ERRORS IN SURVEY REPORTS OF CONSUMPTION EXPENDITURES

Erich Battistin

THE INSTITUTE FOR FISCAL STUDIES
WP03/07

Errors in Survey Reports of Consumption Expenditures

Erich Battistin*
Institute for Fiscal Studies, London

6th May 2003

Abstract

This paper considers data quality issues for the analysis of consumption inequality exploiting two complementary datasets from the Consumer Expenditure Survey for the United States. The Interview sample follows survey households over four calendar quarters and consists of retrospectively asked information about monthly expenditures on durable and non-durable goods. The Diary sample interviews household for two consecutive weeks and includes detailed information about frequently purchased items (food, personal cares and household supplies). Each survey has its own questionnaire and sample. Information from one sample is exploited as an instrument for the other sample to derive a correction for the measurement error affecting observed measures of consumption inequality. Implications of our findings are used as a test for the permanent income hypothesis.

Keywords: Consumption Inequality; Measurement Error; Permanent Income Hypothesis

JEL Classification: C13, C42, D12, D91

*First draft 11^{th} February 2002. This paper benefited from useful discussions with Gordon Anderson, Orazio Attanasio, James Banks, Richard Blundell, Martin Browning, Hide Ichimura, Arthur Kennickell, Costas Meghir, Bruce Meyer, Enrico Rettore, Guglielmo Weber and from comments by audiences at Padova University, Cemmap (London), ESPE 2002, 10^{th} International Conference on Panel Data and NBER Summer Institute 2002. Address for correspondence: Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE - UK. E-mail: erich_b@ifs.org.uk.

EXECUTIVE SUMMARY

This paper aims to quantify the effect of reporting errors affecting diary-based and recall-based data on non-durable expenditure. It is likely that not all the commodities entering non-durable expenditure are well reported exploiting only one of these two survey methodologies. Expenditures on frequently purchased, smaller items are presumably more accurate using diaries while recall data are more appropriate for large expenditures or expenditures occurring on a regular basis.

It turns out that neither diary nor recall-based data alone provide a reliable aggregate measure of total expenditure on non-durables. Ideally, the estimation of totals at micro-level would require information on different consumption categories obtained with the most appropriate methodology. The Family Expenditure Survey for the United Kingdom represents a notable implementation of this strategy.

One might argue that for any practical purpose these alternative data collection strategies lead to consistent results in the estimation of economic models of consumption behavior. Unfortunately, evidence from the literature suggests that conclusions are strongly related to the information being used. This problem is addressed by looking at micro data from two independent samples of households from the Consumer Expenditure Survey for the Unites States. This survey represent a unique source of data because it consists of diary an recall information collected on the same set of items, although it refers to separate samples of households.

The comparison of the two surveys based on the overlap in coverage of expenditures offers insights for the effects of collection modes on data quality. Pictures emerging from the two surveys are very different, both with respect to mean expenditure and, more importantly, with respect to indices of inequality. More precisely, while the differences in mean expenditure are roughly constant over time, consumption inequality presents different levels and trends in the two surveys. This evidence is reconciled using integrated diary and recall information to characterize the most likely pattern of consumption for cohorts of individuals defined by their age.

This 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. The implications of our findings for the estimation of inequality indices are discussed, with an application to permanent income models. In particular, we show that using diary and recall data to improve the quality of household consumption the permanent income hypothesis cannot be rejected.

1 INTRODUCTION

Data quality is an issue of longstanding concern among researchers interested in testing the implications of theoretical models of consumption over the life cycle.

The empirical analysis of these models requires reliable micro-data on expenditures at household or individual level. In many countries expenditure data are regularly collected either by diaries covering purchases made within a short period of time (typically one or two weeks) or by means of retrospectively asked questions on the *usual* spending over a longer period.

There is a consensus that the time-consuming task required by diaries produces good quality expenditure data for small items, while recall questions should be asked 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 (such as transportation or rent).

For this reason diary surveys are designed to obtain detailed recordings of expenditures on small, frequently purchased items which are normally difficult to recall. Such an idea is not only intuitively clear, but it is also supported with evidence from cognitive studies and from the comparison of aggregated consumption measures with national account data.

On the other hand, the drawback of such an evidence is that neither diary nor recall based data alone provide a reliable aggregate measure of *total* consumption. One might argue that for any practical purpose these two alternative designs lead to consistent results, but unfortunately this is not the case. For example, there is some evidence that recall consumption data lead to potentially misleading results in analyzing household saving behavior (Battistin *et al.* 2003). Other studies demonstrate how available consumption data can be unsuitable for the analysis of the permanent income (life cycle) hypotheses and how adjustments provide greater consistency concerning the time series properties of consumption (Wilcox 1992 and Slesnick 1998). Browning *et al.* (2002) provide a detailed discussion of these problems and review alternative survey methodologies that have been suggested to obtain reliable measures of consumption.

Ideally, the estimation of totals at micro-level would require information on different consumption categories obtained with the most appropriate methodology. The Family Expenditure Survey for the United Kingdom represents a notable implementation of this strategy. It consists of a comprehensive household questionnaire which asks about regular household bills and expenditure on major but infrequent purchases and a diary of all personal expenditure kept by each household member (including children) for two weeks. There is some evidence - at least for the United Kingdom - that consumption measures obtained from such a design are comparable to aggregated values from national accounts (Banks and Johnson 1998).

Because of time constraints and survey practice, questionnaires can not cover all the aspects of consumer behavior with the same level of accuracy. Since the collection of diary records is highly time-consuming in itself, the issue arises of whether consumption information based on recall questions is of comparable quality to information based on diary records.

A related problem is the characterization of the response error across different consumption categories in diary and recall data. Retrospectively collected information of household surveys are typically characterized by recall errors. For example, respondents might round off the true measure causing abnormal concentrations of values in the empirical distribution. Neter and Waksberg (1964) discuss the relative

importance of forgetfulness and telescoping in expenditure data from household interviews. See also Browning *et al.* (2002) for a review of data quality issues related to the collection of consumption information.

It is often assumed that response errors in the variable of interest are of classical form, so that they have limited impact on parameter estimates (if consumption is used as a dependent variable) or can be accounted for by using suitable econometric techniques depending on the model specification (see for example Lewbel 1996 and Hausman *et al.* 1995). However, this assumption is made for ease of estimation rather than for any theoretical conviction. In fact, Bound *et al.* (2001) review the state of the art about measurement error in surveys across a wide range of areas in economics and provide evidence that standard assumptions are likely to be violated to various extents.

In the presence of validation data, one might be able to correct biases and derive consistent estimates from primary data without further assumptions on the error structure (see Lee and Sepanski 1995 and the references therein). But again this would require the availability of complementary datasets, ideally with the same individuals or, at least, with a set of common information rich enough to motivate matching procedures (see Ziliak 1998 and Battistin *et al.* 2003 for recent applications).

The aim of this paper is to shed light on the comparison between recall based and diary based data on household consumption using micro-level data from the Consumer Expenditure Survey for the Unites States (CEX in the following). This survey consists of two different components: a quarterly Interview Survey (IS) and a weekly Diary Survey (DS), each with its own questionnaire and sample. The most interesting feature that makes the CEX an extremely appealing source of data is that the IS and the DS overlap for many categories of consumption for which information is collected using different methodologies.

According to the line of thinking discussed above, neither of these two surveys provides accurate estimates of total consumption at household level. In fact, the two survey components are explicitly designed to collect information on different types of expenditures (see Bureau of Labor Statistics 2002). The IS aims to obtain data on the types of expenditures respondents can recall for a period of three months or longer; the DS is instead designed to obtain data on frequently purchased smaller items. Accordingly, the Bureau of Labor Statistics (BLS in the following) publishes data integrated from the two components to provide a complete accounting of consumer expenditures, which neither survey component alone is designed to do.

Inconsistencies between the CEX and national accounts have been already pointed out in the literature by several papers (notably by Slesnick 1998, 2001), suggesting that the quality of these data might have deteriorated over the last decade. Figure 1 presents aggregate expenditure on non-durable goods using published tables from the CEX and from the Personal Consumption Expenditures (PCE).[1] The two series follow a similar trend until 1992, with CEX expenditure peaking in 1989. However, CEX data do not appear to pick up the trend in total consumer spending found in the PCE over the 1990s. Slesnick (2001) finds that only part of this discrepancy can be explained by definitional differences, concluding that "*the remaining gap is a mystery that can be resolved only by further investigation*" (page 52).

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

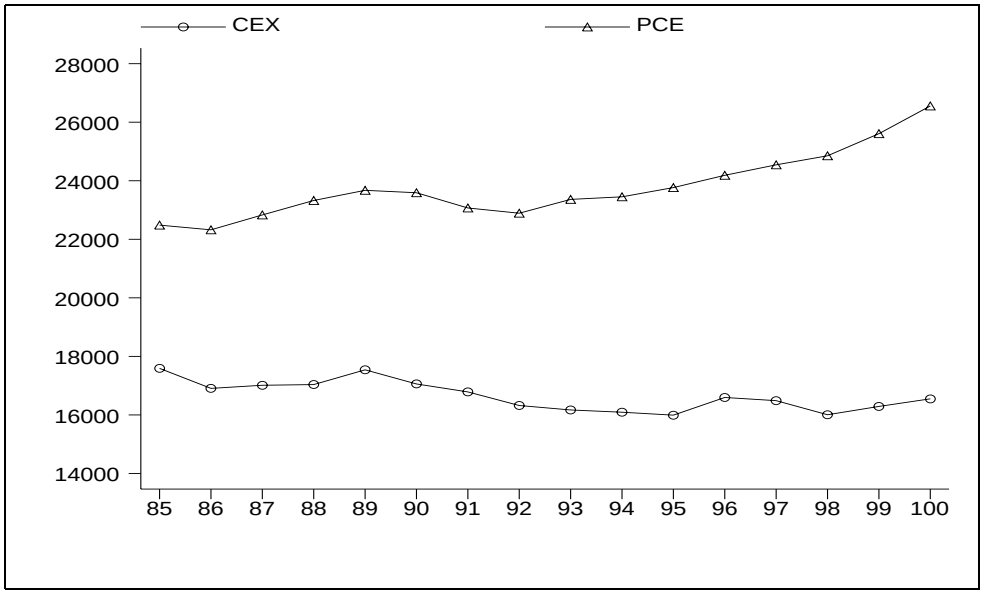

Figure 1: Non-durable expenditures in 2000 dollars - Consumer Expenditure Survey (CEX) and Personal Consumption Expenditures (PCE)

Despite of the puzzle implied by these findings, the IS is currently the most widely used source of consumption data for the United States: very many authors have used it to validate different theoretical constructs of economic behavior over the years (see, for example, Attanasio and Weber 1995 and Krueger and Perri 2001). The most appealing features of the IS are primarily its panel component (individuals are interviewed every three months over five calendar quarters) and the richer amount of household information collected with respect to the DS. On the other hand, although the DS exists in its current format since 1986, it hasn't received so much attention by researchers so far (the only example we are aware of is Sabelhaus 1996). Actually, the DS is still a relatively unknown source of data.

Given the overlap between the IS and the DS for many categories of consumption, and given that the IS is explicitly designed to collect good quality information only on a subset of these categories, the question then arises of whether we can jointly exploit DS and IS data to derive a superior measure of total consumption. The answer to this question has very many empirical implications. As already pointed out by Wilcox (1992), the imperfections of micro-data on consumption expenditures may be important enough to influence the conclusions of empirical work. Are our data relevant to the theory? Is the economic model really in error? Should research be directed towards alternative models of economic behavior or is data itself not suitable to validate existing models?

We will address these issues presenting three sets of results. Firstly, *new* evidence on the evolution of consumption inequality for the United States in the last twenty years is presented using data from the DS. Consumption inequality for the United States has recently received much attention amongst researchers, since according to IS data it does not appear to have grown much during a period characterized by a marked increase in income inequality. This result has also generated

discussion on the appropriate measure of economic well-being in the evaluation of inequality.

Secondly, the amount of misreporting of total non-durable expenditure characterizing the IS is discussed exploiting information from the DS. Since the two surveys refer to different samples, individual level consumption cannot be straightforwardly defined from combined IS and DS data. For this reason, we will mainly focus on figures for mean expenditure on different commodities for cohorts of people identified by their year of birth or, equivalently, by their age in a base year.

Finally, the available information from the IS and the DS allows us to examine to what extent tests of the permanent income (life cycle) hypothesis are sensitive to the choice of consumption measure. Our motivating example is the work by Deaton and Paxson (1994), where they examine the evolution over time of the variance of total non-durable consumption at cohort level as a test for the original formulation of permanent income theory (Hall 1978).

The main findings of this paper can be summarized as follows. Firstly, the pattern of consumption inequality obtained from DS data is very different from the pattern largely discussed in the literature using IS data (see for example the discussion in Krueger and Perri 2001). DS inequality seems to grow over time, and particularly over the last ten years when there is no evidence of increasing inequality using IS data. A correction procedure is proposed to reconcile the evidence from the CEX and to characterized the most likely pattern of consumption inequality. This issue is further developed by Attanasio *et al.* (2003).

Secondly, we show that the quality of IS information has worsened with respect to frequently purchased smaller items, housekeeping supplies and personal care products and services (that is those items the DS is designed for). This decline is particularly accentuated for the last ten years. On the other hand, we show that DS data quality has improved also for all those non-durable items which presumably are better described using IS data.

Finally, we point out that data quality issues in the IS should deserve more attention in the evaluation of alternative constructs of economic behavior. We emphasize this point showing that the permanent income model as formulated by Hall (1978) cannot be rejected combining IS and DS data to improve the quality of reported consumption. The same test applied to IS data over the period covered by our analysis leads to different conclusions.

The remaining of this paper is organized as follows. Section 2 describes the two surveys and compares descriptive statistics of household characteristics already found to be relevant for data quality in previous studies of expenditure surveys. Section 3 discusses the economic model used as a motivating example. Section 4 presents a puzzle implied by the comparison of means and inequality indicators of non-durable expenditure exploiting information from the two surveys. Section 5 analyzes such discrepancies looking at the contribution of different non-durable goods. Also, the identification restrictions needed to combine IS and DS information are presented. Section 6 presents results on the reporting errors affecting non-durable components. Section 7 discusses the testing procedure to validate the permanent income model using combined information from the IS and the DS. Results from this test are presented in Section 8. Some more technical comments on data collection issues and data problems are discussed in the Appendix.

2 DATA

The main characteristics of the two survey components from the CEX are summarized in what follows. In particular, Section 2.1 describes diary and recall questionnaires. Section 2.2 discusses the extent to which the IS and the DS are comparable with respect to sample designs, population coverage and information collected; also, the definition of household total consumption is presented. Finally, Section 2.3 discusses the working sample considered in this paper. The reader interested in more specific details on the survey methodology in the CEX is referred to Bureau of Labor Statistics (2002).

2.1 The consumer expenditure surveys

The CEX is currently the only micro-level data set reporting comprehensive measures of consumption expenditures for a large cross-section of households in the United States. Essentially, sample consumer units are households (literally, "all members of a particular housing unit who are related by blood, marriage, adoption, or some other legal arrangement", Bureau of Labor Statistics 2002), whose buying habits provide the basis for revising weights and associated pricing samples for the Consumer Price Index. Only one person responds for the whole consumer unit, typically the most knowledgeable of expenditures in the family.

The survey consists of two separate components, each of them with its own questionnaire addressing a *different* sample. The IS sample is selected on a rotating panel basis targeted at 5000 units each quarter; DS data refer to repeated cross sections of households (around 4500 per year) interviewed over a two-week period. Response rates for the two components are reasonably good (around 80 percent).

In the IS, households are interviewed about their expenditures every three months over five consecutive quarters. After the last interview households are dropped and replaced by a new unit, so that - by design - 20 percent of the sample is tossed out every quarter. Expenditure information is collected in the second through the fifth interview; one month recall expenditures are asked in the first interview only for bounding purposes. The percentage of households completing all five interviews is about 75 percent, with single persons more likely to attrit.

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?*". For non-durable goods households are asked to report their usual weekly expenditure only for tobacco products and for food and non-alcoholic beverages consumed at home. The expenditure on the latter category is obtained as the difference between the usual weekly total expenditure at grocery stores or supermarkets and how much of this amount was for non-food items (specified as '*paper products, detergents, home cleaning supplies, pet foods, and alcoholic beverages*'). Expenditures on alcoholic beverages and food away from home (but not food consumed on vacation) are referred to the usual monthly amount.

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

The DS is instead a cross-section of consumer units asked to self-report their daily purchases for two consecutive one-week periods by means of product-oriented diaries. Each diary is organized by day of purchase and by broad classifications of goods and services. Respondents are assisted by printed cues and - whether it is needed - by interviewers at pick-up. The percentage of households completing both diaries is about 92 percent. Interestingly, the DS also collects recall information for food and non-alcoholic beverages consumed at home as in the IS, implying that for this category of consumption recall and diary information is available for the same households.

Both IS and DS collect information on a very large set of household characteristics (demographics, work-related variables, education and race) as well as on income and assets (using a twelve-month recall period). Income figures refer to total before-tax family income in the last year. This information is subject to top-coding in the CEX, but only for a small proportion of household in the two samples (less than 1 percent in any year). However, income and assets data are known to be not as reliable as the expenditure data: the amount of incomplete income reporters is about 20 percent in the two surveys and missing values are currently not imputed. For this reason many applications in the literature have combined consumption information from the CEX to income information from complementary data sets (see, for example, Lusardi 1996 and Blundell *et al.* 2002).

