Errors in Survey Reports of Consumption Expenditures

Erich Battistin*
Institute for Fiscal Studies, London

28th January 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 it 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 produce some evidence of non-classical measurement error affecting the aggregate measure of consumption both for diary and recall based data; we also show the implications of our findings to 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, looking at the inequality in the distribution of household expenditures using US micro-level data. 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.

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

*This paper benefited from useful discussions with O. Attanasio, R. Blundell, C. Meghir, H. Low, E. Rettore, U. Trivellato, G. Weber and J. Eltinge, T. Garner and W. Weber at the BLS. Address for correspondence: Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE - UK. E-mail: erich_b@ifs.org.uk.

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).

In what follows we jointly exploit diary and recall data from two independent samples from the US Consumer Expenditure Survey to define inequality measures for 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. Diary surveys are designed to obtain detailed recordings of expenditures on small, frequently purchased items which are normally difficult to recall.

According to the evidence from different countries, we look at diaries as a more reliable and accurate benchmark for the true underlying household expenditure on non-durable goods and services. We explicitly model measurement error sources exploiting information from one sample as an instrument for the other to derive a correction for the bias affecting observed inequality.

One of the main points in discussing welfare indices is the identification of permanent and transitory aspects; a constant inequality pattern over time would mean that a policy maker should mainly account for the variation induced by transitory shocks. 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. Diverging saving rates across the permanent income distribution due to changing in tastes, in social norms or in liquidity constraints would also be an explanation for such a pattern (Clark, 2001). 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).

The remaining of this paper is organized as follows. Section 2 describes the two data sources and compares descriptive statistics of household characteristics already found to be relevant for data quality in previous studies of expenditure surveys. Section 3 presents a puzzle implied by the comparison of means (Section 3.1) and inequality indicators (Section 3.2) of total non-durable expenditure exploiting the information in the two surveys. Section 4 analyzes such discrepancies looking at the contribution of different non-durable goods. The identification restrictions to define an improved measure of expenditure combining recall and diary information on each non-durable commodity are discussed in Section 5 and Section 6. Section 7 and Section 8 present the estimation procedures for total non-durable expenditure and the reporting error processes affecting non-durable components, respectively. Results are presented in Section 9 and Section 10 concludes. Some more technical comments on our empirical findings are reported in the Appendix.

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 different from the ones in the Interview sample (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 of the instruments, there is a set of items (basically non-durable goods) for which expenditures are captured by both instruments. In any case the direct 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.

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 the one consumed on vacation) are referred to the usual monthly amount. The second type of question asks for the amount of expenditures 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.

Clearly household characteristics (occupation and economic activity of the head, household composition, region of residence) affect the share of spending and the quality in 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 annual available family income (before taxes), work-related information and characteristics found to be relevant for data quality in previous analysis of CEX data (Tucker, 1992). 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 particularly significant).

The data exploited in this analysis cover ten years between 1988 and 1998. The main difference in the two surveys is confirmed to lie in the diary relative over-sampling of higher educated households with young children (probably leading to a significative difference in available family income). The amount of weeks worked per year by the reference person[1] is higher in the Diary sample; however, significant differences are found along several dimensions and with a different pattern over time.

3 Evidence

3.1 Expenditure levels

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 for total non-durable consumption using average cohort techniques[2].

Following Attanasio and Weber (1995) we include amongst components of non-durable expenditure food (at home and away from home), alcohol, tobacco, clothing and footwear, heating fuel and electricity, public and private transport (including gasoline) and expenditure on goods and services for personal care. We account for the difference in reference period defining monthly expenditures in the Diary sample as $26/12 = 2.16$ times the expenditure observed over two weeks, assuming equally complete reporting.

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

[1] In what follows the head of the family 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).

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

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	-2.36	1.62	-0.79	0.33	0.10	0.35	-0.02	-0.59	1.64	1.05	1.23
Number of components 18+	1.28	-0.38	1.25	1.56	0.63	1.51	-0.87	1.76	-0.56	0.14	-1.43
Number of components 18−	0.05	1.87	1.98	-0.15	-2.36	-1.25	-0.41	-0.78	-2.87	-0.33	0.05
Proportion of children $0-3$	1.68	0.68	0.64	0.06	-2.60	2.25	0.89	0.07	1.17	1.19	0.54
Proportion of children $4-7$	1.06	0.43	0.40	0.38	1.59	1.17	2.10	-1.05	-0.19	0.24	2.07
Proportion of children $8-12$	2.65	-0.83	-0.30	-0.69	-0.69	0.50	1.97	0.55	-0.82	-0.12	0.03
Proportion of children $13-18$	2.29	3.67	0.40	1.16	-1.16	0.25	0.57	-1.07	-1.40	1.44	1.44
Age of the reference person	-1.68	0.33	-1.07	0.27	-0.65	-2.50	-0.37	-0.76	-0.97	-0.55	-3.07
Dummy for retired head	-1.83	-2.40	-1.48	-0.81	-1.52	-4.44	-1.75	-1.17	2.88	1.71	-2.05
Weeks worked per year	-5.53	-4.62	-5.69	-4.05	-5.44	-6.61	-1.86	-4.02	-4.61	-4.52	-3.98
Total amount of income before taxes	4.83	3.85	3.88	4.36	6.33	6.35	2.76	2.36	-3.98	-0.83	3.79
Dummy for Midwest region	-1.39	-1.27	-0.86	-0.02	-0.42	-0.28	-0.20	0.93	0.10	1.83	0.34
Dummy for South region	-0.33	-0.77	-1.15	-1.93	-1.50	-0.93	0.71	2.25	0.63	2.53	1.40
Dummy for West region	-1.79	-0.31	-1.72	-1.40	-0.56	0.05	1.10	1.06	-1.13	2.28	-0.11
Husband and wife (H/W) only	-1.38	-1.95	-2.01	-1.70	-0.40	1.05	-0.27	-0.52	-0.08	-0.00	-0.48
H/W, own children only, oldest child $0-5$	-3.33	-2.90	-1.28	-0.83	0.15	-2.48	-2.22	-2.37	-0.77	-1.33	-2.15
H/W, own children only, oldest child $6-17$	-2.66	-2.63	-0.91	-1.91	0.32	-0.42	-0.70	-1.55	-1.50	-1.36	-1.73
H/W, own children only, oldest child over 18	-0.08	-2.56	1.58	-0.60	-0.37	0.92	-0.27	-0.02	-0.56	-1.53	-1.55
All other H/W households	0.52	-1.78	-1.42	-0.82	0.69	0.13	1.40	-0.48	-0.45	-1.54	-0.91
One parent (male) at least one child $0-18$	-0.52	1.83	-0.87	-0.91	1.79	-0.24	-0.05	0.42	-0.37	1.31	0.20
One parent (female) at least one child $0-18$	-2.13	0.57	0.21	0.39	1.58	0.19	-1.08	-0.23	-0.21	0.36	-2.31
Single persons	-1.03	0.25	0.18	0.73	1.32	2.46	1.47	0.46	1.88	2.40	1.36
Dummy for Black	1.23	2.91	2.28	1.62	3.98	2.31	3.38	-0.09	0.23	1.46	1.35
Dummy for American Indian	1.33	-0.31	-0.40	0.61	-0.25	0.98	-0.51	-0.23	1.98	-0.93	1.56
Dummy for Asian or Pacific Islander	1.84	0.20	-0.23	-0.35	-0.14	0.27	-0.24	-0.41	-1.94	0.08	-1.09
Less than High School	-1.89	0.29	-2.37	-2.57	-2.30	-1.49	-1.82	-1.21	3.92	-1.30	-2.25
High School	-3.77	1.42	-0.95	-1.33	-1.74	-1.38	-1.69	-1.97	1.04	-1.50	0.20
Post High School	-3.25	-1.53	-3.01	-2.60	-2.94	-3.36	-3.43	-3.68	0.69	-2.08	-1.26
Intercept	10.14	6.14	9.38	6.73	7.75	7.93	6.74	7.99	7.99	7.16	8.36

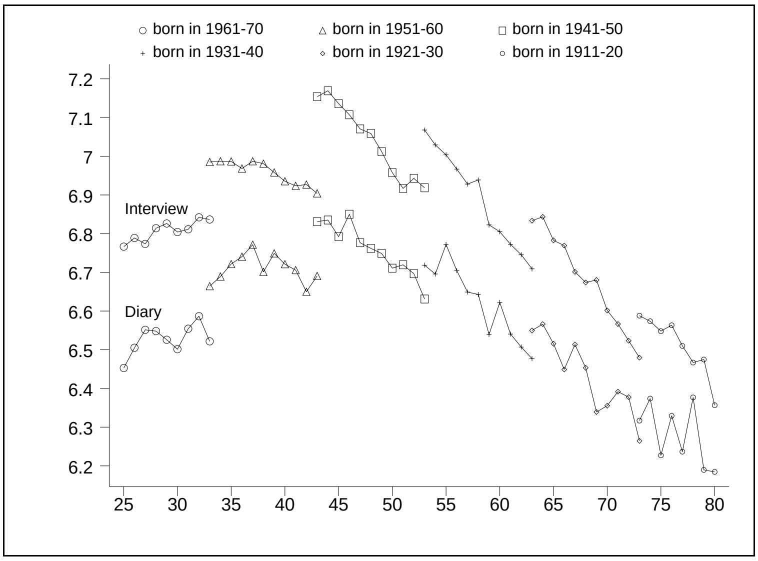

Figure 1: Age profiles of log non-durable expenditure by cohort

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 propensity score based weights derived from the regressions in Table 1 (see for example Battistin, Miniaci and Weber, 2001). 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).

Figure 1 reports age profiles of log non-durable consumption using the Consumer Price Index published monthly by the Bureau of Labor Statistics to equivalize expenditures for Diary and Interview data. Each data point represents the mean expenditure of the cohort in the generic year over the period $1988-1998$; the age of the head is the mid-age of the cohort[3]. Since the weighting scheme adopted reweighs 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.

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 remain positive for both the

[3]The youngest cohort enters the graph only in 1990 (when its mid-age is 25); the oldest cohort is followed until 1995 (when its mid-age is 80). Note also 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).

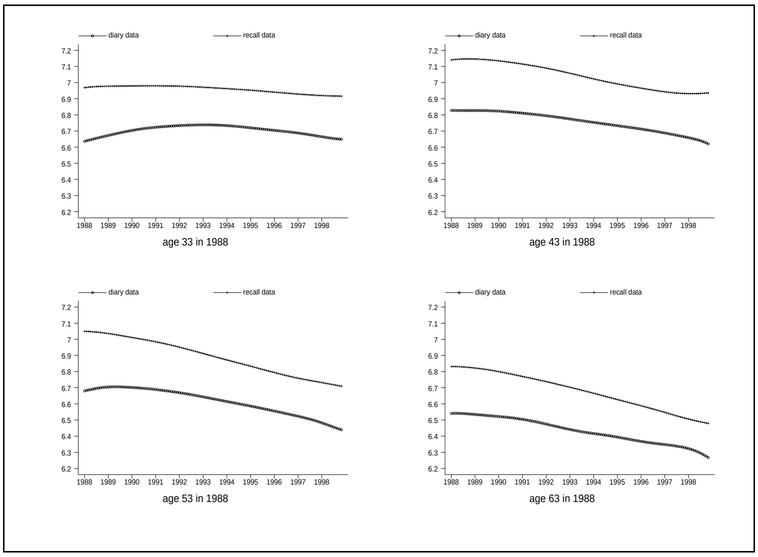

Figure 2: Mean of monthly log consumption by cohort (thinner lines refer to Interview data)

surveys but they look 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. The same graph for *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.

