

SURVEY OF HOUSEHOLD INCOME AND WEALTH 2008

Summary

1. Introduction	1
2. Data format.....	1
3. Information contained in the datasets	2
4. Aggregate variables	4

1. Introduction

The Survey on Household Income and Wealth (SHIW) covers 7,977 households composed of 19,907 individuals and 13,266 income-earners. The number of households interviewed in previous surveys (*panel households*) is 4,345.

Panel households have the same questionnaire number (NQUEST) as in 2006.

Given that sampling design involves unequal stratum sampling fractions, the use of sampling weights (variable PESOFIT) is required to obtain unbiased estimates. Weights are given at household level since all members of the household have the same weight.

Furthermore it is available the dataset PESIJACK08 containing replication weights that can be used to calculate sampling variance¹.

2. Data format

All the data are available in three different formats: **ASCII**, **SAS** and **STATA** (version 7 and above) formats. Data are compressed in one single zip archive named **INDxx_format**, where **xx** denotes the last two digit of the survey year and **format** indicates whether the file is ASCII, SAS or STATA file. Data-files can be decompressed by using file manager (Windows XP) or any compression program such as **7-Zip** (downloadable at www.7-zip.org).

All the ASCII files are "comma separated" (file CSV), the first row of the file containing the variable names. This feature reduces loading time for almost all statistical packages. These statistical packages load the data naming the variables according to the information in the first record.

¹ The dataset contains 319 replication weights calculated following the Jackknife method. Weights are called PWTx where x indicates the progressive number of the replication, and goes from 1 to 319. Further information about the construction of replication weights and on their use for the estimate of sampling variance in the SHIW can be found in the methodological appendix of the Supplements to the Statistical Bulletin of the 2008 survey. With respect to the problems connected to statistical inference using complex survey design see Faiella, I. (2008) "Accounting for sampling design in the SHIW", Bank of Italy Working Papers (Temi di discussione) n. 662 – April.

3. Information contained in the datasets

The variable names and the answer codes are reported in the questionnaire contained in the documentation. The variables are organized in the dataset following the grouping reported in table 1. Variables carrying an asterisk on the questionnaire are not available to external users.

Table. 1

Datasets available in the 2008 annual database

<i>Dataset</i>	<i>Content</i>	<i>Primary key</i>
	QUESTIONANAIRE DATASETS	
CARCOM08	Characteristics of the individuals	NQUEST NORD
Q08A	Section A (Households' composition)	NQUEST
LAVORO	Section B (Employment)(*)	NQUEST NORD
Q08C1	Section C (Payment Instruments - <i>Questions from C01 to C22</i>)	NQUEST
Q08C2	Section C (Financial Assets and financial information- <i>Questions from C23 to C48</i>)	NQUEST
Q08D	Section D (Properties and debts)	NQUEST
Q08E	Section E (Expenditures)	NQUEST
Q08F	Section F (Insurance)	NQUEST
Q08G	Section G (Information provided by the interviewer)	NQUEST
USCITI	Individuals that left the panel household	NQUEST NORDP
ROTAZ1_1	I round, I part (Family of origin)	NQUEST
ROTAZ1_2	I round, II part (Happiness and family decisions)	NQUEST
ROTAZ2	II round (Payment instruments and financial information)	NQUEST
ALLB1	Annex B1 (Payroll employees)	NQUEST NORD
ALLB2	Annex B2 (Self-employed worker)	NQUEST NORD
ALLB3	Annex B3 (Family business)	NQUEST
ALLB4	Annex B4 (Working shareholder/partner)	NQUEST NORD
ALLB5	Annex B5 (Pensions)	NQUEST NORD
ALLB6	Annex B6 (Other income sources)	NQUEST NORD
ALLD1	Annex D1 (Property, other than principal residence)	NQUEST
ALLD2	Annex D2 (Loans)	NQUEST
	DERIVED DATASETS	
RFAM08	Household Incomes	NQUEST
RISFAM08	Household Expenditure and Savings	NQUEST
RICFAM08	Household Wealth	NQUEST
RPER08	Individual Incomes	NQUEST NORD
PESIACK08	Replication weights	NQUEST

(*) Information on working prevalent status are contained in CARCOM08.

The primary key to merge household level information is NQUEST (household ID). NQUEST must be considered together with NORD (ID of each household member) to merge individual level information.

It is possible to link panel household information using NQUEST. At individual level, NQUEST must be used together with NORDP, contained in CARCOM08, that represents the ID of each household member in the previous wave.

The file CARCOM08 contains all the social-demographic characteristics of each household member and other important information:

PESOFIT	= unit sampling weight (defined at household level)
CFRED	= head of household, defined as the major income earner
ETA	= age (years)
CLETA	= age class (Up to 30 years, 31-40, 41-50, 51-65, more than 65 years)
NCOMP	= N° of household members
NPERC	= N° of household income earners
PERC	= income earner
Q	= working status (1=employee, 2=self-employed, 3=not-employed)
QUAL	= employment status (1= blue-collar worker, 2= office worker or school teacher, 3= cadre or manager, 4= sole proprietor/member of the arts or professions, 5=other self-employed, 6=pensioner, 7=other not-employed)
SETT	= branch of activity (1=agriculture, 2= industry, 3= public administration, 4= other sector, 5= not employed)
AREA3	= geographical area (1=North, 2= Centre, 3=South and Islands)
AREA5	= geographical area (1=North-east, 2= North-west, 3=Centre, 4=South, 5=Islands)
IREG	= Istat code for region of residence (1=Piemonte, 2=Valle d'Aosta, 3=Lombardia, 4=Trentino, 5=Veneto, 6=Friuli, 7=Liguria, 8=Emilia Romagna, 9=Toscana, 10=Umbria, 11=Marche, 12=Lazio, 13=Abruzzo, 14=Molise, 15=Campania, 16=Puglia, 17=Basilicata, 18=Calabria, 19=Sicilia, 20=Sardegna)
NASCREG	= region of birth (Istat code)
NASCAREA	= geographical area of birth (1=North, 2= Centre, 3=South and Islands)
ACOM4C	= town size (0-20.000 inhabitants, 20.000-40.000, 40.000-500.000, more than 500.000 inhabitants).

