzzles of consumption and income distribution explained: Gibrat's law for permanent income*, unpublished manuscript
- [11] Blundell, R. and Preston, I. (1998), *Consumption Inequality and Income Uncertainty*, Quarterly Journal of Economics, Vol. , No. , 603-640

- [12] Bound, J. Brown, C. and Mathiowetz, N. (2001), *Measurement error in survey data*, in Heckman, J.J. and Leamer, E. (eds), *Handbook of Econometrics*, Volume 5, 3707-3843
- [13] Branch, R. and Jayasuriya, B. (1997), *Consumer Expenditure Interview and Diary Survey Data Selection: a New Method*, mimeo, Bureau of Labor Statistics
- [14] Chesher, A. and Schluter, C. (2001), *Welfare measurement and measurement error*, forthcoming *Review of Economic Studies*
- [15] Deaton, A. (1992), *Understanding Consumption*, Oxford: Clarendon
- [16] Deaton, A. and Paxon, C. (1994), *Intertemporal Choice and Inequality*, *Journal of Political Economy*, Vol. 102, n.3, 437-467
- [17] Garner, T.I. (1993), *Consumer Expenditures and Inequality: An Analysis Based on Decomposition of the Gini Coefficient*, *The Review of Economics and Statistics*, Vol. 75, No. 1, 134-138
- [18] Gieseman, R. (1987), *The Consumer Expenditure Survey: Quality Control by Comparative Analysis*, *Monthly Labor Review*, March, 8-14
- [19] Johnson, D. and Shipp, S. (1995), *Trends In Inequality Using Consumer Expenditures: 1960 To 1993*, BLS Economic Working Paper Series
- [20] Kim, J.O. and Mueller, C.W. (1978), *Introduction to Factor Analysis*, Beverly Hills: Sage
- [21] Krueger, D. and Perri, F. (2001), *Does Income Inequality Lead to Consumption Inequality? Empirical Findings and a Theoretical Explanation*, mimeo, Stanford University
- [22] Hall, R.E. (1978), *Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence*, *Journal of Political Economy*, Vol. 86, No. 6, 971-987
- [23] Horowitz, J.L. and Markatou, M. (1996), *Semiparametric estimation of regression models for panel data*, *Review of Economic Studies*, Vol. 63, 145-168
- [24] Li, T. and Vuong, Q. (1998), *Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators*, *Journal of Multivariate Analysis*, Vol. 65, 139-165
- [25] Lyberg, L., Biemer, P., Collins, M., de Leeuw, E., Dippo, C., Schwarz, N. and Trewin, D. (eds.) (1997), *Survey Measurement and Process Quality*, New York: Wiley
- [26] Manning, A. and Dickens, P. (2001), *The national minimum wage and wage inequality*, unpublished manuscript, London School of Economics
- [27] Pischke, J.S. (1995), *Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study*, *Journal of Business and Economic Statistics*, Vol. 13, No. 3, 305-314

- [28] Poikolainen, K. and Karkkainen, P. (1983), *Diary Gives More Accurate Information About Alcohol Consumption than Questionnaire*, Drug and Alcohol Dependence, Vol.11, 209-216
- [29] Rodgers, W.L. Brown, C. and Duncan, G.J. (1993), *Errors in Survey Reports of Earnings, Hours Worked and Hourly Wages*, Journal of the American Statistical Association, Vol. 88, n. 424, 1208-1218
- [30] Rosati, N. (2001), *Poverty in Italy in the 1980s and the 1990s: a measurement error approach*, working paper no. 4, Department of Statistics, University of Padova
- [31] Shorrocks, A.F. (1982), *Inequality Decomposition by Factor Components*, Econometrica, Vol. 50, No. 1, 193-211
- [32] Shorrocks, A.F. (1983), *The Impact of Income Components on the Distribution of Family Incomes*, Quarterly Journal of Economics, Vol. 98, No. 2, 311-326
- [33] Silberstein, A.R. and Jacobs, C.A. (1989), *Symptoms of Repeated Interview Effects in the Consumer Expenditure Survey*, in Kasprzyk, D. Duncan, G. Kalton, G. and Singh, M.P. (eds.) *Panel Surveys*, New York: Wiley, 289-303
- [34] Silberstein, A.R. and Scott, S. (1991), *Expenditure Diary Surveys and Their Associated Errors*, in Biemer, P.P. Groves, R.M. Lyberg, L.E. Mathiowetz, N.A. and Sudman, S. (eds.) *Measurement Errors in Surveys*, New York: Wiley, 303-326
- [35] Slesnick, D.T. (1998), *Are our Data Relevant to the Theory? The Case of Aggregate Consumption*, Journal of Business and Economic Statistics, Vol. 16, No. 1, 52-61
- [36] Stanton, J.L. and Tucci, L.A. (1982), *The Measurement of Consumption: A comparison of Surveys and Diaries*, Journal of Marketing Research, Vol. 19, 274- 277
- [37] Torelli, N. and Trivellato, U. (1993) *Modelling Inaccuracies in Job-Search Duration Data*, Journal of Econometrics, Vol. 59, No. 1/2, 187-211
- [38] Tucker, C. (1992), *The Estimation of Instrument Effects on Data Quality in the Consumer Expenditure Survey*, Journal of Official Statistics, Vol. 8, No. 1, 41-61
- [39] Turner, R. (1961), *Inter- Week Variations in Expenditure Recorded During a Two-Week Survey of Family Expenditure*, Applied Statistics, Vol. 10, No. 3, 136-146
- [40] Van Praag, B.M.S. and Vermeulen, E.M. (1993), *A count-amount model with endogenous recording of observations*, Journal of Applied Econometrics, Vol. 8, 383-395
- [41] Wilcox, D.W. (1992), *The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical work*, American Economic Review, Vol. 82, 922-941