Figure 2 presents the same information contained in Figure 1 smoothing consumption profiles for the central cohorts; the first panel of Figure 4 reports the evolution of expenditure levels for the full sample over the ten years of data. Figure 3 and the first panel of Figure 5 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 (considering *per capita* income doesn't modify this pattern). It turns out that Diary data lead to considerably higher values of saving rates over the life cycle.

3.2 Aggregate inequality

In this section we discuss observed differences across the two samples with respect to the dispersion of non-durable expenditure over time. 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.

According to the well known Permanent Income Hypothesis (see Deaton, 1992) the variance of consumption within a fixed-membership group of people

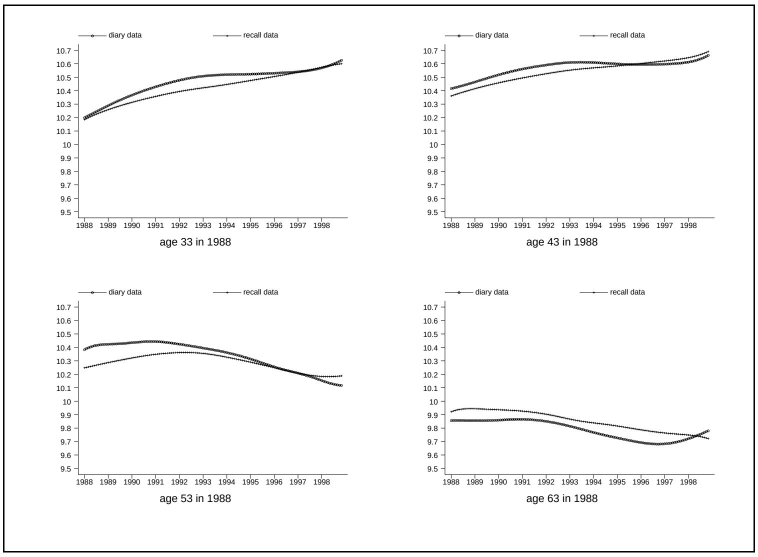

Figure 3: Mean of monthly log income by cohort (thinner lines refer to Interview data)

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 income[4].

Changes in the age structure of the population and their impact on the non-durable distribution need to be examined systematically before interpreting the observed trend in overall inequality. If the reference population changes significantly over time, these differences feed into many aspects of applied microeconomic and policy analysis. The Permanent Income model does not necessarily imply that aggregate inequality increases over time since age composition effects can compensate for within group inequalities. Inequality is greater among older cohorts and less among the young, and since young people are continually replacing the old, there is no automatic presumption that overall dispersion should increase. According to this model, if the reference population is ageing over time the aggregate inequality is more likely to increase because young people are not replacing the old.

The second panel in Figure 4 shows the pattern of aggregate inequality both for diary and recall data exploiting all households in each survey year. 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 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). The second panel in Figure 5 reports

[4]However, 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.

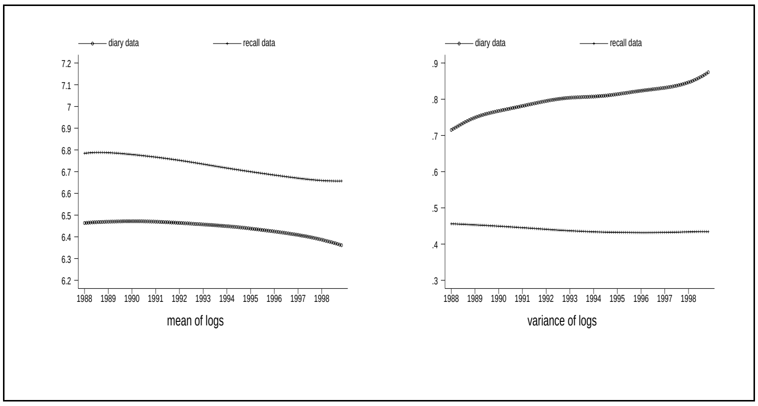

Figure 4: Mean and variance of monthly log consumption (thinner lines refer to Interview data)

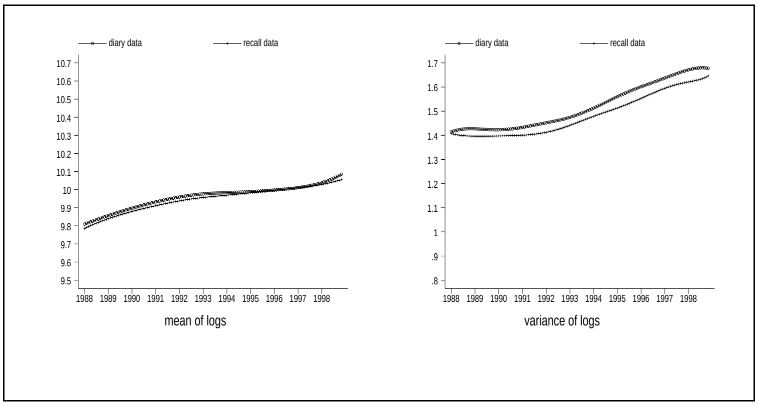

Figure 5: Mean and variance of monthly log income (thinner lines refer to Interview data)

the same inequality measure referred to total disposable income.

Available data are used to investigate whether the time series properties of consumption are consistent with the Permanent Income (life-cycle) Hypotheses. Note that the weighting scheme adopted leads the distribution of income and household composition to be the same across the two samples over time; both these variables could appreciably affect the shape of age inequality profiles because of an increasing dispersion of household size (and, as a consequence, available family income) as cohort ages[5].

The information contained in each sample leads to contradictory results with respect to the inequality pattern: Diary inequality increases over time while the recall counterpart remains pretty flat over the entire period. We checked the

[5] It might be interesting considering how much robust our results are to variations in head's age definition (Deaton and Paxon, 2000, and Miniaci, Monfardini and Weber, 2001). However, there is not particular reason to believe that any bias arising from such problem affects the two instruments in a different way and/or with a different sign.

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. The class of indexes we considered is given by

$$
\begin{aligned}
GE_0 &= \frac{1}{n}\sum_{i=1}^{n} ln(\frac{\mu_y}{y_i}), \\
GE_1 &= \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{\mu_y} ln(\frac{y_i}{\mu_y}), \\
GE_c &= \frac{1}{n}\frac{1}{c(c-1)}\sum_{i=1}^{n}[(\frac{y_i}{\mu_y})^c - 1], \qquad c \neq 0, 1;
\end{aligned}
$$

where the parameter c reflects different perceptions of inequality, with lower values indicating a higher degree of inequality aversion. The more positive c is, the more sensitive the corresponding index is to differences at the top of the distribution.

Results are reported in Table 2 where the first, the third and the fifth columns are referred to the Theil mean log deviation measure (GE_0), the Theil coefficient (GE_1) and half the squared coefficient of variation (GE_2), respectively. According to the Interview sample, it follows that the rising income inequality presented in the second panel of Figure 5 is not accompanied by a corresponding rise in consumption inequality. The result we find about an inequality fall (both for Diary and Recall) during the early 1990's - probably due to the strong business cycle - has already been pointed out by Johnson and Shipp (1995).

We also analyzed how the reference population ages over the observed period. The proportion of households whose head is aged less than 30 decreases over time while the class $40-50$ presents a mild upward trend (this result is not documented here but it is available on request). According to the Permanent Income Hypothesis, this mild aging over time could spuriously induce an increasing pattern of the overall inequality. However it is hard to believe that this mild population ageing totally explains the differences in the overall inequality for the two samples.

Indeed, we decompose the inequality indices presented in Table 2 into within- and between-group components in a manner similar to the decomposition defined exploiting the classical analysis of variance. We disaggregate by age on the grounds that age acts as a proxy for most of the important structural changes over the considered period. This allows us to investigate how the observed patterns reflect differences in inequality due to changes in age structure over time. It follows that, besides comparing inequality levels within subgroups, we can evaluate the proportion of total inequality due to differences across subgroups.

The decomposition equations for the generalized entropy measures considered above are the following

$$
\begin{aligned}
GE_0 &= \sum_a \frac{n_a}{n} GE_{0,a} + \sum_a \frac{n_a}{n} ln(\frac{\mu_y}{\mu_{y,a}}), \\
GE_1 &= \sum_a \frac{n_a}{n}\frac{\mu_{y,a}}{\mu_y} GE_{1,a} + \sum_a \frac{n_a}{n}\frac{\mu_{y,a}}{\mu_y} ln(\frac{\mu_{y,a}}{\mu_y}), \\
GE_c &= \sum_a \frac{n_a}{n}(\frac{\mu_{y,a}}{\mu_y})^c GE_{c,a} + \frac{1}{c(c-1)}\sum_a \frac{n_a}{n}[(\frac{\mu_{y,a}}{\mu_y})^c - 1], \qquad c \neq 0, 1
\end{aligned}
$$

Table 2: Age decomposition of aggregate inequality

	GE_0		GE_1		GE_2	
Interview	total	between	total	between	total	between
1988	0.23030	0.02687	0.23759	0.02678	0.34269	0.02705
1989	0.22846	0.02713	0.23939	0.02716	0.36230	0.02753
1990	0.22773	0.02839	0.24057	0.02838	0.39213	0.02878
1991	0.22203	0.02572	0.22579	0.02462	0.30639	0.02581
1992	0.21911	0.02217	0.22365	0.02228	0.30295	0.02262
1993	0.21301	0.02476	0.21445	0.02465	0.28035	0.02481
1994	0.22053	0.02208	0.22944	0.02203	0.32949	0.02220
1995	0.21251	0.01995	0.21606	0.01986	0.28606	0.01994
1996	0.22047	0.01984	0.23629	0.01980	0.42703	0.01992
1997	0.21752	0.02221	0.22040	0.02211	0.29940	0.02221
1998	0.21392	0.02238	0.22182	0.02207	0.32006	0.02196
	GE_0		GE_1		GE_2	
Diary	total	between	total	between	total	between
1988	0.28006	0.01267	0.25061	0.01239	0.30271	0.01220
1989	0.29827	0.01314	0.29036	0.01283	0.47754	0.01259
1990	0.31796	0.01520	0.30609	0.01475	0.49305	0.01440
1991	0.29519	0.01112	0.27388	0.01086	0.37212	0.01068
1992	0.27825	0.01081	0.24570	0.01041	0.28903	0.01009
1993	0.31542	0.01041	0.28594	0.01017	0.38751	0.00999
1994	0.31502	0.01311	0.27627	0.01253	0.33589	0.01209
1995	0.32313	0.01229	0.28989	0.01182	0.39070	0.01145
1996	0.30915	0.01281	0.28113	0.01259	0.37906	0.01246
1997	0.34446	0.01242	0.34914	0.01209	0.76899	0.01184
1998	0.34354	0.01307	0.31746	0.01262	0.46389	0.01228

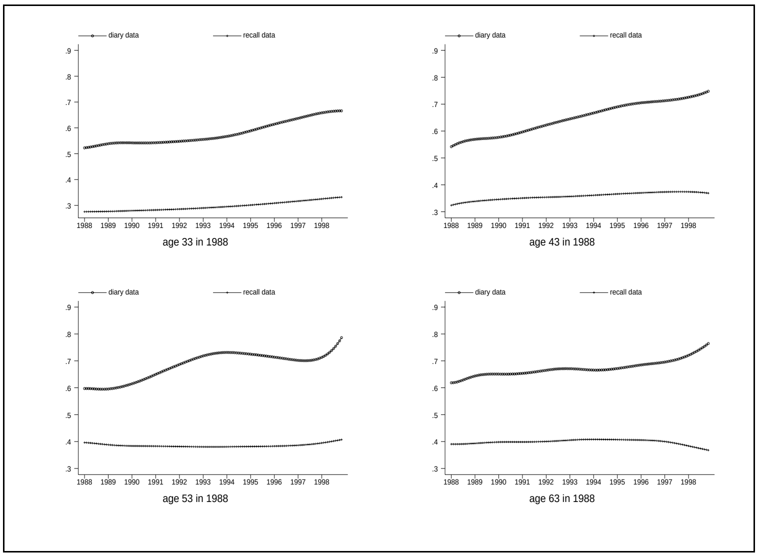

Figure 6: Variance of monthly log consumption by cohort (thinner lines refer to Interview data)

where the index a identifies the age group we considered (family head aged below 30, $31-40$, $41-50$, $51-60$, $61-65$ and over 65). The first term in these equations (the within-group component) is a simple weighted sum of the subgroup inequality values. The second term is the between-group component, reflecting the inequality contribution due solely to differences in the subgroup means (in a decomposition by age, this term corresponds to a pure 'age effect').

