



Representativity Indicators for Survey Quality

RISQ data set documentation
Header document

Deliverables 1.0-1.11

CBS

23 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

The following data sets are documented. Details can be found in documents RISQ deliverable 1.1 to 1.11.

<i>Data set</i>	<i>Year</i>	<i>Beneficiary</i>	<i>Deliverable no.</i>	<i>Document</i>
Health survey	2005	CBS	1.1	RISQ deliverable 1.1
Consumer satisfaction survey	2005	CBS	1.2	RISQ deliverable 1.2
Short-term statistic Industry	2007	CBS	1.3	RISQ deliverable 1.3
Short-term statistic Retail	2007	CBS	1.4	RISQ deliverable 1.4
Census link study	2001	SOTON	1.5	RISQ deliverable 1.5
ESS	2006	SSB	1.6	RISQ deliverable 1.6
Level of living	2004	SSB	1.7	RISQ deliverable 1.7
ESS	2006	KUL	1.8	RISQ deliverable 1.8
Flemish Housing survey	2005	KUL	1.9	RISQ deliverable 1.9
ICT	2007	SURS	1.10	RISQ deliverable 1.10
LFS	2007	SURS	1.11	RISQ deliverable 1.11



Representativity Indicators for Survey Quality

RISQ data set documentation
Dutch Health survey 2005

Deliverable 1.1 (WP2)

José Gouweleeuw, Barry Schouten
CBS

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Gezondheidsenquête In English: Health Survey
Abstract:	The Health Survey is a continuous survey with questions about health, life style and use of medical care. It consists of three questionnaires; a CAPI base module, a CAPI topical module about health and a supplementary paper questionnaire.
Topic classification:	Health, life style, use of medical care
Type of survey	Cross section
Unit of study	Individuals
Target population	All persons of 4 years and older in The Netherlands
Producer:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Data file name:	RISQ - Gezo 2005.sav
File contents:	Data file contains response indicator, several auxiliary variables from administrative data.
File structure:	Rectangular
Number of cases:	15411
Number of variables	50
Design weighting:	Equal design weights for all persons
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	Missing items are set at SPSS system missing
Data collector:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Mode of data collection:	CAPI and paper
Field work period(s):	Monthly samples. For each sample fieldwork period is one month, i.e. total fieldwork period is 01.01.02 – 31.12.02.
Sampling procedure:	The survey is a two-stage sample, in which the clusters in the first stage are formed by municipalities. From the clusters simple random samples without replacement are drawn consisting of persons.
Interviewers:	250 interviewers working in 13 interviewer regions. Experience and age of interviewers varies.

Advance information	Persons receive an individual pre-notification letter
Call schedules	Not applicable
Incentives	No
Refusal conversion	No explicit refusal conversion procedures are applied

Response rates

		Frequency	Percent
Valid	Respondent	10378	67,3%
	Refusal	2952	19,2%
	Unable	895	5,8%
	Language problems	257	1,7%
	Non-contact	555	3,6%
	Moved	338	2,2%
	Other	36	0,2%
	Total	15411	100,0%

Codebook of variables

Variable number	1
Variable name	Rinpersoon
Variable label	RINPerso
Variable type	String
Values and value labels	Unique anonymised identification key

Variable number	2
Variable name	Ethnic background - generation
Variable label	Etn
Variable type	Nominal
Values and value labels	

		Frequency	Percent
Valid	Native	12664	82,2
	1st generation non-native	1369	8,9
	2nd generation non-native – one parent	817	5,3
	2nd generation non-native – two parents	540	3,5
	Total	15390	99,9

Missing	System	21	,1
Total		15411	100,0

Variable number 3
Variable name Ethnic background - group
Variable label Etn
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Native	12664	82,2
	Moroccan	266	1,7
	Turkish	305	2,0
	Surinam	268	1,7
	Netherlands Antilles and Aruba	83	,5
	Other non-westerners	494	3,2
	Other westerners	1310	8,5
	Total	15390	99,9
Missing	System	21	,1
Total		15411	100,0

Variable number 4
Variable name Average house value (WOZ) at zip-code
Variable label Wozgem
Variable type Continuous
Values and value labels 64 missings

	N	Minimum	Maximum	Mean	Std. Deviation
Average house value	15347	3	1983	213,58	109,769

Variable number 5
Variable name Percentage non-western non-native at zip-code
Variable label Percnwest
Variable type Continuous
Values and value labels 3 missings

	N	Minimum	Maximum	Mean	Std. Deviation
Perc. non-western non-native	15408	,00	1,00	,0960	,16824

Variable number 6
Variable name Percentage non-native at zip-code
Variable label Percall
Variable type Continuous
Values and value labels 3 missings

	N	Minimum	Maximum	Mean	Std. Deviation
Perc non-native	15408	,00	1,00	,1812	,18854

Variable number 7
Variable name Appointment and re-allocation during fieldwork status
Variable label Voorgang
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	No re-allocation, no appointment	9778	63,4
	No-re-allocation, appointment	5303	34,4
	Re-allocation, no appointment	223	1,4
	Re-allocation, appointment	107	,7
	Total	15411	100,0

Variable number 8
Variable name Number of contact attempts including contact
Variable label Benaderi
Variable type Discrete
Values and value labels

		Frequency	Percent
Valid	0	223	1,4
	1	4311	28,0
	2	5453	35,4
	3	2858	18,5
	4	1343	8,7
	5	501	3,3
	6	371	2,4
	7	116	,8
	Total	15176	98,5

Missing	System	235	1,5
Total		15411	100,0

Variable number

9

Variable name

Interviewer district

Variable label

Regio

Variable type

Nominal

Values and value labels

		Frequency	Percent
Valid	1	1142	7,4
	2	1540	10,0
	3	1521	9,9
	4	1147	7,4
	5	551	3,6
	6	1392	9,0
	7	970	6,3
	8	734	4,8
	9	1232	8,0
	10	1170	7,6
	11	1463	9,5
	12	1064	6,9
	13	1447	9,4
	Unknown	38	,2
	Total	15411	100,0

Variable number

10

Variable name

Type of nonresponse

Variable label

Rescode

Variable type

Nominal

Values and value labels

See first section

Variable number

11 - 17

Variable name

Date of contact I (i=1,2,...,7)

Variable label

Dag*i* with i=1,2,3,...,7

Variable type

Date

Values and value labels

Between January 1 and December 31, 2005

Variable number

18 - 24

Variable name

Time of contact i (i=1,2,...,7)

Variable label

Tijd*i* with i=1,2,3,...,7

Variable type

Values and value labels

Time of contact

<12h, 12-17h, 17-19h, 19-21h, >21h

Variable number

Variable name

Variable label

Variable type

Values and value labels

25

Response

respons

Nominal

		Frequency	Percent
Valid	Respondent	10378	67,3%
	Non-respondent	5033	32,7%
	Total	15411	100,0%

Variable number

Variable name

Variable label

Variable type

Values and value labels

26

Number of persons in household

Aantper

Discrete

Aantal personen

		Frequency	Percent
Valid	1	2420	15,7
	2	4511	29,3
	3	2597	16,9
	4	3602	23,4
	5	1591	10,3
	6	395	2,6
	7	106	,7
	8	43	,3
	9	21	,1
	10	7	,0
	11	6	,0
	12	3	,0
	13	4	,0
	14	1	,0
	15	1	,0
	Total	15308	99,3
Missing	System	103	,7
Total		15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

27

Date of subscription to Center of Work and Income (CWI)

Aanvcwi

Date

Dates between November 1979 and December 2005, 716 persons have subscription at date of interview

Variable number

Variable name

Variable label

Variable type

Values and value labels

28

Social security allowance at month of interview

abw

Ordinal

		Frequency	Percent
Valid	No	15013	97,4
	Yes	398	2,6
	Total	15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

29

Disability allowance at month of interview

ao

Ordinal

		Frequency	Percent
Valid	No	14585	94,6
	Yes	826	5,4
	Total	15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

30

Paid job at month of interview

Baan

Ordinal

		Frequency	Percent
Valid	No	8780	57,0

Yes	6631	43,0
Total	15411	100,0

Variable number 31

Variable name Marital status

Variable label burgst

Variable type Nominal

Values and value labels

		Frequency	Percent
Valid	Not married	6841	44,4
	Married	6817	44,2
	Widow	758	4,9
	Divorced	918	6,0
	Partnership (formerly married)	46	,3
	Total	15380	99,8
Missing	System	31	,2
Total		15411	100,0

Variable number 32

Variable name Size of municipality

Variable label Gemgrootte

Variable type Ordinal

Values and value labels

		Frequency	Percent
Valid	< 5 000 inhabitants	39	,3
	5 000 - < 10 000 inhabitants	570	3,7
	10 000 - < 20 000 inhabitants	2157	14,0
	20 000 - < 50 000 inhabitants	5384	34,9
	50 000 - < 100 000 inhabitants	2764	17,9
	100 000 - < 150 000 inhabitants	1463	9,5
	150 000 - < 250 000 inhabitants	1340	8,7
	250 000 inhabitants or more	1663	10,8
	Total	15380	99,8

Missing	System	31	,2
Total		15411	100,0

Variable number 33
Variable name Municipality code
Variable label Gem2005
Variable type Nominal
Values and value labels Codes ranging from 0003 to 1955, 31 missing values

Variable number 34
Variable name Dwelling is owned by household
Variable label Huurkoop
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Owned	9659	62,7
	Rented	5634	36,6
	Total	15293	99,2
Missing	System	118	,8
Total		15411	100,0

Variable number 35
Variable name Position in household
Variable label Plhhn
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Child	4191	27,2
	Single	2384	15,5
	Partner in unmarried couple without children	909	5,9
	Partner in married couple without children	3107	20,2

	Partner in unmarried couple with children	464	3,0
	Partner in married couple with children	3552	23,0
	Parent in single parent household	443	2,9
	Reference person other types of households	47	,3
	Other member household	187	1,2
	Member institutionalized household	24	,2
	Total	15308	99,3
Missing	System	103	,7
Total		15411	100,0

Variable number 36
Variable name Pension
Variable label Pensioen
Variable type Ordinal
Values and value labels

		Frequency	Percent
Valid	No	12703	82,4
	Yes	2708	17,6
	Total	15411	100,0

Variable number 37
Variable name Province of residence
Variable label Provincie
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Groningen	522	3,4
	Friesland	623	4,0
	Drenthe	503	3,3
	Overijssel	1102	7,2
	Flevoland	330	2,1
	Gelderland	1743	11,3
	Utrecht	1170	7,6

	Noord-Holland	2210	14,3
	Zuid-Holland	3214	20,9
	Zeeland	354	2,3
	Noord-Brabant	2442	15,8
	Limburg	1167	7,6
	Total	15380	99,8
Missing	System	31	,2
Total		15411	100,0

Variable number 38
Variable name Social-economic status
Variable label Sec
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Employee	5966	38,7
	Self-employed	639	4,1
	Disable (AO)	575	3,7
	Unemployed (WW)	224	1,5
	Social security (ABW)	345	2,2
	Other allowance	177	1,1
	Pension	2487	16,1
	Student	3089	20,0
	Other non-active	1882	12,2
	Total	15384	99,8
Missing	System	27	,2
Total		15411	100,0

Variable number 39
Variable name Degree of urbanization
Variable label stedbuurt
Variable type Ordinal
Values and value labels

		Frequency	Percent
Valid	Very strong	2596	16,8

	Strong	3680	23,9
	Moderate	2996	19,4
	Little	3042	19,7
	Not	3065	19,9
	Total	15379	99,8
Missing	System	32	,2
Total		15411	100,0