The two survey components are based on a *common* sampling frame: the 1980 Census for those households sampled in the 1980s and the 1990 Census for households sampled in the 1990s. Sample designs differ only in terms of frequency and over sampling of DS households during the peak shopping period of Christmas and New Year holidays. The BLS constructs weights to control for systematic aspects of the sampling, to post-stratify by region, home ownership, household size and race and to compensate for under coverage (relative to the Census-based estimates) of people by age, race and gender.

2.2 On the comparability of the two surveys

As far as consumption is concerned, the BLS follows the standard international procedure of exploiting both information from recall questions for more durable items bought in the quarter prior to the interview and diary-based records of purchases carried out within a two-week period. In fact, as discussed above, the IS and the DS are explicitly designed to collect different types of goods and services and neither survey is expected to represent all aspects of consumption.

Since aggregate consumption is obtained from integrated survey data, the question then arises of which survey component provides more accurate information on different items. When data are available from both surveys, the procedure followed by the BLS determines the most reliable survey by comparing CEX figures to those from other data sources - typically from the Personal Consumption Expenditures (see Branch and Jayasuriya 1997 and McCarthy *et al.* 2002). Of course, a potentially interesting question is the extent to which external sources provide an accurate representation of total consumption (Slesnick 2001).

However, some expenditure items are collected only by either the IS or the DS. The IS excludes expenditures on housekeeping supplies (e.g. postage stamps), personal care products and non-prescription drugs, which are instead collected in

the DS.[2] On the other hand, the DS excludes expenditures incurred by members while away from home overnight or longer and information on reimbursements (such as for medical care costs or automobile repairs), which are collected in the IS.

Throughout the analysis, only figures for expenditure on *non-durable goods and services* will be considered.[3] The expenditure categories considered have been defined so that IS and DS definitions are comparable and consistent over time. The definition of non-durable expenditure closely follows the one already given 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 clothing and footwear). In particular, expenditure on health, education (which can be considered as an investment in human capital) and mortgage/rent payments are excluded. To improve the comparability of the two surveys, DS figures for total consumption are considered only after 1986 (see the Appendix for a discussion of data collection problems in the CEX affecting the definition of non-durable expenditures).

Changes in survey instruments characterize the CEX over the time covered by our analysis. For example, a new diary form with more categories and expanded use of cues for respondents was introduced in 1991, based on results from earlier field and laboratory studies. Moreover, changes in the wording of the IS question on food consumption heavily affect the mean of reported expenditure on this category over time (data have been adjusted to solve for this problem; see the Appendix for more details).

Only expenditure figures for the month preceding the interview are considered for the IS sample, thus leaving four observations for each household (one observation for each interview).[4] Monthly expenditure in the DS is defined as $26/12 = 2.16$ times the expenditure observed over two weeks, assuming equally complete reporting.

2.3 The working sample

The CEX has a long history: the first survey was ran in 1917-18. Before the new ongoing structure initiated by the BLS in 1980, previous surveys were conducted only every 10 to 12 years. The 1980 survey is the first year providing consumption information on a continual yearly basis. Differences in the survey methodology between the surveys conducted after 1980 and those conducted before result in visible inconsistencies when compared to the national accounts (see Slesnick 2001).

The information exploited in this paper covers twenty years of data from the IS and the DS between 1982 and 2001. However, public use tapes permit to integrate data on non-durable consumption from both surveys only after 1986, since only selected expenditure and income data from the DS were published before then.

In what follows, the family head is conventionally fixed to be the male in all H/W families (representing the 56 percent and 53 percent of the whole sample for

[2] This expenditures contribute about 5 to 15 percent of total monthly expenditures in our data.

[3] Assuming preference separability between durables and non-durables, expenditure on non-durable goods and services is a relevant consumption measure.

[4] It has been found that expenditures for many items are reported more frequently for this month than for earlier months (see 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.

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

Variable	1982-85	1986-89	1990-92	1993-95	1996-98	1999-2001
Proportion of components 18−						
Proportion of components 64+				2.22	-3.15	2.11
Proportion of children $0-3$						
Proportion of children $4-7$					1.94	
Proportion of children $8-12$			-3.00	2.86		
Proportion of children $13-18$		3.43				
Age of the reference person				2.11		
Dummy for retired head	-6.13	-11.94	-13.01	-8.38	-4.90	
Weeks worked per year	-6.00	-10.73	-12.04	-9.34	-8.20	-3.60
Total amount of income before taxes	10.03	-7.78	9.10	6.35		4.05
Total amount of income before taxes squared	-4.25	-4.01	-6.12	-3.58		-3.78
Dummy for Midwest region				2.57		
Dummy for South region	4.15			2.39		
Dummy for West region						2.44
Husband and wife (H/W) only	-2.93	-4.24	-2.64			
H/W, own children only, oldest child $0-5$		-6.45	-2.65	-3.67	-2.06	
H/W, own children only, oldest child $6-17$	-2.75	-2.16	-2.04	-2.56	-3.16	
H/W, own children only, oldest child over 18		-2.62			-2.41	2.04
All other H/W households		-2.96				
One parent (male) at least one child $0-18$						
Single persons						
Dummy for Black		-2.31				
Dummy for American Indian	-2.37		1.95			
Dummy for Asian or Pacific Islander		-2.14		2.08		
High School Graduate		-2.95	-5.01	-1.98		
College dropout		-2.61	-3.63	-2.26		
At least College graduate		-3.15	-4.75	-3.30		

IS and DS data, respectively). Furthermore, only households headed by individuals aged at least 23 and no more than 73 and not self-employed are considered. These restrictions leave us with a sample of 190.080 and 46.244 households over the considered period of time, for IS and DS data respectively. Sample sizes by year and a detailed description of less important selection criteria and of data problems are presented in the Appendix.[5]

Although the two surveys are designed to be representative of the same population, significant differences in the two samples are found along several dimensions and with a different pattern over time. Table 1 presents t-statistics from a logistic regression of the binary indicator IS/DS household over a set of variables including work-related information and characteristics found to be relevant for data quality in previous analysis of CEX data (Tucker 1992). Weighted results are presented, using population weights from each survey. The specification adopted includes polynomial terms in the age of the reference person and in the proportion of children and members within certain age bands (these terms are not reported because not statistically significant).

For ease of exposition, only statistically significant differences at the 95 percent confidence level or above are reported. Moreover, pooled results for the following six time periods are presented: 1982-85, 1986-89, 1990-92, 1993-95, 1996-98, 1999-2001. A negative (positive) value in the table should be interpreted as an higher concentration of households with that characteristic in the DS sample (IS sample, respectively) with respect to the other sample. The main difference in the composition of the two surveys is confirmed to lie in the DS relative over sampling of more educated households (particularly between 1986 and 1996), and generally of H/W households with retired heads. The amount of weeks worked per year by the reference person and the distribution of total family income are lower in the IS sample. However, significant differences are found along several other dimensions and with a different pattern over time.

3 THE MOTIVATING EXAMPLE

This section follows closely Deaton and Paxson (1994) to review the economic reasoning that motivates our exercise. The simplest formulation of Permanent Income Hypothesis (PIH) implies that, for any cohort of people born at the same time, consumption inequality should grow with age (see, for example, Deaton 1992). The conventional model of consumption under uncertainty assumes that, in each period t, individuals maximize expectation of a time-separable utility function

$$U(c_t) + \sum_{s=t+1}^{T} \delta_s U(c_s),$$

subject to expectation of an intertemporal budget constraint

$$(c_t - y_t) + \sum_{s=t+1}^{T} r_s(c_s - y_s) = \omega_t.$$

Throughout this section U will denote an utility function invariant over the life cycle, δ_t the individual's rate of subjective preference, r_t the real interest rate and

[5]Note that, due to the criteria used to define the working sample, we might observe systematic movements in/out of the sample for IS households over their one-year interview period.

ω_t the accumulated wealth al time t. The length of the life cycle is T, while c_t and y_t represents log real consumption and income in each period.

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

$$U'(c_t) = \beta_t U'(c_{t-1}) + \nu_t,$$

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

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

$$c_t = \beta_t c_{t-1} + \upsilon_t.$$

The last expression has some testable implications that might be used to check the validity of this set-up. First, since the lagged value of consumption incorporate all information about consumers' decision at time t, then consumption or any other related variable (income, in particular) lagged more than one period shouldn't have any explanatory power to predict the current value of consumption. This can be tested by looking at the coefficients in the regression of actual consumption on lagged values of consumption and income (see Hall 1978).

Additional implications become available once we are willing to make assumptions on the evolution of the β_t terms over time. To see that, consider the case where $\beta_t = 1$, as in the original model. Under this condition, individual's variance must increase over time (Deaton and Paxson 1994), since

$$Var(c_t) = Var(c_{t-1}) + Var(\upsilon_t). \tag{1}$$

Secondly, since

$$\frac{c_t}{t} = \frac{1}{t} \sum_{s=0}^{t} \upsilon_s, \tag{2}$$

individual's consumption is a sample average of independent shocks and hence asymptotically normal by means of the central limit theorem. Note, however, that this result applies when t grows to infinity; since individuals have finite life-spans, the distribution of consumption is expected to be normally distributed only amongst older people (see Blundell and Lewbel 1999).

When β_t varies over time, the implications of the model might be different. The distribution of individual's consumption at time t is more disperse than the distribution of consumption at time $t-1$ when β_t is greater that one. If the rate of interest is greater than the rate of time preference (i.e. if incentives to postpone consumption dominate impatience), the distribution of individual's consumption can either concentrate or disperse depending on the variance of υ_t. On the other hand, consumption in older age can be normally distributed only if β_t is always centered around one, to avoid having the sum of shocks in (2) be zero or infinity as t grows to infinity.

Of course, the validity of the previous implications of the PIH rests on several *caveats* that have been already pointed out in the literature (see the discussion in Browning and Lusardi 1996 and the references therein).[6] In the remaining of this paper, we will build on Deaton and Paxson (1994) to investigate the validity of (1) by looking at the cross-sectional dispersion of IS and DS non-durable consumption within cohorts as they age.

4 EVIDENCE ON CONSUMPTION BEHAVIOR

A standard way to analyze the dynamic properties of consumption with repeated cross-sections is to rely on cohort analysis. This section investigates the effects of different data collection methodologies by comparing expenditures levels (Section 4.1) and expenditure variances (Section 4.2) from IS and DS data.

In what follows we will group IS and DS 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. We will focus on four cohorts of individuals born between 1930 and 1969; sample sizes for the IS and the DS are reported in the Appendix. Throughout this paper we will tend to use 'expenditure' and 'consumption' as two synonyms, since the distinction is not relevant in this context.[7]

Clearly household characteristics (occupation and economic activity of the head, household composition, region of residence) affect the share of spending and the quality of reporting own expenditures (Tucker 1992). 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 DS households exploiting a weighting scheme based on the regressions in Table 1 (see Battistin *et al.* 2003). Intuitively, the aim is to down-weigh (up-weigh) those households in the IS sample exhibiting characteristics over represented (under represented) with respect to the DS sample. Such a weighting scheme depends on the conditional probability of observing those characteristics in the population represented by the IS (the so-called *propensity score*), that is on the binary regressions reported in Table 1. 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).

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 profiles because of an increasing dispersion of household size (and, as a consequence, available family income) as the cohort ages.[8]

To summarize the evidence from this section, we find that the evolution of consumption means and variances as cohorts age over the life-cycle turns out very

[6]For example, the evolution of within cohort inequality depends on people's attitude toward risk and on the mechanisms that are available for sharing risks between people and periods.

[7]The panel component of the IS survey is not exploited in this paper, that is quarterly observations for IS households are counted separately as if the four observations over the one year interview referred to different households.

[8]It might be interesting considering how much robust our results are to variations in head's age definition (Deaton and Paxson, 2000). 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.

different depending on the source we consider.[9] In particular, although most of the published research for the United States has noticed that using IS data consumption inequality has not grown significantly over the last twenty years, we show that the picture emerging using DS data is very different. Differences in the two surveys with respect to this point are discussed in the remaining of this section. Section 5 below aims to reconcile the evidence from the IS and the DS and to characterize the most likely pattern of consumption inequality.

4.1 Expenditure levels

Figure 2 presents total expenditure on non-durable consumption by cohort as obtained from raw IS and DS data. Each point in the graph represents mean expenditure of the cohort in a generic year over the period 1982-2001 (1986-2001 for the DS survey). Family consumption is adjusted using an equivalence scale which depends on the number of adults and children in various age ranges (see the Appendix for more details).

There are several discrepancies between estimated profiles for the two surveys. Perhaps the most striking feature is that DS cohorts consume less than IS cohorts, consistently over time. Overall, DS data appear to pick up most of the trends in total spending found in IS data. Non-durable consumption declines in the last part of the life cycle in line with similar drops reported in other countries (the retirement puzzle). However, this decline seems to be more pronounced for IS data.

A further point worth stressing is the sharp drop of the IS expenditure level in the early 1990s. Apparently, differences in reported levels of consumption between the two surveys converge to zero after this time, and both data sources suggest a declining pattern of total consumption in the last ten years.

The implications of these findings for the whole population are presented in Figure 3. Annual growth rates in the two surveys are positive until 1989, then the two series present negative growth in real consumer spending (although this decline occurs later in time for the DS). As already pointed out in the Section 1, the level of expenditure from integrated IS and DS data is decreasing over time, particularly in the 1990s. This finding strongly contradict the pattern of expenditure on non-durables as implied by looking at data from the national accounts (see Figure 1).

The observed pattern of consumption in the two surveys can be further analyzed by breaking down total expenditure into the contribution of different categories of non-durable consumption. The relationship between mean expenditures in the two surveys varies a great deal considering different commodities. Thus, the overall figure for total expenditure appears to be the aggregated outcome of a large number of positive and negative mean differences on non-durable commodities. Figures A.2-A.10 in the Appendix present cohort profiles for expenditure on the nine non-durable groups considered in this paper (as discussed in Section 2.2).

Trends in expenditure on food related items and transportation in IS and DS data probably presents the most striking differences. IS households realized a sharp decline in price-adjusted food expenditure over the 1990s, both at home and away from home. This figure is not consistent with DS data about food at home expenditure, whose values present a less pronounced decline over time. Expenditure for transportation exhibits different trends over time for the two surveys, with DS figures increasing in the 1990s contrary to those from the IS.

[9] We removed the households with the highest and lowest 2% of expenditures in each year so as to enhance robustness of results.

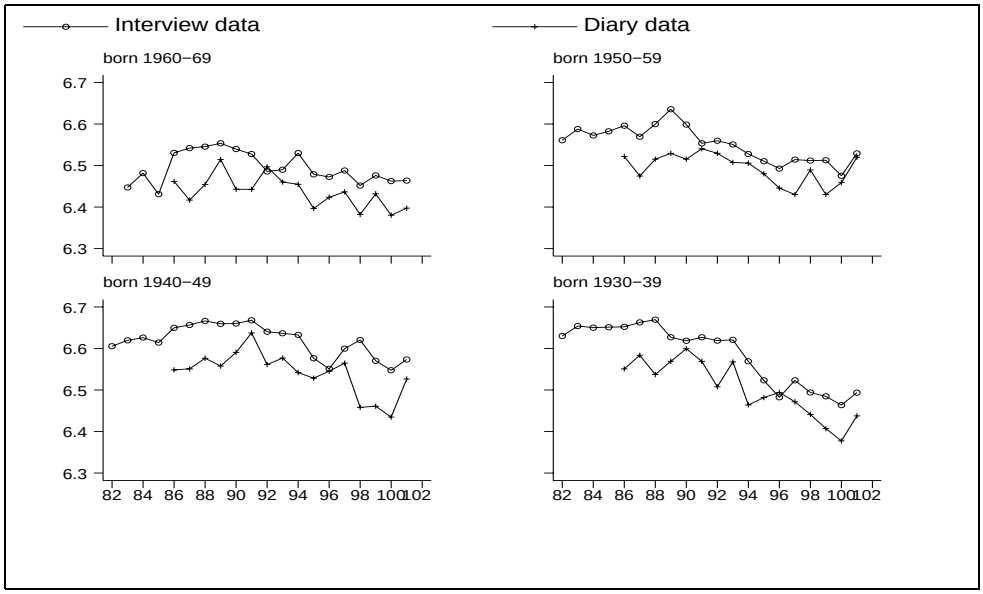

Figure 2: Mean of log family expenditure on non-durable goods by cohort (2001 dollars)

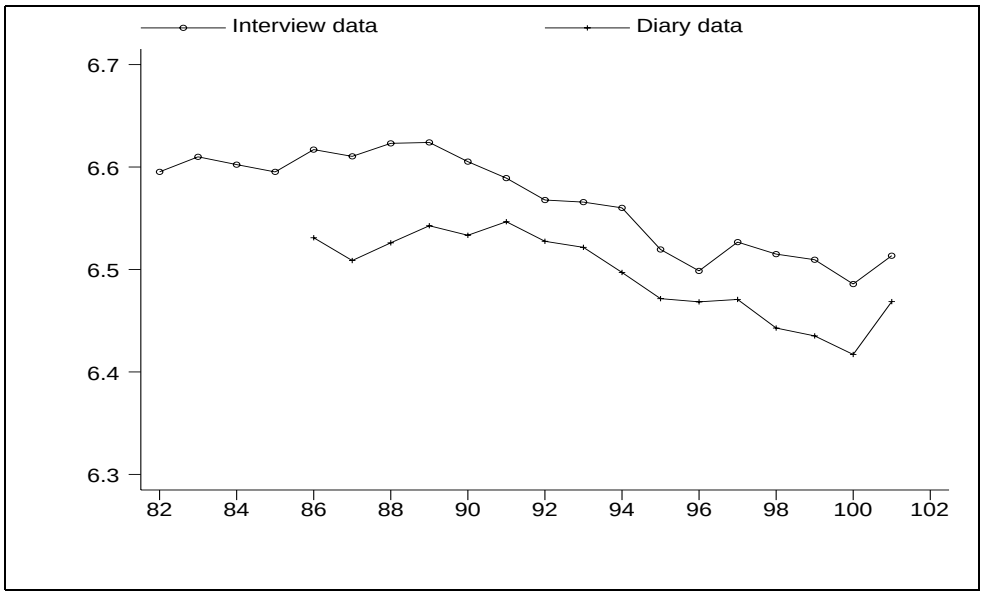

Figure 3: Mean of log family expenditure on non-durable goods (2001 dollars)

Expenditure budgets (that is, the expenditure on each commodity as percentage of expenditure on non-durables) follow the same pattern over time in the two surveys, although they have different levels (see Battistin 2002). The percentage of total expenditure attributable to clothing and footwear, tobacco and alcohol is decreasing (particularly over the 1990s) and is compensated by an increase of expenditure on housing and public services.

