SURVEY OF HOUSEHOLD INCOME AND WEALTH 2020

Summary

1. Introduction

The Survey on Household Income and Wealth (SHIW) covers 6,239 households composed of 15,198 individuals. The number of households interviewed in previous surveys (*panel households*) is 2,986.

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

In order to improve sample representativeness of some population groups, in the 2020 survey the traditional sampling design in use since the 1989 edition was partly amended. The new design improves the quality of the estimators produced for economic analysis, but it makes comparison with previous waves more difficult. In order to overcome this limitation, two different weighting schemes were devised: one to be used for cross-section analysis, which includes the PESOFIT (sampling weight) and PESOFIT2 (population weight) variables in the CARCOM20 annual database and, for calculating sample variance, the PWTx variables in the PESIJACK20 database[1]; one to be used for the purposes of historical comparison, which includes the PESO (sampling weight) and PESOPOP (population weight) variables in the PESO[2] historical database and, for calculating sample variance, PWTx variables in the PESIJACK file from the historical database, version 11.1.

Using sampling weights is recommended to obtain unbiased estimates[3].

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 IND*xx_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**).

---

[1] The dataset contains 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 over 300. 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.

[2] Weights for historical comparison were constructed using a statistical rebalancing technique (raking) that adjusts the new design weights so as to assign each household interviewed in 2020 the same probability of being interviewed as under the previous design. For more information, see R. Gambacorta and E. Porreca, 'Bridging techniques in the redesign of the Italian survey on household income and wealth', Banca d'Italia, Questioni di Economia e Finanza (Occasional Papers), forthcoming.

[3] For a detailed description of the weighting scheme in the IBF, see I. Faiella and R. Gambacorta, 'The weighting process in the SHIW', Banca d'Italia, Temi di Discussione (Working Papers), 636, 2007.

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.

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 2020 annual database

Dataset	Content	Primary key
	QUESTIONANAIRE DATASETS	
Q20A	Section A (Households' composition)	NQUEST
Q20B	Section B (Employment and Income)	NQUEST
Q20C1	Section C (Payment Instruments - *Questions from C01 to C29*)	NQUEST
Q20C2	Section C (Financial Assets and financial information- *Questions from C30 to C47*)	NQUEST
Q20D	Section D (Properties and debts)	NQUEST
Q20E	Section E (Expenditures)	NQUEST
Q20F	Section F (Insurance)	NQUEST
Q20G	Section G (Information provided by the interviewer)	NQUEST
CARCOM20	Characteristics of the individuals	NQUEST NORD
USCITI	Individuals that left the panel household	NQUEST NORDP
ALLB1	Annex B1 (Payroll employees)	NQUEST NORD NATTIVITA
ALLB2	Annex B2 (Self-employed worker)	NQUEST NORD NATTIVITA
ALLB3	Annex B3 (Atypical contracts)	NQUEST NORD NATTIVITA
ALLB4	Annex B4 (Pensions)	NQUEST NORD NATTIVITA
ALLB5	Annex B5 (Other income sources)	NQUEST NORD
ALLD1	Annex D1 (Property- information on 3 properties other than principal residence)	NQUEST NORDIMM
ALLD1B	Annex D1B (Other Properties)	NQUEST
ALLD2_RES	Annex D2_RES (Loans for main residence)	NQUEST NORDEB1
ALLD2_AIMM	Annex D2_AIMM (Loans for properties other than principal residence)	NQUEST NORDEB2
ALLD2_FAM	Annex D2_FAM (Loans for consumer credit)	NQUEST NORDEB3
ALLD2_PROF	Annex D2_PROF (Loans for business purposes of self-employed)	NQUEST NORDEB4
ALLF1	Annex F1 (Supplementary pension plans)	NQUEST NORD
ALLF2	Annex F2 (Insurance policies)	NQUEST NORD
	DERIVED DATASETS	
RFAM20	Household Incomes	NQUEST
RISFAM20	Household Expenditure and Savings	NQUEST
RICFAM20	Household Wealth	NQUEST
RPER20	Individual Incomes	NQUEST NORD
DEBITI20	Household Debts	NQUEST
PESIJACK20	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 CARCOM20, that represents the ID of each household member in the previous wave.

The file CARCOM20 contains all the social-demographic characteristics of each household member and other important information:

PARENTP = Position of the ex-panel household member compared with the R.P. of the household of origin. This variable is set only where no ex-panel household member was part of the household of origin (NORDP is missing for all household members), but there are family ties.
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)
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 Q20A also includes:

TIPOI = face to face interview (code 1) or phone interview (code 2).

The file Q20C1 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 Q20C2 the following variables are no longer available: POS_A1, … POS_H, AMM_A1 … AMM_H, SCL_A1, … SCL_H. They have been replaced by the following set of variables