Table 2 reports the amount of total inequality explained by the between-group component (age) separately for years 1988-98, both for Diary and Interview, for the three inequality indexes considered. The proportion attributable to between-group differences presents a modest downward trend over time and it is always less then 9% (depending on the inequality measure exploited). It follows that vast majority of inequality in each sample is not attributable to age effects over time; the finding that most of the inequality is due to within-group rather than between-group variation is also documented by Johnson and Shipp (1995).

3.3 Within cohort inequality

To shed more light on the nature of the differences so far discussed, Figure 6 and Figure 7 presents how much the within-cohort distribution of expenditures disperses as cohort ages, where cohorts are still identified by year of birth of the family head[6]. Interview inequality is mainly flat over time for all the considered

[6]The theoretical behavior of inequality within cohorts depends on people's attitude toward risk and on the mechanisms that are available 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, and one of the main reasons for measuring consumption

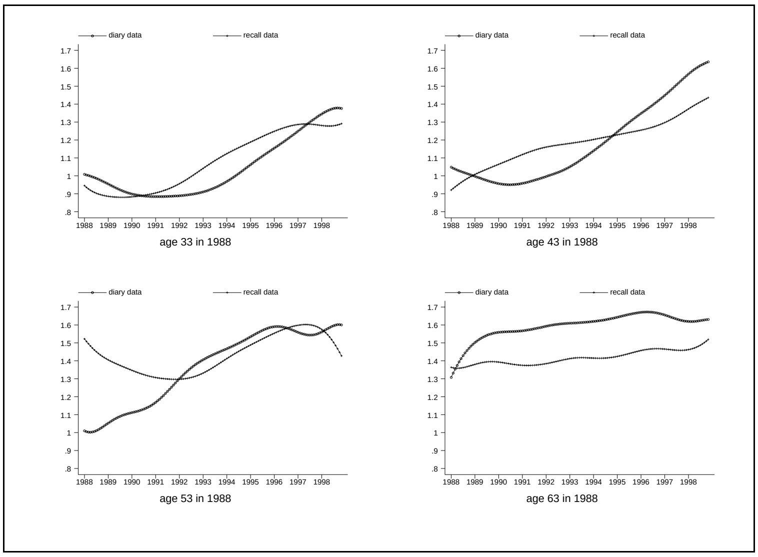

Figure 7: Variance of monthly log income by cohort (thinner lines refer to Interview data)

cohorts, with the exception of the ones referred to head aged 33 and 43 in 1988 where the increasing pattern is not well marked (the inequality values for January 1988, 1993 and 1998 are 0.27, 0.28 and 0.32 for the first cohort and 0.32, 0.35 and 0.37 for the second cohort, respectively). Inequality seems to be most pronounced exploiting Diary data for all groups. The bump-shaped pattern peaking before retirement for those who are 53 in 1988 could be explained by the effect of an increasing leisure time due to the retirement age[7].

The figure discussed above doesn't change if we consider inequality indexes for *per capita* expenditures; the relative constancy of Interview inequality within cohorts would represent an empirical evidence against the classical formulation of the Permanent Income Hypothesis.

4 Inequality decomposition by factors

In this section we seek to describe how marginal changes in expenditures for specific commodities can affect the inequality with respect to total non-durable expenditure. The components of non-durable expenditure we consider are the ones already analyzed by 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

inequality and its evolution is to help understand and calibrate the way in which economy handles risk (see Deaton and Paxon, 1984).

[7]The observed differences between income mean and variance in the two samples are not statistically different from zero (with the exception of years $1988-1990$ for the third cohort in Figure 7). The result follows by definition of the weighting scheme adopted which depends on binary regressions including total available income as an explanatory variable.

clothing and footwear). Health and education expenditures do not enter total expenditure on non-durables.

Table 3 and Table 4 looks into the relative magnitudes of mean expenditures based on the Diary and on the Interview sample, respectively. Year-to-year changes in the ratios of aggregate expenditures provides useful monitors of survey performance over time. We present reporting rates for the Diary and the Interview samples - that is the proportion of non-zero expenditures for a specific item (see Table 3) - and the ratio of spending means - that is conditional to a positive expenditure (see Table 4). By 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).

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 simple 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 3 presumably reflect only two of the three sources of zero 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 samples.

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 samples[8].

Table 4 presents the ratio of spending means in the Interview sample to spending means in the Diary sample, separately for each year and for each good. With the exception of expenditure on public and private transports, the relationship between mean levels for consumers in the two surveys remain constant over time. Figure 8 and Figure 9 reports - separately for each commodity - the analog of what presented for total non-durable expenditure in Section 3.2. To account for the frequency of purchasing problem (i.e. the proportion of non-consumers in each sample) which affects seriously certain goods, we consider the coefficient of variation rather than the variance of logs as a inequality measure. Indeed, the squared coefficient of variation is a first order approximation to the variance of logs and it is extensively used because of its invariance to

[8]Excluding 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 3.