4.2 Expenditure inequality

While the pattern of income inequality in the United States during the last twenty years is well documented, the evidence on the evolution of consumption inequality is much less clear. Several researchers have pointed out a rise in consumption inequality over the 1980s using IS data, both within age-cohorts and for the overall population (see, amongst others, Deaton and Paxson 1994). However, during the first half of the 1990s, inequality partially receded for consumer expenditures while for income it continued to rise (see Johnson and Shipp 1995).[10] For this reason, the issue of what happened to consumption inequality during a period characterized by marked increases in income inequality has recently received much attention. This section shows that the inequality pattern emerging from the DS is different from the one obtained using the IS.

We will at first discuss the evidence by cohort and then consider inequality for the all population. Figure 4 presents the evolution of consumption inequality by cohort both for IS and DS data exploiting all observations in each survey year. We find inequality (defined as the variance of log monthly non-durable expenditure) to be higher for DS as compared to IS data. This may be due to respondent issues, but is definitely related to the interview time periods of the two surveys being different: the shorter time period results in data with greater volatility (the DS reference period is one week).

However, the information contained in each sample leads to contradictory results with respect to the trend of inequality over time. As stated above, within cohort inequality from the IS survey presents a mildly increasing pattern in the 1980s for those born between 1930 and 1949, but there is no evidence of any increase after 1990. This result can be directly compared to the numbers reported in Deaton and Paxson (1994) and Blundell *et al.* (2002) where IS information for the 1980s is used. DS inequality is instead increasing over time uniformly for all cohorts. According to what discussed in Section 3, raw information from the two surveys leads to contradictory conclusions in the validation of the PIH.

To improve the readability of the information contained in Figure 4, Table 2 presents the values of the Gini coefficient for the same data. IS inequality remains flat over time for all the considered cohorts, with the exception of the cohort defined by heads born in $1940-49$. Moreover, the cohort born in $1930-39$ presents a mildly increasing pattern during the 1980s. Inequality seems to be most pronounced exploiting DS data for all cohorts.[11] The bump-shaped pattern before retirement

[10] Table A.3 in the Appendix presents values of the Gini coefficient for total family income over the last twenty years, exploiting additional information from the Current Population Survey (CPS) and the Panel Survey of Income Dynamics (PSID). The picture is consistent with the one already reported in the literature, showing a general increase in income inequality over the time covered by this analysis.

[11] The robustness of this result has been further investigated exploiting additional measures of inequality selected from the generalized entropy family (see for example Shorrocks, 1982). The resulting picture is consistent with the one presented here.

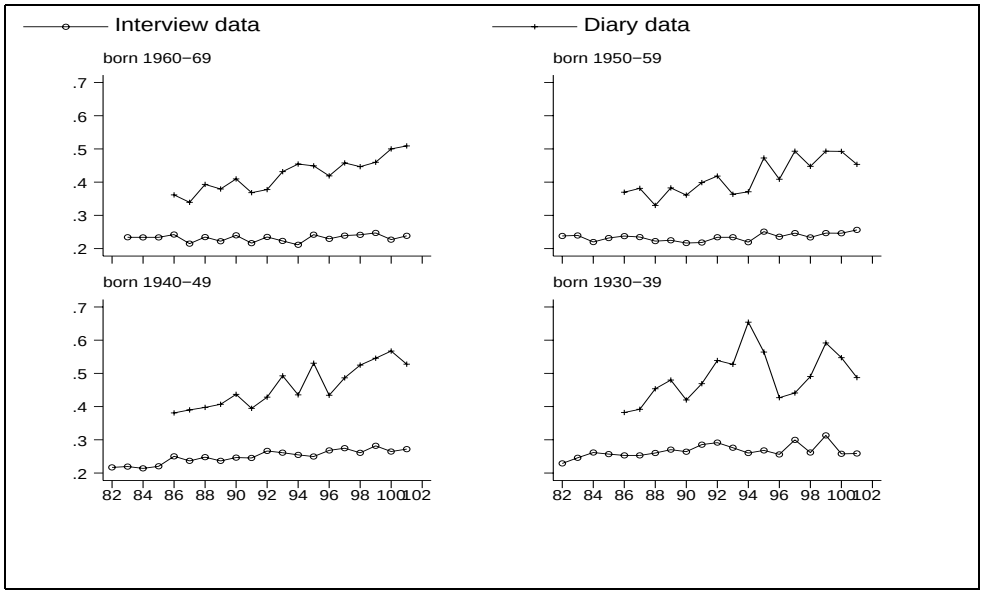

Figure 4: Variance of log family expenditure on non-durable goods by cohort (2001 dollars)

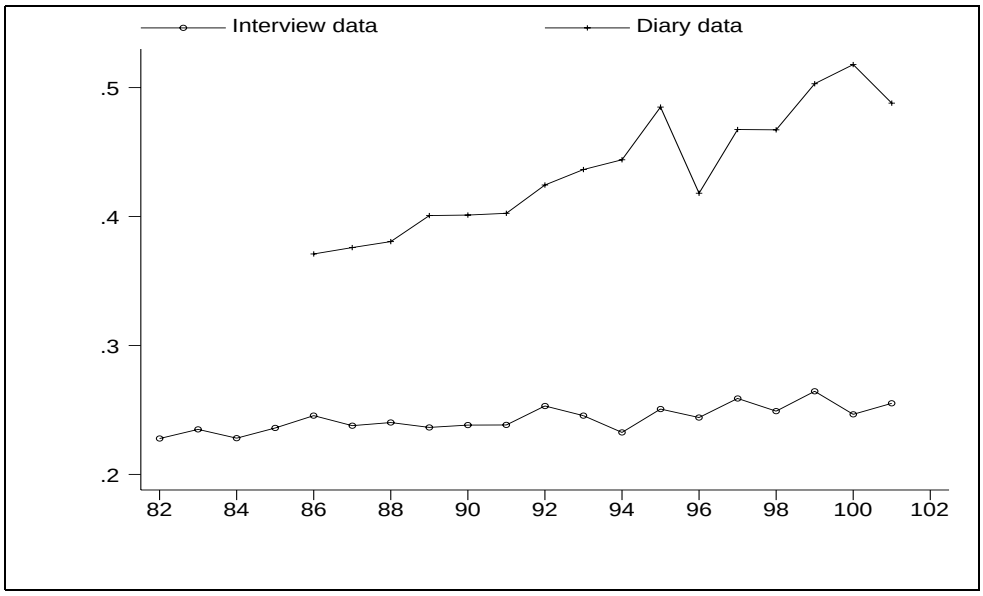

Figure 5: Variance of log family expenditure on non-durable goods (2001 dollars)

Table 2: Gini coefficient for total non-durable expenditure (2001 dollars)

Diary sample					
year	**born 1930-39**	**born 1940-49**	**born 1950-59**	**born 1960-69**	**all**
1986	0.34	0.31	0.31	0.31	0.31
1987	0.33	0.32	0.32	0.29	0.32
1988	0.34	0.32	0.31	0.34	0.33
1989	0.36	0.35	0.32	0.30	0.33
1990	0.35	0.34	0.31	0.32	0.33
1991	0.35	0.33	0.34	0.33	0.34
1992	0.37	0.33	0.32	0.33	0.34
1993	0.38	0.36	0.31	0.33	0.34
1994	0.39	0.35	0.33	0.34	0.35
1995	0.36	0.35	0.36	0.33	0.35
1996	0.37	0.34	0.34	0.34	0.34
1997	0.35	0.37	0.36	0.33	0.35
1998	0.36	0.38	0.34	0.34	0.35
1999	0.37	0.36	0.35	0.34	0.36
2000	0.38	0.37	0.35	0.36	0.36
2001	0.36	0.38	0.35	0.37	0.37

Interview sample					
year	**born 1930-39**	**born 1940-49**	**born 1950-59**	**born 1960-69**	**all**
1982	0.27	0.26	0.26		0.26
1983	0.27	0.26	0.27	0.28	0.27
1984	0.28	0.26	0.26	0.26	0.27
1985	0.28	0.26	0.27	0.27	0.27
1986	0.28	0.27	0.27	0.28	0.27
1987	0.28	0.27	0.26	0.25	0.27
1988	0.28	0.27	0.26	0.27	0.27
1989	0.29	0.27	0.26	0.26	0.27
1990	0.28	0.28	0.27	0.27	0.27
1991	0.30	0.27	0.26	0.26	0.28
1992	0.30	0.28	0.27	0.27	0.28
1993	0.30	0.28	0.27	0.26	0.28
1994	0.29	0.28	0.26	0.26	0.27
1995	0.30	0.28	0.28	0.27	0.27
1996	0.28	0.28	0.27	0.26	0.27
1997	0.31	0.28	0.28	0.27	0.28
1998	0.28	0.28	0.27	0.28	0.28
1999	0.31	0.29	0.27	0.28	0.29
2000	0.29	0.29	0.27	0.26	0.28
2001	0.29	0.29	0.28	0.27	0.28

for those born in $1930-39$ could reflect the effect of an increasing leisure time due to the retirement age.

The corresponding indeces of consumption inequality in the population are presented in Figure 5 and in the last column of Table 2. The resulting pattern for the IS survey is the one already documented in several other papers expoiting this data source: see for example the graphs in Slesnick (2001; Chapter 6) or Krueger and Perri (2001; Figure 1). The variance of log-consumption is 0.22 in 1982, 0.25 in 1993 and 0.25 in 2001. On the other hand, DS inequality presents a markedly increasing pattern over time. A formal test reject the null hyphothesis of constant inequality for the DS over time, while it fails to reject this hyphothesis using IS information for the 1990s.

A first attempt to explain this difference is to describe how marginal changes in expenditures for specific commodities can affect the inequality of total expenditure. This would required the identification of the contribution in overall inequality attributable to each group entering the definition of non-durable consumption. The problem is related to an unique decomposition rule as suggested by Shorrocks (1982), since the inequality contribution assigned to each source can vary arbitrarily depending on the choice of decomposition rule. Particularly important for our purposes is the ability to meaningfully decompose the index into inequality between and within different commodities. The decomposition must be consistent, in the sense that commodities' contribution should add up to the overall amount of inequality.

Table 3 reports non-durable commodities and their percentage contribution to total inequality using the 'natural' decomposition rule

$$Var(Y) = \sum_j Cov(X_j, Y),$$

both for IS and DS 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.[12]

The contribution of food at home and housing and public services is increasing over time for both IS and DS data; the increasing weight of non-durable services for IS inequality is not observed in the DS. The trend for the remaining figures is comparable across the two surveys, although different levels are observed. Food away from home, alcohol and clothing are those commodities presenting a decreasing contribution over time.

5 ACCOUNTING FOR INACCURACIES

Can IS and DS data be exploited together to derive a superior measure of non-durable consumption? The goal of this section is to address this issue by discussing the nature of survey errors that are likely to affect the CEX. In fact, a natural explanation for the different time pattern of means and variances in the two sur-

[12]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 3: Factors contribution as percentage of total inequality

Interview	1982-85	1986-89	1990-92	1993-95	1996-98	1999-2001
Food and non-alcoholic beverages at home	10.62	10.37	11.17	10.43	10.87	11.09
Food and non-alcoholic beverages away	10.14	9.71	8.73	9.11	9.01	8.64
Alcoholic beverages (at home and away)	4.23	3.22	2.65	2.46	2.61	2.38
Non-durable goods and services	11.97	12.77	12.62	13.70	13.27	14.09
Housing and public services	7.72	7.88	9.14	10.22	11.39	12.34
Tobacco and smoking accessories	0.61	0.43	0.55	0.54	0.51	0.76
Clothing and footwear	16.96	18.35	17.66	15.21	13.85	12.97
Heating fuel, light and power	4.80	4.27	4.19	4.40	4.26	4.45
Transport (including gasoline)	32.95	33.00	33.30	33.93	34.24	33.28

Diary	1982-85	1986-89	1990-92	1993-95	1996-98	1999-2001
Food and non-alcoholic beverages at home		9.23	9.22	9.52	9.61	10.20
Food and non-alcoholic beverages away		11.85	10.71	9.57	9.37	10.46
Alcoholic beverages (at home and away)		2.58	2.10	1.83	1.62	1.86
Non-durable goods and services		18.25	18.59	17.54	16.95	17.73
Housing and public services		9.91	10.80	12.19	13.25	15.49
Tobacco and smoking accessories		0.52	0.37	0.48	0.52	0.75
Clothing and footwear		15.54	15.21	13.78	12.11	11.31
Heating fuel, light and power		8.92	8.75	7.98	8.08	7.97
Transport (including gasoline)		23.20	24.27	27.12	28.49	24.22

veys is the aggregate result of inaccuracies affecting the IS and DS reporting of expenditures.

Lyberg *et al.* (1997), Bound *et al.* (2001) and Browning *et al.* (2002) review data quality problems characterizing survey measurements. The main lesson from their findings is that inaccuracies mainly come from those non-durable commodities each survey is not targeted to: frequently purchased smaller items and services (IS) and large expenditures occurring on a regular basis (DS). The aggregate effect of these inaccuracies is likely to vary over time, because it depends on significative changes in the structure of the two surveys (i.e. design and collecting strategies) and on time-in-sample effects (i.e. people might change their disposition to answer accurately or answer at all). As a matter of fact, there has been a deterioration of the correspondence between CEX aggregates and PCE over the 1990s (see McCarthy *et al.* 2002).

Modelling the effects of survey errors generally requires strong assumptions. In fact, although the error on a given variable is often assumed to be independent of the true level of that and of all other variables, this assumption reflects convenience rather than conviction. In what follows we will depart from any model-based assessment of the error. Instead, we will use information from the most reliable survey as validation data to assess the quality of the other survey.

Internal validation data are to be preferred over validation data coming from external surveys. In this sense, the two survey components from the CEX represent a unique example for the United States. However, since the IS and the DS address independent samples of households, diary and recall values are not observable for the same survey households. The drawbacks of this design discussed in Section 6.

5.1 Nature and consequences of survey errors

Collection methodology differences between the two survey components of the CEX certainly represent the main explanation for the evidence presented so far. While the DS collects detailed disaggregated data and then sums these up to obtain total spending, the IS asks a global retrospective question about totals. Differences in levels and inequality indices might be determined by different expenditure estimates on each commodity as a result of this aggregation.

Respondents' partial recollection of past events (mainly related to less important purchases) and/or telescoping effects for the month nearest to the interview are factors likely to affect the accuracy of available information in the IS. In fact, recall is certainly a complex cognitive process in the collection of consumer expenditures data through household interviews. The important implications of forgetfulness and telescoping in large-scale recall surveys have been largely discussed in the literature (see, among others, Neter and Waksberg 1964). Moreover, measurement effects in self vs proxy responses, differences in the interpretation of questions, inability or unwillingness of respondents to provide full information or interviewers' effect on data collection might play a non-negligible role in determining the quality of IS information.[13]

Under reporting of expenditures is likely to be limited in diary surveys, although additional data collection effects might play an important role in the definition of monthly aggregates. Household's expenditures recorded during a limited period of

[13]Both surveys accept proxy responses from any eligible household member who is at least 16 years old if an adult is not available after a few attempts to contact that person.

time (two weeks) might give a misleading impression of its underlying consumption pattern over a longer period (a month). Since commodities entering total non-durable consumption are purchased with a different frequency, zero recorded expenditures might reflect preferences in the frequency of purchasing rather than preferences in consumption behavior.

Besides, other conjectures on the sources of zero reported expenditures include under reporting in an acknowledged purchase and item non-response (i.e. not reporting a purchase that was made). Several studies 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. For example, Turner (1961) and Silberstein and Scott (1991) find that the average of reported expenditure using 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 IS.[14]

An additional explanation for the evidence of the previous section are periodic changes in the survey instruments over the years, both for the IS and the DS. For example, a new diary form with more categories and expanded use of cues for respondents has been introduced in the DS since 1991. The diary instrument for the DS is respondent-filled, and reporting levels and accuracy are known to depend on the adopted diary format (Tucker 1992). On the other hand, the definition of expenditure categories collected in the IS has changed over time (notably for food, as discussed in the Appendix) because of newly collected items or because some of them have been supplanted by new ones.

It what follows, we will refer to *measurement error* as the difference between the actual value of expenditure and the value reported by respondents. From what discussed above, data collection effects and respondents' partial recollection of past events are likely to make such a difference not identically distributed across households and over time.[15] Moreover, the assumption of classical measurement error in survey measurements (i.e. zero mean error independent of the true unobserved variable and of all other variables) has been largely criticized and is usually made for ease of estimation rather than for any theoretical conviction (see Rodgers *et al.*

[14]Table A.4 in the Appendix presents reporting rates for IS and DS data, that is the proportion of non-zero expenditures for a specific commodity. Year-to-year changes in this indicator provide useful monitors of survey performance over time. The frequency of purchasing is generally lower in the DS sample, with the only exception of expenditure on food away from home (which by definition does not include expenditures on vacation). However, 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.

