Italy 1987: Survey Information

Summary table

Generic information	
Name of survey	Survey on Household Income and Wealth – SHIW *(Indagine sui Bilanci delle Famiglie Italiane)*
Institution responsible	Bank of Italy
Frequency	Every two years
Survey year / Wave	1987
Collection period	April to June 1988
Survey structure	Cross-sectional
Coverage	Private households in the whole territory
Geographic information	20 administrative regions (more detailed info on provinces is collected but not available for external users)
Files delivered	Only the data from the historical database are available for public access; it comprises several files at different levels (household, individual, pensions, employment activities, real estates, etc.).
Sample size	
Households	8,027 households
Individuals	25,098 individuals (of which 13,549 income earners and 20,857 aged over 15)
Sampling	
Sampling design	Two-stage stratified sampling, with the stratification of the PSUs (municipalities) by region and demographic size.
Sampling frame	Municipal registry office records.
Questionnaires	Paper-based questionnaire with a modular structure: general part addressing aspects relevant to all households and a series of annexes with questions relevant to specific subsets of households.
Standard classifications	
Education	6 categories
Occupation	6 categories for employees, and 6 for self-employed, used as labour force status and not occupation
Industry	9 sectors
Income	
Reference period	Income in the preceding calendar year (which coincides with the fiscal year)
Unit of collection	Mostly at the individual level, except for property income (household level)
Period of collection	Mostly monthly income with number of months, some annual.
Gross/net	All variables are recorded net of taxes and contributions.
Data editing / processing	
Consistency checks	Standard post-survey consistency checking procedure by the data collection company.
Weighting	Survey data can be grossed up to aggregate values thanks to appropriate weights assigned to each household according to its probability to be included in the survey.
Imputation	All the aggregated variables are checked against their means and eventually corrected.

This document draws extensively upon the methodological Annex to the "I bilanci delle familglie italiane nell'anno 1987", *Supplemento al Bollettino Statistico*, Bank of Italy, Year XLII, No. 5, January 1989.

Table of contents:

A. General characteristics

Official name of the survey/data source:
Survey on Household Income and Wealth – SHIW *(Indagine sui Bilanci delle Famiglie Italiane)*

Administrative Unit responsible for the survey:

Bank of Italy
Research Department
Divisione Rilevazioni e Metodi Statistici - R.M.S.
Address: Via Nazionale 91, 00184 ROMA.
WWW: http://www.bancaditalia.it
e-mail: studi.indagini@ insedia.interbusiness.it

The Survey on Household Income and Wealth (SHIW) began in the 1960s within the Research Department of the Bank of Italy with the aim of gathering data on the incomes and savings of Italian households. Over the years, the scope of the survey has grown and now includes wealth and other aspects of households' economic and financial behaviour such as, for example, which payment methods are used.

The sample used in the most recent surveys (carried out generally every other year) comprises about 8,000 households (24,000 individuals), distributed over about 300 Italian municipalities and it is representative of the whole Italian population. The 1987 survey was the last one to be conducted with a time-independent sample of households. Starting with the previous year's survey, substantial modifications were introduced in order to reorganize the sampling design, the questionnaire and the surveying technique. Already with the 1984 survey the sample framework used switched from the electoral lists to the municipal lists, which had already been a first substantial improvement for the representativeness of the sample. But the most significant improvements were carried out in the 1986 survey with the modification of the stratification criteria and the sampling of the primary units, and in the 1987 survey with the collection of some important

information on non-respondents, the oversampling of less numerous and high income households, and a new methodology for the collection of information on financial activities. All those modifications allowed the collection of more reliable information with respect to the past, especially as far as the financial activities are concerned.

The survey results are regularly published in the Bank's Supplements to the Statistical Bulletin. The data on the households is freely available, in an anonymous form, for further elaboration and research. Other than the annual datasets (only available for external use from the 1989 survey), a historical database has been put together comprising a slightly restricted selection of variables available in the normal annual files but consistent over all the years since 1977.

B. Population, sampling size and sampling methods

Coverage

The sample is representative of the whole Italian population.

Sample size

Table 1a shows the sample size used between 1987 and 2000, indicating the number of households interviewed in more than one survey. The overall size of the sample for the 1987 survey was 8,027 households.

Table 1a
Households interviewed in the 1987-2000 surveys

Year of first interview	Year of survey						
	1987	1989	1991	1993	1995	1998	2000
1987	8,027	1,206	350	173	126	85	61
1989		7,068	1,837	877	701	459	343
1991			6,001	2,420	1,752	1,169	832
1993				4,619	1,066	583	399
1995					4,490	373	245
1998						4,478	1,993
2000							4,128
Sample size	**8,027**	**8,274**	**8,188**	**8,089**	**8,135**	**7,147**	**8,001**

Sampling design

As already mentioned, the 1986 and 1987 surveys mark the beginning of a new series of surveys with substantial modifications, among other things, to the sampling design.