Variable number 40
Variable name Type of household
Variable label Typhhn
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Single person	2384	15,5
	Unmarried couple without children	912	5,9
	Married couple without children	3138	20,4
	Unmarried couple with children	793	5,1
	Married couple with children	6873	44,6
	Single parent household	1066	6,9
	Other type	118	,8
	Institutional household	24	,2
	Total	15308	99,3
Missing	System	103	,7
Total		15411	100,0

Variable number 41
Variable name Self-employment
Variable label Zelfst
Variable type Ordinal
Values and value labels

		Frequency	Percent
Valid	No	14597	94,7
	Yes	814	5,3
	Total	15411	100,0

Variable number 42
Variable name Income in interview month
Variable label Mndbedr
Variable type Continuous
Values and value labels 4256 persons without income

	N	Minimum	Maximum	Mean	Std. Deviation
Income	11157	-6924	71655	2110,72	2138,493

Variable number 43
Variable name House value
Variable label woz
Variable type Continuous
Values and value labels 119 persons with missing house value

	N	Minimum	Maximum	Mean	Std. Deviation
House value	15293	0	68249000	221472,98	579502,955

Variable number 44
Variable name Self-employed at least one month in 2005
Variable label ooitzelfst
Variable type Ordinal
Values and value labels

		Frequency	Percent
Valid	No	14597	94,7
	Yes	814	5,3
	Total	15411	100,0

Variable number 45
Variable name Social security at least one month in 2005
Variable label ooitabw
Variable type Ordinal

Values and value labels

		Frequency	Percent
Valid	No	14953	97,0
	Yes	458	3,0
	Total	15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

46

Disablity allowance at least one month in 2005

ooitao

Ordinal

		Frequency	Percent
Valid	No	14541	94,4
	Yes	870	5,6
	Total	15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

47

Paid job at least one month in 2005

ooitwerknemer

Ordinal

		Frequency	Percent
Valid	No	8286	53,8
	Yes	7125	46,2
	Total	15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

48

Age

leeftijd

Discrete

Age ranging from 0 to 99, one missing

Variable number

49

Variable name

Variable label

Variable type

Values and value labels

Month of interview

Intmnd

Discrete

		Frequency	Percent
Valid	Jan	1084	7,0
	Feb	1219	7,9
	Mar	1316	8,5
	Apr	1366	8,9
	May	1223	7,9
	Jun	1358	8,8
	Jul	1265	8,2
	Aug	1381	9,0
	Sep	1297	8,4
	Okt	1316	8,5
	Nov	1381	9,0
	Dec	1205	7,8
	Total	15411	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

50

Duration in months subscription to Center for Work and Income (CWI)

Cwiduur

Discrete

Ranging from -2 to 308, 715 persons with CWI subscription



Representativity Indicators for Survey Quality

RISQ data set documentation
Dutch Consumer Satisfaction survey 2005

Deliverable 1.2 (WP2)

Annemieke Luiten, José Gouweleeuw
CBS

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Consumer Confidence Survey
Abstract:	The Consumer Confidence Survey is a continuous survey with questions about general economic development, and the financial situation of the household. The survey is meant to provide insight into short term economic development, and early indicators of differences in consumer trends.
Topic classification:	Consumer confidence, short term economic development, financial situation of household
Type of survey	Cross section
Unit of study	Household
Target population	All households in The Netherlands
Producer:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Data file name:	RISQ - CCO 2005.sav
File contents:	Data file contains response indicator, several auxiliary variables from administrative data.
File structure:	Rectangular
Number of cases:	17.908
Number of variables	183
Design weighting:	Equal design weights for all households
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	Missing items are set at SPSS system missing
Data collector:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Mode of data collection:	CATI
Field work period(s):	Monthly samples. Each sample fieldwork period is the first ten workdays of the month.
Sampling procedure:	The survey is a two-stage sample, in which the clusters in the first stage are formed by municipalities. From the clusters simple random samples without replacement are drawn consisting of addresses.
Interviewers:	50 interviewers working in a centralised telephone unit.

Experience and age of interviewers varies.

Advance information

Persons receive an individual pre-notification letter

Call schedules

First two call attempts in the evening, after that spreading over time and days.

Incentives

No

Refusal conversion

No explicit refusal conversion procedures are applied

Response rates

		Frequency	Percent
Valid	response	11.870	66,9
	Refusal	4.905	27,7
	noncontact	958	5,4
	Total	17.733	100,0

Codebook of variables

Variable number	1
Variable name	volgnum
Variable label	VOLGNUMM
Variable type	String
Values and value labels	Unique anonymised identification key

Variable number	2
Variable name	Ntot
Variable label	totaal aantal contactpogingen (number of contact attempts)
Variable type	Continuous
Values and value labels	

Variable number	3
Variable name	Ntocont
Variable label	aantal pogingen tot aan 1e contact (number of contact attempts to first contact)
Variable type	Discrete
Values and value labels	

Variable number	4
Variable name	eindres
Variable label	endresult
Variable type	Nominal
Values and value labels	See table with response rates

Variable number	5 (also var number 11, 18, 26, 33, 40, 47, 54, 61, 68, 75, 82, 89, 96, 103, 110, 117, 124, 131, 138, 145) for consecutive call attempts
Variable name	dialdat1, dialdat2...dialdat21
Variable label	DialDate
Variable type	Date
Values and value labels	Between January 1 and December 31, 2005

Variable number	6 (also var number 12, 19, 27, 34, 41, 48, 55, 62, 69, 76, 83, 90, 97, 104, 111, 118, 125, 132, 139, 146) for consecutive call attempts
Variable name	dialtime1, dialtime2...dialtime21
Variable label	Dialtime
Variable type	Date
Values and value labels	

Variable number	7 (also var number 13, 20, 28, 35, 42, 49, 56, 63, 70, 77, 84, 91, 98, 105, 112, 119, 126, 133, 140, 147) for consecutive call attempts
Variable name	result1
Variable label	Resultaat (result; result of call attempt)
Variable type	Nominal
Values and value labels	

Result 1		Frequency	Percent
Valid	Response	5.578	31,1
	No contact	4.119	23,0
	Line busy	472	2,6
	Appointment	3.414	19,1
	Refusal	2.457	13,7

	Answering Machine	1.718	9,6
	Disconnected	141	0,8
	Other	7	0,0
	Total	17.906	100,0
Missing	System	2	0,0
Total		17.908	100,0

Variable number 8 (also var number 14, 21, 29, 36, 43, 50, 57, 64, 71, 78, 85, 92, 99, 106, 113, 120, 127, 134, 141, 148) for consecutive call attempts

Variable name weekday1

Variable label Dag van de week van laatste contact (weekday of contact attempt)

Variable type Discrete

Values and value labels Monday 2
Tuesday 3
Wednesday 4
Thursday 5
Friday 6

		Frequency	Percent
Valid	Monday	3.208	17,9
	Tuesday	3.623	20,2
	Wednesday	3.557	19,9
	Thursday	3.221	18,0
	friday	4.297	24,0
	Total	17.906	100,0
Missing	System	2	0,0
Total		17.908	100,0

Variable number 9 (also var number 15, 22, 30, 37, 44, 51, 58, 65, 72, 79, 86, 93, 100, 107, 114, 121, 128, 135, 142, 149) for consecutive call attempts

Variable name dagdeel1

Variable label dagdeel laatste contactpoging (part of the day of contact attempt)

Variable type Nominal

Values and value labels

		Frequency	Percent
Valid	1 morning	914	5,1
	2 afternoon	1.407	7,9
	3 evening	15.585	87,0
	Total	17.906	100,0
Missing	System	2	0,0
Total		17.908	100,0

Variable number 10 (also var number 16, 23, 31, 38, 45, 52, 59, 66, 73, 80, 87, 94, 101, 108, 115, 122, 129, 136, 143, 150) for consecutive call attempts

Variable name Belpogin2

Variable label Belpoging (number of next contact attempt)

Variable type Discrete

Values and value labels

Variable number 151

Variable name GESLACHT_mean

Variable label gemiddelde sekse hh (mean household composition: sex)

Variable type nominal

Values and value labels 1 = man / men 2 = woman / women 3 = mixed 99 multiple household, therefore undetermined

	Frequency	Percent
Valid Men	2.058	11,5
Mixed	11.419	63,8
Women	3.826	21,4
Multiple households	605	3,4
Total	17.908	100,0

Variable number 152

Variable name period

Variable label Waarneemperiode (fieldwork period in months)

Variable type nominal

Values and value labels 1 = january, 2 = February etc

Variable number 153

Variable name NbewonersPC

Variable label aantal personen op dit adres (number of persons on this address)

Variable type Discrete

Values and value labels

	Frequency	Percent
Valid 1	4.719	26,4
2	6.541	36,5
3	2.297	12,8
4	2.629	14,7
5	971	5,4

6	231	1,3
7	72	0,4
8	30	0,2
9	11	0,1
10	4	0,0
11	3	0,0
12	3	0,0
13	3	0,0
14	1	0,0
15	1	0,0
17	1	0,0
20	1	0,0
21	1	0,0
23	1	0,0
67	1	0,0
Total	17.521	97,8
Missing System	387	2,2
Total	17.908	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

154

abw_sum

HH HEEFT WEL/GEEN BIJSTANDSUITKERING
(sum of persons in household on social security
in month of interview)

Discrete

There are 297 households with one or more
persons on social security

		Frequency	Percent
Valid	1	239	1,3
	2	58	0,3
	Total	297	1,7
Missing	System	17.611	98,3
Total		17.908	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

155

ao_sum

WEL/GEEN AO-UITKERING (sum of persons in
household with disability allowance in month of
interview)

Discrete

There are 1650 households with one or more
persons on disability allowance at least one
month in 2005; label >2 means that there are
multiple households at the address; recode to
missing is an option

		Frequency	Percent
Valid	1	1.516	8,5
	2	127	0,7
	3	3	0,0
	4	1	0,0
	5	1	0,0
	6	1	0,0
	17	1	0,0
	Total	1.650	9,2
Missing	System	16.258	90,8
Total		17.908	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

156

baan_sum

WEL/GEEN WERKNEMER (number of persons in household that are employed at time of interview)

Discrete

There are 10.245 households with one or more persons employed at the time of the interview; Label >2 means that there are multiple households at the address; recode to missing is an option

		Frequency	Percent
Valid	1	5.386	30,1
	2	4.698	26,2
	3	127	0,7
	4	15	0,1
	5	8	0,0
	6	3	0,0
	7	1	0,0
	8	2	0,0
	10	1	0,0
	11	2	0,0
	15	1	0,0
	16	1	0,0
	Total	10.245	57,2
Missing	System	7.663	42,8
Total		17.908	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

157

burgst_mean

Marital status of household

Nominal

	Frequency	Percent
Valid Not married	2.989	16,7
Married	10.100	56,4
Widow	2.508	14,0
Divorced	1.322	7,4
Partnership (formerly not married)	65	0,4
multiple household	924	5,2
Total	17.908	100,0

Variable number 158
Variable name gemgrootte_mean
Variable label Size of municipality
Variable type Ordinal
Values and value labels

	Frequency	Percent
Valid < 5 000 inhabitants	44	0,2
5 000 - < 10 000 inhabitants	591	3,3
10 000 - < 20 000 inhabitants	2.809	15,7
20 000 - < 50 000 inhabitants	6.686	37,3
50 000 - < 100 000 inhabitants	3.009	16,8
100 000 - < 150 000 inhabitants	1.677	9,4
150 000 - < 250 000 inhabitants	1.334	7,4
250 000 inhabitants of more	1.720	9,6
undetermined	17	0,1
Total	17.887	99,9
Missing System	21	0,1
Total	17.908	100,0

Variable number 159
Variable name huurkoop_mean
Variable label HUUR- OF KOOPWONING (Dwelling is owned by household)
Variable type Nominal
Values and value labels

