

SURVEY OF HOUSEHOLD INCOME AND WEALTH 2016

Summary

1. Introduction	1
2. Data format	1
3. Information contained in the datasets	2
4. Aggregate variables	5

1. Introduction

The Survey on Household Income and Wealth (SHIW) covers 7,420 households composed of 16,462 individuals. The number of households interviewed in previous surveys (*panel households*) is 3,804.

Panel households have the same questionnaire number (NQUEST) as the previous surveys.

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 PESIJACK16 containing replication weights that can be used to calculate sampling variance¹.

The datasets are periodically revised; users are advised to use the latest available version.

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 334 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 347. 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 notes. 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 2016 annual database

<i>Dataset</i>	<i>Content</i>	<i>Primary key</i>
	QUESTIONANAIRE DATASETS	
Q16A	Section A (Households' composition)	NQUEST
LAVORO	Section B (Employment)(*)	NQUEST NORD
Q16C1	Section C (Payment Instruments - <i>Questions from C01 to C23</i>)	NQUEST
Q16C2	Section C (Financial Assets and financial information- <i>Questions from C24 to C35</i>)	NQUEST
Q16D	Section D (Properties and debts)	NQUEST
Q16E	Section E (Expenditures)	NQUEST
Q16F	Section F (Insurance)	NQUEST
Q16G	Section G (Information provided by the interviewer)	NQUEST
CARCOM16	Characteristics of the individuals	NQUEST NORD
USCITI	Individuals that left the panel household	NQUEST NORDP
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_RES	Annex D2_RES (Loans for main residence)	NQUEST
ALLD2_AIMM	Annex D2_AIMM (Loans for properties other than principal residence)	NQUEST
ALLD2_FAM	Annex D2_FAM (Loans for consumer credit)	NQUEST
ALLD2_PROF1	Annex D2_PROF1 (Loans for business purposes of family businesses)	NQUEST
ALLD2_PROF2	Annex D2_PROF1 (Loans for business purposes of self-employed)	NQUEST NORD
	DERIVED DATASETS	
RFAM16	Household Incomes	NQUEST
RISFAM16	Household Expenditure and Savings	NQUEST
RICFAM16	Household Wealth	NQUEST
RPER16	Individual Incomes	NQUEST NORD
DEBITI16	Household Debts	NQUEST
PESIACK16	Replication Weights	NQUEST

(*) Information on working prevalent status are contained in CARCOM16.

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 CARCOM16, that represents the ID of each household member in the previous wave.

The file CARCOM16 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)
CLETA5	= age class (Up to 34 years, 35-44, 45-54, 55-64, more than 64 years)
NCOMP	= N° of household members
NPERC	= N° of household income earners
PERC	= income earner
NPERL	= N° of household work income earners
PERL	= work 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)
ISCO	= International Standard Classification of Occupations
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).
ACOM5	= town size (0-5.000 inhabitants, 5.000-20.000, 20.000-50.000, 50.000-200.000 more than 200.000 inhabitants).

The file Q16A also includes:

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

The file Q16C1 does not include DEPBANC, NDEPBANC, IDEPBANC, DEPBANR, NDEPBANR, IDEPBANR, DEPPOSC, NDEPPOSC, IDEPPOSC, DEPPOSR, NDEPPOSR, IDEPPOSR, ABIBA1..8, ABIBAP. The file includes the following variables:

DEPOSIT	= Ownership of at least one bank or postal deposit
NDEPOSIT	= Total number of bank or postal deposits

In the file Q16C2 the following variables are no longer available: POS_A1, POS_A3, POS_A4, POS_B1, CLA_A1 ... H, AMM_A1 ... H, SCL_A1 ... H. They have been replaced by the following set of variables:

Table 1

	Ownership at end 2016	Amount held at end 2016 ⁽¹⁾
FORMS OF SAVING (Section C of questionnaire)	Variables in Q16C2	
Bank or postal deposits (current, saving accounts or deposit books)	= PDEPOS	AFDEPOS
Certificates of deposit	= POS_A3	AFA3
Repos	= POS_A4	AFA4
PO savings certificates	= POS_B3	AFB3
BOTs (T-bills)	= POS_C1	AFC1
CCTs (T-certificates)	= POS_C2	AFC2
BTPs (T-bonds)	= POS_C3	AFC3
inflation-indexed BTPs (T-bonds)	= POS_C4	AFC4
CTZs (zero coupon)	= POS_C5	AFC5
Other (CTEs, CTOs et al.)	= POS_C6	AFC6
Bonds issued by Italian firms	= POS_D1	AFD1
Bonds issues by Italian banks	= POS_D2	AFD2
Funds or ETFs in money market or in liquidity	= POS_D3	AFD3
Funds or ETFs in bonds	= POS_D4	AFD4
Balanced (or mixed) funds or ETFs	= POS_D5	AFD5
Funds or ETFs in equities	= POS_D6	AFD6
Funds or ETFs in foreign currencies	= POS_D7	AFD7
Shares of listed companies (at their market value at end-2016)	= POS_E1	AFE1
Shares of unlisted companies (at their estimated realizable value at end-2016)	= POS_E3	AFE3
Shares in companies limited by shares - srl (at their estimated realisable value at end-2016)	= POS_E4	AFE4
Shares of partnerships (at their estimated realizable value at end-2016)	= POS_E5	AFE5
Managed savings	= POS_F	AFF
Foreign government securities	= POS_G1	AFG1
Foreign bonds	= POS_G2	AFG2
Foreign Shares and equities	= POS_G3	AFG3
Other foreign securities	= POS_G4	AFG4
Loans to cooperatives	= POS_H	AFH
Other financial assets (options, futures, royalties, etc.)	= POS_I	AFI
(1) For those households who have not reported the exact amount held, the figure is imputed using the information on the size class of holding.		