[15]Note that if the distribution of the measurement error 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. To give a flavor of such a problem, assume that the error affecting reports of spending on commodity X is multiplicative and that its intensity is given by a parameter σ (see Chesher and Schluter 2001). If we assume independence between X and the reporting error process, a second-order approximation for the Gini coefficient of the error-contaminated consumption is given by

$$G_X + \sigma^2 \frac{E[X^2 f_X(x)]}{E[X]},$$

where f_X and G_X are the density and the Gini coefficient associated to X, respectively. It follows that the 'distance' between the true and the observed Gini coefficient 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 the distribution of X is heavily right skewed).

1993, Torelli and Trivellato 1993 and Pischke 1995). In particular, Battistin *et al.* (2003) provide some evidence for the case in which the magnitude of the measurement error is endogenously determined by the real amount of expenditure (with higher expenditure levels associated with larger errors), so that the independence assumption is no longer valid.

It follows that modelling measurement errors affecting IS and DS information would require relatively strong assumptions. We instead depart from any model-based approach and exploit information from the most reliable survey as a validation source for the other survey.

Our procedure develops along the following lines. Since the IS and the DS are explicitly designed to obtain reliable measures of expenditure on different commodities, the error affecting total non-durable expenditure results from those categories of each survey component that are considered 'less reliable'. Those commodities either having regular periodic billing or involving major outlays easily recalled for a period of three months or longer are better described using IS data. On the other hand, those non-durable commodities referring to frequently purchased and smaller items are presumably more reliable in the DS survey. The next section discusses the validity of such an assumption by presenting the evidence from a broad range of studies.

5.2 Choosing among alternative collection methodology

Throughout the analysis we will make the following assumptions.

Condition 1. Either IS or DS data identify correctly (i.e. report without measurement error) the actual spending on non-durable commodities.

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

Condition 1 builds on the well-established conviction that diary and recall surveys provide reliable consumption information on different commodities. Neither survey component alone can be exploited to get an accurate measure of non-durable spending: it is the joint use of IS and DS information that leads to accurate totals.

Condition 2 defines the aggregation rule one should follow in pooling IS and DS information, and it is therefore more debatable. As discussed earlier, the question of which survey component provides more accurate estimates for non-durable items is an issue of longstanding concern in the design of expenditure surveys (see Browning *et al.* 2002). Reliability of expenditure data has been assessed either by examining how aggregate spending on a certain commodity compares with aggregate spending from national accounts (see for example Banks and Johnson 1998 and Slesnick 2001), or by looking at results from controlled experiments (see for example Winter 2002). The remaining of this section discusses the implications of these findings for the aggregation rule used in this paper.

Even if potentially the sign of the bias in recall and diary data could be in both directions depending on different commodities (over or under reporting of true expenditures), 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.

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 IS survey. Indeed, exploiting validation data from the national accounts IS expenditures for *transports* and *fuel* 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. Diaries were found to give more reliable information about alcohol consumption than recall data (see Poikolainen and Kakkainen 1983 and Atkinson *et al.* 1990); comparisons of tobacco expenditures based on mean squared error methods exploiting national accounts 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, IS data seem to be more reliable for costly and salient apparel items (with quite variable results exploiting different methods of source selection), but DS data generally capture more apparel spending (Silberstein and Scott 1991).

There is some evidence that diaries are the most reliable source to measure purchases on *food* made away from home (Stanton and Tucci 1982). However, it is also well documented that the quality of reporting is higher in the first week of the diary (see Turner 1961 and Figure A.11 in the Appendix). On the other hand, as explained in Section 2.1, IS information of food at home is derived as the difference between the usual spending at grocery stores and how much of this amount was for non-food items.

According to the evidence from the studies summarized in this section, Table 4 presents the aggregation rule we will follow to obtain aggregate values of total expenditure on non-durables. Branch and Jayasuriya (1997) discusses how consumer expenditures from the two surveys are chosen by the BLS for publication. It is worth noting that, although the level of aggregation considered in their paper is finer than the one exploited here, the classification procedure suggested in Table 4 broadly reflects the one currently being used by the BLS. Not surprisingly, DS data are exploited as the reference source for expenditures on grocery items and personal care, entertainments and other services; IS data to identify expenditures on those components having regular periodic billing or involving major outlays. In what follows these two sets of commodities will be denoted by $\mathcal{I}$ and $\mathcal{D}$, respectively.[16]

6 ERROR CORRECTION

The classification procedure discussed in the previous section provides a rule to define a superior measure of total consumption by exploiting together IS and DS information. Nevertheless, straightforward pooling cannot be implemented since diary and recall expenditures are not observed for the same survey households.

The aim of this section is to formalize the restrictions presented in Table 4. Although by sampling design Conditions 1 and 2 do not allow us to fully identify the

[16] According to the classification procedure suggested, the 'true' unobserved value of expenditure is a mixture of observed expenditures from the IS and the DS. Conditions 1 and 2 impose zero/one restrictions on the weights of this mixture, so that either IS or IS expenditures are considered.

Table 4: Survey selection

Commodities 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
Non-Durable Goods and Services	Diary

Commodities in $\mathcal{I}$	**Survey**
Housing and Public Services	Interview
Tobacco and Smoking Accessories	Interview
Clothing, Footwear and Services	Interview
Heating Fuel, Light and Power	Interview
Transportation (including gasoline)	Interview

distribution of total non-durable consumption, they are identifying restrictions concerning the first moment of this distribution. The identification of higher moments (and percentiles) would require to model the relationship between true expenditures and reporting errors. However, the second moment of this distribution can be bounded without imposing any restriction on the nature of the error by means of the Cauchy-Schwartz inequality. To a certain extent, this turns out to be an informative test for the relationship in (1) implied by the PIH.

Throughout the remaining of this paper we will focus on within cohort means and variances. The issue of how IS and DS information can be jointly used to characterized the most likely pattern of consumption inequality for the entire population is further elaborated in Attanasio *et al.* (2003).

6.1 A potential outcomes approach

Let

$$(X_j^r, X_j^d)$$

be the two outcomes that result from being interviewed on commodity X_j via a recall or a diary based questionnaire, respectively, and let X_j^* be the true expenditure on the same commodity (which is not observed). Commodities X_j's are those that have been described in the previous section and reported in Table 4.

Clearly, the difference between X_j^r and X_j^d is informative about the effect of reporting expenditures exploiting recall rather than diary based questionnaires. Since the two surveys refer to separate samples, an identification problem arises from the fact that - by design - only one of these measurements is observed on each household. In fact, they represent potential outcomes from using alternative survey instruments to collect information on household consumption. The measurement error on each commodity would be identified if we could observe the counterfactual expenditure for each household, that is what the same household would have reported had it participated the other survey.

Let C^* be total expenditure on all non-durable commodities X_j's

$$C^* \quad = \quad C_1^* + C_2^*,$$

where C_1^* and C_2^* represent expenditure on commodities in $\mathcal{I}$ and in $\mathcal{D}$, respectively, as defined by Table 4

$$C_1^* = \sum_{j \in \mathcal{I}} X_j^*, \qquad C_2^* = \sum_{j \in \mathcal{D}} X_j^*.$$

Total expenditure is not observable to the analyst. Rather two error affected measurements of C^* are observed, representing the aggregate expenditures that result from using IS or DS data

$$\begin{aligned} C^r &= C_1^r + C_2^r, \\ C^d &= C_1^d + C_2^d, \end{aligned}$$

where (C_1^r, C_1^d) and (C_2^r, C_2^d) are IS and DS expenditures on commodities in $\mathcal{I}$ and in $\mathcal{D}$

$$\begin{aligned} C_1^r &= \sum_{j \in \mathcal{I}} X_j^r, & C_2^r &= \sum_{j \in \mathcal{D}} X_j^r, \\ C_1^d &= \sum_{j \in \mathcal{I}} X_j^d, & C_2^d &= \sum_{j \in \mathcal{D}} X_j^d. \end{aligned}$$

By means of Conditions 1 and 2, the measurement error affecting the aggregate expenditures C^r and C^d depends on the measurement error on different subsets of commodities entering total non-durable expenditure (i.e. those commodities either in $\mathcal{D}$ or in $\mathcal{I}$, respectively). That is, total expenditure from IS data can be written as

$$C^r = C_1^* + C_2^r,$$

where the last expression follows since expenditures on commodities in $\mathcal{I}$ are not affected by measurement error by assumption ($X_j^r = X_j^*, \forall j \in \mathcal{I}$). By analogy, it follows that

$$C^d = C_1^d + C_2^*,$$

since $X_j^d = X_j^*, \forall j \in \mathcal{D}$. Accordingly, the error due to recall and diary interviews can be written as

$$\varepsilon^r = C_2^r - C_2^*, \tag{3}$$
$$\varepsilon^d = C_1^d - C_1^*, \tag{4}$$

respectively. Any difference in the distribution of these errors over time is responsible for the different pattern of means and variances observed in raw IS and DS data.

The main points arising from the last two expressions can be summarized as follows. Firstly, the mean of the IS (DS) error can be written as a linear combination of means of errors for commodities in $\mathcal{D}$ ($\mathcal{I}$, respectively). Since error means on all non-durable commodities are identifiable because of Conditions 1 and 2, the mean values of (3) and (4) are identified by

$$E(\varepsilon^r) = E(C_2^r) - E(C_2^d), \tag{5}$$
$$E(\varepsilon^d) = E(C_1^d) - E(C_1^r), \tag{6}$$

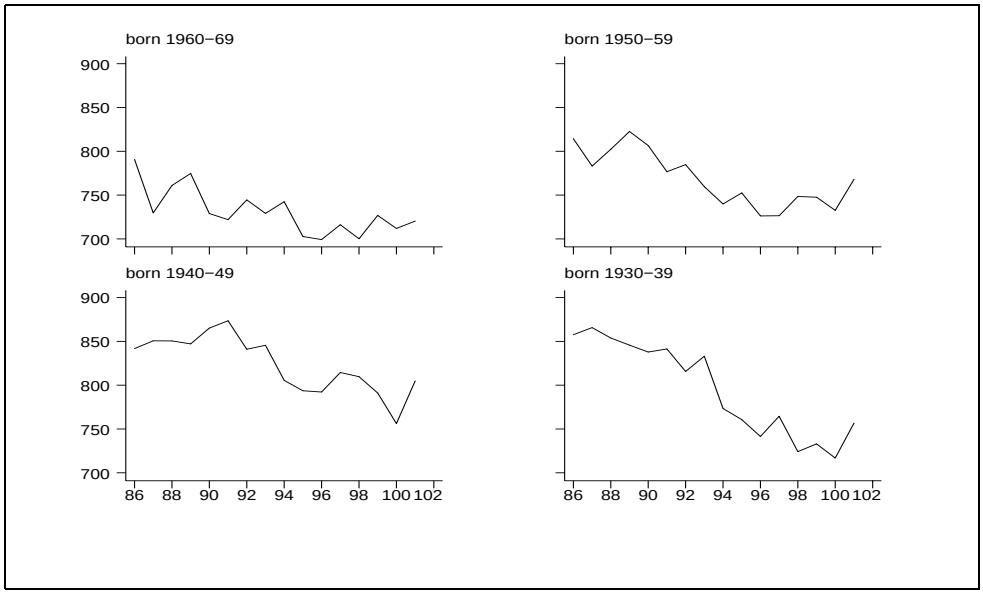

Figure 6: Mean of family expenditure on non-durable goods by cohort after correction (2001 dollars)

respectively. Accordingly, Conditions 1 and 2 are identifying restrictions concerning the first moment of the distribution of total non-durable consumption (see Section 6.2).

Secondly, recall and diary errors are likely to be correlated with C^*, since they depend on a common set of commodities. This implies that the difference between the real variance of consumption and the observed values of this variance in IS and DS data depends on the variance of the terms in (3) and (4) and on their correlation with C^*, which are not observable. Therefore the variance of C^* cannot be estimated without imposing additional restrictions on the error structure (see Section 6.3).

6.2 Consumption levels

Differences in mean expenditure values for all commodities entering total non-durable expenditure have been already discussed in Section 4.1 (see also Figures A.2-A.10 in the Appendix). By means of Conditions 1 and 2, they can be interpreted as the effect of collecting expenditure information using the less appropriate methodology, depending on whether the considered commodity belongs to $\mathcal{D}$ or $\mathcal{I}$.

Mean aggregate errors for IS and DS data (that is the quantities in (5) and (6), respectively) are presented in Table 5, separately by cohort and over time. More precisely, figures for IS and DS errors as proportion of total non-durable expenditure are reported, that is

$$E(\varepsilon^r)/E(C^*) \qquad E(\varepsilon^d)/E(C^*)$$

Table 5: Survey errors as a proportion of total non-durable expenditure

year	**born 1960-69**		**born 1950-59**		**born 1940-49**		**born 1930-39**		**all**	
	IS	DS	IS	DS	IS	DS	IS	DS	IS	DS
1986	-0.05	-0.12	-0.03	-0.10	-0.01	-0.10	-0.02	-0.11	-0.02	-0.11
1987	0.02	-0.11	-0.02	-0.11	-0.03	-0.09	-0.01	-0.08	-0.01	-0.10
1988	-0.05	-0.09	-0.06	-0.10	-0.05	-0.08	-0.04	-0.08	-0.05	-0.09
1989	-0.07	-0.07	-0.05	-0.10	-0.05	-0.10	-0.07	-0.06	-0.06	-0.08
1990	-0.02	-0.06	-0.07	-0.09	-0.06	-0.06	-0.07	-0.05	-0.06	-0.07
1991	-0.03	-0.06	-0.08	-0.03	-0.07	-0.05	-0.06	-0.06	-0.06	-0.05
1992	-0.09	-0.04	-0.08	-0.05	-0.05	-0.08	-0.03	-0.04	-0.07	-0.05
1993	-0.07	-0.03	-0.05	-0.05	-0.06	-0.05	-0.06	-0.03	-0.06	-0.04
1994	-0.06	-0.04	-0.05	-0.03	-0.03	-0.05	-0.04	0.01	-0.04	-0.03
1995	-0.04	-0.06	-0.07	-0.03	-0.07	-0.02	-0.08	0.02	-0.06	-0.02
1996	-0.05	-0.02	-0.06	-0.05	-0.07	-0.04	-0.08	0.02	-0.06	-0.03
1997	-0.05	-0.02	-0.04	-0.03	-0.06	-0.01	-0.07	-0.06	-0.05	-0.03
1998	-0.07	-0.06	-0.08	-0.01	-0.04	-0.08	-0.06	-0.02	-0.06	-0.04
1999	-0.08	-0.03	-0.07	-0.06	-0.06	-0.06	-0.05	-0.04	-0.06	-0.05
2000	-0.07	-0.05	-0.08	-0.02	-0.04	-0.04	-0.08	-0.05	-0.07	-0.04
2001	-0.08	-0.05	-0.07	-0.01	-0.08	-0.02	-0.10	-0.06	-0.08	-0.03

respectively, where the mean of total expenditure is estimated by

$$\begin{aligned} E(C^*) &= E(C_1^*) + E(C_2^*) \\ &= E(C_1^r) + E(C_2^d). \end{aligned} \tag{7}$$

Negative (positive) numbers in the table can be interpreted as the proportion of expenditure under (over) reported in the IS or in the DS. Apparently, this proportion varies across cohorts, with younger cohorts more likely to have higher errors than older cohorts, both in the IS and in the DS. The amount of misreporting has an almost stationary distribution over time for the IS, while it varies for the DS. In particular, the DS appears to become more accurate over time.[17]

Figure 6 presents estimated mean expenditures by cohort using (7) and can be compared to the results presented in Figure 2. It is evident that, even after the correction, consumption profiles are still sharply decreasing during the first half of the 1990s.[18]

6.3 Consumption inequality

The goal of this section is to derive the analogue of Figure 4 once the effect of reporting errors is accounted for. Throughout our analysis, we will consider figures for the squared coefficient of variation of total expenditure instead of figures for the variance of logs. The reason for this choice will be clear from what follows.

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

$$Var(C^*) = Var(C_1^*) + Var(C_2^*) + 2Cov(C_1^*, C_2^*). \tag{8}$$

The ratio of the previous quantity to the squared mean of total consumption represents a first order approximation for the variance of log consumption.[19] Conditions 1 and 2 are identifying restrictions only for the first two terms of the previous expression, using either IS or DS data. On the other hand, the covariance term cannot be identified from the available information since by definition $\mathcal{D} \cap \mathcal{I} = \emptyset$

[17] Given the restrictions imposed so far, the measurement error affecting the reporting of non-durable commodities cannot be further characterized. If we observed both recall and diary outcomes for commodity X on the same household, according to Condition 1 and Condition 2 the error on that commodity would be (non-parametrically) identified. In fact, by writing

$$\xi = x - x^*,$$

we could identify the distribution of ξ by taking the difference between observed diary and recall outcomes depending on the rule discussed in Table 4. In our case, although the first moment of this distribution can be easily recovered, the identification of higher moments and percentiles would require to know the structure of correlation between x and x^*, which is not observable.

[18] Following the same lines of the previous footnote, it can be shown that the distribution of C^* is not identified. Additional restrictions are required. Alternatively, bounds on this distribution could be derived exploiting Frechet's or Markov's inequalities.

[19] A Taylor expansion of lnC^* in a neighborhood of its mean gives

$$lnC^* = lnE(C^*) + \frac{C^* - E(C^*)}{E(C^*)} - \frac{[C^* - E(C^*)]^2}{2E(C^*)^2} + error,$$

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

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

so that true expenditures on commodities in $\mathcal{D}$ and $\mathcal{I}$ cannot be observed for the same individuals.