Table 3: 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

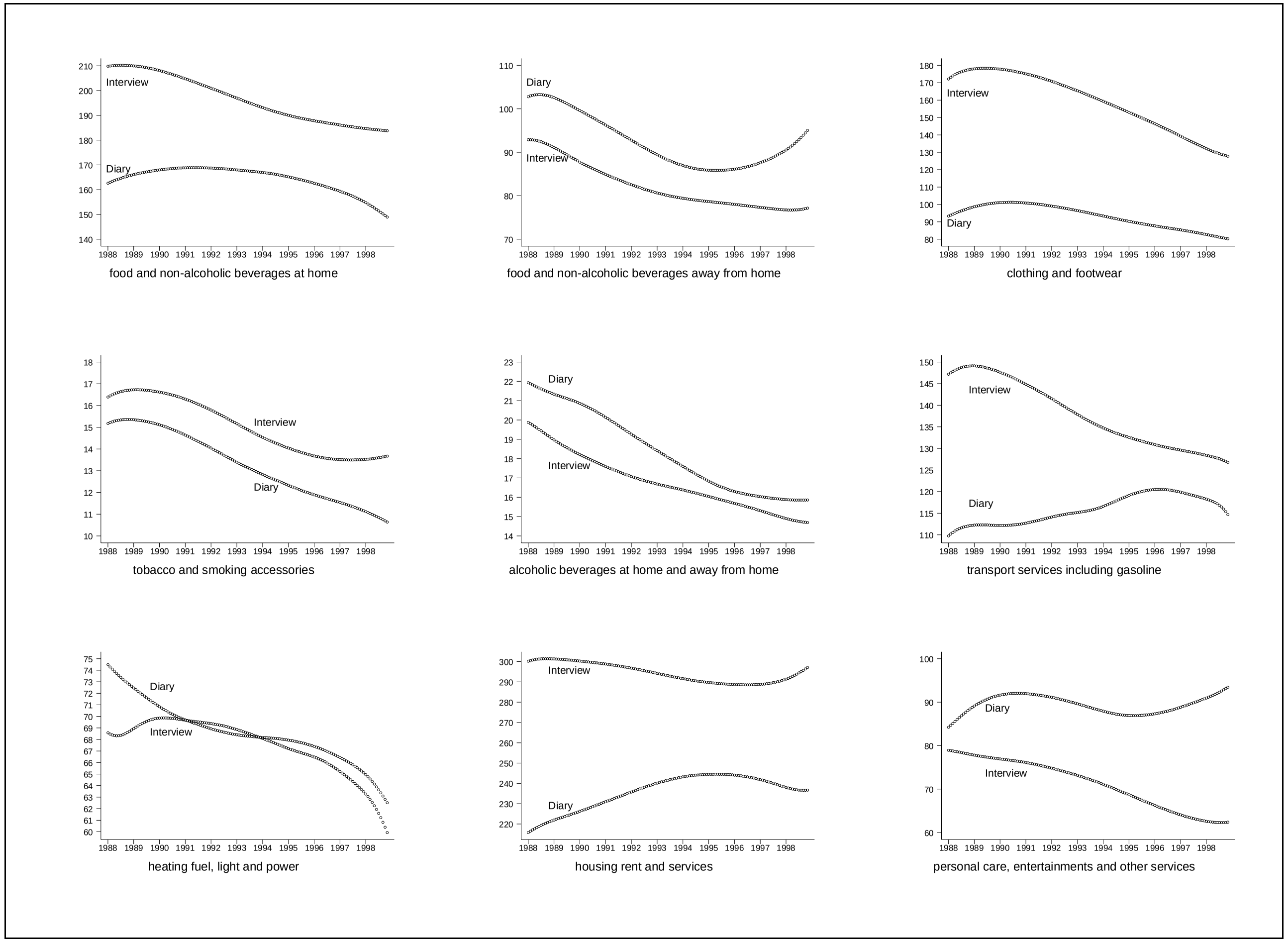

Figure 8: Mean of monthly log consumption by factors

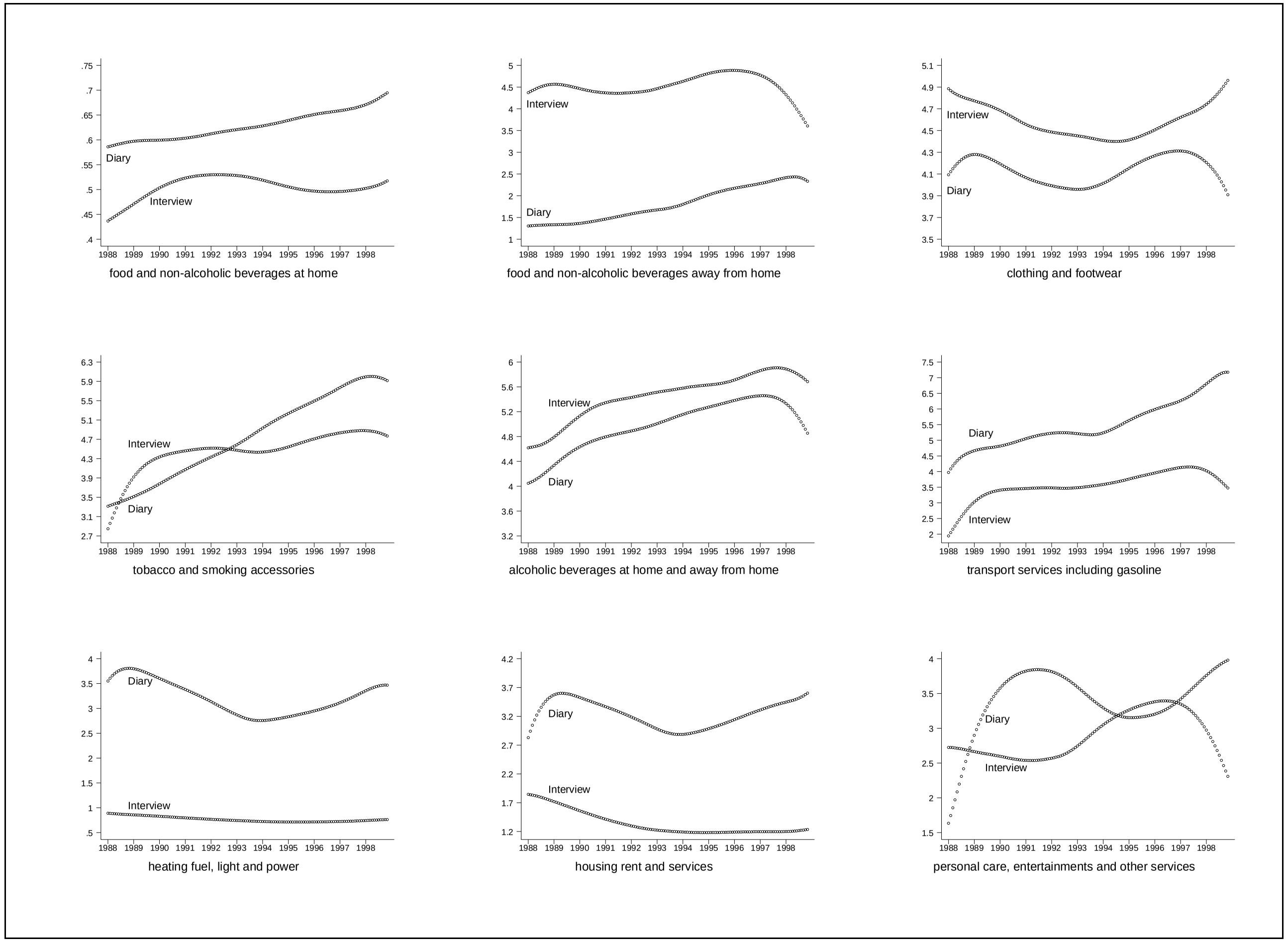

Figure 9: Coefficient of variation squared of monthly consumption by factors

Table 4: Spending mean ratios: $E(X|X > 0,\text{Interview})/E(X|X > 0,\text{Diary})$

Category	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	1.262	1.214	1.228	1.180	1.173	1.112	1.157	1.117	1.118	1.143	1.171
Food and non-alcoholic beverages away	0.966	0.905	0.919	0.913	0.908	0.914	0.937	0.962	0.935	0.902	0.849
Clothing and footwear	1.727	1.644	1.569	1.589	1.643	1.615	1.681	1.467	1.471	1.609	1.418
Tobacco and smoking accessories	1.093	1.062	1.138	1.119	1.104	1.112	1.118	1.078	1.098	1.140	1.129
Alcoholic beverages (at home and away)	0.812	0.835	0.794	0.777	0.789	0.857	0.891	0.842	0.867	0.806	0.872
Transport services (including gasoline)	1.324	1.209	1.292	1.251	1.138	1.018	1.023	1.000	1.079	0.895	1.105
Heating fuel, light and power	0.541	0.499	0.532	0.542	0.523	0.501	0.509	0.501	0.491	0.495	0.504
Housing (rent and services)	1.267	1.209	1.172	1.174	1.166	1.076	1.048	1.008	1.059	1.095	1.084
Personal care, entertainments and services	0.841	0.816	0.737	0.732	0.778	0.758	0.781	0.717	0.671	0.671	0.622

proportional changes in expenditures for all the units[9]. Note also that half the squared coefficient of variation is the GE_2 member of the Generalized Entropy class of inequality measures already presented in Table 2.

The relationship between mean expenditures varies a great deal considering different commodities. Thus, the phenomenon we observe about inequality measures in Section 3 would appear to be the aggregated outcome of a large number of positive and negative mean differences at the considered level. Table 5 presents the relationship between various expenditure budget components and total non-durable expenditure (the share of each category out of the total expenditure). The share of expenditure for food away from home purchases is greater for Diary data and represents (averaging over the considered period) around the 11% of non-durable expenditure, against the 8% exploiting Interview data. Not surprisingly, recall data weight much more clothing and housing expenditures; remaining categories look more comparable over time across the two samples. The increasing pattern for Diary variances presented in Figure 4 is probably determined by food at home, housing and transports components whose corresponding variances increase over time and whose weight in determining total expenditure is high.

Results in Figure 9 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 6 reports commodity categories determining total non-durable expenditure and their percentage contribution to total inequality using the 'natural' decomposition rule $Var(Y) = \sum_j Cov(X_j, Y)$, both for Diary and Interview data. The contribution of each commodity X_j to total inequality is then expressed as the slope coefficient of the Engel regression of X_j on non-durable expenditure Y. Alternative procedures based on decompositions of the Gini coefficient (see for example Garner, 1993) lead to the same result[10]. It turns out that the contribution of each commodity can also be written as

$$\frac{Cov(X_j, Y)}{Var(Y)} = \varrho_j \frac{E(X_j)}{E(Y)} [\frac{GE_2(X_j)}{GE_2(Y)}]^{1/2}, \qquad j = 1, \ldots, K$$

that is as the product of commodity correlations with total expenditure ϱ_j, the commodity shares already presented in Table 5 and how much each commodity inequality moves with respect to total inequality (i.e. the root of the ratio between what reported in Figure 9 to the last but one column in Table 2).

[9]Note that for the generic commodity X the s-th moment is defined as $E(X^s) = E(X^s|X > 0)Pr(X > 0)$, that is as the s-th moment of the spending distribution times the proportion of consumers in the sample. It follows that the means reported in Figure 8 are a smoothed version of the product of each numerator (denominator) in Table 4 and the probability of a positive Interview (Diary) spending in Table 3.

[10]However, 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 5: Factor shares as percentage of total non-durable expenditure: $\mathrm{E}(X_j)/\mathrm{E}(Y)$

Interview	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	18.55	18.75	18.88	19.12	18.83	18.46	18.63	19.21	19.12	19.04	19.24
Food and non-alcoholic beverages away	8.15	8.08	7.98	7.40	7.50	7.71	7.74	8.22	7.80	7.76	8.11
Clothing and footwear	16.81	16.37	16.03	16.09	15.93	15.60	15.63	13.67	14.13	14.23	12.27
Tobacco and smoking accessories	1.45	1.50	1.55	1.49	1.49	1.41	1.35	1.39	1.38	1.38	1.41
Alcoholic beverages (at home and away)	1.72	1.68	1.57	1.56	1.60	1.58	1.61	1.60	1.56	1.52	1.56
Transport services (including gasoline)	13.21	13.25	13.71	13.30	13.07	13.00	12.99	13.05	13.67	13.24	13.22
Heating fuel, light and power	6.40	6.38	6.30	6.50	6.26	6.58	6.75	6.64	6.96	6.85	6.69
Housing (rent and services)	26.75	26.96	27.14	27.48	28.09	28.38	28.54	29.27	28.94	29.54	30.81
Personal care, entertainments and services	6.90	6.97	6.81	7.01	7.18	7.22	6.73	6.92	6.39	6.39	6.63

Diary	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	18.51	18.50	18.49	18.82	18.72	19.01	18.35	18.68	18.71	18.20	17.85
Food and non-alcoholic beverages away	11.25	11.37	11.03	9.86	10.07	9.75	9.49	9.35	9.48	9.67	11.09
Clothing and footwear	11.70	11.48	11.51	11.85	11.27	10.56	10.15	9.94	10.48	9.57	9.71
Tobacco and smoking accessories	1.72	1.68	1.68	1.56	1.54	1.38	1.40	1.33	1.40	1.29	1.30
Alcoholic beverages (at home and away)	2.52	2.30	2.26	2.19	2.24	1.90	1.87	1.80	1.86	1.80	1.88
Transport services (including gasoline)	12.27	12.71	12.27	11.55	12.54	13.16	13.56	13.31	13.20	15.25	12.43
Heating fuel, light and power	7.76	7.55	7.23	8.57	7.34	7.38	7.81	7.33	7.82	7.64	6.99
Housing (rent and services)	24.30	24.58	24.99	25.13	26.15	26.96	28.10	28.47	27.07	26.69	27.77
Personal care, entertainments and services	9.92	9.79	10.50	10.43	10.09	9.86	9.24	9.74	9.95	9.84	10.94

Table 6: Factors contribution as percentage of total inequality

Interview	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	6.89	7.26	6.29	8.86	8.61	8.58	7.39	8.17	5.38	8.17	7.68
Food and non-alcoholic beverages away	11.09	10.94	12.30	9.86	9.93	11.62	10.32	12.84	11.93	11.58	11.57
Clothing and footwear	29.89	26.28	22.81	28.95	29.79	25.11	26.87	25.79	17.05	24.58	20.12
Tobacco and smoking accessories	0.31	0.29	0.34	0.40	0.42	0.31	0.29	0.34	0.21	0.45	0.31
Alcoholic beverages (at home and away)	1.24	1.37	1.14	1.56	1.51	1.44	1.42	1.53	1.79	1.58	1.53
Transport services (including gasoline)	14.12	15.71	26.53	16.13	15.85	17.03	18.56	16.37	31.52	18.08	20.23
Heating fuel, light and power	2.35	2.20	2.13	2.52	2.21	2.50	2.16	2.38	1.50	2.11	1.99
Housing (rent and services)	27.01	27.30	21.75	23.41	23.56	25.18	24.43	23.88	18.65	25.49	28.41
Personal care, entertainments and services	7.11	8.64	6.71	8.31	8.12	8.23	8.56	8.71	11.96	7.95	8.16

Diary	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	7.86	5.18	5.49	8.00	7.68	7.52	7.68	7.57	9.13	4.09	6.83
Food and non-alcoholic beverages away	8.00	5.37	5.80	6.06	7.42	6.41	7.78	5.50	5.78	5.98	6.94
Clothing and footwear	15.97	9.06	14.79	16.09	13.50	10.51	12.65	10.48	21.70	5.66	9.23
Tobacco and smoking accessories	0.44	0.29	0.25	0.38	0.36	0.45	0.35	0.31	0.48	0.20	0.40
Alcoholic beverages (at home and away)	1.75	1.05	1.21	1.69	2.05	1.01	1.53	1.52	1.45	0.69	1.12
Transport services (including gasoline)	14.77	41.81	10.82	15.70	20.25	22.16	20.39	18.18	15.18	56.78	13.69
Heating fuel, light and power	7.37	3.86	4.97	6.90	5.33	4.57	5.58	4.64	5.02	3.04	4.66
Housing (rent and services)	32.79	24.90	43.05	30.38	30.60	34.30	35.22	42.49	29.90	18.01	43.49
Personal care, entertainments and services	11.04	8.46	13.62	14.79	12.82	13.07	8.83	9.31	11.36	5.56	13.64

Table 7: Factor correlation with total non-durable expenditure: ϱ_j

Interview	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	0.454	0.448	0.416	0.486	0.459	0.469	0.452	0.455	0.374	0.474	0.441
Food and non-alcoholic beverages away	0.522	0.564	0.528	0.540	0.529	0.526	0.510	0.532	0.528	0.513	0.594
Clothing and footwear	0.671	0.630	0.601	0.667	0.669	0.630	0.625	0.579	0.531	0.622	0.568
Tobacco and smoking accessories	0.095	0.074	0.081	0.094	0.105	0.086	0.086	0.085	0.067	0.104	0.082
Alcoholic beverages (at home and away)	0.296	0.324	0.305	0.327	0.337	0.308	0.328	0.311	0.419	0.339	0.347
Transport services (including gasoline)	0.557	0.610	0.658	0.573	0.558	0.574	0.598	0.581	0.741	0.608	0.608
Heating fuel, light and power	0.305	0.333	0.321	0.338	0.313	0.344	0.295	0.307	0.236	0.273	0.278
Housing (rent and services)	0.620	0.645	0.585	0.597	0.595	0.603	0.610	0.601	0.558	0.619	0.655
Personal care, entertainments and services	0.577	0.581	0.559	0.589	0.570	0.585	0.535	0.534	0.707	0.580	0.586