	Frequency	Percent
Valid Owned	10.983	61,3
Rented	6.827	38,1
undetermined	98	0,5

Total	17.908	100,0
-------	--------	-------

Variable number 160

Variable name pensioen_sum

Variable label WEL/GEEN PENSIOEN (number of persons in household on pension)

Variable type Discrete

Values and value labels There are 6.943 households with one or more persons pensioned at the time of the interview; Label >2 means that there are multiple households at the address; recode to missing is an option

	Frequency	Percent
Valid 1	4.651	26,0
2	2.271	12,7
3	16	0,1
4	4	0,0
5	1	0,0
Total	6.943	38,8
Missing System	10.965	61,2
Total	17.908	100,0

Variable number 161

Variable name provincie

Variable label Province of residence

Variable type Nominal

Values and value labels

	Frequency	Percent
Valid Groningen	637	3,6
Friesland	726	4,1
Drenthe	642	3,6
Overijssel	1.299	7,3
Flevoland	283	1,6
Gelderland	2.253	12,6
Utrecht	1.310	7,3
Noord-Holland	2.598	14,5
Zuid-Holland	3.611	20,2
Zeeland	500	2,8
Noord-Brabant	2.663	14,9
Limburg	1.365	7,6
Total	17.887	99,9
Missing System	21	0,1
Total	17.908	100,0

Variable number 162
Variable name stedbuurt_mean
Variable label Degree of urbanization
Variable type Ordinal
Values and value labels

		Frequency	Percent
Valid	Very strong (≥ 2500 omgevingsadressen/km ²)	2.844	15,9
	Strong (1500 tot 2500 omgevingsadressen/km ²)	3.961	22,1
	Moderate (1000 tot 1500 omgevingsadressen/km ²)	3.455	19,3
	Little (500 tot 1000 omgevingsadressen/km ²)	3.692	20,6
	Not (< 500 omgevingsadressen/km ²)	3.913	21,9
	undetermined	43	0,2
	Total	17.908	100,0

Variable number 163
Variable name typhhn_mean
Variable label Type of household
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Single person	4.906	27,4
	Unmarried couple without children	1.031	5,8
	Married couple without children	5.023	28,0
	Unmarried couple with children	517	2,9
	Married couple with children	5.047	28,2
	Single parent household	756	4,2
	Other type	91	0,5
	Institutional household	46	0,3
	undetermined	491	2,7
	Total	17.908	100,0

Variable number 164
Variable name zelfst_sum
Variable label Self-employment

Variable type

Discrete

Values and value labels

There are 1.943 households with one or more persons self-employed at the time of the interview; Label >2 means that there are multiple households at the address; recode to missing is an option

		Frequency	Percent
Valid	1	1.497	8,4
	2	439	2,5
	3	5	0,0
	4	1	0,0
	5	1	0,0
	Total	1.943	10,8
Missing	System	15.965	89,2
Total		17.908	100,0

Variable number

165

Variable name

Gem2005

Variable label

Municipality code

Variable type

Nominal

Values and value labels

Codes ranging from 0003 to 1987, 21 missing values

Variable number

166

Variable name

aantper

Variable label

Number of persons in household

Variable type

Discrete

Values and value labels

		Frequency	Percent
Valid	1,00	4.963	27,7
	2,00	6.493	36,3
	3,00	2.220	12,4
	4,00	2.551	14,2
	5,00	911	5,1
	6,00	199	1,1
	7,00	50	0,3
	8,00	26	0,1
	9,00	6	0,0
	10,00	1	0,0
	11,00	1	0,0
	onbepaald	464	2,6
	Total	17.884	99,9
Missing	System	24	0,1
Total		17.908	100,0

Variable number 167
Variable name mndbedr_sum
Variable label Household income in interview month (kernel only)
Variable type Continuous
Values and value labels 204 households without income

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Income	17704	-24930	86065	3621,72	2946,709

Variable number 168
Variable name woz_mean
Variable label House value
Variable type Continuous
Values and value labels 70 households with missing house value

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
WOZ-WAARDE in euro	17838	0	9403000	220569,95	164775,322

Variable number 169
Variable name Ooitabw_sum
Variable label Social security at least one month in 2005
Variable type Discrete
Values and value labels Number of persons in household that were at least one month on social security in 2005. Label > 3 means that there are multiple households at address. Recode to missing is an option.

	Frequency	Percent
Valid 1	269	1,5
2	66	0,4
3	2	0,0
4	1	0,0
5	1	0,0
15	1	0,0

Total	340	1,9
Missing System	17.568	98,1
Total	17.908	100,0

Variable number 170

Variable name ooitao _sum

Variable label Disability allowance at least one month in 2005

Variable type Discrete

Values and value labels Number of persons in household that were at least one month on disability allowance in 2005. Label > 3 means that there are multiple households at address. Recode to missing is an option.

	Frequency	Percent
Valid 1	1.601	8,9
2	134	0,7
3	3	0,0
4	1	0,0
5	1	0,0
6	1	0,0
17	1	0,0
Total	1.742	9,7
Missing System	16.166	90,3
Total	17.908	100,0

Variable number 171

Variable name ooitwerknemer _sum

Variable label Paid job at least one month in 2005

Variable type Discrete

Values and value labels Number of persons in household that were at least one month employed in 2005. Label > 3 means that there are multiple households at address. Recode to missing is an option.

	Frequency	Percent
Valid 1	5.330	29,8
2	4.981	27,8
3	151	0,8
4	27	0,2
5	11	0,1
6	3	0,0
7	2	0,0
8	1	0,0
9	1	0,0

11	1	0,0
12	1	0,0
13	1	0,0
14	1	0,0
17	1	0,0
19	1	0,0
Total	10.513	58,7
Missing System	7.395	41,3
Total	17.908	100,0

Variable number 172

Variable name ooitzelfst _sum

Variable label Self-employed at least one month in 2005

Variable type Discrete

Values and value labels Number of persons in household that were at least one month self employed in 2005. Label > 3 means that there are multiple households at address. Recode to missing is an option

	Frequency	Percent
Valid 1	1.497	8,4
2	439	2,5
3	5	0,0
4	1	0,0
5	1	0,0
Total	1.943	10,8
Missing System	15.965	89,2
Total	17.908	100,0

Variable number 173

Variable name Cwidoor _max

Variable label Maximal duration in months subscription to Center for Work and Income (CWI) in household

Variable type Discrete

Values and value labels Ranging from -2 to 366, 737 (persons in households with CWI subscription. (-2 means that a person inscribed with CWI, two months after the interview).

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
inschrijving cwi in maanden	737	-2	366	23,17	38,908
Valid N (listwise)	737				

Variable number 174

Variable name etngroep_mean
Variable label Ethnic background - household
Variable type Nominal
Values and value labels NB Deze moet nog even opnieuw; veel te veel
surinamers = mix die toevallig op drie uitkomt.

		Frequency	Percent
Valid	Native	14.695	83,9
	Moroccan	40	0,2
	Turkish	163	0,9
	3,00	1.452	8,3
	Netherlands		
	Antilles and	47	0,3
	Aruba		
	Other non- westers	138	0,8
	Other westers	719	4,1
	mixed	314	1,8
	Total	17.521	100,0

Variable number 175
Variable name etngen_mean
Variable label Ethnic background - generation
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Native	15.019	83,9
	1st generation non-native	1.560	8,7
	mixed	179	1,0
	2nd generation non-native – one parent born abroad	348	1,9
	mixed	13	0,1
	2nd generation non-native – two parents born abroad	35	0,2
	Undetermined; multiple households	754	4,2
	Total	17.908	100,0

Variable number 176
Variable name kern2_mean

Variable label	Number of persons in kernel
Variable type	Discrete
Values and value labels	Number of persons in kernel; If N>2, multiple households at address.

		Frequency	Percent
Valid	0,00	38	0,2
	1,00	5.539	30,9
	2,00	11.505	64,2
	3,00	272	1,5
	4,00	79	0,4
	5,00	23	0,1
	6,00	8	0,0
	7,00	3	0,0
	8,00	2	0,0
	9,00	6	0,0
	10,00	1	0,0
	11,00	1	0,0
	12,00	1	0,0
	13,00	1	0,0
	14,00	2	0,0
	20,00	1	0,0
	21,00	1	0,0
	Total	17.483	97,6
Missing	System	425	2,4
Total		17.908	100,0

Variable number	177
Variable name	leeftijd_mean
Variable label	Mean age of household kernel
Variable type	Discrete
Values and value labels	Age 1 should not be possible; kernel should at least be about 16 y.o.a.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
gemiddelde leeftijd hh	17908	1	103	54,28	16,610
Valid N (listwise)	17908				

Variable number	178
Variable name	meervbew_mean
Variable label	Multiple household
Variable type	Nominal
Values and value labels	

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	One household	17504	97,7	97,7	97,7
	Multiple households	404	2,3	2,3	100,0
	Total	17908	100,0	100,0	

Variable number 179
Variable name Average house value (WOZ) at zip-code
Variable label Wozgem
Variable type Continuous
Values and value labels 67 missings

	N	Minimum	Maximum	Mean	Std. Deviation
Gemiddelde woz op postcode	17.841	9	1.733	218,41	111,465

Variable number 180
Variable name Number of persons at zip-code
Variable label aantpers
Variable type Discrete
Values and value labels 1 missing

	N	Minimum	Maximum	Mean	Std. Deviation
Aantal personen op postcode	17.907	1	1.088	53,34	43,817

Variable number 181
Variable name percaut
Variable label Percentage of natives at zip-code
Variable type Continuous
Values and value labels 1 missing

	N	Minimum	Maximum	Mean	Std. Deviation
Percentage autochtoon	17.907	0,00	1,00	0,8478	0,15275

Variable number 182
Variable name percwest
Variable label Percentage of non western non - natives at zip-code

Variable type	Continuous
Values and value labels	1 missing

	N	Minimum	Maximum	Mean	Std. Deviation
Percentage niet-westers allochtoon	17907	,00	1,00	,0669	,12392

Variable number	183
Variable name	percall
Variable label	Percentage of non- natives at zip-code
Variable type	Continuous
Values and value labels	1 missing

	N	Minimum	Maximum	Mean	Std. Deviation
Percentage allochtoon	17.907	0,00	1,00	0,1522	0,15275



Representativity Indicators for Survey Quality

RISQ data set documentation
Dutch Short Term Statistic Industry 2007

Deliverable 1.3 (WP2)

Barry Schouten
CBS

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Korte Termijn Statistiek (KS) Industrie. In English: Short Term Statistics (STS) Industry
Abstract:	STS is a monthly survey for Eurostat. It measures turnover for businesses in The Netherlands
Topic classification:	Turnover
Type of survey	Cross section
Unit of study	Business units according to Stat Netherlands business frame ABR
Target population	Business units in Industry (SBI 15 – 37)
Producer:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Data file name:	RISQ - KS 2007 - Industry.sav
File contents:	Data file contains response indicator, several auxiliary variables from administrative data and one target variable (reported turnover).
File structure:	Rectangular
Number of cases:	64413
Number of variables	15
Design weighting:	Inclusion probabilities are in file
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	Missing items are set at SPSS system missing
Data collector:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Mode of data collection:	STS is collected by a concurrent mixed-mode design consisting of three modes: paper, electronic and questionnaire send to unit by e-mail, electronic and questionnaire send to unit by post.
Field work period(s):	Monthly samples. There is no restriction to length of fieldwork period. However, statistics need to be produced after 30 days.
Sampling procedure:	Stratified simple random sampling without replacement
Interviewers:	Not applicable

Advance information	Advance letter plus paper or electronic questionnaire
Call schedules	Not applicable
Incentives	No
Refusal conversion	Two reminders are conducted, one after approximately 14 days and one after approximately 28 days. Exact number of days depend on timing advance letter (weekend or weekdays).