The derivation of the aggregate variables contained in RFAM16, RISFAM16, RICFAM16, RPER16 and DEBITI16 is explained in the following section (see tables 2, 3, 4 and 5).

All the amounts are expressed in euros

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

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	B5.7
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.8(1,2)
YM	Net self-employment income	
YMA1	Self-employment income	B2.10 + B3.9
YMA2	Entrepreneurial income	B4.7 + B4.8
YC	Property income	
YCA	Income from real-estate	
YCA1	Actual rents	D1.12 + D1.12b
YCA2	Imputed rents ⁽³⁾	(D.16a + D16b)*12 + (D1.13a + D1.13b)
YCF	Income from financial assets ⁽⁴⁾	
YCF1	Interest on deposits	Rate1*(1-Tax1)*C.25(A,B)
YCF2	Interest on government securities	Rate2*(1-Tax2)*C.25(C)
YCF3	Income from other securities	Rate2*(1-Tax1)*C.25(D1,D3,D4,D5,D6,D7,E3,E4 ,E5,F,G,H,I) + Rate3*(1-Tax1)*C.25(E1) + Rate4*(1-Tax1)*C.25(D2)
YCF4	Interest payments (-)	Rate5*(D2_RES.7+D2_AIMM.7)
CLY	Income deciles	
CLT2	Income quintiles	

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

(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*(1-Tax imposed)*capital stock. The rates are the following: rate1=0,4953%, rate2=0,4529%, rate3=2,74%, rate4=2,8466%, rate5=3,4585%, Tax1=26%, Tax2=12,5%

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.2(3)-E.4(2)-E.4(3)
CD2	Expenditure for furniture, etc.	E.2(4)
CN	Non-durables	(E.09a+E.09b+E.10b+E.11b/12 + E.13b/12+E.14b)*12+D.11*12 +YL2+YCA2
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.21*D.4 + D.25
AR2	Business equity	B2.15 + B3.14 + B4.10
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,I)
AF4	Trade credit or credit due from other households	B2.13(2) + B3.12(2) + D.37(1)
PF	Financial liabilities (-)	
PF1	Liabilities to banks and financial companies	D2_RES.7+D2_AIMM.7+D2_FAM.6 +D2_PROF.6+C.16+C.12
PF2	Trade debt	B2.13(1) + B3.12(1)
PF3	Liabilities to other households	D.37(2)
CLW	Wealth deciles	
CLW2	Wealth quintiles	
Memorandum item:		
BD	Durables	
BD1	Transport equipment	E.5(2) + E.5(3)
BD2	Furniture, etc.	E.5(4)

$$W = AR + AF - PF$$

(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).

Tav.5**Aggregation of variables: debts**

Nome	Descrizione	Riferimenti al questionario (1)
RATADEB	Installment debt paid in the current year	
RATADEB_RES	for the principal residence	D2_RES.11
RATADEB_AIMM	for other properties	D2_AIMM.11
RATADEB_FAM	for other household needs	D2_FAM.7
RATADEB_PROF	for business purposes	D2_PROF.7
PFIMM	Loans for real properties	D2_RES.7 + D2_AIMM.7
PFCONS	Loans for consumption	
TDEBITC	motor vehicles	D2_FAM.6 se D2_FAM.1 = 1
TDEBITD	furniture, appliances, etc	D2_FAM.6 se D2_FAM.1 = 2
TDEBITE	non-durable goods	D2_FAM.6 se D2_FAM.1 = 3
TDEBITF	other purchases or daily expenses	D2_FAM.6 se D2_FAM.1 = 4
TDEBITG	education expenses	D2_FAM.6 se D2_FAM.1 = 5
PFAZ	Loans for business purposes.	D2_PROF.6
PFCARTE	Outstanding balance on credit cards	C.16
PFCC	Overdrafts	C.12
PFCOLL	Collateralized loans	D2_RES.7 [se D2_RES.12 in (1,2)] + D2_AIMM.7 [se D2_AIMM.12 in (1,2)] + D2_FAM.6 [se D2_FAM.10 in (1,2,3)] + D2_PROF.6 [se D2_PROF.9 in (1,2,3)]
PFNOCOLL	Non collateralized loans	D2_RES.7 (se D2_RES.12=3) + D2_AIMM.7 (se D2_AIMM.12 = 3) + D2_FAM.6 [se D2_FAM.10 in (4,5)] + D2_PROF.6 (se D2_PROF.9 = 4)+C.16+C.12

RATADEB = RATADEB_RES+ RATADEB_AIMM+ RATADEB_FAM+ RATADEB_PROF

PFCONS = TDEBITC+ TDEBITD+ TDEBITE+ TDEBITF+ TDEBITG

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