Diary	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	0.467	0.375	0.421	0.511	0.433	0.476	0.470	0.462	0.533	0.376	0.470
Food and non-alcoholic beverages away	0.520	0.409	0.473	0.478	0.475	0.415	0.437	0.435	0.478	0.357	0.474
Clothing and footwear	0.561	0.451	0.570	0.569	0.515	0.468	0.525	0.499	0.606	0.375	0.485
Tobacco and smoking accessories	0.136	0.104	0.107	0.130	0.113	0.161	0.113	0.137	0.118	0.100	0.130
Alcoholic beverages (at home and away)	0.291	0.228	0.272	0.281	0.264	0.242	0.298	0.296	0.265	0.192	0.259
Transport services (including gasoline)	0.511	0.581	0.467	0.529	0.544	0.532	0.528	0.515	0.488	0.753	0.495
Heating fuel, light and power	0.411	0.474	0.349	0.418	0.399	0.364	0.363	0.376	0.350	0.270	0.382
Housing (rent and services)	0.659	0.661	0.713	0.691	0.668	0.687	0.699	0.727	0.671	0.516	0.761
Personal care, entertainments and services	0.567	0.453	0.548	0.563	0.549	0.506	0.513	0.484	0.542	0.417	0.585

Table 8: Factor inequalities: $GE_2(X_j)/GE_2(Y)$

Interview	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	0.631	0.720	0.609	0.891	0.962	0.966	0.734	0.896	0.555	0.818	0.816
Food and non-alcoholic beverages away	6.943	5.683	7.903	6.128	6.316	8.464	6.923	10.65	8.433	8.421	5.752
Clothing and footwear	7.045	6.508	5.572	7.296	8.109	6.605	7.127	7.527	5.320	7.690	8.325
Tobacco and smoking accessories	4.472	6.383	7.108	7.747	6.498	7.361	6.735	8.198	5.564	9.592	6.846
Alcoholic beverages (at home and away)	5.895	6.348	5.811	9.479	7.785	8.852	7.440	9.760	7.239	9.313	7.892
Transport services (including gasoline)	3.635	3.758	9.230	4.498	4.663	5.309	5.885	5.345	9.590	5.040	6.324
Heating fuel, light and power	1.417	1.064	1.073	1.291	1.253	1.249	1.177	1.205	0.786	1.271	1.142
Housing (rent and services)	2.648	2.469	1.824	2.008	1.908	2.041	2.032	1.967	1.307	1.938	1.977
Personal care, entertainments and services	3.190	4.645	3.138	4.090	3.948	3.905	5.763	6.347	7.035	4.594	4.403

Diary	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food and non-alcoholic beverages at home	0.901	0.577	0.599	0.764	0.789	0.795	0.853	0.769	0.774	0.452	0.719
Food and non-alcoholic beverages away	2.005	1.234	1.373	1.767	2.295	2.687	3.044	1.954	1.596	3.351	1.898
Clothing and footwear	5.761	4.035	5.293	5.136	5.313	4.574	5.437	4.618	9.629	2.787	4.301
Tobacco and smoking accessories	4.867	3.803	3.496	5.354	6.351	5.988	6.512	5.594	9.615	3.439	6.210
Alcoholic beverages (at home and away)	7.085	5.211	4.392	7.807	11.27	6.529	6.674	6.765	7.813	4.511	5.654
Transport services (including gasoline)	5.398	15.05	3.770	5.929	7.591	8.970	6.647	7.162	4.656	20.50	4.686
Heating fuel, light and power	4.886	4.904	3.813	3.722	3.927	3.512	3.599	2.910	3.486	2.506	3.119
Housing (rent and services)	4.211	4.165	5.314	3.377	3.636	3.511	3.655	4.197	3.501	2.128	4.101
Personal care, entertainments and services	3.610	3.051	5.813	5.692	4.496	5.963	3.635	3.391	3.979	2.618	4.347

Table 7 and Table 8 reports coefficients ϱ_j and commodity inequalities $GE_2(X_j)/GE_2(Y)$, respectively. The contribution of food, tobacco, transports and housing expenditures looks constant over time but with different levels for each survey. The weight of expenditures on alcoholic beverages (after 1993) and on services (after 1991) increases for Interview data and decreases for Diary data; clothing's contribution is slightly decreasing for both the sources as well as the one of fuel (but more for Diary data).

5 Modelling inaccuracies

The goal of this section is to derive error-adjusted measurements of the inequality indices presented in Section 3. Exploiting jointly the two surveys we (i) discuss on the quality characterizing each commodity component of non-durable expenditure and (ii) define an improved measure of overall consumption. As a side result, step (i) enables us to study whether the distribution of reporting errors for certain commodities changes over time and across waves (i.e with the interview number).

A first possible explanation for the evidence we found in Section 3 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 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 Survey[11].

Let r and d 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

[11]In 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 (for the last point see Banks and Blow, 2001).

problem arises from the fact that - by design - we do observe only one of these measurements on each household.

In what follows we proceed as if we knew the counterfactual measurement for each household, that is what recall (diary) households would have reported had they been interviewed using diary (recall) based questions. We will discuss on the estimation of this counterfactual in Section 6. Under this assumption r and d are two measurements of the same latent variable of interest (non-durable expenditure) for each household.

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 3 about differences in means and variances across the two surveys suggests that the error processes leading to r and d 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 reporting[12].

A possible approach is to represent the association between r and d by a suitable model that embodies latent expenditures in such a way that r is independent of d given the true value of consumption (the resulting analysis is known as *factor analysis*; see for example Kim and Mueller, 1978). If the latent expenditure and the measurement errors affecting each of the two surveys were mutually independent and we observed both r and d on the same household, the joint distribution (r, d) would uniquely determine the density functions (i) of the 'true' unobserved non-durable expenditure and (ii) of the measurement errors in r and d (for example, under the additional assumption of their symmetry about the origin)[13]. 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

[12]To 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).

[13]The 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.

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 for recall based 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), 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.

5.1 Identifying 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 (r, d) 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 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 Melemberg, 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-)monthly expenditure derived from diaries (as the sum of detailed food data) and (log-)monthly '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 weeks) diary period and its accuracy is therefore not influenced by how respondents learn about own expenditures during the interview[14].

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 ranging from -0.3493 to -0.0896 (with mean -0.1988 and median -0.1969) and generally decreasing in absolute value over time[15]. There is a mild effect of the interview month on the magnitude of under-reporting, since households interviewed in December usually present values closer to zero (mean -0.1459 and median -0.0435). This evidence supports the idea that recall questions overestimate the real spending on food-related items probably because in their reporting households include more goods. 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 9 summarizes the discussion of this section.

[14]We have already discussed the quality of information characterizing this recall-based question commenting on Figure 3.2. Note also 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 source of information will be used for future research.

[15]This 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.

Table 9: 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

6 Diary counterfactuals

To summarize the contents of Section 5, 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 not constant 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.

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. Without any loss of generality in this section we exploit the additional Diary information to approximate - even if the samples are different - counterfactual measurements in the Interview survey. The same definitions given below can be applied to obtain Diary counterfactuals using the Interview information.

6.1 Definition

Suppose that for the h-th household in the Interview Survey the (unobserved) counterfactual measurement on the generic good depends additively (i) on an unknown function g of observable characteristics z common to all households and (ii) on a factor e representing unobserved heterogeneity with respect to spending habits

$$d_{hk} = g_k(z_h) + e_{hk}, \qquad k = 1, \ldots, K \tag{1}$$

where k is the generic good[16]. The last expression implies that households with fixed characteristics z present same level of consumption $g_k(z)$ up to individual

[16] It is worth stressing again that the notation (r, d) is referred to the potential outcomes as the result of being retrospectively interviewed or compiling a diary about own expenditures. Therefore only one of these two variables is observed for each household; since in this section we work conditional on the Interview sample, the Diary counterfactual in (1) is not observed.

preferences e. Assume that e is a zero-mean term independent and identically distributed across units and independent of z, so that

$$E(d_k|z_h) = g_k(z_h), \qquad k = 1, \ldots, K.$$

Let $S(z_h)$ be the set of households in the Diary Survey presenting the same characteristics z_h as the ones of the h-th household in the Interview Survey. Clearly $g_k(z_h)$ can be identified averaging reported expenditures for households in the Diary Survey belonging to $S(z_h)$. Expression (1) then implies that the average expenditure over the set of households in the Diary Survey presenting z_h characteristics approximates - up to a (zero mean) unobserved term e_h - the counterfactual measurement for household h in the Interview Survey.

Since for each household in $S(z_h)$ expression (1) holds, the following relationship

$$e_{jk} = d_{jk} - E(d_k|z_h) \qquad j \in S(z_h), \quad k = 1, \ldots, K \tag{2}$$

suggests that exploiting the variability in $S(z)$ we can identify the entire distribution of the random variable e from the Diary sample.

6.2 Identification

To define $S(z_h)$ we matched the h-th Interview household to the set of Diary households in the same income decile presenting same composition (defined by the set of dummy variables considered in Table 1) and whose propensity score was within a 5% distance from the one estimated for household h.

To account for preferences due to price variation over time we matched conditional on the expenditure month, so that for any fixed household h the set $S(z_h)$ depends on the interview number. For Interview households presenting $S(z_h)$ with poor sample size (less than 10 units in nearly 35% of cases) we weakened the condition of exact matching on income decile and consider also households belonging to the closest upper and lower deciles. After this correction the overall proportion of households in the Interview survey presenting less than 10 'twins' units in the Diary survey was nearly 4% and was discarded from the analysis.

6.3 Measurement errors

It follows from the discussion in Section 6.1 that variability in $d|z$ is uniquely determined by unobserved heterogeneity in expenditures habits; in the same way, variability in $r|z$ is due to unobserved heterogeneity plus measurement error in reporting expenditures.

Indeed let u be the difference between reported r and actual spending d (the measurement error we are interested in). Exploiting expression (1) it follows that the difference between r and $g(z)$ - the best prediction of its counterfactual based on characteristics z - identifies a mixture of u and unobserved heterogeneity e

$$r_{hk} - g_k(z_h) = u_{hk} + e_{hk}. \qquad k = 1, \ldots, K. \tag{3}$$

We therefore observe realizations of a random variable whose distribution is given by the sum of the variable of interest u and the variable e. Intuitively since the latter distribution is identified from Diary information exploiting (2), we could 'subtract' it from the distribution of $r-g(z)$ to identify the distribution of u.

The deconvolution of u is feasible if we assume the independence of the reporting error from e, that is if unobserved factors determining household preferences in the amount of spending are independent of factors (different from z) affecting the quality of reporting using recall questions[17]. Under this assumption the entire distribution of u is non-parametrically identified.

7 Estimating error-corrected expenditures

In this and in the following section we will use Diary as the benchmark survey to correct expenditures in the Interview survey, both because the former is widely considered as the best large-scale cross-sectional dataset on consumption for the US economy and because this allows us to assess the robustness of our results with respect to the empirical findings in related works (see amongst others Johnson and Shipp, 1995, and Krueger and Perri, 2001).