Response rates

	Frequency	Percent
Valid Respondent	59596	92,5%
Non-respondent	4817	7,5%
Total	64413	100,0%

Codebook of variables

Variable number	1
Variable name	Identification key
Variable label	Rownames
Variable type	Integer
Values and value labels	Ranging from 1 to 64413

Variable number	2
Variable name	Business register identification number
Variable label	Be_id
Variable type	Integer
Values and value labels	8 digit unique key

Variable number	3
Variable name	Observation identification number; number that changes for every month or quarter of observation
Variable label	Opgave_i
Variable type	Integer
Values and value labels	8 digit unique key per month or quarter

Variable number

4

Variable name

Standard business classification (NACE5)

Variable label

sbi

Variable type

Nominal

Values and value labels

		Frequency	Percent
Valid	1511	509	,8
	1512	287	,4
	1520	390	,6
	1531	132	,2
	1532	43	,1
	1533	371	,6
	1541	23	,0
	1542	83	,1
	1543	61	,1
	1551	336	,5
	1552	75	,1
	1561	128	,2
	1562	80	,1
	1571	686	1,1
	1572	155	,2
	1581	2155	3,3
	1582	632	1,0
	1583	18	,0
	1585	12	,0
	1586	133	,2
	1587	156	,2
	1588	72	,1
	1589	300	,5
	1591	105	,2
	1592	12	,0
	1594	12	,0
	1596	69	,1
	1597	13	,0
	1598	73	,1
	1600	118	,2
	1711	53	,1
	1712	13	,0
	1715	11	,0
	1716	24	,0
	1721	70	,1

1722	25	,0
1723	15	,0
1725	77	,1
1730	145	,2
1740	505	,8
1751	258	,4
1752	36	,1
1753	57	,1
1754	123	,2
1760	36	,1
1771	12	,0
1821	37	,1
1822	170	,3
1823	24	,0
1824	52	,1
1910	47	,1
1920	36	,1
1930	97	,2
2020	45	,1
2040	238	,4
2051	132	,2
2111	5	,0
2122	82	,1
2123	143	,2
2124	24	,0
2125	408	,6
2211	839	1,3
2212	493	,8
2213	763	1,2
2214	60	,1
2215	84	,1
2221	83	,1
2223	426	,7
2224	221	,3
2225	59	,1
2231	38	,1
2232	24	,0
2233	25	,0
2330	19	,0
2411	48	,1
2412	155	,2
2413	191	,3
2415	84	,1
2416	525	,8
2417	36	,1
2420	63	,1
2430	567	,9
2441	72	,1
2442	572	,9
2451	308	,5
2452	132	,2

2462	176	,3
2463	59	,1
2464	33	,1
2465	24	,0
2466	245	,4
2470	65	,1
2511	12	,0
2512	12	,0
2513	212	,3
2521	781	1,2
2522	816	1,3
2523	574	,9
2524	1301	2,0
2611	24	,0
2612	134	,2
2613	37	,1
2614	48	,1
2615	96	,1
2621	71	,1
2622	12	,0
2623	11	,0
2626	36	,1
2630	23	,0
2640	245	,4
2651	12	,0
2652	11	,0
2662	46	,1
2663	199	,3
2664	44	,1
2665	22	,0
2666	35	,1
2670	256	,4
2681	48	,1
2682	145	,2
2710	58	,1
2721	11	,0
2722	84	,1
2731	19	,0
2733	32	,0
2734	17	,0
2742	181	,3
2743	52	,1
2744	12	,0
2745	12	,0
2751	118	,2
2752	24	,0
2753	123	,2
2754	24	,0
2811	4933	7,7
2812	587	,9
2821	232	,4

2822	107	,2
2830	74	,1
2840	1134	1,8
2851	1269	2,0
2852	1871	2,9
2861	8	,0
2862	358	,6
2863	114	,2
2871	23	,0
2872	120	,2
2873	265	,4
2874	158	,2
2875	1043	1,6
2911	444	,7
2912	813	1,3
2913	243	,4
2914	161	,2
2921	118	,2
2922	1410	2,2
2923	1071	1,7
2924	1923	3,0
2932	972	1,5
2941	24	,0
2942	173	,3
2943	72	,1
2951	24	,0
2952	197	,3
2953	731	1,1
2954	107	,2
2955	116	,2
2956	1004	1,6
2960	11	,0
2971	111	,2
2972	183	,3
3001	76	,1
3002	104	,2
3110	316	,5
3120	252	,4
3130	161	,2
3140	19	,0
3150	351	,5
3161	36	,1
3162	416	,6
3210	515	,8
3220	140	,2
3230	80	,1
3320	1012	1,6
3330	208	,3
3340	176	,3
3350	24	,0
3410	112	,2

3430	296	,5
3511	796	1,2
3512	372	,6
3520	65	,1
3530	208	,3
3541	57	,1
3542	90	,1
3543	46	,1
3550	35	,1
3611	381	,6
3613	215	,3
3614	573	,9
3615	65	,1
3621	12	,0
3622	48	,1
3630	49	,1
3640	105	,2
3650	75	,1
3661	12	,0
3662	47	,1
3710	63	,1
3720	401	,6
15131	319	,5
15132	462	,7
15841	59	,1
15842	431	,7
20101	137	,2
20102	60	,1
20301	1281	2,0
20302	504	,8
21121	106	,2
21122	82	,1
21123	34	,1
21211	749	1,2
21212	167	,3
22221	166	,3
22222	183	,3
22223	971	1,5
22224	73	,1
22225	162	,3
22226	1008	1,6
23201	44	,1
23202	87	,1
24141	128	,2
24142	246	,4
26611	927	1,4
26612	13	,0
33101	322	,5
33102	424	,7
34201	682	1,1
34202	448	,7

36121	616	1,0
36122	393	,6
36632	158	,2
Total	64413	100,0

Variable number

5

Variable name

Business size classification (number of employees)

Variable label

gk

Variable type

Nominal

Values and value labels

		Frequency	Percent
Valid	5	36509	56,7
	6	13836	21,5
	7	8043	12,5
	8	4454	6,9
	9	1571	2,4
	Total	64413	100,0

Variable number

6

Variable name

Period of observation

Variable label

Opgave_p

Variable type

Nominal

Values and value labels

		Frequency	Percent
Valid	January	5117	7,9
	February	5137	8,0
	March	5158	8,0
	April	5114	7,9
	May	5101	7,9
	June	5108	7,9
	July	5083	7,9
	August	5102	7,9
	September	5092	7,9
	October	5051	7,8
	November	4988	7,7
	December	4812	7,5
	Weeks 1-4	276	,4

Weeks 5-8	282	,4
Weeks 9-12	281	,4
Weeks 13-16	282	,4
Weeks 17-20	277	,4
Weeks 21-24	273	,4
Weeks 25-28	272	,4
Weeks 29-32	273	,4
Weeks 33-36	269	,4
Weeks 37-40	273	,4
Weeks 41-44	274	,4
Weeks 45-48	266	,4
Weeks 49-52	252	,4
Total	64413	100,0

Variable number 7
Variable name Data collection mode
Variable label Waarneem
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Electronic, invitation by e-mail	41738	64,8
	Electronic, invitation by letter	8605	13,4
	Paper	14070	21,8
	Total	64413	100,0

Variable number 8
Variable name Reported turnover for reference period
Variable label waarde
Variable type Continuous
Values and value labels Missing for 4817 units (nonresponse)

Variable number 9
Variable name Date of receiving questionnaire
Variable label Datum_ti

Variable type
Values and value labels

Date
Ranging from 01.02.07 to 19.04.08

Variable number
Variable name
Variable label
Variable type
Values and value labels

10
Response indicator
Respons
Nominal

Variable number
Variable name
Variable label
Variable type
Values and value labels

11
Standard business classification (NACE2)
Sbisubse
Nominal

		Frequency	Percent
Valid	DA	8510	13,2
	DB	1743	2,7
	DC	180	,3
	DD	2397	3,7
	DE	7478	11,6
	DF	150	,2
	DG	3729	5,8
	DH	3708	5,8
	DI	2495	3,9
	DJ	13063	20,3
	DK	9908	15,4
	DL	4632	7,2
	DM	3207	5,0
	DN	3213	5,0
	Total	64413	100,0

Variable number
Variable name
Variable label
Variable type
Values and value labels

12
Inclusion probability
Insluitk
Continuous
Ranging from 0,5 to 1

	Frequency	Percent
Valid ,50	264	,4
,55	429	,7
,56	270	,4
,58	85	,1
,64	111	,2
,65	2088	3,2
,70	984	1,5
,70	261	,4
,70	1360	2,1
,70	1468	2,3
,71	73	,1
,76	655	1,0
,79	238	,4
,85	218	,3
,89	209	,3
,91	628	1,0
,92	209	,3
1,00	54863	85,2
Total	64413	100,0

Variable number 13
Variable name VAT 2006
Variable label Btw2006
Variable type Continuous
Values and value labels Missing for 24697 units

Variable number 14
Variable name Number of days between end of reporting period and date of receiving questionnaire
Variable label Dagen_na
Variable type Discrete
Values and value labels Ranging from 0 to 120 for respondents

Variable number 15
Variable name VAT 2007
Variable label Btw2007
Variable type Continuous
Values and value labels Missing for 19119 units



Representativity Indicators for Survey Quality

RISQ data set documentation
Dutch Short Term Statistic Retail 2007

Deliverable 1.4 (WP2)

Barry Schouten
CBS

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Korte Termijn Statistiek (KS) Detailhandel. In English: Short Term Statistics (STS) Retail
Abstract:	STS is a monthly survey for Eurostat. It measures turnover for businesses in The Netherlands
Topic classification:	Turnover
Type of survey	Cross section
Unit of study	Business units according to Stat Netherlands business frame ABR
Target population	Business units in Retail (SBI 52)
Producer:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Data file name:	RISQ - KS 2007 Retail.sav
File contents:	Data file contains response indicator, several auxiliary variables from administrative data and one target variable (reported turnover).
File structure:	Rectangular
Number of cases:	93799
Number of variables	14
Design weighting:	Inclusion probabilities are in file
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	Missing items are set at SPSS system missing
Data collector:	Centraal Bureau voor de Statistiek (CBS). In English: Statistics Netherlands
Mode of data collection:	STS is collected by a concurrent mixed-mode design consisting of three modes: paper, electronic and questionnaire send to unit by e-mail, electronic and questionnaire send to unit by post.
Field work period(s):	Monthly samples. There is no restriction to length of fieldwork period. However, statistics need to be produced after 30 days.
Sampling procedure:	Stratified simple random sampling without replacement
Interviewers:	Not applicable

Advance information	Advance letter plus paper or electronic questionnaire
Call schedules	Not applicable
Incentives	No
Refusal conversion	Two reminders are conducted, one after approximately 14 days and one after approximately 28 days. Exact number of days depend on timing advance letter (weekend or weekdays).