The sampling design used was that of multi stage sampling, with stratification of primary sampling units. All the Italian municipalities were split into 51 strata according to 17 geographic repartitions[1] and 3 classes of demographic size.

The 1987 sample was drawn in two stages (municipalities and households), with the stratification of the primary sampling units (municipalities) by region and demographic size.[2] All municipalities with a population over 40,000 were put into the sample; those with a population between 20,000 and 40,000 were selected from regional stratum with probability proportional to the size of the municipality. The other municipalities, were further split into nine subgroups according to the height and the most relevant economic activity of the population (the selection being carried out with probability proportional to the size of the municipality even in this case).

The random selection of first stage units allowed, already since the 1986 survey, to improve the representativity of the sample (in previous surveys, the different interviewers' availability in different geographic areas made poorer regions underrepresented in the sample).

After having selected the first stage units (which amounted to 339 in 1987, see Table 2a), households were selected randomly from municipal registry office records.

Table 2a
Survey municipalities[*] (percentage values)

Size of municipality	Region			Total municipalities
	North	Centre	South and Islands	
Up to 20,000	10.1	5.6	9.2	24.8
20,000 to 40,000	8.5	6.2	7.2	21.9
Over 40,000	20.6	11.1	21.6	53.3
Total	**39.2**	**22.9**	**37.9**	**100.0**

* The total number of municipalities sampled was 339; of those 33 were not included in the survey because the list of sampled households arrived too late for the interviewing.

Such a sampling method, introduced in 1984, allowed the elimination of the distortion due to the preceding criterion based on electoral lists, which produced an oversampling of larger households. Because of the positive correlation between household income and number of household members, the extraction from municipal registry office records also allowed the improvement of the estimate of mean income.

With the purpose of getting more reliable estimates for the small group of high income households, in 1987 an oversampling of those household was carried out, which interested about 700 households of managers, entrepreneurs and free professionals

---

[1] Each of the 20 administrative regions represents a repartition, with the exception of Piemonte and Valle d'Aosta, Abruzzi and Molise, Basilicata and Calabria, which were regrouped in pairs.
[2] Since the 1986 survey, the stratification and selection criteria for the primary sampling units used by ISTAT for the Labour Force Survey have been adopted.

extracted from professional and commercial registers; the member of such categories was determined ex ante on the basis of information collected in 1986, and was corrected ex post with the post-stratification method described below.

C. Data collection and acquisition

Data collection

The interviews for the sample survey of Italian household budgets in 1989 were conducted between April and June 1988.

Interviews were conducted by a specialized company using professional interviewers. The interview stage was preceded by a series of meetings at which Bank of Italy officials and representatives of the company gave instructions directly to the interviewers. The households contacted for interviews, who are guaranteed complete anonymity, receive a booklet describing the purpose of the survey and giving a number of examples of the ways in which the data are used. The participating households may request a copy of the results of a previous survey. Finally, the Research Department of the Bank of Italy is available for any eventual clarification concerning the survey.

In spite of these practices, the percentage of the households not wiling to participate is rather high (see Table 3a). Since the households interviewed in substitution (selected from a ad hoc reserve list, compiled following the same criteria as the main one), could belong to a different socio-economic category than those who felt out of the sample, some distortions could be present in the results.

Table 3a
Households contacted and reason for non-participation[*] (percentages)

	Percentage of total
Households:	
Interviewed	59.7
Not interviewed	40.3
Total	*100.0*
Reasons for non-interview:	
Household unknown at the address	5.0
Head of household deceased or sick	2.1
Refusals	19.8
Absent at all contacts	13.4
Total	*40.3*

* 8,027 households were interviewed in 1987.

Until 1986, in order to evaluate the degree of representativeness of those who agreed to participate in the survey, some information was collected by the interviewer concerning a subjective estimate of the socio-economic condition for the household. Those estimates, even if interesting for the evaluation of the profile of non-interviewed households, were

little useful for ex post correction procedures (because of the subjective characteristic above-mentioned).

For this reason, in 1987 it was decided to introduce a module distinct from the rest of the questionnaire, where to enter generic information on not-interviewed households (such as household size, age and profession of the head, geographical region). This reasoning underlying this data collection was that, the strict distinction between this module and the income and wealth questions, would have allowed to the interviewers to obtain such information. Nevertheless, such information was collected for only 12 per cent of not-interviewed households, and was thus not used for the correction of the sample structure. This information has nevertheless not stressed any substantial difference between the two groups of households, besides household size, which resulted slightly lower for not-interviewed households.