It follows from the discussion of the previous section that we can identify the entire distribution of expenditures conditional on each category entering total non-durable expenditure. Indeed, while such distribution is identified exploiting available information for those categories well-reported by means of Interview data, for the remaining categories it is determined by the convolution of estimated expenditures $g(z)$ and preferences e exploiting Diary information[18].

Let the total expenditure on non-durables be defined as

$$Y = \sum_{k \in \mathcal{I}} X_k + \sum_{k \in \mathcal{D}} X_k,$$

where the two sets of goods over which we are aggregating are those already discussed in Table 9. The overall consumption measure in the Interview sample is likely to be error-affected because of the contribution of goods which are presumably more accurate when collected using diaries. The share of this set of goods out of the total budget is on average around 35% according to what presented in Table 5. An improved measure of total non-durable consumption could therefore be obtained exploiting Diary counterfactuals in such a way that

$$X_k = \begin{cases} r_k & \text{if } k \in \mathcal{I} \\ d_k & \text{if } k \in \mathcal{D} \end{cases}.$$

Since we do not observe such counterfactuals because Interview and Diary samples are referred to different households, we proceed estimating them from

[17]We are aware that this is obviously not always the case: as an example we can take the shadow price of leisure, which clearly affects preferences in buying habits and, at the same time, is likely to be related to the quality of reporting. Another (extreme) example is the case of an alcoholic, whose propensity to alcohol affects his amount of spending on non-durables and - probably - the quality of reporting.

[18]For food and non-alcoholic beverages (at home and away), alcoholic beverages and services (the first four commodities in Table 9) such distribution is given by

$$f_d(t) = \int f_{g(z)}(t-e)dF_e.$$

what derived in Section 6. Since

$$E[X_k] = \begin{cases} E[r_k] & \text{if } k \in \mathcal{I} \\ E[g_k(z)] & \text{if } k \in \mathcal{D} \end{cases},$$

$$Var[X_k] = \begin{cases} Var[r_k] & \text{if } k \in \mathcal{I} \\ Var[g_k(z)] + Var[e_k] & \text{if } k \in \mathcal{D} \end{cases},$$

the sample counterparts of the previous quantities yields a consistent estimator of the first moment and of the variance of each component entering non-durable expenditure.

Note however that while the identification of each commodity mean leads to identify the mean of total non-durable expenditure, the knowledge of each commodity variance is not enough to identify the overall variance. Indeed, the overall expenditure on non-durables can be expressed as

$$Y = [\sum_{k \in \mathcal{I}} r_k + \sum_{k \in \mathcal{D}} g_k(z)] + \sum_{k \in \mathcal{D}} e_k,$$

so that we can identify Y up to an unobserved component given by the aggregate contribution of preferences referred to set $\mathcal{D}$. It turns out from definition (1) that such contribution has zero-mean but unknown variance, since it depends on a term describing the correlation amongst counterfactual preferences

$$\sum_{i \in \mathcal{D}} \sum_{k \in \mathcal{D}} Cov[e_i, e_k]. \tag{4}$$

Moreover, such component is also correlated with the identifiable component of Y because of the correlation between preferences referred to observed and counterfactual measurements which goes through r

$$\sum_{i \in \mathcal{D}} \sum_{k \in \mathcal{I}} Cov[e_i, e_k]. \tag{5}$$

In presenting our estimation results we will assume

Condition 3 $\sum_{i \in \mathcal{D}} \sum_{k \in \mathcal{D}} Cov[e_i, e_k] = 0,$

Condition 4 $\sum_{i \in \mathcal{D}} \sum_{k \in \mathcal{I}} Cov[e_i, e_k] = 0,$

that is we will assume the aggregate contribution of covariances in (4) and (5) to be zero[19].

8 Estimating measurement errors

In this section we present the parametric structure we impose to the process determining observed expenditures in the Interview Survey for those categories better described using diaries. We characterize the measurement error u in Section 6.3 by means of a scale-position transformation of 'true' expenditures to

[19]To derive means of logs from means of levels we exploit the following second-order approximation

$$E(lnX) = lnE(X) - \frac{CV(X)^2}{2}.$$

account for the pattern of means and variances documented by Figure 8 and Figure 9. Our model accounts for the possibility that the coefficients characterizing the reporting error may be heterogeneous across households and provides estimates of these coefficients at different points in time (corresponding to significative changes in the survey instruments over the years) and conditional on different waves (to test for the time-in-sample effect).

It is worth stressing that the purpose of our exercise is merely describing parsimoniously the potential sources of measurement error affecting goods in the set $\mathcal{D}$; since the parametric structure implied by our model is not rejected by our data (see next section) we maintained such a specification even if under the assumptions discussed in Section 6.3 u is non-parametrically identified.

By means of the assumption we made in Section 5.1, the value d represents the 'true' expenditure each household would have reported had it compiled a diary; once again we proceed as if this counterfactual measurement were observed for each household. We assume that reported expenditures are a scale-position transformation of real expenditures, that is observed values r represent a certain share of d plus an additional error (for simplicity we omit the indexes h and k referred to Interview households and goods, respectively)

$$r = \beta d + \varepsilon. \tag{6}$$

Clearly the previous expression imposes a rule in the way expenditures are reported, implying that each household remembers a fraction β of an error affected measure of the real expenditure d. The error structure is completely characterized by the pair (β, ε); in this context the scale parameter captures the correlation between measurement errors $(r - d)$ - i.e. observed minus counterfactual values - and real expenditures d.

Presumably responses of households to different levels of expenditure may be heterogeneous and depend, for instance, on demographic factors that cannot always be observed. To add flexibility to our model we therefore let the coefficient β be subject to random variation and be heterogeneous across households. We then assume that the random variables (β, d, ε) are mutually independent with β and ε independent and identically distributed across units. The multiplicative effect of the random variable β modifies the distribution of d allowing for a different pattern in the variances of observed and real expenditures over time (as we argued from Figure 9).

The relationship in (6) together with the independence assumption amongst the involved variables allows us to define a recursive strategy to estimate simultaneously the moments of β and ε regressing (powers of) r on (powers of) d. Indeed, since

$$E(r^s|d) = \sum_{l=0}^{s} \binom{s}{l} E(\beta^{s-l}) d^{s-l} E(\varepsilon^s), \qquad s \geq 1 \tag{7}$$

the linear regression of the s-th power of r on d identifies the moments of β and ε up to the s-th order; these estimates are plugged in the regression referred to the $(s+1)$-th power of r to identify the $(s+1)$-th moments of β and ε. The root-n consistency and asymptotic normality of the resulting estimates follows straightforwardly under the usual regularity conditions on the existence of the estimated moments; moreover, Beran and Hall (1992) prove the uniform convergence of these estimates to their true values.

Once the moments of β and ε have been estimated we can derive the corresponding distributions in a variety of ways, for example exploiting an orthogonal series - Legendre polynomial expansion or a partial Fourier inversion (see the Appendix). We account for the time-in-sample effect estimating these distributions separately for each wave. This enables us to test whether households report their consumption with the same accuracy over the interview period (one year).

Since the presence of any time-in-sample effect would explain the pattern of means and variances only within the period covered by the interview (one year), we allow for structural breaks referred to changes in the survey instruments over the years estimating separately model (6) for three periods (1988-1991, 1992-1995, 1996-1998). This allows us to check the stationarity of the measurement error process over time for goods in $\mathcal{D}$.

Since we do not identify directly the counterfactual measurement d, instead of (7) we consider the regression function(s) of (powers of) r on the observed variable $g(z)$

$$E[r^s|g(z)] = \sum_{l=0}^{s} \binom{s}{l} E[\beta^{s-l} d^{s-l} \varepsilon^s | g(z)], \qquad s \geq 1$$

which by means of equation (1) clearly depends on powers of e. The interpretation of the previous expression follows straightforwardly: we replace the unobserved value d with its best prediction based on what we actually observe, $g(z)$, introducing in each regression moments of the variable e that are 'exogenously' identified using (2).

Under the following conditions

Condition 5 $\{\beta, \varepsilon, g(z)\}$ are mutually independent,

Condition 6 $\{\beta, \varepsilon\} \perp e|g(z)$,

we can apply the same recursive procedure described above to obtain (β, ε). Condition 6 formalizes the idea that factors different from z determining the amount of spending must not affect the measurement error process in reporting expenditure levels (thus what we have already discussed at the end of Section 6.3).

9 Results

We present estimates of the measurement error process for goods in set $\mathcal{D}$ as the result of the characterization in terms of (β, ε) discussed in Section 8. Note that if we subtract d from both sides of (6) we have a relation explaining the reporting error $r - d$ as a function of the true consumption level d: this is an increasing function if $\beta > 1$, a decreasing function if $\beta < 1$. The classical measurement error assumption obtains when $\beta = 1$.

Under the assumptions stated by Conditions 1-2 and Conditions 5-6 above, all moments of β and ε can be identified from the available information imposing the parametric structure (6). Table 10 reports the first two kumulants of $f(\beta)$ and $f(\varepsilon)$ together with their associated coefficient of skewness controlling for interview number (i.e. the time-in-sample effect) and for time effects. Negative

Table 10: Measurement error profiles by interview and time

	distribution of β			distribution of ε		
Food at home	**Skewness**	**Mean**	**Std. Dev.**	**Skewness**	**Mean**	**Std. Dev.**
1st interview	5.5166	0.8985	0.4230	5.3605	49.365	102.23
2nd interview	3.2517	0.9038	0.3975	5.4637	48.111	101.60
3rd interview	1.9529	0.9157	0.4130	6.6653	47.498	106.02
4th interview	7.2440	0.9190	0.5018	3.9565	46.351	92.713
1988-1991	4.3565	0.9552	0.4288	5.4464	49.394	107.42
1992-1995	5.6370	0.8836	0.4363	5.4277	46.899	99.278
1996-1998	6.9881	0.8601	0.3938	5.0291	50.604	99.518
Food away	**Skewness**	**Mean**	**Std. Dev.**	**Skewness**	**Mean**	**Std. Dev.**
1st interview	4.9239	0.7765	1.2447	70.963	14.217	144.20
2nd interview	6.8624	0.7544	1.2728	56.541	13.548	107.30
3rd interview	6.2974	0.7116	1.0351	38.692	15.927	126.92
4th interview	9.6312	0.7678	1.4327	120.33	10.744	104.43
1988-1991	12.642	0.8224	1.5478	127.20	7.3191	76.190
1992-1995	7.1302	0.7330	1.0458	39.683	17.626	137.76
1996-1998	5.2793	0.6877	1.1235	77.329	16.168	132.89
Alcohol	**Skewness**	**Mean**	**Std. Dev.**	**Skewness**	**Mean**	**Std. Dev.**
1st interview	-0.2080	0.3810	0.4840	8.8273	10.127	36.050
2nd interview	-3.9656	0.3791	0.4102	8.8389	9.4922	34.632
3rd interview	0.4781	0.3931	0.5498	9.0258	9.5566	34.698
4th interview	1.2108	0.3913	0.5021	7.6780	9.3644	34.039
1988-1991	3.7871	0.3750	0.6283	9.3913	10.368	35.186
1992-1995	-9.6041	0.3643	0.2135	7.3545	10.101	35.365
1996-1998	5.2557	0.4160	0.7916	9.4284	8.5015	31.556
Services	**Skewness**	**Mean**	**Std. Dev.**	**Skewness**	**Mean**	**Std. Dev.**
1st interview	-0.6105	0.5188	0.3133	9.6950	25.540	102.26
2nd interview	9.2070	0.5487	0.7703	31.855	21.171	82.519
3rd interview	9.1497	0.5159	0.6865	39.943	23.993	103.80
4th interview	-20.564	0.5169	0.2392	18.533	24.541	113.24
1988-1991	10.761	0.5215	0.5763	19.976	28.297	100.02
1992-1995	-53.609	0.5532	0.1643	17.622	24.127	116.97
1996-1998	10.642	0.4932	0.7594	46.823	18.777	80.062

values for the skewness (which depends on the third moment from the mean) indicate that data are left skewed and positive values for the skewness indicate data that are right skewed[20]. The time-in-sample affects much more the shape of reporting errors than their mean value with a pattern depending on the commodity; the right tail of each distribution turns out heavier in almost all the cases with a strong evidence of non-normality both for β and ε. As suggested by the time profiles, the measurement error is not stationary over time. Moreover, the magnitude of such error depends on the real amount of spending for each commodity. In all the cases we find that the mean of β is less that 1, implying that the higher is d the lower is the error in reporting that good (by assumption ε is not influenced by d). However, the dispersion around this value can be considerably high.