Response rates

	Frequency	Percent
Valid Respondent	86607	92,3%
Non-respondent	7192	7,7%
Total	93799	100,0%

Codebook of variables

Codebook of variables

Variable number	1
Variable name	Identification key
Variable label	Rownames
Variable type	Integer
Values and value labels	Ranging from 1 to 64413
Variable number	2
Variable name	Business register identification number
Variable label	Be_id
Variable type	Integer
Values and value labels	8 digit unique key
Variable number	3
Variable name	Observation identification number; number that changes for every month or quarter of observation
Variable label	Opgave_i

Variable type
Values and value labels

Integer
8 digit unique key per month or quarter

Variable number
Variable name
Variable label
Variable type
Values and value labels

4
Standard business classification (NACE5)
sbi
Nominal

		Frequency	Percent
Valid	5211	2954	3,1
	5221	934	1,0
	5223	615	,7
	5225	1904	2,0
	5226	1213	1,3
	5233	484	,5
	5261	7682	8,2
	5271	189	,2
	5272	258	,3
	5273	91	,1
	5274	366	,4
	52121	281	,3
	52122	212	,2
	52221	1623	1,7
	52222	156	,2
	52241	592	,6
	52242	304	,3
	52271	324	,3
	52272	339	,4
	52273	231	,2
	52274	346	,4
	52321	1226	1,3
	52322	347	,4
	52411	182	,2
	52412	142	,2
	52413	141	,2
	52421	1172	1,2
	52422	2478	2,6
	52423	928	1,0
	52424	3535	3,8
	52425	439	,5
	52426	242	,3
	52427	510	,5

52431	1376	1,5
52432	616	,7
52441	4380	4,7
52442	975	1,0
52443	492	,5
52444	2392	2,6
52445	365	,4
52446	125	,1
52447	712	,8
52451	377	,4
52452	1025	1,1
52453	562	,6
52454	1397	1,5
52455	231	,2
52456	1135	1,2
52457	538	,6
52458	140	,1
52461	791	,8
52462	597	,6
52463	217	,2
52464	374	,4
52465	2585	2,8
52466	772	,8
52467	787	,8
52468	2201	2,3
52471	1271	1,4
52472	528	,6
52473	1061	1,1
52481	708	,8
52482	1151	1,2
52483	1982	2,1
52484	762	,8
52485	1441	1,5
52486	439	,5
52487	2080	2,2
52488	156	,2
52489	562	,6
52491	2253	2,4
52492	1041	1,1
52493	919	1,0
52494	2477	2,6
52495	973	1,0
52496	356	,4
52497	566	,6
52499	3120	3,3
52501	852	,9
52502	241	,3
52503	1139	1,2
52621	769	,8
52622	1690	1,8
52623	425	,5

52624	1588	1,7
52625	657	,7
52626	1485	1,6
52631	145	,2
52632	920	1,0
52633	2440	2,6
Total	93799	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

5

Business size classification (number of employees)

gk

Nominal

	Frequency	Percent
Valid 1	20254	21,6
2	38642	41,2
3	14633	15,6
4	8113	8,6
5	7197	7,7
6	2156	2,3
7	1166	1,2
8	692	,7
9	946	1,0
Total	93799	100,0

Variable number

Variable name

Variable label

Variable type

Values and value labels

6

Period of observation

Opgave_p

Nominal

	Frequency	Percent
Valid January	7694	8,2
February	7726	8,2
March	7621	8,1
April	7585	8,1
May	7542	8,0
June	7719	8,2

July	7580	8,1
August	7588	8,1
September	7715	8,2
October	7685	8,2
November	7625	8,1
December	7702	8,2
Weeks 1-4	158	,2
Weeks 5-8	158	,2
Weeks 9-12	151	,2
Weeks 13-16	151	,2
Weeks 17-20	150	,2
Weeks 21-24	152	,2
Weeks 25-28	147	,2
Weeks 29-32	148	,2
Weeks 33-36	150	,2
Weeks 37-40	168	,2
Weeks 41-44	161	,2
Weeks 45-48	162	,2
Weeks 49-52	161	,2
Total	93799	100,0

Variable number 7
Variable name Data collection mode
Variable label Waarneem
Variable type Nominal
Values and value labels

		Frequency	Percent
Valid	Electronic; invitation by e-mail	23083	24,6
	Electronic; invitation by letter	19795	21,1
	Paper	50921	54,3
	Total	93799	100,0

Variable number 8
Variable name Reported turnover for reference period
Variable label waarde
Variable type Continuous
Values and value labels Missing for 7192 units (nonresponse)

Variable number 9
Variable name Date of receiving questionnaire
Variable label Datum_ti
Variable type Date

Values and value labels	Ranging from 01.02.07 to 07.02.08
Variable number	10
Variable name	Response indicator
Variable label	Respons
Variable type	Nominal
Values and value labels	
Variable number	11
Variable name	Number of days between end of reporting period and date of receiving questionnaire
Variable label	Dagen_na
Variable type	Discrete
Values and value labels	Ranging from 1 to 120 for respondents
Variable number	12
Variable name	Inclusion probability
Variable label	Insluitk
Variable type	Continuous
Values and value labels	Ranging from 0,01 to 1
Variable number	13
Variable name	VAT 2007
Variable label	Btw2007
Variable type	Continuous
Values and value labels	Missing for 11106 units
Variable number	14
Variable name	VAT 2006
Variable label	Btw2006
Variable type	Continuous
Values and value labels	Missing for 23075 units



Representativity Indicators for Survey Quality

RISQ data set documentation
UK 2001 Census Link File

Deliverable 1.5 (WP2)

Ana Marujo, Natalie Shlomo and Gabi Durrant
SOTON

21 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title: UK 2001 Census Link Study

Abstract: The UK 2001 Census Link Study contains the response outcome of six major UK government household surveys linked to 2001 UK census data on a range of household and individual characteristics, interviewer observations about the household and extensive information about the interviewer and area information. All variables are available for both respondents and non-respondents of the six surveys. The study includes only those surveys conducted with face to face interviewing.

We are currently in the process of obtaining a license for the data from the Microdata Review Panel of the ONS, UK. Therefore, the documentation presented is still limited.

Topic classification: Six UK Social Surveys linked to the 2001 UK Census. The surveys are:

- (1) Expenditure and Food Survey (EFS),
- (2) Family Resources Survey (FRS),
- (3) General Household Survey (GHS),
- (4) Omnibus Survey (OMN) (this survey selects and interviews one person in the household),
- (5) National Travel Survey (NTS),
- (6) Labour Force Survey (LFS).

Type of surveys: cross- sectional

Unit of study: households, except for the OS where one person is selected from each household

Target population: All persons of 16 years and older

Producer: UK Office for National Statistics (ONS)

Data file name: UK 2001 Census Link File

File contents: Response outcome of six major UK government household surveys containing the following information:

- 2001 UK census information. Survey records of respondents and nonrespondents are linked to their census record, both for households and individuals within households. This comprises primarily socio-demographic and some attitudinal information about the individuals within a household, and household characteristics;
- Interviewer observation data. The interviewer recorded information about the household at each visit, even if no contact was made, including characteristics of the accommodation (e.g. whether a house or flat, the presence of security measures such as locked gates or burglar alarm), any information about the household composition, the quality of housing and observations of the surrounding neighbourhood.
- field-process and interviewer calling data - also referred to as paradata (Couper, 1998). This comprises primarily information on the frequency of calls to the household, the time and date and the outcome of each call, as well as information about the interaction between the interviewer and the household at the 'doorstep' if contact was made. This information was recorded by the interviewer at the survey data collection stage. Note that at this stage we are unable to document these variables.
- interviewer information. This information was obtained via a separate comprehensive survey (Interviewer Attitude Survey) of face-to-face ONS interviewers during June 2001, at around the time of the survey and census data collection period. Interviewers were asked about their socio-demographic background, work experience, interviewing strategies and behaviours, and attitudes towards their work and towards gaining contact and cooperation.

Households selected for interview in one of the above surveys during May-June 2001, the months immediately following the 2001 Census, were included in the study. The following cases were excluded from the analysis sample: all persons under 16 (to exclude ineligible cases); sample units that were unable to respond due to language problems; individuals and households that were

imputed in the 2001 census (because only basic area information was available for these cases); vacant homes; households that had moved between the census and the survey date (to avoid, for example, a mis-match between interviewer observations and census data); mode switches, where after failing to receive a face-to-face interview a telephone interview was attempted; and re-issues, cases where one interviewer failed to get a positive outcome from a sample unit and subsequently the sample unit was re-issued to another interviewer to attempt conversion. Only households for which all data components could be linked successfully to the survey data were included in the analysis sample.

Structure: rectangular

Number of cases: 18,530 households: 3,683 for EFS, 2,219 for FRS, 3,415 for GHS, 3,318 for OMN, 2,642 for NTS and 3,253 for LFS and 565 interviewers

There are 39,375 individuals from the Census household file and 39,223 individuals from the Census person file for these cases and there is a need to reconcile this difference. The documentation includes both numbers depending on the origin of the variable. This discrepancy will have to be resolved when the file becomes available for use.

Number of variables: The file consists of many variables. The main variables relative to non-response research are documented below. There are other variables that will be added to this document once the file is received from the ONS.

Design weighting: Household surveys generally have equal inclusion probabilities. The OS selects one person in the household for the interview and therefore design weights are proportional to household size.

Adjustment weights: Calibrated to geography, sex and age population distributions.

Imputation: In some cases of missing data, imputation could be carried out by using other information available for the household or interviewer observations.

Missing data: Some missing data remaining and denoted by the category "missing"

Data collector: ONS

Mode of data collection: Face to face interview

Field work period(s): May-June 2001

Sampling procedure: Systematic sample of postcodes within strata defined by geographies

Interviewers: 565 trained interviewers

The following Table contains a summary of the main survey characteristics for the six surveys:

Survey Design Characteristic	EFS	FRS	GHS	OMN	NTS	LFS
---------------------------------	-----	-----	-----	-----	-----	-----

Maximum number of calls to household	No limit	No limit	No limit	No limit	No limit	No limit
Minimum number of calls to household	4	4	4	4	8	4
Length of data collection period	1 month +1 week	1 month	1 month	3 weeks	2.5 to 6.5 weeks	7+7+2 days (spread over 13 week period)
Interviewer workload in number of addresses	18	24	23	30	23	20
ONS initial interviewer training given	Yes	Yes	Yes	Yes	Yes	Yes
Type of additional interviewer training given	1 day	1 day	briefing	postal	1.5 days	4 days (interviewers work only on this survey)
Advance letter	Yes	Yes	Yes	Yes	Yes	Yes
Purpose leaflet available	Yes: in the field	Yes: in the field	Yes: in the field	Yes	Yes: postal (London only)	Yes: postal
Respondent incentives	Stamps; £10/£5 for diary	Stamps	None	Stamps	Pen and fridge magnet	None
Respondent rules	All householders aged 16+	All householders aged 16+	All householders aged 18+	One household holder aged 16+	All householders aged 16+	All householders aged 16+
Proxy response allowed	Yes	Yes	Yes	No	Yes	Yes
Average lengths of interview (in mins)	70	80	70	26	60	30 (for wave 1)
Diary required (in addition to questionnaire)	Yes: 2 weeks	No	No	No	Yes: 1 week	No

(The surveys collect information based on the household as a whole and on the individuals within the households. Further information on the different surveys can be obtained from the ONS website, www.statistics.gov.uk)

Information collected by survey:

EFS: core topics include: household expenditure, rent and mortgage payments, taxes, benefits, detailed information about income of each household member, trends in nutrition.

FRS: aims to provide information on living standards, people's relationship and interaction with the social security system. The questionnaire seeks information on income and benefits, tenure and housing costs, assets and savings, occupation and employment, health and ability to work, pensions and insurance, childcare and carers.

GHS: core topics include: accommodation, consumer durables, housing tenure, migration, employment, pensions, education, health, smoking, drinking, family formation, income.