The questionnaire

The questionnaire used in the survey has a modular structure, so that it can easily be adapted to the several economic structures of the households. It is composed of a general part addressing aspects relevant to all households and a series of annexes with questions relevant to specific subsets.

The questionnaire structure remained substantially the same from 1973 to 1984. In the 1986 and 1987 surveys, substantial modifications were introduced in the sections relative to real properties, income, savings and consumption expenditures.

As far as income is concerned, in order to get to an income definition more directly comparable with national accounts, it was necessary to distinguish between members of the professions, self-employed, sole proprietors and entrepreneurs with fewer than 20 employees on one side and entrepreneurs with 20 or more employees on the other side. Furthermore, the collection of separate information on the interest and capital part of mortgage loans, as well as the redefinition of some transfers of the capital account (inheritances, lottery gains, etc.) and income account, allowed the detraction of some income components not existing in the national accounts.

Concerning financial activities, the 1986 survey represented a turning point; since the variable savings, and especially the part invested in financial activities, had never provided satisfying results because of the reticence of households to release such information, those variables were not collected in 1986. Because of the importance of the subject though, starting from 1987, new collection methods were experimented.

D. Definition of the survey units

Household

The basic survey unit is the household, which is defined as "a group of individual linked by ties of blood, marriage or affection, sharing the same dwelling and pooling all or part of their incomes". Are to be included in the household all persons that normally lived in the dwelling at 31 December 1987 who contributed at least part of their income to the household; this includes any members temporarily absent (e.g. on vacation, temporarily away for study, etc) and any non-relatives that lived stably in the household at 31 December 1987.

Head of household

The head of the household is defined at the survey stage as the person who says he/she has "the most responsibility for family finances", which satisfies the need to determine the best informed person.

E. Contents

The survey contains information about:
- demographic characteristics of the household members (including education and main employment status);
- real properties of the household (with separate modules for each real property owned at end-1987, for each real property sold during 1987, and for each real property only partly owned);
- income from employment (with separate modules for each payroll employment activity, self-employment activity and each activity as owner or manager of societies with more than 20 employees) of the head of household and other household members;
- income from transfers (with a separate module on each pension) of the head of household and other household members;
- financial activities;
- consumption and other family expenditures;
- balance.

F. Quality of data

Quality of estimates

Households that could not be interviewed were replaced by others selected randomly in the same municipality.[3] Obviously, this technique does not eliminate the risk of obtaining samples in which the less-cooperative segments of the population are underrepresented, thus generating biased estimates *(selection bias)*.

Checking data and imputing missing data

---

[3] The substitution of non interviewed households was carried out through reserve lists, compiled following the same criteria of the basic lists.

The questionnaires are checked first of all to verify that no annexes are missing, no questions have been skipped and that there are no editing errors. In this phase the codification of open-answer fields is carried out (i.e. the answer "other – please specify"). Data are subsequently entered into magnetic support and automatic checks are carried out to verify the consistency of single fields or correlated fields.

Once the checks were completed, work began on aggregating elementary variables to get to the measures of interest for research. Finally, once the variables income, consumption expenditure, total and financial savings, propension to consumption and to saving are determined, the questionnaires that report values of those variables very far from the averages are signaled, analyzed and eventually corrected by Bank of Italy officials (together with representatives of the data collection company). For the 1987 survey, this concerned 2.5 per cent of the questionnaires.

Weighting: the sample estimates

In a two-stage sampling design, with stratification of primary sampling units, an estimator of the mean value of variable x is given by:

$$(1) \quad \bar{x} = \frac{\sum_{h}^{L}\sum_{i}^{m_h}\sum_{j}^{n_{hi}} w_{hij} x_{hij}}{\sum_{h}^{L}\sum_{i}^{m_h}\sum_{j}^{n_{hi}} w_{hij}}$$

where x_{hij} is the value of variable x for the jth household sampled in the ith municipality of the hth stratum and w_{hij} is the weight to give to the same household taking the sampling probabilities into account:

$$(2) \quad w_{hij} = \begin{cases} \dfrac{N_{hi}}{n_{hi}} \\[2em] \dfrac{1}{m_h} \dfrac{1}{P_{hi}} \dfrac{N_{hi}}{n_{hi}} \end{cases}$$

respectively for municipalities with more than 40,000 inhabitants and for municipalities with up to 40,000 inhabitants, where P_h, N_{hi} and m_h are respectively the resident population, that of the municipalities in the survey and the number of sample municipalities in the hth stratum, and P_{hi} and n_{hi} are respectively the population and the number of respondents in the ith municipality of the hth stratum.