Two measurements of the covariance between C_1^* and C_2^* are actually observable in our data. The first one is derived by taking the covariance between commodities in $\mathcal{D}$ and commodities in $\mathcal{I}$ from the IS; the second one is derived from the DS. Figure 7 presents the squared coefficient of variation obtained when the unknown covariance is estimated by using IS or DS information and the mean of total consumption is estimated by (7). Apparently, both IS and DS figures suggest increasing within cohort inequality over time.

However, without additional assumptions neither observed covariance is a consistent estimator for the covariance of interest. In fact, using (3) and (4) the observed covariance in the IS sample can be written as

$$\begin{aligned} Cov(C_1^r, C_2^r) &= Cov(C_1^*, C_2^r) \\ &= Cov(C_1^*, C_2^*) + Cov(C_1^*, \varepsilon^r), \end{aligned} \tag{9}$$

and, by analogy, the observed covariance in the DS sample is

$$\begin{aligned} Cov(C_1^d, C_2^d) &= Cov(C_1^d, C_2^*) \\ &= Cov(C_1^*, C_2^*) + Cov(C_2^*, \varepsilon^d). \end{aligned} \tag{10}$$

Figure 8 presents the evolution of IS and DS covariances over time and separately by cohort. The difference between the two lines of each panel reflects the difference between the last terms in (9) and in (10). These two terms are not identified from observed data. Overall there is evidence that DS covariances are higher and present different trends than IS covariances. A formal test for the equality of observed covariances in IS and DS data rejects this hypothesis over the time period 1986-2001. Similarly, the inequality indices presented in Figure 7 are found to be statistically different (but both statistically increasing).

7 BOUNDS ON INEQUALITY

7.1 Definitions

As discussed in the last section, additional structure is required to identify the covariance term in (8). For example, Attanasio *et al.* (2003) consider alternative identification strategies based on the specification of a demand system for C_1^* and C_2^*. In this paper we take a different route and we investigate whether informative bounds on the within cohort variance can be derived from fairly general assumptions on IS and DS errors.

The strategy we will take consists of three different steps. We will at first use Cauchy-Schwartz bounds resulting from the following equality

$$|Cov(C_1^*, C_2^*)| = |\rho^*| \sqrt{Var(C_1^*) Var(C_2^*)}, \tag{11}$$

where ρ^* represents the Pearson correlation index between C_1^* and C_2^*. The well-known relationship

$$|\rho^*| \leq 1 \tag{12}$$

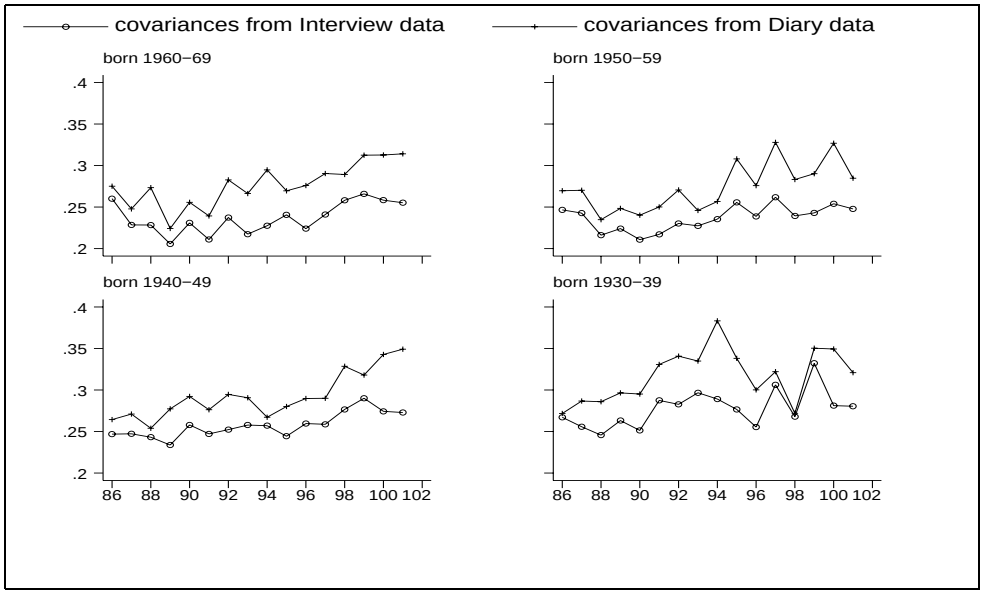

Figure 7: Squared coefficient of variation of family expenditure on non-durable goods by cohort using observed covariances (2001 dollars)

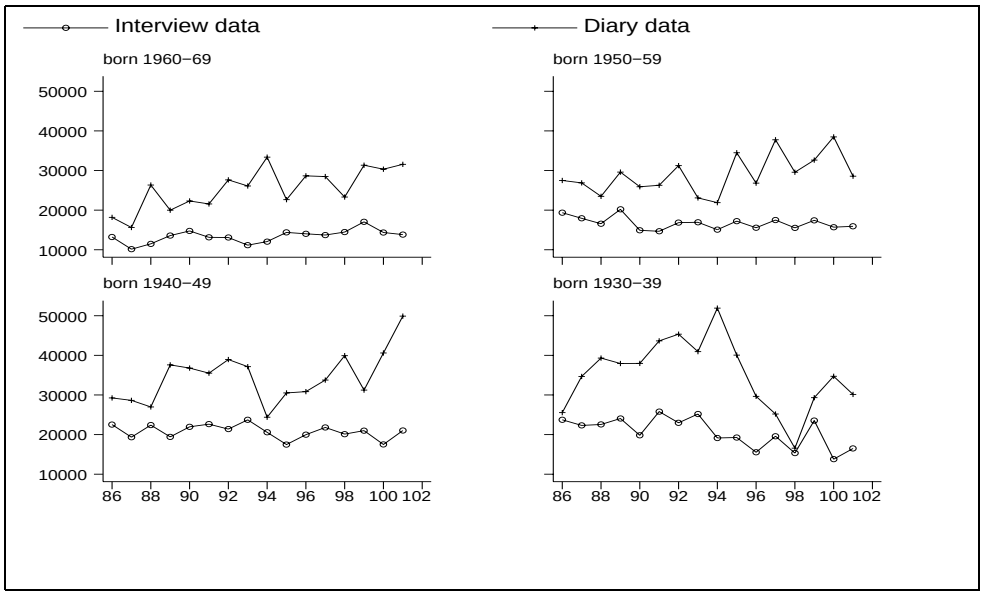

Figure 8: Observed covariances by cohort (2001 dollars)

defines bounds for the within cohort covariance over time, since the product of the variances in (11) is identifiable using the restrictions in Table 4.

However, although the interval $[-1, +1]$ is often considered the reference range for evaluating the strength of empirical correlation, it is well known that the upper and the lower bounds can only be achieved if the marginal distributions of C_1^* and C_2^* are *linearly* related (indeed, the Pearson coefficient detects only linear correlation!). If the two distributions cannot be linearly related, the maximum attainable absolute correlation coefficient is *lower* than one (see for example the discussion by Shih and Huang 1992).

More formally, let F_1^* and F_2^* be the cumulative distribution functions of C_1^* and C_2^*, respectively, and let the corresponding inverse be defined as

$$F_j^{*-1}(t) = inf\{x : F_j^*(x) \geq t\}, \qquad j = 1, 2$$

for $0 \leq t \leq 1$. Then the correlation between C_1^* and C_2^* is bounded by

$$corr(F_1^{*-1}(W), F_2^{*-1}(1-W)) \leq \rho^* \leq corr(F_1^{*-1}(W), F_2^{*-1}(W)), \tag{13}$$

where W is a random variable distributed uniformly between zero and one. Note that, if C_1^* can be written as a linear function of C_2^*, then the previous expression collapses to the Cauchy-Schwartz bound in (12).[20] Bounds in (13) can be estimated by taking the correlation between percentiles of the empirical distributions of C_1^* and C_2^* which, by means of Conditions 1 and 2, are identified. They represent the tightest bounds on the correlation coefficient attainable without imposing further restrictions on the correlation of interest.

Bounds for the within cohort inequality resulting from this procedure are presented in Figure 9. To assess the importance of sampling variability, figures from bootstrapped 95% confidence intervals will be discussed throughout this section. As expected, results are not informative on the pattern of within cohort inequality over time.

Bounds in (13) can be improved by looking at the partial correlation between C_1^* and C_2^* once a set of regressors Z is controlled for. The expression in (8) can be written as the sum of the following within-group and between-group components

$$Var(C^*) \quad = \quad E_Z Var(C^*|Z) + Var_Z E(C^*|Z),$$

which are defined as

$$\begin{aligned} Var_Z E(C^*|Z) &= Var_Z\{E(C_1^*|Z) + E(C_2^*|Z)\}, \\ E_Z Var(C^*|Z) &= E_Z\{Var(C_1^*|Z) + Var(C_2^*|Z) + 2Cov(C_1^*, C_2^*|Z)\}, \end{aligned}$$

where Z represents a set of family characteristics *common* across the two surveys. The first component is identified by assumption, since it represents the variance across groups defined by Z of total mean expenditure as it results from (7). The second component is not identified because it contains the within group covariance between C_1^* and C_2^*.

This covariance can be bounded along the same lines of what discussed above. To fix ideas, in what follows Z will represent *total family income* and households will be split into *five* mutually exclusive groups defined by income deciles (20% or

[20] Shih and Huang (1992) provide several examples where the bounds implied by (13) are very different from the 'natural' bounds implied by (12).

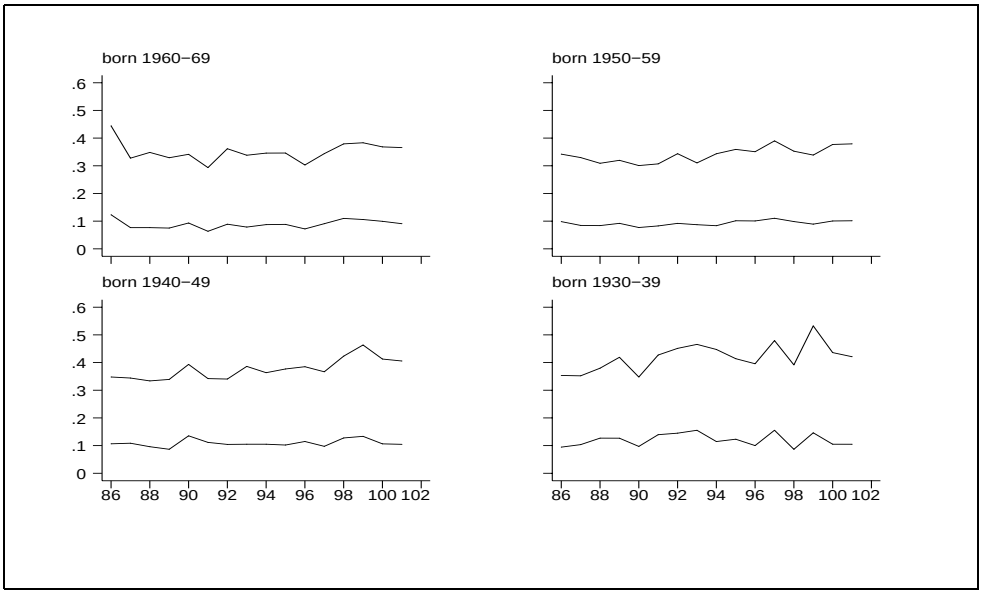

Figure 9: Squared coefficient of variation of family expenditure on non-durable goods by cohort (2001 dollars) - bounds using (13)

less, 20-40%, 40-60%, 60-80%, 80% or more). The within group correlation between C_1^* and C_2^* is then bounded using the analogue of (13) once income is controlled for. However, the resulting bounds are still not informative on the pattern of inequality over time and therefore are not presented here.[21]

Tighter bonds on these covariances can be derived moving from the following evidence. By using the relationships in (9) and (10), observed IS and DS covariances can be first differenced over time to get

$$\begin{aligned} \Delta Cov(C_1^r, C_2^r|Z) &= \Delta Cov(C_1^*, C_2^*|Z) + \Delta Cov(C_1^*, \varepsilon^r|Z), \\ \Delta Cov(C_1^d, C_2^d|Z) &= \Delta Cov(C_1^*, C_2^*|Z) + \Delta Cov(C_2^*, \varepsilon^d|Z). \end{aligned}$$

Table 6 presents results from a bootstrapped test at the 95% confidence level for the null hypothesis

$$\Delta Cov(C_1^r, C_2^r|Z) - \Delta Cov(C_1^d, C_2^d|Z) = 0.$$

For the sake of completeness, results are reported also for the test referred to the covariance in levels. Numbers in the table refer to the proportion of income groups for which the hypothesis is rejected, separately by cohort and over time. Accordingly, 0.20 means that for *one* out of the five income groups considered the null hypothesis is rejected, 0.40 means that for *two* out of five groups the null hypothesis is rejected, and so on.

Although IS and DS covariances appear to have different levels even controlling for Z (see the top panel of the table), overall results support the hypothesis of *stationary* difference between observed IS and DS covariances over time once Z is

[21] Similar results are found by controlling for a richer set of regressors. Results are available on request.

Table 6: Percentage of groups for which the null hypothesis is rejected

$$H_0: \; Cov(C_1^r, C_2^r|Z) - Cov(C_1^d, C_2^d|Z) = 0$$

year	**born 1960-69**	**born 1950-59**	**born 1940-49**	**born 1930-39**
1986		0.40	0.40	
1987		0.60		0.40
1988	0.20	0.20	0.20	0.20
1989		0.20	0.40	0.40
1990	0.20	0.20	0.20	0.40
1991		0.80	0.40	
1992	0.20		0.60	0.20
1993	0.20	0.40	0.20	0.40
1994	0.60	0.60	0.20	0.80
1995	0.60	0.80	0.20	0.20
1996	1.00	0.40	0.20	0.40
1997	0.40	0.40	0.60	0.20
1998	0.20	1.00	0.40	0.20
1999	0.80	1.00	0.40	0.20
2000	1.00	0.80	0.60	0.20
2001	1.00	0.60	0.80	0.20

$$H_0: \; \Delta Cov(C_1^r, C_2^r|Z) - \Delta Cov(C_1^d, C_2^d|Z) = 0$$

year	**born 1960-69**	**born 1950-59**	**born 1940-49**	**born 1930-39**
1987	0.20			
1988				
1989				
1990				
1991				
1992				
1993				
1994		0.20		
1995				
1996				
1997				
1998				
1999				
2000				
2001			0.20	

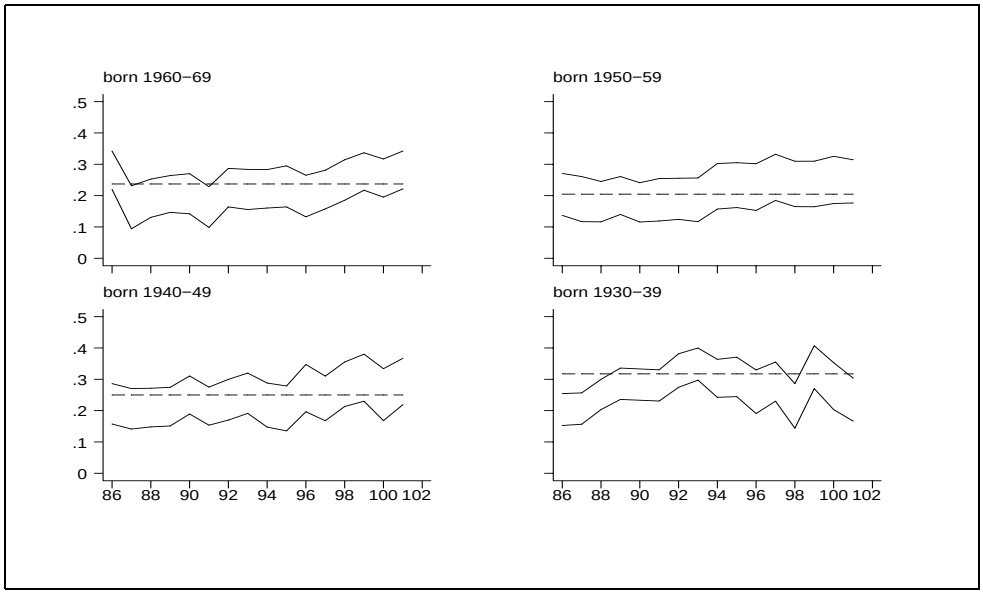

Figure 10: Squared coefficient of variation of family expenditure on non-durable goods by cohort (2001 dollars) - bounds using (13) and Condition 3

controlled for. This represents a necessary condition for the following assumption to be satisfied.

Condition 3. $\Delta Cov(C_2^*, \varepsilon^d | Z) = 0$ and $\Delta Cov(C_1^*, \varepsilon^r | Z) = 0$.

The previous assumption together with the bounds implied by (13) restricts the range of values of the partial covariance between C_1^* and C_2^* given Z. Although the level of the covariance of interest is unknown, the series of its changes over time turns out to be identified by the series of variations in IS and DS data. Condition 3 defines bounds on the inequality index over time. Technical details on how these bounds are derived are presented in the Appendix.

Results are reported in Figure 10. Although Condition 3 turns out to be very powerful in tightening the natural bounds implied by (13), it is not fully informative on the pattern of within cohort inequality over time. Apparently, inequality grows over time for the cohort of those born in 1930-39 and for those born in 1930-39 before 1993 (that is before the retirement age). Inequality for the two remaining cohorts (in particular for those born in 1950-59) is not statistically increasing at the confidence level considered.

7.2 Discussion

A sensitivity analysis of these results with respect to the true value of the correlation coefficient $\rho^*(Z) = corr(C_1^*, C_2^* | Z)$ is presented in Figures 11-13. More precisely, the same procedure described in the previous section is derived exploiting the additional assumption that the correlation between C_1^* and C_2^* net of Z is lower than 0.90, 0.85 and 0.80 in absolute value.