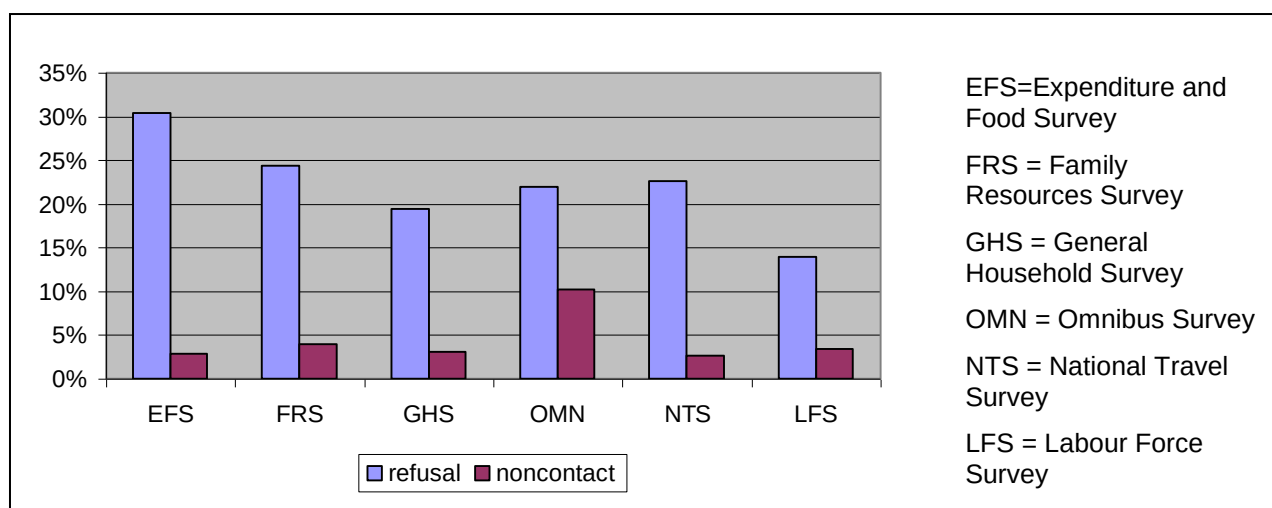
NTS: aims to provide a comprehensive picture of personal travel behaviour. Questions include information about ethnic group, place of work, reliability and frequency of local services such as buses and trains, use of vehicles, long distance journeys and travel outside of Great Britain.

OMN: multi-purpose survey, which aims to obtain information about the general population or about particular groups. The questionnaire is in two parts, including first a set of core classificatory questions and then a series of unrelated modules on varying topics at the request of customers. Core questions include information on demographic details, economic status, job details, employment status, full- or part-time working, tenure, ethnic origin.

LFS: aims to provide information about the UK labour market and unemployment. The survey seeks information on respondent's personal circumstances, their labour market status and income.

Response Outcome:

Figure: Refusal and noncontact rates for the six surveys included in the Census Link Study.



Houtcom6 Indicator – Individual response rates - Total

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	28863	73.6	73.6	73.6
	Refusal	8490	21.6	21.6	95.2
	Non-contact	1870	4.8	4.8	100
	Total	39223	100	100	

Household response rates - Total

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	13621	73.51	73.51	73.51
	Refusal	4097	22.11	22.11	95.62
	Non-contact	812	4.38	4.38	100
	Total	18530	100	100	

Household response rates EFS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	2465	66.88	66.88	66.88
	Refusal	1118	30.34	30.34	97.23
	Non-contact	102	2.78	2.78	100
	Total	3686	100	100	

Household response rates FRS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	1594	71.72	71.72	71.72
	Refusal	541	24.34	24.34	96.06
	Non-contact	88	4	4	100
	Total	2222	100	100	

Household response rates GHS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	2642	77.54	77.54	77.54
	Refusal	660	19.35	19.35	96.89
	Non-contact	106	3.11	3.11	100
	Total	3408	100	100	

Individual response rates OMN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	2247	67.9	67.9	67.9
	Refusal	725	21.91	21.91	89.81
	Non-contact	337	10.19	10.19	100
	Total	3310	100	100	

Household response rates NTS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	1975	74.78	74.78	74.78
	Refusal	598	22.64	22.64	97.42
	Other	68	2.58	2.58	100
	Total	2641	100	100	

Household response rates LFS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	2697	82.75	82.75	82.75
	Refusal	451	13.83	13.83	96.58
	Non-contact	111	3.42	3.42	100
	Total	3259	100	100	

survey Person: SURVEY ID

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 EFS	5862	14.9	14.9	14.9
	2 FRS	8601	21.8	21.8	36.7
	3 GHS	5134	13.0	13.0	49.8
	4 OMN	7394	18.8	18.8	68.5
	5 NTS	3676	9.3	9.3	77.9
	6 LFS	8708	22.1	22.1	100.0
	Total	39375	100.0	100.0	

Codebook of main variables:

Census Person File

Variable name: gender
Variable label: genpuk
Variable type: nominal
Values and value labels: 1,2

genpuk Person: GENDER

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Male	23730	60.5	60.5	60.5
2 Female	15493	39.5	39.5	100.0
Total	39223	100.0	100.0	

Variable name: age
Variable label: agepuk
Variable type: continuous
Values and Value labels: 00-110 (summary of age groups given in the table)

agepuk Person: AGE

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 16-25	1466	3.7	3.7	3.7
26-35	6435	16.4	16.4	20.1
36-45	7732	19.7	19.7	39.8
46-55	7271	18.5	18.5	58.3
56-65	6113	15.6	15.6	73.9
66-75	5754	14.7	14.7	88.6
76-85	3444	8.8	8.8	97.4
86-95	972	2.5	2.5	99.9
96-101	36	0.1	0.1	100.0
Total	39223	100.0	100.0	

Variable name Marital status

Variable label mstpuk

Variable type nominal

Values and value labels 1,2,3,4,5,6

mstpuk Person:MARITAL STATUS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Single (Never married)	7565	19.3	19.3	19.3
2 Married (First marriage)	16904	43.1	43.1	62.4
3 Re-married	3028	7.7	7.7	70.1
4 Separated (but still legally married)	1357	3.5	3.5	73.6
5 Divorced	4760	12.1	12.1	85.7
6 Widowed	5609	14.3	14.3	100.0
Total	39223	100.0	100.0	

Variable name Student indicator

Variable label Stupuk

Variable type Nominal

Values and value labels 1,2

stupuk Person:STUDENT INDICATOR

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Full-time student	352	.9	.9	.9
2 Not a full-time student	38871	99.1	99.1	100.0
Total	39223	100.0	100.0	

Variable name Country of birth

Variable label cobpuk

Variable type Nominal

Values and value labels 3-digit classification (summary table below)

cobpuk Person:Country of birth

code group		Frequency	Percent	Cumulative Percent
	Total	39223	100.00%	
1	Great Britain	35946	91.65%	91.65%
2	Republic of Ireland	494	1.26%	92.90%
3	Europe	722	1.84%	94.75%
4	Africa	514	1.31%	96.06%
5	Asia (Middle and Far East)	991	2.53%	98.58%
6	Central and North America, Canada	359	0.92%	99.50%
7	South America	60	0.15%	99.65%
8	Australia and Pacific Islands	135	0.34%	99.99%
unknown	-	2	0.01%	100.00%

Variable name Health
Variable label heapuk
Variable type nominal
Values and value labels 1-3, missing

heapuk1 Person:HEALTH

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Good	22979	58.6	58.6	58.6
	2 Fairly good	11339	28.9	28.9	87.6
	3 Not good	4870	12.4	12.4	100.0
	Total	39188	99.9	100.0	
Missing	-5	25	.1		
	System	10	.0		
	Total	35	.1		
Total		39223	100.0		

Variable name Long term illness indicator

Variable label illpuk

Variable type Nominal

Values and value labels 1,2, missing

illpuk1 Person:LONG TERM ILLNESS INDICATOR

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	9635	24.6	24.6	24.6
	2 No	29553	75.3	75.4	100.0
	Total	39188	99.9	100.0	
Missing	-5	25	.1		
	System	10	.0		
	Total	35	.1		
Total		39223	100.0		

Variable number

Variable name industry

Variable label indpuk

Variable type Hierarchical ordinal

Values and value labels 3 digit hierarchy (summary table below)

indpuk1 Person:INDUSTRY

		Frequency	Percent	Valid Percent	Cumulative Percent
valid	Agriculture, fishing, mining	508	1.3	1.8	1.8
	Manufacturing	5274	13.4	18.8	20.7
	Electricity, gas and water	257	0.7	0.9	21.6
	Construction	2410	6.1	8.6	30.2
	Trade, Hotels	5208	13.3	18.6	48.8
	Transport, storage	2198	5.6	7.9	56.6
	Finance	4675	11.9	16.7	73.4
	Public Administration	1665	4.2	5.9	79.3
	Education	4517	11.5	16.1	95.4
	Community, Social	1276	3.3	4.6	100.0
	Total	27988	71.4	100.0	
missing	-9	1150	2.9		
	-5	2276	5.8		
	System	7809	19.9		
	Total	39223	100.0		

Variable number

Variable name occupation

Variable label occpuk

Variable type Hierarchical ordinal

Values and value labels 4 digit hierarchy (first digit summary in table)

occ puk1 Person:OCCUPATION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Managers	4749	12.1	16.4	16.4
	Professionals	3429	8.7	11.8	28.2
	Technical	3705	9.4	12.8	41.0
	Administrative	2856	7.3	9.9	50.9
	Skilled Trades	4167	10.6	14.4	65.3
	Personal Services	1564	4.0	5.4	70.7
	Sales	1460	3.7	5.0	75.7
	Process and machines	3419	8.7	11.8	87.5
	Elementary	3620	9.4	12.5	100.0
	Total	28969	73.9	100.0	
Missing	-9 No code required	1052	2.7		
	-7	18	0.0		
	-5	2080	5.3		
	System	7104	18.1		
	Total	10254	26.1		
Total		39223	100.0		

Variable name Economic activity (several grouped variables)

Variable label Ecopuk

Variable type Nominal

Values and value labels 1-6, missing

ECOPUK2 Person:ECONOMIC ACTIVITY

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00 Employee	19649	50.1	57.1	57.1
	2.00 Self employed	3591	9.2	10.4	67.6
	3.00 Unemployed	947	2.4	2.8	70.3
	4.00 Retired	6304	16.1	18.3	88.6
	5.00 Looking after home or family	1136	2.9	3.3	91.9
	6.00 Other (include student, permanently sick)	2772	7.1	8.1	100.0
	Total	34399	87.7	100.0	
Missing	-9.00 NO CODE REQUIRED	387	1.0		
	-5.00 No census records	4437	11.3		
	Total	4824	12.3		
Total		39223	100.0		

Variable name	Highest qualification
Variable label	hlqpuk
Variable type	nominal
Values and value labels	2-digit classification

hlqpuk1 Person: HIGHEST QUALIFICATION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10 No academic or professional qualifications	10065	25.7	29.3	29.3
	11 1+O levels/CSE/GCSE (any grades)/NVQ level/Foundation GNVQ	4969	12.7	14.4	43.7
	12 5+ O levels/5+CSEs (grade1)/5+ GCSEs (grades A-C) etc/1+ A levels/ AS levels/NVQ level 2/Intermediate GNVQ	5000	12.7	14.5	58.2
	13 2+ A levels/4+ AS levels/Higher School Certificate/NVQ level 3/Advanced GNVQ	1896	4.8	5.5	63.8
	14 First degree/Higher degree/NVQ levels 4-5/HNC/HND/Qualified Teacher... Doctor... Dentist... Nurse... Midwife... Health Visitor	6553	16.7	19.0	82.8
	15 Other qualifications (eg City and Guilds etc)/Other Professional qualifications	2770	7.1	8.1	90.9
	20 No qualifications	1171	3.0	3.4	94.3
	21 'O' Grade/Standard grade/GCSE/CSE etc/GSVQ/SVQ Level 1 or 2/SCOTVEC module etc	721	1.8	2.1	96.4
	22 Higher grade/CSYS/'A' level, etc/GSVQ/SVQ Level 3/ONC/OND etc	365	0.9	1.1	97.4
	23 HNC/HND/SVQ level 4 or 5 etc	230	0.6	0.7	98.1
	24 First degree/higher degree/Professional qualifications	659	1.7	1.9	100.0
	Total	34399	87.7	100.0	
Missing	-9	387	1.0		
	-5	987	2.5		
	System	3450	8.8		
	Total	4824	12.3		
Total		39223	100.0		