It is worth noting that the correlation between errors in reporting commodities in $\mathcal{D}$ is set to zero, that is we assume that the error affecting the reporting of food expenditure doesn't depend on the error on different commodities entering non-durable expenditure. Such assumption might look quite restrictive; however, it seems very difficult to extend the moment method proposed in Section 8 to the general case of dependent errors amongst goods, since too many moments would be required to approximate (β, ε) reasonably. We will not discuss further on this point here and we will not consider any alternative estimation procedure to overcome this problem[21].

Figures 10-13 present non-parametric estimates of reporting error distributions for the same commodities controlling for education and race effects (the main two characteristics already found to be relevant for data quality in CEX data by Tucker, 1992). The estimation procedure is described in the Appendix and it identifies each density from estimated moments exploiting an orthonormal series approximation[22].

Figures 14-16 present results for error-corrected measures of mean and variance of non-durable expenditure as described in Section 7. As expected, the

[20]Each moment is estimated separately from the remaining moments regardless of the relationship amongst them, thus - for example - without imposing that the estimated variance (obtained as the difference between the second moment and the squared first moment) must be positive. In 3 out of 40 cases (i.e. first five moments of β and ε for each of the four commodities in $\mathcal{D}$) some of these relationships involving higher moments are violated (in particular, the estimated forth moment turned out negative in one case for 'alcohol' and two cases for 'services'); this small proportion we take as a evidence that the model imposed is not rejected by our data.

[21]As an example, Beran and Millar (1994) propose a minimum distance estimator for the following multivariate model resulting from the (four) error-affected components of Interview non-durable expenditures

$$\begin{pmatrix} r_{1,h} \\ r_{2,h} \\ r_{3,h} \\ r_{4,h} \end{pmatrix} = \begin{pmatrix} d_{1,h} & 0 & 0 & 0 \\ 0 & d_{2,h} & 0 & 0 \\ 0 & 0 & d_{3,h} & 0 \\ 0 & 0 & 0 & d_{4,h} \end{pmatrix} \begin{pmatrix} \beta_{1,h} \\ \beta_{2,h} \\ \beta_{3,h} \\ \beta_{4,h} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,h} \\ \varepsilon_{2,h} \\ \varepsilon_{3,h} \\ \varepsilon_{4,h} \end{pmatrix}.$$

The fit of this kind of models to our data is left for future research.

[22]Each density estimate is obtained removing small bumps in the estimated distributions and considering positive values of the quantity

$$\hat{f}_X(x) - \tau,$$

where $\hat{f}_X(x)$ is the orthonormal series approximation for the density of factor X using a Legendre basis up to the 5th order and τ is a constant such that the resulting estimate integrates to one.

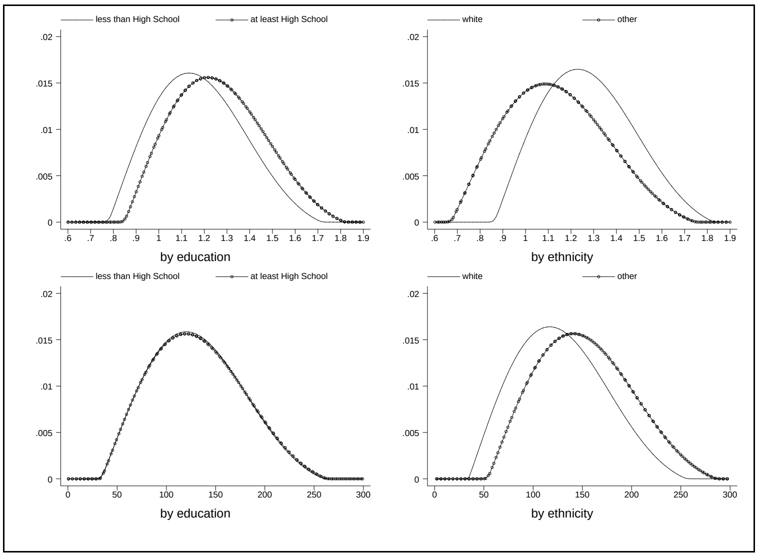

Figure 10: Food and non-alcoholic beverages at home. Reporting error distributions by education and race for β (first two panels) and ε (last two panels)

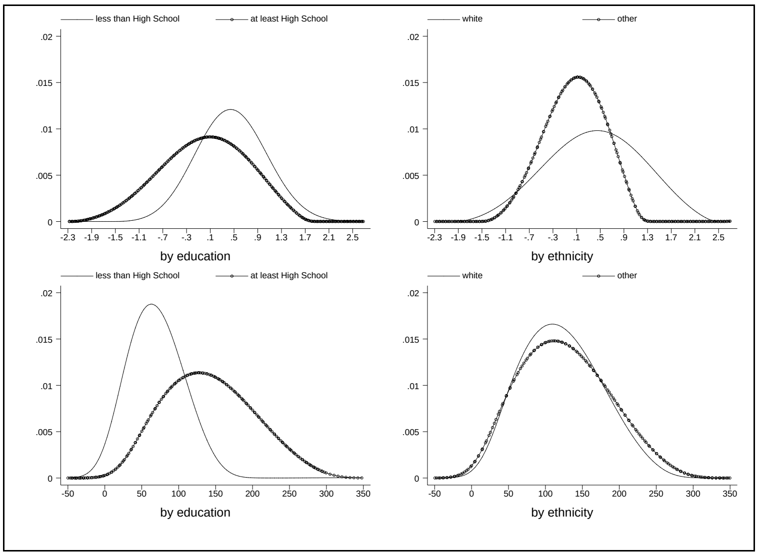

Figure 11: Food and non-alcoholic beverages away from home. Reporting error distributions by education and race for β (first two panels) and ε (last two panels)

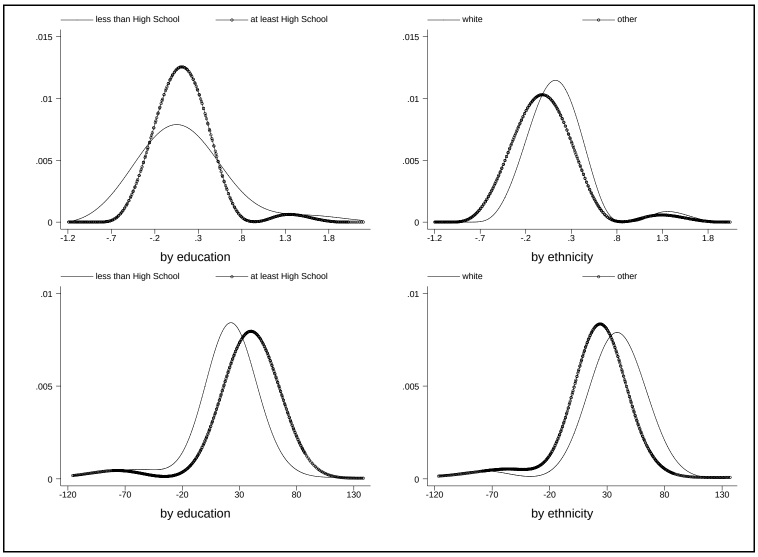

Figure 12: Alcoholic beverages. Reporting error distributions by education and race for β (first two panels) and ε (last two panels)

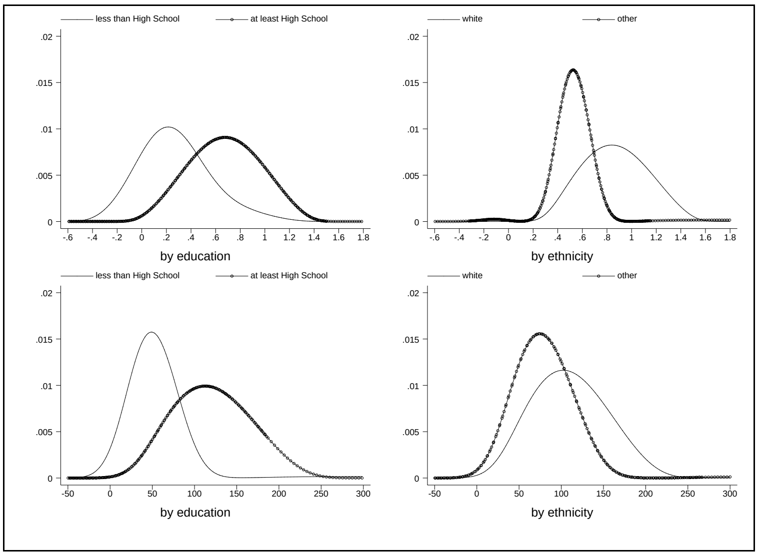

Figure 13: Personal care, entertainments and other services. Reporting error distributions by education and race for β (first two panels) and ε (last two panels)

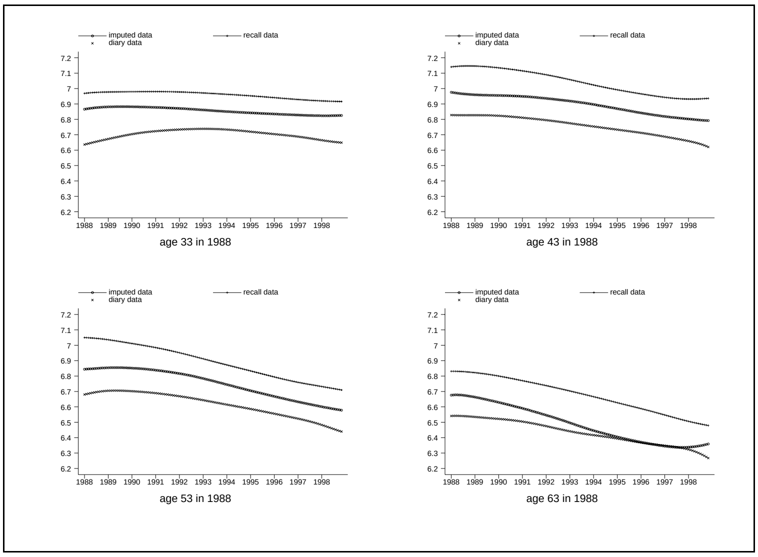

Figure 14: Mean of monthly log consumption by cohort after correction (Diary < Imputed < Interview)