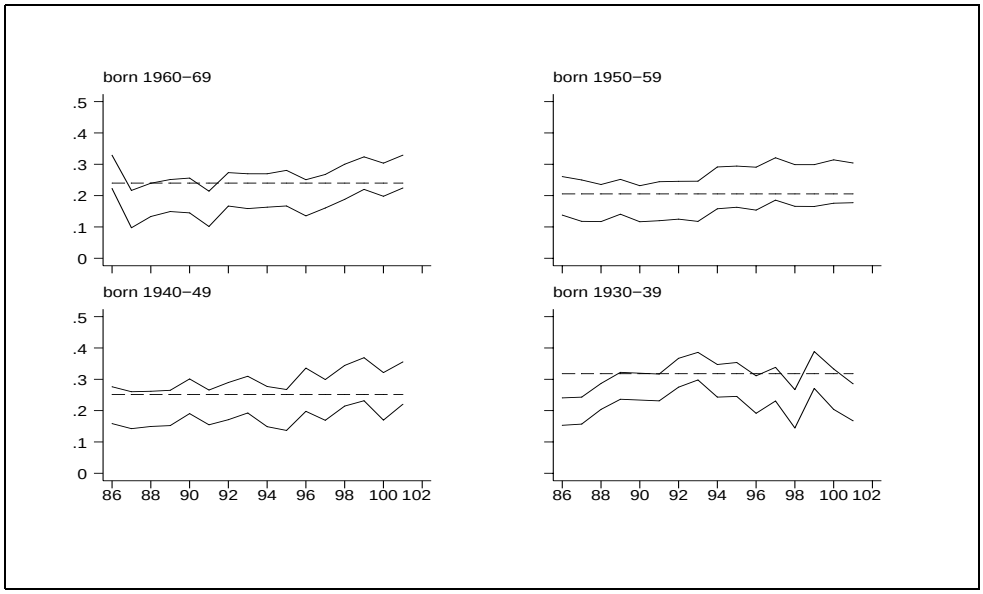

Figure 11: Squared coefficient of variation of family expenditure on non-durable goods by cohort (2001 dollars) - bounds using (13), Condition 3 and assuming $|\rho^*| \leq 0.90$

By means of a grid search on the support of $\rho^*(Z)$ the following question is addressed: what is the maximum correlation required such that the inequality index for the two 'central' cohorts in Figure 10 is statistically increasing over time? Apparently, a maximum value of $|\rho^*(Z)|$ between 0.85 and 0.90 is enough to detect increasing inequality for those born in 1940-49, while 0.80 is required for those born in 1950-59.

As discussed above, the range of possible values of $\rho^*(Z)$ implied by (13) is different from the interval $[-1, +1]$. Table 7 presents the distribution over time of the minimum and the maximum values attainable by this correlation by cohort and income group as a result of (13) once Z is controlled for. Figures for the 5th, 50th and 95th percentiles of the minimum and maximum value of $\rho^*(Z)$ between 1986 and 2001 are reported. The median of the minimum correlation attainable is always below -0.90, thus the restrictions exploited in Figures 11-13 rule out high and positive values of $\rho^*(Z)$.

Table 8 presents the distribution over time of the observed values of $\rho^*(Z)$ for IS and DS data. Apparently both IS and DS data suggest that the strength of empirical correlation net of Z is lower than the values considered to derive bounds in Figures 11-13.

8 CONCLUSIONS

This paper has discussed how to account for reporting errors affecting diary-based and recall-based data on non-durable consumption. In fact, it is likely that not all the commodities defining non-durable consumption are well reported exploiting

Table 7: Bounds implied by (13)

income group		**born 1960-69**		**born 1950-59**		**born 1940-49**		**born 1930-39**	
		min	max	min	max	min	max	min	max
0-20%	5%	-0.88	0.91	-0.87	0.98	-0.91	0.92	-0.86	0.96
	50%	-0.81	0.99	-0.81	0.99	-0.78	0.99	-0.76	0.98
	95%	-0.70	1.00	-0.76	1.00	-0.65	1.00	-0.68	1.00
20-40%	5%	-0.93	0.96	-0.94	0.98	-0.90	0.98	-0.92	0.96
	50%	-0.89	0.99	-0.90	0.99	-0.85	0.99	-0.82	0.98
	95%	-0.77	1.00	-0.83	1.00	-0.78	1.00	-0.67	1.00
40-60%	5%	-0.95	0.97	-0.96	0.98	-0.95	0.98	-0.93	0.93
	50%	-0.90	0.99	-0.90	0.99	-0.90	0.99	-0.85	0.98
	95%	-0.81	1.00	-0.87	1.00	-0.84	1.00	-0.76	0.99
60-80%	5%	-0.94	0.97	-0.96	0.97	-0.95	0.96	-0.92	0.97
	50%	-0.89	0.99	-0.91	0.99	-0.88	0.99	-0.86	0.98
	95%	-0.83	1.00	-0.87	1.00	-0.81	1.00	-0.76	0.99
80-100%	5%	-0.95	0.95	-0.89	0.98	-0.91	0.97	-0.89	0.95
	50%	-0.87	0.99	-0.85	0.99	-0.83	0.99	-0.81	0.98
	95%	-0.82	1.00	-0.79	1.00	-0.78	1.00	-0.65	0.99

Table 8: Observed correlations

income group		**born 1960-69**		**born 1950-59**		**born 1940-49**		**born 1930-39**	
		IS	DS	IS	DS	IS	DS	IS	DS
0-20%	5%	0.34	0.26	0.39	0.46	0.41	0.51	0.42	0.45
	50%	0.42	0.51	0.46	0.61	0.48	0.56	0.50	0.58
	95%	0.51	0.68	0.52	0.71	0.57	0.66	0.63	0.71
20-40%	5%	0.23	0.22	0.35	0.39	0.28	0.28	0.26	0.31
	50%	0.39	0.50	0.44	0.48	0.42	0.53	0.38	0.46
	95%	0.49	0.73	0.48	0.60	0.54	0.63	0.45	0.68
40-60%	5%	0.11	0.19	0.28	0.38	0.23	0.21	0.26	0.27
	50%	0.30	0.43	0.34	0.46	0.36	0.49	0.37	0.47
	95%	0.41	0.58	0.40	0.61	0.44	0.57	0.45	0.61
60-80%	5%	0.14	0.25	0.19	0.33	0.23	0.27	0.17	0.30
	50%	0.27	0.42	0.34	0.46	0.32	0.43	0.31	0.43
	95%	0.35	0.51	0.38	0.58	0.44	0.67	0.35	0.64
80-100%	5%	0.18	0.18	0.23	0.25	0.23	0.28	0.26	0.25
	50%	0.29	0.37	0.33	0.40	0.30	0.46	0.39	0.45
	95%	0.40	0.55	0.42	0.56	0.39	0.55	0.46	0.66

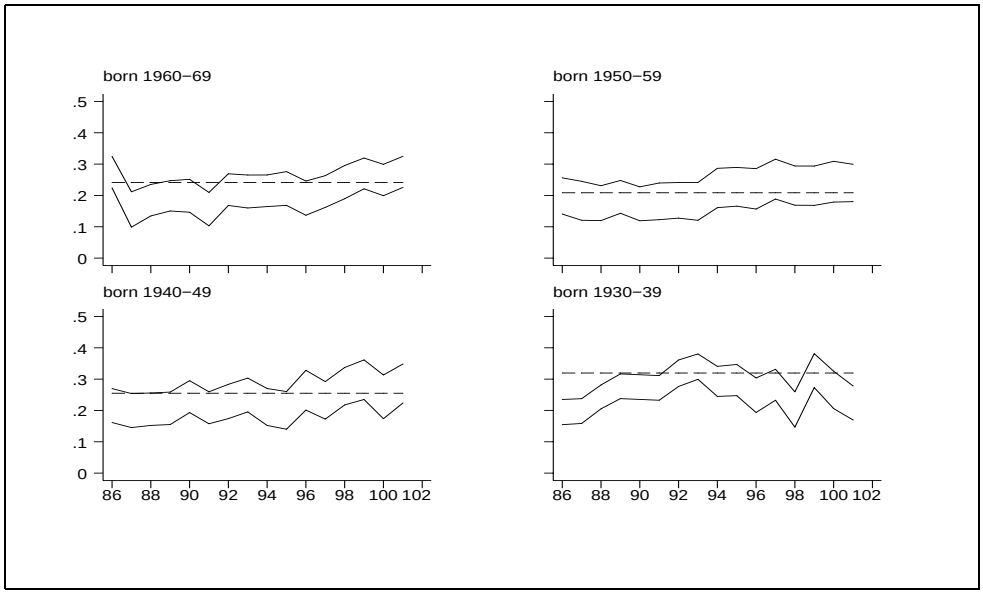

Figure 12: Squared coefficient of variation of family expenditure on non-durable goods by cohort (2001 dollars) - bounds using (13), Condition 3 and assuming $|\rho^*| \leq 0.85$

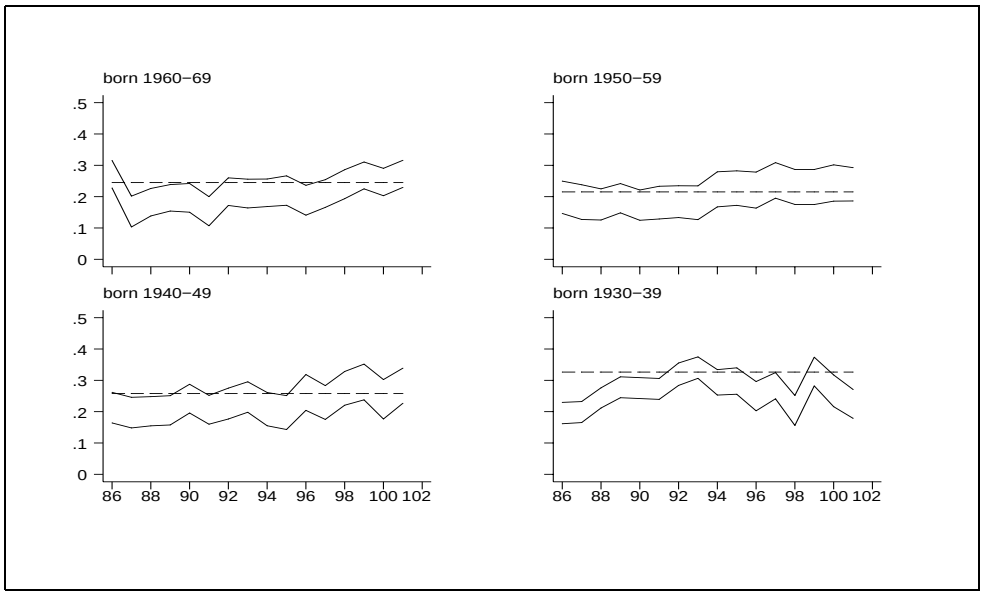

Figure 13: Squared coefficient of variation of family expenditure on non-durable goods by cohort (2001 dollars) - bounds using (13), Condition 3 and assuming $|\rho^*| \leq 0.80$

only one of these two survey methodologies. Expenditures on frequently purchased, smaller items are presumably more accurate using diaries while recall data are more appropriate for large expenditures or expenditures occurring on a regular basis.

It turns out that neither diary nor recall-based data alone provide a reliable aggregate measure of total expenditure on non-durables. One might argue that for any practical purpose these alternative data collection strategies lead to consistent results in the estimation of economic models of consumption behavior. Unfortunately, evidence from the literature suggests that data themselves might not be suitable to validate existing models and that conclusions are strongly related to the information being used.

This issue has been addressed by looking at micro data from two independent samples of households from the Consumer Expenditure Survey for the Unites States. This survey represent a unique source of data because it consists of diary an recall information collected on the same set of items, although it refers to separate samples of households. The integration of these datasets presents the problem of determining the appropriate survey component from which to select expenditure items. On the basis on evidence reported in a number of previous studies, we have 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.

The comparison of the two surveys based on the overlap in coverage of expenditures offers insights for the effects of collection modes on data quality. Estimating the true expenditure level on frequently purchased items for the recall sample (and on bulky items in the diary sample) can be seen as a problem of inferring counterfactuals: what is the counterfactual diary (recall) expenditure measure for recall (diary) respondents?

We have shown that the pictures emerging from the two surveys are very different, both with respect to mean expenditure and, more importantly, with respect to indices of inequality. More precisely, while the differences in mean expenditure are roughly constant over time, consumption inequality presents different levels and trends in the two surveys. We have reconciled this evidence using integrated diary and recall information to characterize the most likely pattern of consumption for cohorts of individuals defined by their age.

This 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. The implications of our findings for the estimation of inequality indices have been discussed, with an application to permanent income models. In particular, we have shown that using diary and recall data to improve the quality of household consumption the permanent income hypothesis cannot be rejected.

One general lesson can be deducted from this paper. Surveys are large measurement machines where many points might influence the ultimate error distribution. The collection of consumption data may be better structured as a set of overlapping questionnaires asking about only a small subset of consumption (as for the Family Expenditure Survey for the United Kingdom). In this way, more accurate information on different components of consumption would be obtained by the most appropriate survey methodology. Moreover, information from diary and recall questions collected for the same individuals would permit to shed more light into the effects of collection modes on the quality of available information. The evidence

on this point currently available in the literature is very much limited to a small groups of commodities (mainly food).

References

[1] Atkinson, A. B., Gomulka, J., and Stern, N.H. (1990), "Spending on Alcohol: Evidence from the Family Expenditure Survey 1970-1983," *The Economic Journal*, 100, 402, 808-827.

[2] Attanasio, O.P., Battistin, E. and Ichimura, H. (2003), unpublished manuscript, University College London.

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

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

[5] Battistin, E. (2002), "Measurement error issues in modelling economic behaviours - four case studies," Ph.D thesis in Statistics, University of Padova.

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

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

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

[9] Bound, J. Brown, C. and Mathiowetz, N. (2001), "Measurement error in survey data," in *Handbook of Econometrics*, eds. Heckman, J.J., and Leamer, E., Volume 5, 3707-3843.

[10] Branch, R., and Jayasuriya, B. (1997), "Consumer Expenditure Interview and Diary Survey Data Selection: a New Method," unpublished manuscript, Bureau of Labor Statistics.

[11] Browning, M. and Lusardi, A. (1996), "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, XXXIV, December, 1797-1855.

[12] Browning, M., Crossley, T.F., and Weber, G. (2002), "Asking Consumption Questions in General Purpose Surveys," unpublished manuscript.

[13] Bureau of Labor Statistics (2002), *Handbook of Methods, Chapter 16, Consumer Expenditures and Income*, available at www.bls.gov.

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

[15] Costa, D.L. (2001), "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves," *Journal of Political Economy*, 109, 6, 1288-1310

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

[17] Deaton, A., and Paxson, C. (1994), "Intertemporal Choice and Inequality," *Journal of Political Economy*, 102, 3, 437-467.

[18] Deaton, A. and Paxson, C. (2000), "Growth and savings among individuals and households," *The Review of Economics and Statistics*, 82, 2, 212-225.

[19] Garner, T.I. (1993), "Consumer Expenditures and Inequality: An Analysis Based on Decomposition of the Gini Coefficient," *The Review of Economics and Statistics*, 75, 1, 134-138.

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

[21] Hall, R.E. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 6, 971-987.

[22] Hausman, J., Newey, W.K., and Powell, J. L. (1995), "Nonlinear Errors in Variables: Estimation of Some Engel Curves," *Journal of Econometrics*, 65, 1, 205-233.

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

[24] Krueger, D., and Perri, F. (2001), "Does Income Inequality Lead to Consumption Inequality? Empirical Findings and a Theoretical Explanation," unpublished manuscript.

[25] Lee, L., and Sepanski, J.H. (1995) "Consistent estimation of linear and nonlinear errors-in-variables models with validation information," *Journal of the American Statistical Association*, 90, 429, 130-140.

[26] Lewbel, A. (1996), "Demand Estimation with Expenditure Measurement Errors on the Left and Right Hand Side," *Review of Economics and Statistics*, 78, 4, 718-725.

[27] Lusardi, A. (1996), "Permanent Income, Current Income and Consumption: Evidence from Two Panel Data Sets," Journal of Business and Economic Statistics, 14, 1, 81-90.

[28] 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.

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

[30] Neter, J., and Waksberg, J. (1964), "A Study of Response Errors in Expenditures Data from Household Interviews," *Journal of the American Statistical Association*, 59, 305, 18-55.

[31] Pischke, J.S. (1995), "Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study," *Journal of Business and Economic Statistics*, 13, 3, 305-314.

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

[33] 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*, 88, 424, 1208-1218.

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

[35] Shih, W.J., and Huang, W.M. (1992), "Evaluating Correlation with Proper Bounds Biometrics," *Biometrics*, 48, 4, 1207-1213.

[36] Shorrocks, A.F. (1982), "Inequality Decomposition by Factor Components," *Econometrica*, 50, 1, 193-211.

[37] Silberstein, A.R., and Jacobs, C.A. (1989), "Symptoms of Repeated Interview Effects in the Consumer Expenditure Survey," in *Panel Surveys*, eds. Kasprzyk, D., Duncan, G., Kalton, G., and Singh, M.P., New York: Wiley, 289-303.

[38] Silberstein, A.R., and Scott, S. (1991), "Expenditure Diary Surveys and Their Associated Errors," in *Measurement Errors in Surveys*, eds. Biemer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S., New York: Wiley, 303-326.

[39] Slesnick, D.T. (1998), "Are our Data Relevant to the Theory? The Case of Aggregate Consumption," *Journal of Business and Economic Statistics*, 16, 1, 52-61.