Variable name	Family status
Variable label	fmspuk
Variable type	Nominal
Values and value labels	1-8, missing

fmspuk1 Person:FAMILY STATUS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Not in a family - pens	6065	15.5	15.5	15.5
	2 Not in a family - other	5952	15.2	15.2	30.7
	3 In couple family - member of couple	23025	58.7	58.8	89.4
	5 In couple family - other child of couple	4	.0	.0	89.4
	6 In lone parent family - parent	4111	10.5	10.5	99.9
	7 In lone parent family - dependent child of parent	1	.0	.0	99.9
	8 In lone parent family - other child of parent	30	.1	.1	100.0
	Total	39188	99.9	100.0	
Missing	-5	25	.1		
	Sy stem	10	.0		
	Total	35	.1		
Total		39223	100.0		

Variable name:	Household reference person
Variable label:	Frppuk
Variable type	Nominal
Values and value labels	0,1, missing

frppuk1 Person:FAMILY REFERENCE PERSON

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 Not family reference person	1314	3.4	4.8	4.8
	1 Family reference person	25872	66.0	95.2	100.0
	Total	27186	69.3	100.0	
Missing	-9 No code required	1185	3.0		
	-5	2413	6.2		
	System	8439	21.5		
	Total	12037	30.7		
Total		39223	100.0		

Variable name: Ethnicity
Variable label: ethpuk
Variable type: nominal
Values and value labels: 3 digit classification (summary table below)

ethpukx Person:Ethnic group_5 categories

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 White group	37144	94.7	94.8	94.8
	2 Mixed group	203	.5	.5	95.3
	3 Asian group	804	2.0	2.1	97.4
	4 Black Group	596	1.5	1.5	98.9
	9 Other group	441	1.1	1.1	100.0
	Total	39188	99.9	100.0	
Missing	System	35	.1		
Total		39223	100.0		

Variable name: Relationships
Variable label: relations
Variable type: nominal
Values and value labels: 1-9, A,B,X (arranged in a matrix)

1	Husband or wife
2	Partner
3	Son or daughter
4	Step child
5	Brother or Sister
6	Mother or father
7	Stepmother/Stepfather
8	Grandchild
9	Grandparent
A	Other related
B	Other unrelated
X	No code required

No frequencies provided

Variable name: Hours worked
Variable label: houpuk
Variable type: numerical
Values and value labels 1-99, missing

houpuk1 Person:HOURS WORKED

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 10	594	1.5	0.0	0.0
	11 20	1513	3.9	2.5	2.5
	21-30	1326	3.4	6.5	9.0
	31-40	11921	30.4	5.3	14.3
	41-50	6022	15.4	49.5	63.8
	51-60	2078	5.3	24.4	88.2
	61-70	461	1.2	8.5	96.7
	71-80	224	0.6	1.9	98.6
	82-90	99	0.3	0.9	99.5
	91-99	27	0.1	0.4	99.9
	Total	24265	61.9	100.0	
Missing	-9 No code required	444	1.1		
	-5	3156	8.0		
	System	11358	29.0		
	Total	14958	38.1		
Total		39223	100.0		

Variable name:	National socio-economic classification
Variable label:	nsspuk
Variable type:	numerical
Values and value labels	01-40, XX

No frequencies provided

Census Household File

Variable name Tenure
Variable label Tenhuk
Variable type Nominal
Values and value labels 1-9

tenhuk1 Household:TENURE

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 Owned outright	12004	30.5	30.5	30.5
1 Buying with a mortgage	15917	40.4	40.4	70.9
2 Shared ownership	224	.6	.6	71.5
3 Rent from council or Scottish Homes	5397	13.7	13.7	85.2
4 Rent from RSL or HA	2173	5.5	5.5	90.7
5 Private landlord or letting agency	2533	6.4	6.4	97.1
6 Employer of a hhld member	80	.2	.2	97.3
7 Relative or friend of a hhld member	218	.6	.6	97.9
8 Other	64	.2	.2	98.1
9 Lives rent free	765	1.9	1.9	100.0
Total	39375	100.0	100.0	

Variable name Household composition
Variable label Ahthuk1
Variable type nominal
Values and value labels 0-10

ahthuk1 Household:ALTERNATIVE HOUSEHOLD TYPE

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 Mar cple hhld dep child	7575	19.2	19.2	19.2
1 Mar cple hhld no dep child	12087	30.7	30.7	49.9
2 Cohabit cple hhld dep child	1342	3.4	3.4	53.3
3 Cohabit cple hhld no dep child	2052	5.2	5.2	58.6
4 Lone par hhld dep child	2769	7.0	7.0	65.6
5 Lone par hhld no dep child	1386	3.5	3.5	69.1
6 Same-sex cple hhld dep child	16	.0	.0	69.1
7 Same-sex cple hhld no dep child	93	.2	.2	69.4
8 1 pers hhld	11076	28.1	28.1	97.5
9 Multi pers hhld all student	114	.3	.3	97.8
10 Multi pers hhld other	865	2.2	2.2	100.0
Total	39375	100.0	100.0	

Variable name Household size
Variable label sizhuk
Variable type numerical
Values and value labels 1-15

sizhuk Household:HOUSEHOLD SIZE

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	11076	28.1	28.1	28.1
2	13769	35.0	35.0	63.1
3	6246	15.9	15.9	79.0
4	5544	14.1	14.1	93.0
5	1957	5.0	5.0	98.0
6	582	1.5	1.5	99.5
7	117	.3	.3	99.8
8	53	.1	.1	99.9
9	18	.0	.0	100.0
10	9	.0	.0	100.0
11	1	.0	.0	100.0
12	2	.0	.0	100.0
15	1	.0	.0	100.0
Total	39375	100.0	100.0	

Variable name	Number of cars
Variable label	Cavhuk1
Variable type	numerical
Values and value labels	0-6

cavhuk1 Household:NUMBER OF CARS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 None	10.398	26,4	26,4	26,4
	1 One	17.477	44,4	44,4	70,8
	2 Two	9.277	23,6	23,6	94,4
	3 Three	1.709	4,3	4,3	98,7
	4 Four	384	1,0	1,0	99,7
	5 Five	85	0,2	0,2	99,9
	6 Six or more	45	0,1	0,1	100,0
	Total	39.375	100,0	100,0	

Variable name	Number of rooms
Variable label	Norhuk1
Variable type	numerical
Values and value labels	1-98 (summary table below)

norhuk1 Household: NUMBER OF ROOMS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	235	0,6	0,6	0,6
2	855	2,2	2,2	2,8
3	3.282	8,3	8,3	11,1
4	7.961	20,2	20,2	31,3
5	11.092	28,2	28,2	59,5
6	8.195	20,8	20,8	80,3
7	3.633	9,2	9,2	89,5
8	2.170	5,5	5,5	95,0
9-98	1.952	4,9	4,9	100,0
Total	39.375	100,0	100,0	

Variable name:	Type of accommodation
Variable label:	Acchuk
Variable type:	Nominal
Values and value labels	1-9, missing

a1 IO:What type of accomodation is it?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Detached	4758	12.1	22.7	22.7
	2 Semi-detached	6425	16.3	30.7	53.4
	3 Terrace/end of terrace	6165	15.7	29.5	82.9
	4 In a purpose built block	2762	7.0	13.2	96.1
	5 Part of a converted house/some other kin	563	1.4	2.7	98.8
	6 Room or rooms	32	.1	.2	98.9
	7 Caravan, mobile home or houseboat	47	.1	.2	99.1
	8 Some other kind of accomodation	23	.1	.1	99.3
	9 Don t know/not applicable/unable to code	155	.4	.7	100.0
	Total	20930	53.2	100.0	
Missing	Sy stem	18445	46.8		
Total		39375	100.0		

Variable name: Number of dependent children
Variable label: Dpchuk
Variable type: Nominal
Values and value labels: Xx,01-19

No frequencies provided

Interviewer Observation File

Variable name: Physical barriers to entry
 Variable label: A41IO
 Variable type: Nominal
 Values and value labels 1-6, missing

a41 IO:Are there any physical barriers to entry

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Locked common entrance	1861	4.7	8.9	8.9
	2 Locked gates	154	.4	.7	9.6
	3 Security staff or other gatekeeper	49	.1	.2	9.9
	4 Entry phone access	556	1.4	2.7	12.5
	5 None	18021	45.8	86.1	98.6
	6 Don t know	289	.7	1.4	100.0
	Total	20930	53.2	100.0	
Missing	Sy stem	18445	46.8		
Total		39375	100.0		

Variable name: Houses/blocks in the area
 Variable label: N2_2IO
 Variable type: Nominal
 Values and value labels 0-3, missing

n2_2 IO: are the houses/blocks in this area

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 good	15167	38.5	72.5	72.5
	1.00 fair	5212	13.2	24.9	97.4
	2.00 bad	334	.8	1.6	99.0
	3.00 unable to code	217	.6	1.0	100.0
	Total	20930	53.2	100.0	
Missing	Sy stem	18445	46.8		
Total		39375	100.0		

Variable name: Safety
 Variable label: N6_IO
 Variable type: Nominal
 Values and value labels 1-4, missing

n6 IO:[*] How safe would you feel walking alon

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Very safe	8397	21.3	40.1	40.1
	2 Fairly safe	10175	25.8	48.6	88.7
	3 A bit unsafe	1977	5.0	9.4	98.2
	4 Very unsafe	313	.8	1.5	99.7
	9	68	.2	.3	100.0
	Total	20930	53.2	100.0	
Missing	Sy stem	18445	46.8		
Total		39375	100.0		

Call Schedules, response outcomes

No tables provided

Area Classifications

Variable name: ACORN classification
 Variable label: Sacorn
 Variable type: Nominal
 Values and value labels 2 digit classification (summary in table below)

sacorn Household: ACORN CODE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-12 wealthy achievers	9,054	23.0	23.6	23.6
	13-23 urban prosperity	5,420	13.8	14.2	37.8
	24-36 comfortable off	14,340	36.4	37.4	75.2
	37-43 moderate means	5,163	13.1	13.5	88.7
	44-55 hard pressed	4,265	10.8	11.1	99.9
	XX	53	0.1	0.1	100.0
	Total	38,295	97.3	100.0	
missing		1,052	2.7		
	0	28	0.1		
	Total	39,375	100.0		

Variable name: Government Office Region 2001
 Variable label: Gor2001
 Variable type: nominal
 Values and value labels 1 digit classification

gor2001 Region (AFPD):Government Office Region 2001

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A North East	1791	4.5	4.5	4.5
	B North West	4762	12.1	12.1	16.6
	D Yorkshire and The Humber	3481	8.8	8.8	25.5
	E East Midlands	2905	7.4	7.4	32.9
	F West Midlands	3557	9.0	9.0	41.9
	G East of England	3830	9.7	9.7	51.6
	H London	4470	11.4	11.4	63.0
	J South East	5543	14.1	14.1	77.1
	K South West	3461	8.8	8.8	85.8
	W Wales	1976	5.0	5.0	90.9
	X Scotland	3599	9.1	9.1	100.0
	Total	39375	100.0	100.0	

Variable name: Area type
Variable label: Sla1
Variable type: Nominal
Values and value labels: 1-6

sla1 Household: AREA TYPE 6 CATEGORIES

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 LONDON	4470	11.4	11.4	11.4
	2 METROPOLITAN	7555	19.2	19.2	30.5
	3 ENGLISH UAs	5885	14.9	14.9	45.5
	4 WELSH UAs	1974	5.0	5.0	50.5
	5 SCOTTISH COUNCIL AREAS	3598	9.1	9.1	59.6
	6 NON METROPOLITAN	15892	40.4	40.4	100.0
	Total	39374	100.0	100.0	
Missing	System	1	.0		
Total		39375	100.0		