The calculation of weights (2) is determined on the basis of the number of households assigned to each sample municipality, but the number of households effectively interviewed is usually lower (for example, because of the refusal to participate or the impossibility of contact of the household); it is thus necessary to apply a correction, by multiplying the weighs by the inverse of the response rate:

$$(3) \qquad w'_{hij} = w_{hij} \frac{n_{hi}}{n'_{hi}}$$

Expression (3) was used for the 1986 survey, but was not usable for the 1987 one because of the oversampling. Therefore, it was necessary to redefine the estimator (1) on the basis of appropriate correction weighs. More specifically, the correctives c_{ag} for the oversampling were defined at the level of geographical area (strata aggregation) and professional status:

$$(4) \qquad c_{ag} = \frac{\sum_h^{L_a} \sum_i^{m_h} \sum_j^{n'_{hig}} w'_{hij}}{\sum_h^{L_a} \sum_i^{m_h} \sum_j^{(n'_{hij}+I_{hij})} w'_{hij}}$$

so that the estimator of the mean value of variable x becomes:

$$(5) \qquad \bar{x} = \frac{\sum_a^{A} \sum_h^{L_a} \sum_i^{m_h} \sum_g^{G} \sum_j^{(n'_{hij}+I_{hij})} w'_{hij} \, c_{ag} \, x_{higj}}{\sum_a^{A} \sum_h^{L_a} \sum_i^{m_h} \sum_g^{G} \sum_j^{(n'_{hij}+I_{hij})} w'_{hij} \, c_{ag}}$$

Since some oversampled households (members of the professions and entrepreneurs) resulted characterised by a higher mean income than the corresponding categories in the basis sample, a regression analysis was carried out to evaluate whether such effect was determined by structural variables (age and education of income earner, number of employees in the firm). As a result, for members of the professions and the entrepreneurs it was necessary to carry out a post-stratification correction on the basis of the variable number of employees, which contributed to explain the largest part of the income differential between households of the oversample and of the basic sample.

Standard errors
The standard errors of the means of the main variables, calculated taking the sampling design into account, are shown in Table 4a.

Table 4a
Standard errors in the estimation of the means for the main variables
(thousands of lire, percentages)

Variables	Standard error[*]
Household income	2.5
Household consumption	2.3
Household net wealth	5.1
Mean consumption propension	0.7
Rate net wealth/income	4.7

* As a percentage of the mean value.

Because of the strict correspondence between the first stage units and the interviewers used for the collection, such estimates take into account also biases due to non-sampling errors (the so-called interviewer bias).

G. Uses of the survey

Publications

The results are regularly published in the Supplements to the Statistical Bulletin of the Bank of Italy. A whole bibliography of the research carried out using data from the SHIW is available (in Italian) from the Bank of Italy web-site (http://www.bancaditalia.it/statistiche/ibf/statistiche/ibf/pubblicazioni/altre/biblio.pdf).   A list of the Economic Research Papers of the Bank of Italy concerning the SHIW is reported here:

E. Battistin, R. Miniaci and G. Weber (2003), What do we learn from recall consumption data?, Bank of Italy, *Temi di Discussione*, N. 466.

Giovanni D'Alessio and Ivan Faiella (2002), Non-response behaviour in the Bank of Italy's Survey of Household Income and Wealth, Bank of Italy, *Temi di Discussione*, N. 462.

Silvia Magri (2002), Italian households' debt: determinants of demand and supply, Bank of Italy, *Temi di Discussione*, N. 454.

Guido de Blasio and Sabrina Di Addario (2002), Labor market pooling: evidence from Italian industrial districts, Bank of Italy, *Temi di Discussione*, N. 453.

A. Brandolini, P. Cipollone and P. Sestito (2001), Earnings dispersion, low pay and household poverty in Italy, 1977-1998, Bank of Italy, *Temi di Discussione*, N. 427.

Andrea Brandolini and Piero Cipollone (2001), Multifactor Productivity and Labour Quality in Italy, 1981-2000, Bank of Italy, *Temi di Discussione*, N. 422.

Piero Cipollone (2001), Is the Italian Labour Market Segmented?, Bank of Italy, *Temi di Discussione*, N. 400.

G. D'Alessio e L. F. Signorini (2000), Disuguaglianza dei redditi individuali e ruolo della famiglia in Italia, Bank of Italy, *Temi di Discussione*, N. 390.

Andrea Brandolini (1999), The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality, Bank of Italy, *Temi di Discussione*, N. 350.

Poverty and income distribution

According to the publication "I bilanci delle familglie italiane nell'anno 1987", *Supplemento al Bollettino Statistico*, Bank of Italy, Year XLII, No. 5, January 1989, the Gini coefficient of concentration is 0.34 for the distribution of household income and 0.36 for the distribution of individual income.