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

[41] Stanton, J.L., and Tucci, L.A. (1982), "The Measurement of Consumption: A comparison of Surveys and Diaries," *Journal of Marketing Research*, 19, 274-277.

[42] Torelli, N., and Trivellato, U. (1993), "Modelling Inaccuracies in Job-Search Duration Data," *Journal of Econometrics*, 59, 1/2, 187-211.

[43] Tucker, C. (1992), "The Estimation of Instrument Effects on Data Quality in the Consumer Expenditure Survey," *Journal of Official Statistics*, 8, 1, 41-61.

[44] Turner, R. (1961), "Inter- Week Variations in Expenditure Recorded During a Two-Week Survey of Family Expenditure," *Applied Statistics*, 10, 3, 136-146.

[45] Wilcox, D.W. (1992), "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical work," *American Economic Review*, 82, 922-941.

[46] Winter, J.K. (2002), "Design effects in survey-based measures of household consumption," unpublished manuscript.

[47] Ziliak, J.P. (1998), "Does the Choice of Consumption Measure Matter? An Application to the Permanent-Income Hypothesis," *Journal of Monetary Economics*, 41, 1, 201-216.

TECHNICAL APPENDIX

Bounds on $Cov(C_1^*, C_2^*|Z)$ using Condition 3 presented in Section 7 are derived in what follows. Let

$$\eta_t = Cov(C_1^*, C_2^*|Z)$$

be the quantity of interest and let

$$\widetilde{\eta}_t = \eta_t - \eta_{t-1}$$

be its changes over time. It is assumed that η_t can vary between *known* values l_t and u_t defined by the Cauchy-Schwartz bounds. Moreover, $\widetilde{\eta}_t$ is also identified by observed changes in IS and DS covariances if Condition 3 holds true.

Let

$$s_t = \{\dots, \eta_t - \widetilde{\eta}_t - \widetilde{\eta}_{t-1}, \eta_t - \widetilde{\eta}_t, \eta_t, \eta_t + \widetilde{\eta}_{t+1}, \eta_t + \widetilde{\eta}_{t+1} + \widetilde{\eta}_{t+2}, \dots\}$$

be the the time series of covariances given η_t. In light of Condition 3, such a sequence would be fully determined if η_t were observed.

Let

$$S_t = \{s_t : l_t \leq \eta_t \leq u_t\}$$

the set of sequences s_t consistent with the assumption that η_t is bounded between l_t and u_t. The set

$$\mathbf{B} = \bigcap_t S_t$$

defines all the sequences that *jointly* satisfy the Cauchy-Schwartz bounds over time and are consistent with the time series of observed changes $\widetilde{\eta}_t$. This set represents the basis to derive the bounds exploited in Figures 10-13.

DATA APPENDIX

This appendix describes the selection criteria used to derived the working sample for the analysis presented in this paper. It also discusses how methodological changes in the definition of consumption categories over time are handled.

Sample information from the the two survey components of the CEX is used for the period 1982-2001. All those households satisfying *at least* one of the following criteria

1. living in rural areas
2. with single females
3. residing in student housing
4. whose head is self-employed
5. whose head is aged below 22 and above 74
6. whose total expenditure on food (at home at away from home) is zero

are not considered in the final sample. Non-urban households are excluded to make information comparable over time, since they were discontinuously sampled before 1984. Additionally, households

7. not completing the diary for two weeks
8. not belonging to the cohorts analyzed in the paper

are excluded. Sample sizes for each cohort are given in Table A.1. Throughout the analysis, the family head is conventionally fixed to be the male in all H/W families (representing 56 percent and 53 percent of the whole sample for IS and DS data, respectively). Households presenting null expenditure on total food (both at home and away from home) are dropped from the analysis (less than 1 percent in each sample). These restrictions leave us with a sample of 190.080 and 46.244 units, for IS and DS data respectively.

Because of the small within-quarter variation in reported IS expenditures (less than 2 percent of the total variation in our sample), monthly consumption figures for IS households refer to the month *before* the interview. Monthly expenditures in the DS are instead defined as $26/12 = 2.16$ times the expenditure observed over two weeks, thus assuming equally complete reporting. All expenditure data have been seasonally adjusted, by taking residuals from regressions on zero-sum monthly dummies. Real expenditures are obtained using the the Current Price Index (CPI) published by the BLS. Although CPI bias has been of considerable concern to policy makers and researchers in recent years (see, for example, Costa 2001), we do not deal with this issue in the paper.

Expenditures of different family types are adjusted onto a comparable basis using an equivalence scale which depends on the number of adults and children in various age ranges. Specifically, the scale assigns weight 1 to the first adult and 0.670 to all remaining adults in the household. Children are weighted 0.233 if aged 3 or below, 0.333 if aged 4 to 7, 0.400 if aged 8 to 12 and 0.533 if aged 13 to 18. Sensitivity of main findings to alternative equivalence scales was investigated, and results were found to be qualitatively similar.

Table A.1: Sample sizes

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

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

Table A.2: Definitions of expenditure categories

Commodities in $\mathcal{D}$	**Diary**	**Interview**
Food and Non-Alcoholic Beverages at Home		83-87
Food and Non-Alcoholic Beverages Away from Home		
Alcoholic Beverages (at home and away from home)		
Non-Durable Goods and Services	82-85	
Newspapers and Magazines		
Non-durable Entertainment Expenses		
Housekeeping Services		
Personal Care		

Commodities in $\mathcal{I}$	**Diary**	**Interview**
Housing and Public Services	82-85	
Home Maintenance Services		
Public Utilities		
Miscellaneous Home Services		
Tobacco and Smoking Accessories		
Clothing, Footwear and Services		
Clothing, Footwear		
Services		
Heating Fuel, Light and Power	82-85	
Transportation (including gasoline)	82-85	
Fuel for Transportation		
Transportation Equipment Maintenance and Repair		
Public Transportation		
Vehicle Rental and Misc. Transportation Expenses		

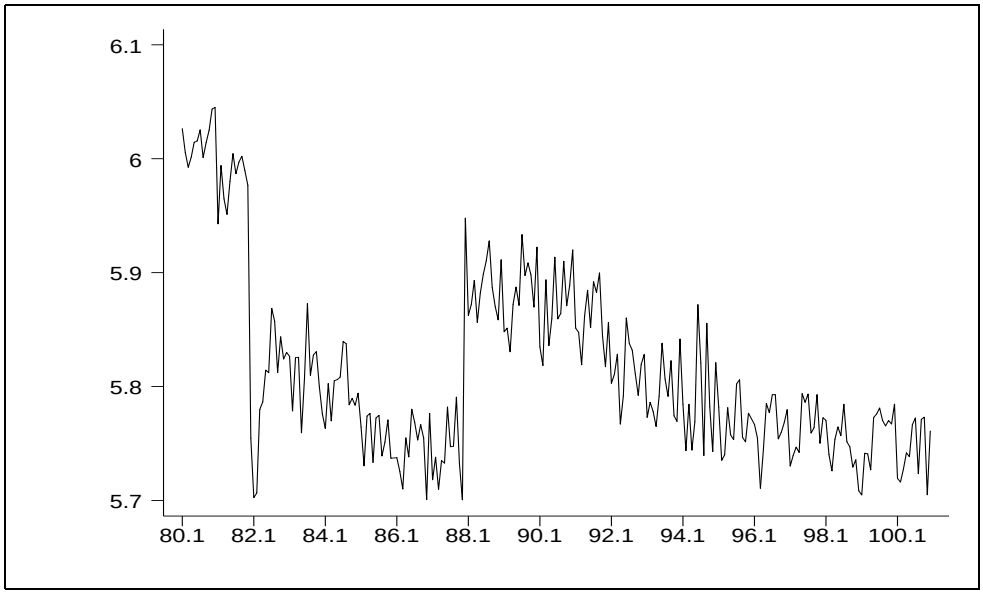

Figure A.1: Logs of monthly Food at Home expenditure in 2000 dollars - Interview Survey

Detailed expenditure categories included in total non-durable consumption are reported in Table A.2. The top panel refers to those commodities labelled in the text as $\mathcal{D}$, while the bottom panel to commodities in $\mathcal{I}$. The second and the third columns report the time intervals for which expenditure on non-durable commodities cannot be directly derived from public use tapes because of definitional problems, separately for IS and DS data. Accordingly, expenditures referring to "Housing and Public Services" and "Non-durable Services" have been introduced in the DS only after 1986, with the exception of very few items for "Home Maintenance Services" and "Non-durable Entertainment Expenses". Similarly, information on "Fuel" and "Transportation" expenses is not available from public tapes between 1982 and 1985.

As for IS data, the time series of food at home expenditure presents discontinuities introduced by changes in survey design in 1982 and 1987. In fact, recall information on food was derived in different ways over the years and the average spending appears to be heavily affected by the exact wording, as shown in Figure A.1. Spikes downwards in 1982 and upwards in 1988 are the effect of asking respondents for usual spending on a monthly basis or on a weekly basis. In 1980-1981 the question was on usual weekly expenditure on food over the past three months. In 1982-1987 the question was on how often and how much was spent in food over the previous month. In 1988, the 1980-1981 question was resumed.

Discontinuities in levels are accounted for by running a regression of real food expenditure on a quadratic time trend and a time dummy for the period 1982-1987, a polynomial in family income and additional household controls. Reported expenditure are then scaled up for all households between 1982 and 1987 by the same amount. Accordingly, the underlying assumption made is that changes in the wording of food-related questions only affect the level of reporting, but *not* the

variance.

Because of the definitional problems summarized in Table A.2, it follows that total expenditure figures for IS and DS data are not fully comparable over time. For this reason, figures reported in the paper for total spending on non-durables exploit 16 years of data for the DS (i.e. 1986-2001) and 20 years for the IS (i.e. 1982-2001). The aggregation rule suggested in Table 4 is therefore partially implemented over time. As a first approximation, only aggregate values of IS expenditures are considered before 1986, while integrated information for the DS is used after then.

The comparison of mean expenditure for the commodities defined in Table A.2 is presented in Figures A.2-A.10. A detailed description of the items used to define the categories of non-durable consumption can be downloaded at

http://www.stat.unipd.it/∼erich/papers.html,

separately for IS and DS data.

Diary and recall information is available from the DS on expenditure for food items. As noted earlier, food and non-alcoholic beverages is the only commodity entering non-durable expenditure for which recall and diary-based measurements on the *same* household are available. The information on the 'usual' spending is collected for each household at the beginning of the (two-week) diary period. Therefore, its accuracy is presumably not influenced by how respondents learn about own expenditures during the interview.

Figure A.11 presents the difference between weekly expenditure derived from diaries (as the sum of detailed food data) and weekly 'usual' expenditure for food and non-alcoholic beverages at grocery stores given by respondents. As already pointed out by several papers using similar data sources, diary expenditure levels in the first week tend to be higher that those for the second week. However, information from recall data leads to higher values of reported expenditures.[1]

It is worth noting that the difference between diary and recall figures conditional on the expenditure month (to control for seasonal effects) is always negative for the period of time covered by this analysis, with values generally decreasing in absolute value over time. There is a mild effect of the interview month on the magnitude of diary under reporting, since households interviewed in December usually present values closer to zero. This evidence supports the idea that recall questions overstate the real spending on food-related items possibly because households' reporting includes more than just food-related goods. Indeed the recall question about food expenditure in the IS is derived subtracting to the usual amount spent at the grocery store the usual amount on non-food items.

[1] We found a very similar pattern comparing recall and diary information from the National Food Survey for the United Kingdom. Results are available on request.

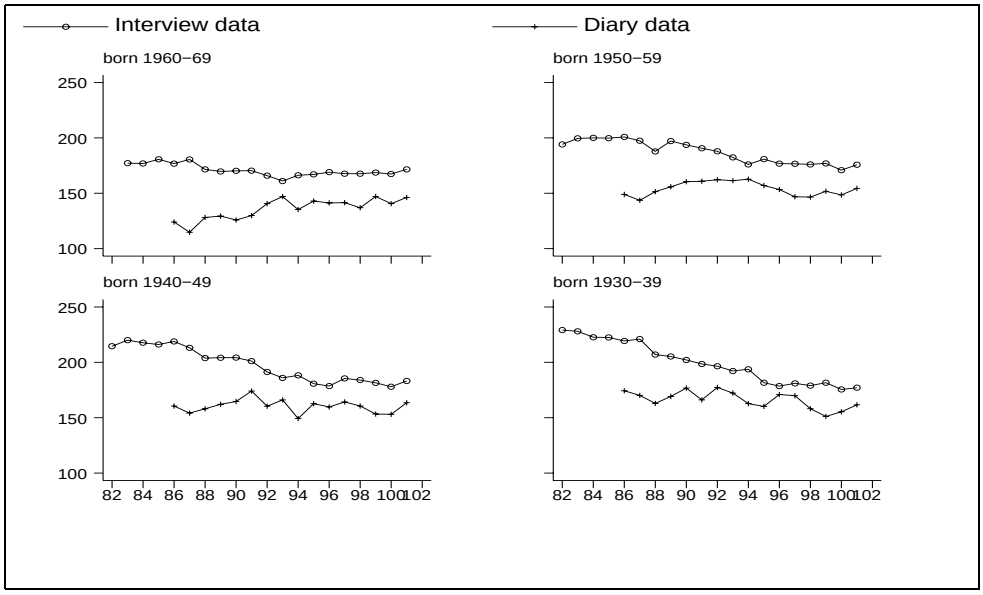

Figure A.2: Mean of family expenditure on food and non-alcoholic beverages at home (2000 dollars)

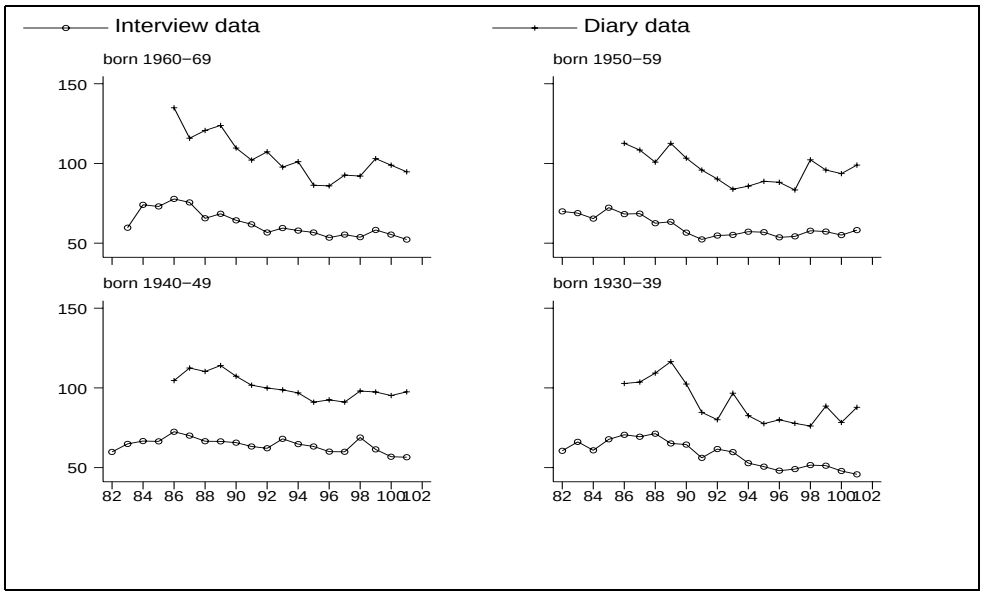

Figure A.3: Mean of family expenditure on food and non-alcoholic beverages away from home (2000 dollars)

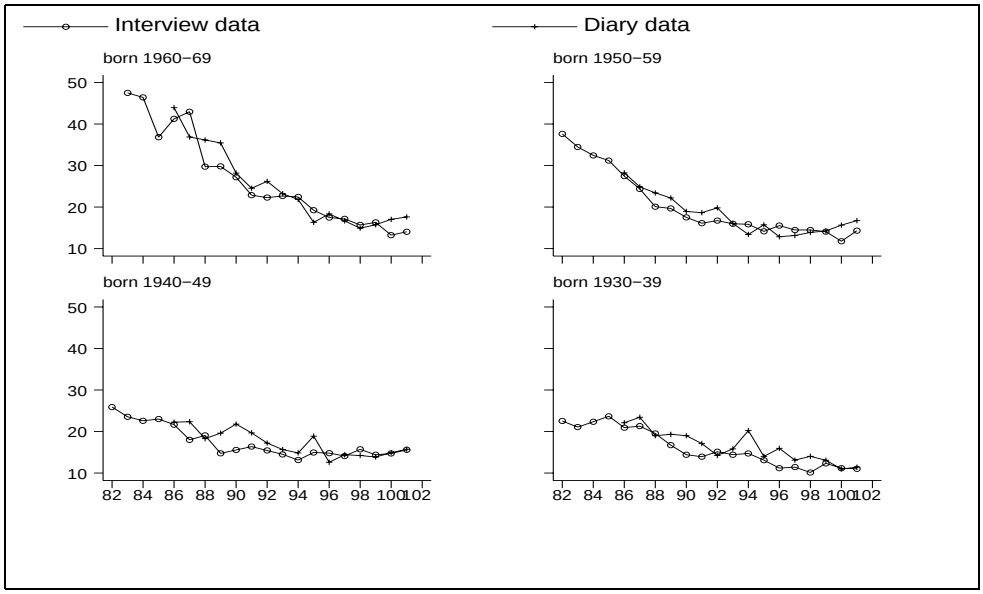

Figure A.4: Mean of family expenditure on alcoholic beverages (2000 dollars)