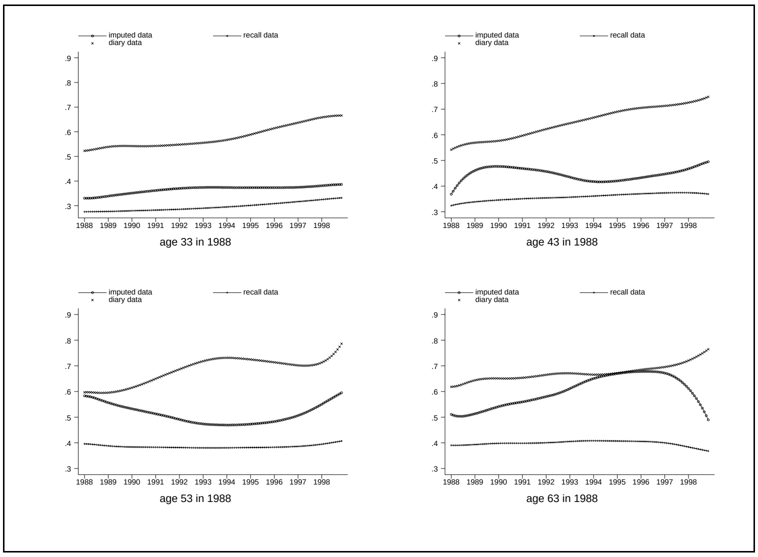

Figure 15: Variance of monthly log consumption by cohort after correction (Interview < Imputed < Diary)

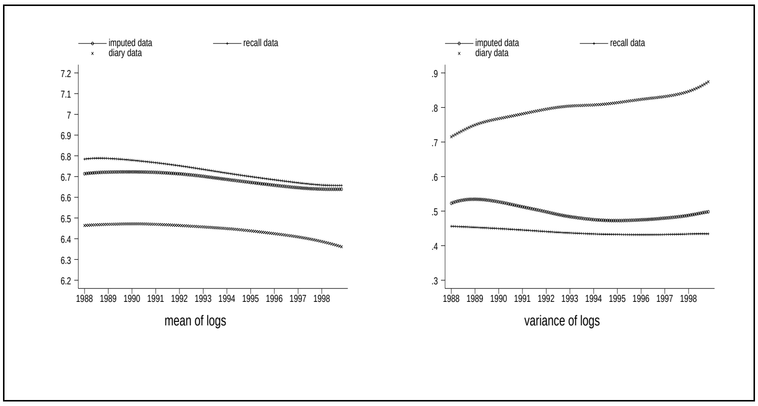

Figure 16: Mean and variance of monthly log consumption after correction (means: Diary < Imputed < Interview; variances: Interview < Imputed < Diary)

mean of the imputed measure of non-durable consumption lies between the observed Interview and Diary curves, with the exception of the life-time profile in late sixties as reported in the last panel of Figure 14. Imputed variances within each cohort generally increase over time, presenting an 'u-shaped' pattern before the retirement period ($55 - 60$). The variance for the whole sample has a peak in 1989, decreases in early 1990s (negative business cycle for the US economy) and increases mildly after 1994.

10 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 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 means and - under suitable conditions on the covariance between commodities - variances for macro-units (i.e. cohorts of households presenting the same observable characteristics) or for the whole population.

In this chapterpaper 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 of 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*.

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 produce some evidence of non-classical error affecting the aggregate measure of non-durable consumption both for diary and recall-based data; we also discuss on the implications of our findings for the estimation of several inequality indices. We finally show that - exploiting jointly diary and recall data to account such measurement problem - we come out with an inequality pattern over the 1990s that is not against the permanent income hypothesis.

11 Appendix

In what follows we show how to identify the distributions of β and ε from their moments estimated as described in Section 8. Both the methods we suggest lead to similar results when applied to our data: the first one is based on a rescaled and truncated orthogonal series approximation of the unknown densities; the second one derives these densities applying the Levy's inversion theorem approximating the characteristic function of each variable as a truncated series involving its moments.

The class of orthogonal series estimators can be derived describing the unknown density via a series expansion

$$f_X(x) = \sum_{k \in Z} a_k \psi_k(x), \tag{8}$$

where $\{\psi_k, k \in Z\}$ is an orthogonal system in $L_2([a,b])$ and $[a,b]$ is the domain of the variable X. Each coefficient a_k represents the orthogonal projection of $f_X(x)$ in the space spanned by $\{\psi_0, \psi_1, \ldots, \psi_k\}$ and is defined as

$$a_k = \int_a^b \psi_k(x) f_X(x) dx = E[\psi_k(x)]. \tag{9}$$

An estimator of the unknown density is derived taking a truncated version of the series approximation (8); the choice of the cut-off point is determined - as in every non-parametric procedure - by a tradeoff between efficiency and bias of the resulting estimate (for example minimizing the MISE). See Efromovich (1999) for further details.

Given a particular space, different orthogonal systems define different orthogonal series estimators of the unknown density $f_X(x)$. If the variable X

is square integrable over the support $[-1, 1]$ a possible orthonormal basis is represented by the following functions

$$\psi_k(x) = \frac{1}{2^k k!}\sqrt{k + 1/2}\frac{d^k}{dx^k}(x^2 - 1)^k,$$

that is an orthonormalized version of the well-known Legendre polynomials (which are orthogonal but not orthonormal over $[-1, 1]$). The resulting approximation is known as Fourier-Legendre Series Expansion.

The first few polynomials of this basis are the following

$$\begin{aligned}
\psi_0(x) &= \sqrt{1/2}, \\
\psi_1(x) &= \sqrt{3/2}x, \\
\psi_2(x) &= \sqrt{5/8}(3x^2 - 1), \\
\psi_3(x) &= \sqrt{7/8}(5x^3 - 3x),
\end{aligned}$$

so that the coefficients in (9) are only function of moments of X.

Suppose to identify the moments of a variable Z which we assume to be supported on the compact interval $[\mu - c, \mu + c]$, where μ and c are known constants; this is exactly the case discussed in Section 8, where the variables of interest were those characterizing the measurement error process (β and ε). Define the following one-to-one transformation of the support set and of the moments of Z to the interval $[-1, 1]$

$$X = \frac{Z - \mu}{c}.$$

The previous transformation maps the domain of Z into $[-1, 1]$ so that the moments of X are identified from the moments of Z

$$E(X^s) = \frac{E[(Z - \mu)^s]}{c^s}, \qquad s \geq 1.$$

It follows that fitting a truncated and rescaled Fourier-Legendre polynomials expansion we estimate both the density and the cdf of Z by means of

$$\begin{aligned}
\hat{f}_Z(z) &= \frac{1}{c}\hat{f}_X(\frac{z - \mu}{c}), \\
\hat{F}_Z(z) &= \hat{F}_X(\frac{z - \mu}{c}),
\end{aligned}$$

where the latter result follows integrating (8) with respect to X.

An alternative strategy of estimation for $f_Z(z)$ - or $F_Z(z)$ - builds on the characteristic function of Z which can always be expressed as

$$\sum_{k=0}^{+\infty} \frac{(it)^k}{k!} E(Z^k).$$

Since all the moments of Z are identified by assumption, we can invert (a truncated version of) this series to obtain an estimate of the corresponding density function.

References

[1] Alessie, R. Gradus, R. and Melemberg, 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. (1998), *Cohort Analysis of Saving Behavior by U.S. Households*, Journal of Human Resources, 33, 3, 575-609

[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. and Blow, L. (2001), *Trends in household expenditure in the UK, 1978-1999: aggregation issues and the bias of cost-of-living indices*, unpublished manuscript, Institute for Fiscal Studies, London

[7] Banks, J. Blundell, R. and Tanner, S. (1998), *Is There a Retirement-Savings Puzzle?*, American Economic Review, Vol. 88, n.4, 769-788

[8] Battistin, E. Miniaci, R. and Weber, G. (2001), *What do we learn from recall consumption data?*, working paper n. 10/00, Institute for Fiscal Studies, London

[9] Beran, R. and Hall, P. (1992), *Estimating Coefficient Distributions in Random Coefficient Regressions*, The Annals of Statistics, Vol. 20, No. 4, 1970-1984

[10] Beran, R. and Millar, P.W. (1994), *Minimum Distance Estimation in Random Coefficient Regression Models*, The Annals of Statistics, Vol. 22, No. 4, 1976-1992

[11] Blair, J. Menon, G. and Bickart, B. (1991), *Measurement effects in self vs proxy responses to survey questions: an information-processing perspective*, in Biemer, P.P. Groves, R.M. Lyberg, L.E. Mathiowetz, N.A. and Sudman, S. (eds.) *Measurement Errors in Surveys*, New York: Wiley, 145-166

[12] Blundell, R. and Preston, I. (1998), *Consumption Inequality and Income Uncertainty*, Quarterly Journal of Economics, Vol. , No. , 603-640

[13] 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

[14] Branch, R. and Jayasuriya, B. (1997), *Consumer Expenditure Interview and Diary Survey Data Selection: a New Method*, mimeo, Bureau of Labor Statistics

[15] Chesher, A. (1991), *The effect of measurement error*, Biometrika, 78, 3, 451-462

[16] Chesher, A. and Schluter, C. (2001), *Welfare measurement and measurement error*, forthcoming Review of Economic Studies

[17] Clark, T. (2001), *Which inequality? Interpreting contrasting stories*, unpublished manuscript, Institute for Fiscal Studies, London

[18] Deaton, A. (1992), *Understanding Consumption*, Oxford: Clarendon

[19] Deaton, A. and Muellbauer, J. (1980), *Economics and Consumer Behavior*, Cambridge: Cambridge University Press

[20] Deaton, A. and Paxon, C. (1994), *Intertemporal Choice and Inequality*, Journal of Political Economy , Vol. 102, n.3, 437-467

[21] Deaton, A. and Paxon, C. (2000), *Growth and savings among individuals and households*, The Review of Economics and Statistics, Vol. 82, n.2, 212-225

[22] Efromovich, S. (1999), *Nonparametric Curve Estimation*, New York: Sringer-Verlag

[23] Fan, J. and Gijbels, I. (1995), *Local Polynomial Modeling and Its Application - Theories and Methodologies*, New York: Chapman and Hall

[24] 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

[25] Gieseman, R. (1987), *The Consumer Expenditure Survey: Quality Control by Comparative Analysis*, Monthly Labor Review, March, 8-14

[26] Johnson, D. and Shipp, S. (1995), *Trends In Inequality Using Consumer Expenditures: 1960 To 1993*, BLS Economic Working Paper Series

[27] Kim, J.O. and Mueller, C.W. (1978), *Introduction to Factor Analysis*, Beverly Hills: Sage

[28] Krueger, D. and Perri, F. (2001), *Does Income Inequality Lead to Consumption Inequality? Empirical Findings and a Theoretical Explanation*, mimeo, Stanford University

[29] Horowitz, J.L. and Markatou, M. (1996), *Semiparametric estimation of regression models for panel data*, Review of Economic Studies, Vol. 63, 145-168

[30] Li, T. and Vuong, Q. (1998), *Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators*, Journal of Multivariate Analysis, Vol. 65, 139-165

[31] 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

[32] Manning, A. and Dickens, P. (2001), *The national minimum wage and wage inequality*, unpublished manuscript, London School of Economics

[33] Miniaci, R. Monfardini, C. and Weber, G. (2001), *Changing Consumption Patterns*, unpublished manuscript, Department of Economics, University of Padova

[34] 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

[35] Poikolainen, K. and Karkkainen, P. (1983), *Diary Gives More Accurate Information About Alcohol Consumption than Questionnaire*, Drug and Alcohol Dependence, Vol.11, 209-216

[36] 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

[37] 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

[38] Schennach, S. (2000), *Estimation of nonlinear models with measurement error*, unpublished manuscript, Dept. of Economics, University of Chicago

[39] Shorrocks, A.F. (1982), *Inequality Decomposition by Factor Components*, Econometrica, Vol. 50, No. 1, 193-211

[40] 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

[41] 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

[42] 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

[43] Simonoff, J.S. (1996), *Smoothing Methods in Statistics*, Springer Verlag, New York.

[44] 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

[45] 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

[46] 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

[47] 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

[48] 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

[49] 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