		Ownership at end 2020	Amount held at end 2020 [1]
FORMS OF SAVING (Section C of questionnaire)		**Variables in Q20C2**	
Bank or postal deposits (current, saving accounts or deposit books)	=	PDEPOS	AFDEPOS
Repos / PO savings certificates	=	POS_A3	AFA3
Managed savings	=	POS_B	AFB
BOTs (T-bills)	=	POS_C1	AFC1
BTPs (T-bonds) / inflation-indexed BTPs (T-bonds)	=	POS_C2	AFC2
Other (CTZs, CCTs et al.)	=	POS_C3	AFC3
Bonds issued by Italian firms	=	POS_D1	AFD1
Bonds issues by Italian banks	=	POS_D2	AFD2
Funds or ETFs	=	POS_D3	AFD3
Shares of listed companies (at their market value at end-2020)	=	POS_E1	AFE1
Shares in companies limited by shares – srl - and shares of unlisted companies (at their estimated realizable value at end-2020)	=	POS_E2	AFE2
Foreign deposits	=	POS_F1	AFF1
Other foreign financial assets (Foreign government securities, Foreign bonds, foreign shares, etc..)	=	POS_F2	AFF2
Loans to coperatives (social loans, etc.)	=	POS_G	AFG
Other financial assets (options, futures, royalties, etc.)	=	POS_H	AFH

(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 RFAM20, RISFAM20, RICFAM20, RPER20 and DEBITI20 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	B1.6
YL2	Fringe benefits	B1.8
YT	Pensions and net transfers	
YTP	Pensions and arrears	
YTP1	Pensions	B4.4 * B4.5
YTP2	Arrears	B4.7
YTA	Other transfers	
YTA1	Financial assistance (wage suppl. etc.)	
YTA11	CIG	B5.a1 + CIG component in B1.6[3]
YTA12	Altra assistenza economica	B5.(a2,a3,b1,b2,b3)
YTA2	Scholarships	B5.c1
YTA3	Alimony and gifts	
YTA31	Received	B5.(c2a,c3,c4,c5)
YTA32	paid (-)	E.9 (1,2)
YTA4	COVID payments (excl. CIG)[4]	B.30a B.41
YTA5	Minimum income/pension scheme[4]	B.38 B.39
YM	Net self-employment income	B2.12 + B2.19
YC	Property income	
YCA	Income from real-estate	D1.13 + D1B.4
YCA1	Actual rents	D.16 * 12 + D1.14 + D1B.6
YCA2	Imputed rents [5]	
YCF	Income from financial assets [6]	Rate1*(1-Tax1)*C.32A
YCF1	Interest on deposits	Rate2*(1-Tax2)*C.32C
YCF2	Interest on government securities	Rate2*(1-Tax1)*C.32(B,D1,D3, E2,F,G,H) + Rate3*(1-Tax1)*C.32(E1) + Rate4*(1-Tax1)*C.32(D2)
YCF3	Income from other securities	Rate5*(D2_RES.19+D2_AIMM.19)
YCF4	Interest payments (-)	B2.12 + B2.19
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) The CIG component of the income from employment was estimated when the respondent was not able to isolate it.
(4) In RPER20 the amount is given to the R.P. (NORD=1).
(5) Excludes buildings used for self-employment.
(6) Interest rate*(1-Tax imposed)*capital stock. The rates are the following: rate1=0,4188%, rate2=0,5328%, rate3=3,55%, rate4=2,038%, rate5=1,7943%, 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 +D1B.2+ D.21*D.4 + D.25
AR2	Business equity	B2.17 + B2.20
AR3	Valuables	E.5(1)
AF	Financial assets	
AF1	Deposits	C.32A
AF2	Government securities	C.32C
AF3	Other securities	C.32 (B,D,E,F,G,H)
AF4	Trade credit or credit due from other households	B2.14(2) + D.32(1)
PF	Financial liabilities (-)	
PF1	Liabilities to banks and financial companies	D2_RES.19+D2_AIMM.19+D2_FAM.12 +D2_PROF.8 + C.13 + C.17
PF2	Trade debt	B2.14(1)
PF3	Liabilities to other households	D.32(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).

<div align="right">**Table 5**</div>

<div align="center">**Aggregation of variables: debts**</div>

Name	Description	Questionnaire reference[1]
RATADEB	Installment debt paid in the current year	
RATADEB_RES	for the principal residence	D2_RES.18
RATADEB_AIMM	for other properties	D2_AIMM.18
RATADEB_FAM	for other household needs	D2_FAM.11
RATADEB_PROF	for business purposes	D2_PROF.7
PFIMM	Loans for real properties	D2_RES.19 + D2_AIMM.19
PFCONS	Loans for consumption	
TDEBITC	motor vehicles	D2_FAM.12 se D2_FAM.1 = 1
TDEBITD	furniture, appliances, etc	D2_FAM.12 se D2_FAM.1 = 2
TDEBITE	non-durable goods	D2_FAM.12 se D2_FAM.1 = 3
TDEBITF	other purchases or daily expenses	D2_FAM.12 se D2_FAM.1 = 4
TDEBITG	education expenses	D2_FAM.12 se D2_FAM.1 = 5
PFAZ	Loans for business purposes.	D2_PROF.8
PFCARTE	Outstanding balance on credit cards	C.17
PFCC	Overdrafts	C.13
PFCOLL	Collateralized loans	D2_RES.19 [se D2_RES.8 in (1,2)] + D2_AIMM.19 [se D2_AIMM.9 in (1,2,3)] + D2_FAM.12 [se D2_FAM.2 in (1,2,3)] + D2_PROF.8 [se D2_PROF.10 in (1,2,3)]
PFNOCOLL	Non collateralized loans	D2_RES.19 (se D2_RES.8 = 3) + D2_AIMM.19 (se D2_AIMM.9 = 4) + D2_FAM.12 [se D2_FAM.2 in (4,5)] + D2_PROF.8 (se D2_PROF.10= 4) + C.13 + C.17

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).