The file Q08A also includes:

CAPI	= information collected through the CAPI (<i>Computer Assisted Personal interviewing</i>) (code 1) or not (code 0).
------	---

The file Q08C2 also includes the variables AFXX² that indicate the amount held from the family at the end of 2008 in each form of saving.³

The derivation of the aggregate variables contained in RFAM08, RISFAM08, RICFAM08 and RPER08 is explained in the following section (see tables 2, 3 and 4).

All the amounts are expressed in euros

Missing values (doesn't know, unwilling to answer, inapplicable) are indicated with ".".

² Where XX=A1...H indicates the form of saving as reported in the corresponding table in the questionnaire. For example the variable AFC2 contains the amount held from the household at the end of 2008 in CCTs (T-certificates) while AFD1 the amount held in bonds and so on.

³ For those households who have not reported the exact amount held, the figure is imputed using the information on the size class of holding.

4. Aggregate variables

Table 2

Aggregation of variables: income account

Variable name	Description ⁽¹⁾	Questionnaire reference ⁽²⁾
Y	Net disposable income	
YL	Payroll income	
YL1	Net wages and salaries	B1.7
YL2	Fringe benefits	B1.9
YT	Pensions and net transfers	
YTP	Pensions and arrears	
YTP1	Pensions	B5.4 * B5.5
YTP2	Arrears	B5.6
YTA	Other transfers	
YTA1	Financial assistance (wage suppl. etc.)	B6.(a1,a2,a3,b1,b2,b3,b4,b5, b6,b7)
YTA2	Scholarships	B6.c1
YTA3	Alimony and gifts	
YTA31	received	B6.(c2a,c3,c4)
YTA32	paid (-)	E.6(1,2)
YM	Net self-employment income	
YMA1	Self-employment income	B2.11 + B3.10
YMA2	Entrepreneurial income	B4.8 + B4.9
YC	Property income	
YCA	Income from real-estate	
YCA1	Actual rents	D1.11
YCA2	Imputed rents ⁽³⁾	(D.12*12) + D1.12
YCF	Income from financial assets ⁽⁴⁾	
YCF1	Interest on deposits	Rate1*C.26(A,B)
YCF2	Interest on government securities	Rate2*C.26(C)
YCF3	Income from other securities	Rate2*C.26(D,E,F,G,H)
YCF4	Interest payments (-)	Rate3*(D.28(a,b) + D2(1))

$$Y = YL + YT + YM + YC$$

(1) A minus sign indicates the item is included with a negative sign in calculating the aggregate of which it is a component.

(2) The questionnaire reference is coded as follows: section or annex.question (where necessary, line of question).

(3) Excludes buildings used for self-employment.

(4) Interest rate * capital stock.

Table 3**Aggregation of variables: use of income account**

Variable name	Description	Questionnaire reference ⁽¹⁾
Y	Net disposable income	
C	Consumption	
CD	Durables	
CD1	Expenditure for transport equipment	E.2(2) – E.4(2)
CD2	Expenditure for furniture, etc.	E.2(3)
CN	Non-durables	$((E.7 + D.9) * 12) + B1.9 + (D.12 * 12) + D1.12$
S	Saving ⁽²⁾	
$Y = C + S$		

(1) The questionnaire reference is coded as follows: section or annex.question (where necessary, line of question).

(2) Determined as a residual.

Table 4**Aggregation of variables: capital account**

Variable name	Description ⁽¹⁾	Questionnaire reference ⁽²⁾
W	Net wealth	
AR	Real assets	
AR1	Real estate	$D1.9 * D1.2 + D.17 * D.4 + D.22$
AR2	Business equity	$B2.15 + B3.14 + B4.11$
AR3	Valuables	E.5(1)
AF	Financial assets	
AF1	Deposits	C.26 (A,B)
AF2	Government securities	C.26 (C)
AF3	Other securities	C.26 (D,E,F,G,H)
AF4	Trade credit or credit due from other households	$B2.12(5) + B3.11(5) + D.32(1)$
PF	Financial liabilities (-)	
PF1	Liabilities to banks and financial companies	$D.28(1,2,3,4) + B2.12(1,2,3) + B3.11(1,2,3) + D2(1)$
PF2	Trade debt	$B2.12(4) + B3.11(4)$
PF3	Liabilities to other households	D.32(2)
Memorandum item:		
BD	Durables	
BD1	Transport equipment	E.5(2)
BD2	Furniture, etc.	E.5(3)
$W = AR + AF - PF$		

(1) A minus sign indicates the item is included with a negative sign in calculating the aggregate of which it is a component.

(2) The questionnaire reference is coded as follows: section or annex.question (where necessary, line of question).