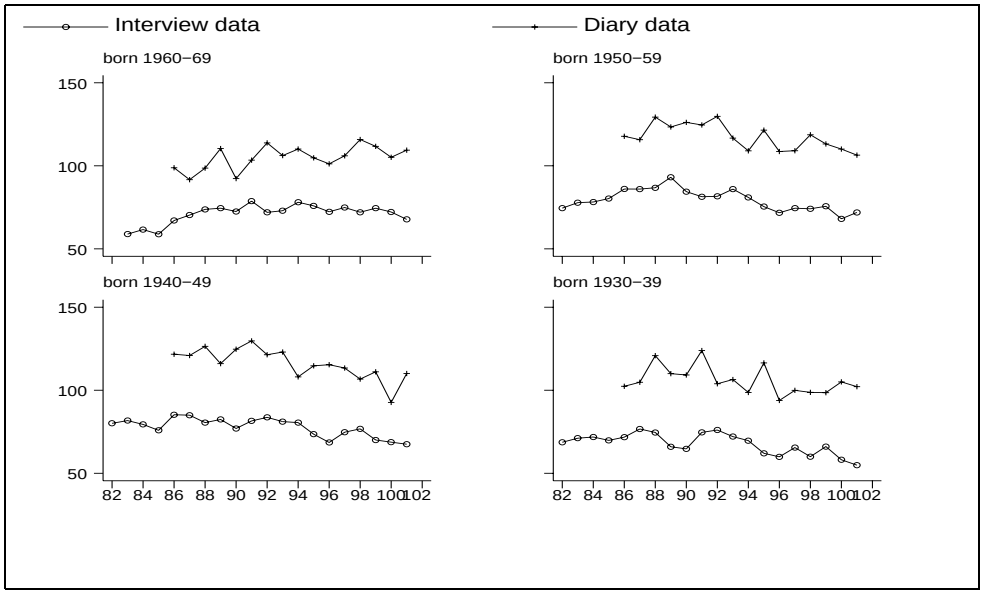

Figure A.5: Mean of family expenditure on non-durable goods and services (2000 dollars)

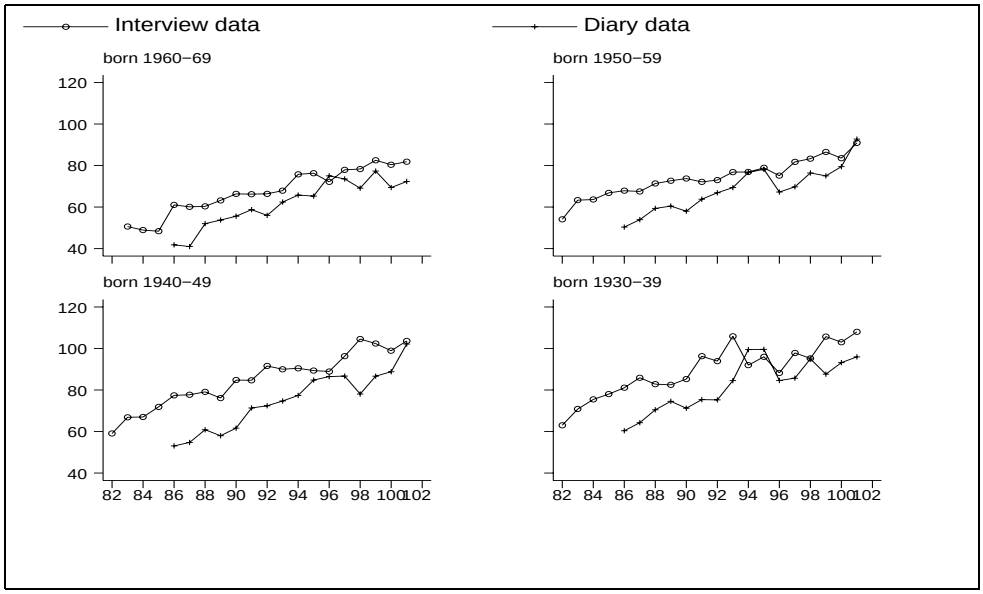

Figure A.6: Mean of family expenditure on housing and public services (2000 dollars)

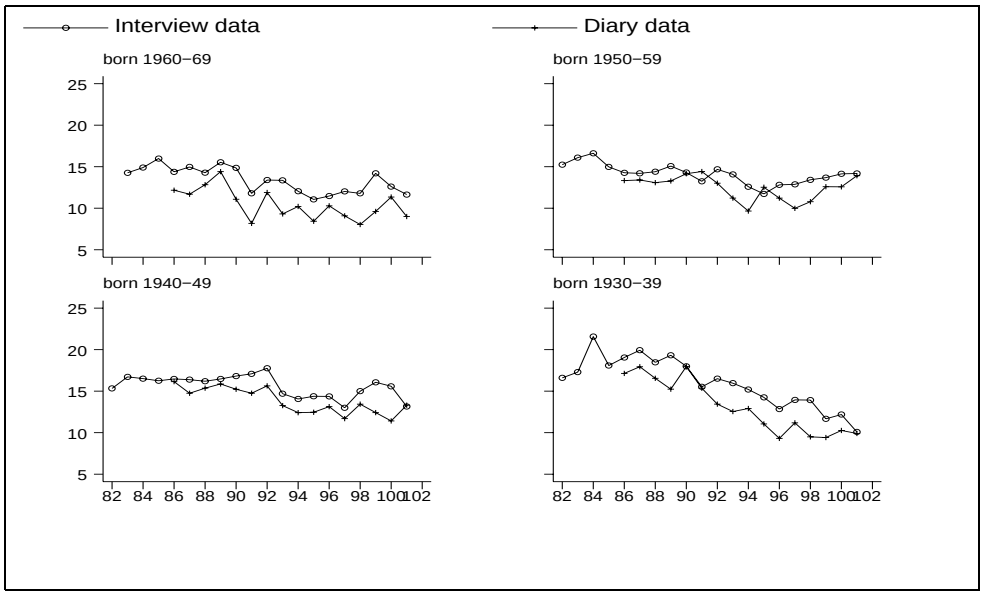

Figure A.7: Mean of family expenditure on tobacco (2000 dollars)

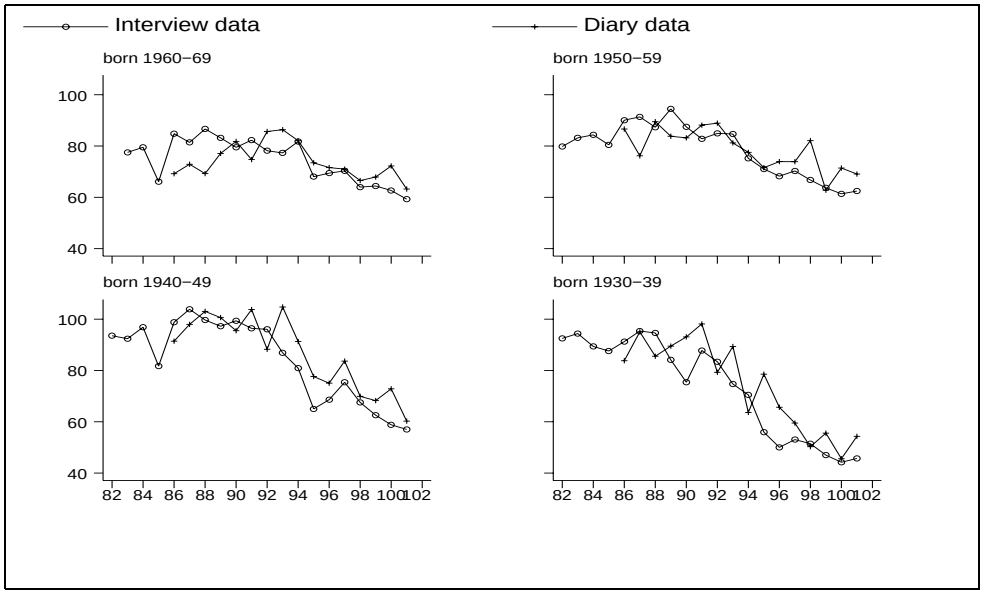

Figure A.8: Mean of family expenditure on clothing, footwear and services (2000 dollars)

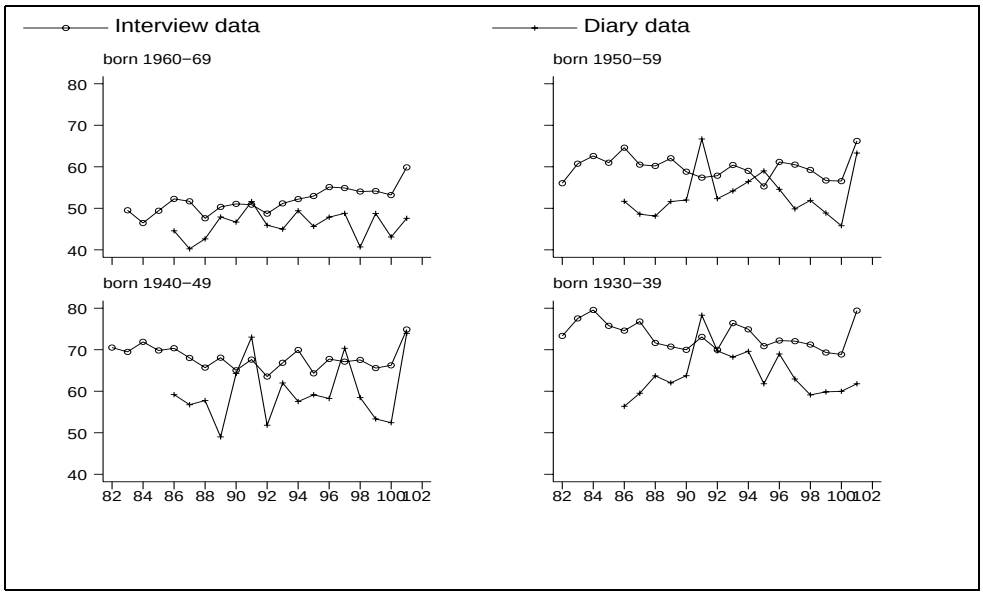

Figure A.9: Mean of family expenditure on heating fuel, light and power (2000 dollars)

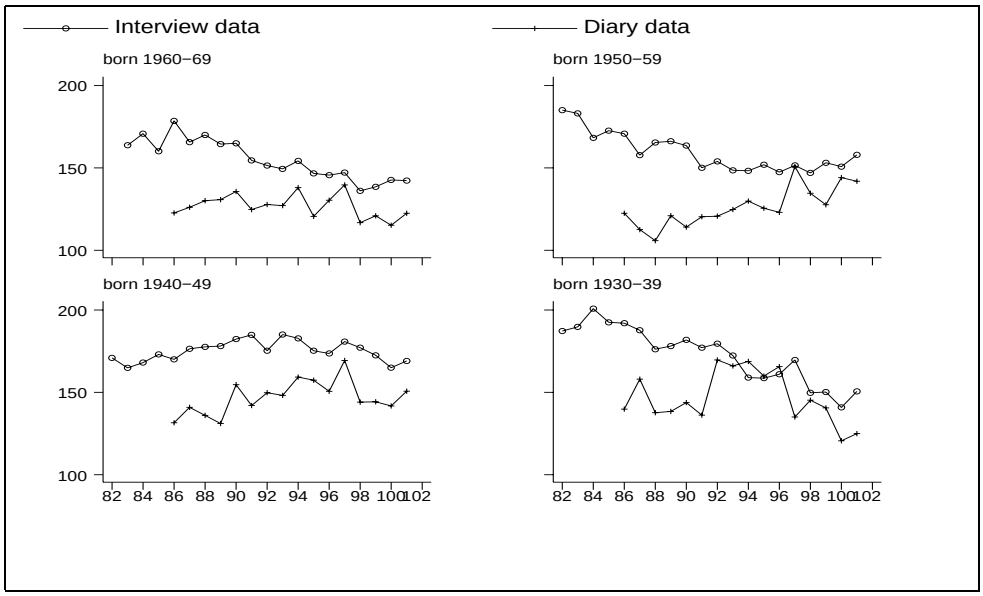

Figure A.10: Mean of family expenditure on transportation (2000 dollars)

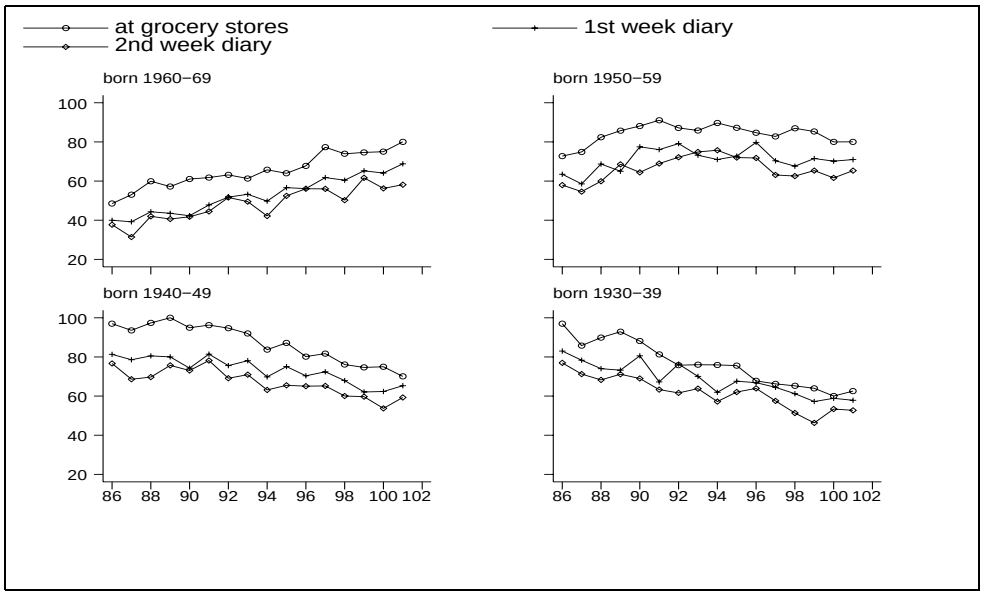

Figure A.11: Median weekly expenditure on food and non-alcoholic beverages at home (2000 dollars): usual expenditure at grocery stores and expenditure by diary week

Table A.3: Gini coefficient for total family income

	born 1930-39			born 1940-49			born 1950-59			born 1960-69			all		
year	**CEX**	**CPS**	**PSID**	**CEX**	**CPS**	**PSID**	**CEX**	**CPS**	**PSID**	**CEX**	**CPS**	**PSID**	**CEX**	**CPS**	**PSID**
1982	0.34			0.31			0.31						0.33	0.41	
1983	0.34			0.31			0.32			0.30			0.34	0.41	
1984	0.36			0.32			0.32			0.32			0.34	0.42	
1985	0.34			0.32			0.31			0.34			0.34	0.42	
1986	0.32			0.32			0.32			0.32			0.33	0.43	
1987	0.35			0.32			0.31			0.31			0.34	0.43	
1988	0.37			0.32			0.30			0.32			0.34	0.43	
1989	0.37			0.32			0.31			0.31			0.34	0.43	
1990	0.36			0.33			0.31			0.32			0.34	0.43	
1991	0.38			0.32			0.32			0.32			0.34	0.43	
1992	0.41			0.34			0.33			0.34			0.36	0.43	
1993	0.39			0.36			0.33			0.33			0.36	0.45	
1994	0.41			0.35			0.34			0.34			0.36	0.46	
1995	0.40			0.34			0.35			0.34			0.36	0.45	
1996	0.44			0.38			0.34			0.32			0.37	0.46	
1997	0.43			0.37			0.34			0.32			0.36	0.46	
1998	0.42			0.37			0.35			0.34			0.37	0.46	
1999	0.46			0.38			0.35			0.36			0.39	0.46	
2000	0.44			0.38			0.36			0.34			0.38	0.46	
2001	0.43			0.38			0.35			0.34			0.38	0.47	

source for CPS data: http://www.census.gov/hhes/income/histinc/ie1.html

Table A.4: Percentage of consumers: $\Pr(X > 0|\text{Interview})$ and $\Pr(X > 0|\text{Diary})$

Interview	1982-85	1986-89	1990-92	1993-95	1996-98	1999-2001
Food and non-alcoholic beverages at home	1.00	1.00	1.00	1.00	1.00	1.00
Food and non-alcoholic beverages away	0.89	0.90	0.88	0.88	0.87	0.85
Alcoholic beverages (at home and away)	0.67	0.63	0.57	0.56	0.51	0.48
Non-durable goods and services	0.97	0.97	0.97	0.97	0.95	0.95
Housing and public services	0.97	0.97	0.98	0.98	0.98	0.98
Tobacco and smoking accessories	0.50	0.44	0.39	0.36	0.34	0.29
Clothing and footwear	0.88	0.87	0.85	0.82	0.78	0.72
Heating fuel, light and power	0.88	0.89	0.91	0.92	0.93	0.93
Transport (including gasoline)	0.98	0.98	0.98	0.98	0.98	0.97

Diary	1982-85	1986-89	1990-92	1993-95	1996-98	1999-2001
Food and non-alcoholic beverages at home		0.98	0.99	0.98	0.98	0.98
Food and non-alcoholic beverages away		0.94	0.92	0.89	0.90	0.89
Alcoholic beverages (at home and away)		0.59	0.53	0.50	0.45	0.45
Non-durable goods and services		0.94	0.94	0.92	0.91	0.89
Housing and public services		0.66	0.65	0.64	0.63	0.63
Tobacco and smoking accessories		0.43	0.38	0.34	0.33	0.29
Clothing and footwear		0.73	0.73	0.70	0.67	0.64
Heating fuel, light and power		0.46	0.51	0.46	0.46	0.45
Transport (including gasoline)		0.96	0.94	0.93	0.94	0.93