Variable name: Urban/rural indicator
Variable label: Ur2001a
Variable type: Nominal
Values and value labels: 1,2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Urban	34855	88.5	89.3	89.3
	2 Rural	4165	10.6	10.7	100.0
	Total	39020	99.1	100.0	
Missing	Sy stem	355	.9		
Total		39375	100.0		

Variable name	Urban/rural indicator in regions
Variable label	Ur2001
Variable type	nominal
Values and value labels	-

ur2001 Region (AFPD):Urban Rural Indicator

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	19Z Eng/Wales: Urban with pop above 10k & sparse surrounds	74	.2	.2	.2
	29Z Eng/Wales: Small Town and Fringe & sparse surrounds	277	.7	.7	.9
	39Z Eng/Wales: Village & sparse surrounds	355	.9	.9	1.8
	49Z Eng/Wales: Hamlet/Isolated Dwelling & sparse surrounds	159	.4	.4	2.2
	59Z Eng/Wales: Urban with pop above 10k & less sparse surrounds	27706	70.4	71.0	73.2
	69Z Eng/Wales: Small Town and Fringe & less sparse surrounds	3704	9.4	9.5	82.7
	79Z Eng/Wales: Village & less sparse surrounds	2263	5.7	5.8	88.5
	89Z Eng/Wales: Hamlet/Isolated Dwelling & less sparse surrounds	885	2.2	2.3	90.8
	91Z Scotland: Large Urban Area: 125,000 people	1445	3.7	3.7	94.5
	92Z Scotland: Other Urban Area: 10,000 to 125,000 people	1054	2.7	2.7	97.2
	93Z Scotland: Accessible Small Town: 3,000 to 10,000 people	410	1.0	1.1	98.2
	94Z Scotland: Remote Small Town: 3,000 to 10,000 people	113	.3	.3	98.5
	95Z Scotland: Very Remote Small Town: 3,000 to 10,000 people	72	.2	.2	98.7
	96Z Scotland: Accessible Rural: under 3,000 people	350	.9	.9	99.6
	97Z Scotland: Remote Rural: under 3,000 people	87	.2	.2	99.8
	98Z Scotland: Very Remote Rural: under than 3,000 people	66	.2	.2	100.0
	Total	39020	99.1	100.0	
Missing	no information available	355	.9		
Total		39375	100.0		



Representativity Indicators for Survey Quality

RISQ data set documentation
Norwegian European Social Survey 2006

Deliverable no. 1.6 (WP 2)

Øyvind Kleven, Li-Chun Zhang
SSB

27 June, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Den europeiske samfunnsundersøkelsen (2006) In English: The European Social Survey (2006)
Abstract:	ESS is a biennial multi-country survey covering over 30 nations. It is an academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations. The survey has been funded through the European Commission's Fifth and Sixth Framework Programmes, the European Science Foundation and national funding bodies in each country.
Topic classification:	Media use, social and public trust; political interest and participation; social values; social exclusion, national, ethnic and religious allegiances; health and security; demographics.
Type of survey	Cross section
Unit of study	Individuals
Target population	All persons of 15 years and older in Norway
Producer:	Statistisk Sentralbyrå (SSB). In English: Statistics Norway
Data file name:	RISQ_dataset_ESS_Norway.sav
File contents:	Data file contains response indicator, several auxiliary variables from administrative data and four survey item from the base questionnaire.
File structure:	Rectangular
Number of cases:	2 673
Number of variables	19
Design weighting:	Equal design weights for all persons
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	Missing items are set at SPSS system missing
Data collector:	Statistisk Sentralbyrå (SSB). In English: Statistics Norway
Mode of data collection:	Computer assisted personal interview.
Field work period(s):	21.08.06 – 19.12.06.
Sampling procedure:	Simpel random sample from the Norwegian population register
Interviewers:	123 interviewers. Experience and age of interviewers varies.

Advance information	Persons receive an individual pre-notification letter
Call schedules	Not applicable
Incentives	A lottery ticket in the advance letter
Refusal conversion	As far as possible, refusal conversion was made by new interviewers. Refusals received a "motivation letter" where the purpose of the survey was emphasised, before the interviewer made contact. In addition to the letter, they received two lottery tickets. The decision on which refusals to re-contact was made by the NC, PL and the field staff. All refusals with the exception of those marked "Will definitely not cooperate in the future", were re-contacted. The interviewers had been instructed in advance to use this category with care. 339 refusals were re-contacted.

Response rates

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Respondent	1750	65,5	65,5	65,5
Refusal	735	27,5	27,5	93,0
Unable	61	2,3	2,3	95,2
Language problem	28	1,0	1,0	96,3
Non contact	62	2,3	2,3	98,6
Other	37	1,4	1,4	100,0
Total	2673	100,0	100,0	

Codeboek of variables

Variable number	1
Variable name	Ref_no
Variable label	
Variable type	Nominal

Variable number	2
Variable name	Pool
Variable label	1 or 2 batch of cases
Variable type	Nominal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 batch	1774	66,4	66,4	66,4
	2 batch	899	33,6	33,6	100,0
	Total	2673	100,0	100,0	

Variable number 3
Variable name Gender
Variable label Gender of respondent
Variable type Nominal
Values and value labels

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	1303	48,7	48,7	48,7
	Woman	1370	51,3	51,3	100,0
	Total	2673	100,0	100,0	

Variable number 4
Variable name Age
Variable label Age of respondent
Variable type Interval
Values and value labels

N	Valid	2673
	Missing	0
Median		45,00
Minimum		15
Maximum		101

Variable number 5
Variable name Household
Variable label Number of persons in household from register
Variable type Interval (ordinal)
Values and value labels

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	787	29,4	29,4	29,4

2	731	27,3	27,3	56,8
3	434	16,2	16,2	73,0
4 or more	721	27,0	27,0	100,0
Total	2673	100,0	100,0	

Variable number 6
Variable name County
Variable label County - Regional identification
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Østfold	154	5,8	5,8	5,8
Akershus	291	10,9	10,9	16,6
Oslo	311	11,6	11,6	28,3
Hedmark	115	4,3	4,3	32,6
Oppland	103	3,9	3,9	36,4
Buskerud	145	5,4	5,4	41,9
Vestfold	129	4,8	4,8	46,7
Telemark	95	3,6	3,6	50,2
Aust-Agder	54	2,0	2,0	52,3
Vest-Agder	95	3,6	3,6	55,8
Rogaland	214	8,0	8,0	63,8
Hordaland	259	9,7	9,7	73,5
Sogn og Fjordane	65	2,4	2,4	75,9
Møre og Romsdal	144	5,4	5,4	81,3
Sør-Trøndelag	154	5,8	5,8	87,1
Nord-Trøndelag	73	2,7	2,7	89,8
Nordland	133	5,0	5,0	94,8
Troms	94	3,5	3,5	98,3
Finnmark	45	1,7	1,7	100,0
Total	2673	100,0	100,0	

Variable number 7
Variable name Centrality
Variable label Centrality of municipality
Variable type Ordinal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Less central municipalities	356	13,3	13,3	13,3

Less remote municipalities	189	7,1	7,1	20,4
Fairly central municipalities	665	24,9	24,9	45,3
Central municipalities	1463	54,7	54,7	100,0
Total	2673	100,0	100,0	

Variable number 8
Variable name Education
Variable label Level of education
Variable type Ordinal
Values and value labels

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low	599	22,4	22,4	22,4
	Middle	1478	55,3	55,3	77,7
	High (Univ)	596	22,3	22,3	100,0
	Total	2673	100,0	100,0	

Variable number 9
Variable name Reassigned
Variable label Ordinary or re-assigned case
Variable type Nominal
Values and value labels

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Ordinary	2155	80,6	80,8	80,8
	Re-assigned	512	19,2	19,2	100,0
	Total	2667	99,8	100,0	
Missing	System	6	,2		
Total		2673	100,0		

Variable number 10
Variable name TELNUM
Variable label Telephone number on somebody in the household present
Variable type Nominal
Values and value labels

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Present	2599	97,2	97,2	97,2
	No phone	74	2,8	2,8	100,0
	Total	2673	100,0	100,0	

Variable number 11
Variable name Day
Variable label Day in fieldwork period where final status (restype) was determined
Variable type Interval
Values and value labels

N	Valid	2664
	Missing	9
Median		44,50
Minimum		1
Maximum		114

Variable number 12
Variable name Week
Variable label Week in fieldwork period where final status (restype) was determined
Variable type Interval
Values and value labels

N	Valid	2664
	Missing	9
Mean		7,77
Median		7,00
Minimum		1
Maximum		18

Variable number 13
Variable name Contacts
Variable label Number of contacts with household
Variable type Interval
Values and value labels

N	Valid	2673
	Missing	0
Mean		3,01
Median		2,00
Minimum		0
Maximum		10

Variable number 14
Variable name Resptype
Variable label Type of response to survey
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Respondent	1750	65,5	65,5	65,5
Refusal	735	27,5	27,5	93,0
Unable	61	2,3	2,3	95,2
Language problem	28	1,0	1,0	96,3
Non contact	62	2,3	2,3	98,6
Other	37	1,4	1,4	100,0
Total	2673	100,0	100,0	

Variable number 15
Variable name polintr
Variable label How interested in politics
Variable type Ordinal
Values and value labels

How interested in politics

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Very interested	165	6,2	9,4	9,4
Quite interested	674	25,2	38,5	48,0
Hardly interested	775	29,0	44,3	92,3
Not at all interested	135	5,1	7,7	100,0
Total	1749	65,4	100,0	
Missing Don't know	1	,0		
System	923	34,5		
Total	924	34,6		
Total	2673	100,0		

Variable number 16
Variable name Vote
Variable label Voted last national election

Values and value labels

Voted last national election

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	1364	51,0	78,0	78,0
	No	230	8,6	13,2	91,2
	Not eligible to vote	154	5,8	8,8	100,0
	Total	1748	65,4	100,0	
Missing	Don't know	2	,1		
	System	923	34,5		
	Total	925	34,6		
Total		2673	100,0		

Variable number

17

Variable name

prtvtno

Variable label

Party voted for in last national election, Norway

Variable type

Nominal

Values and value labels

Party voted for in last national election, Norway

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Red Electoral Alliance (RV)	18	,7	1,4	1,4
	Socialist left party (SV)	149	5,6	11,5	12,8
	Labour Party (A)	409	15,3	31,4	44,3
	Liberal Party (V)	62	2,3	4,8	49,0
	Christian Democratic Party (Krf)	95	3,6	7,3	56,3
	Centre Party (Sp)	90	3,4	6,9	63,3
	Conservative Party (H)	235	8,8	18,1	81,3
	Progress Party (FrP)	227	8,5	17,4	98,8
	Coast Party (KYST)	1	,0	,1	98,8
	Other	15	,6	1,2	100,0
	Total	1301	48,7	100,0	
Missing	Not applicable	386	14,4		
	Refusal	49	1,8		
	Don't know	14	,5		

System	923	34,5		
Total	1372	51,3		
Total	2673	100,0		

Variable number 18

Variable name mnactic

Variable label Main activity, last 7 days. All respondents. Post coded

Variable type Nominal

Values and value labels

Main activity, last 7 days. All respondents. Post coded

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Paid work	1083	40,5	62,0	62,0
	Education	196	7,3	11,2	73,2
	Unemployed, looking for job	16	,6	,9	74,1
	Unemployed, not looking for job	14	,5	,8	74,9
	Permanently sick or disabled	44	1,6	2,5	77,4
	Retired	275	10,3	15,7	93,1
	Community or military service	3	,1	,2	93,3
	Housework, looking after children, others	105	3,9	6,0	99,3
	Other	12	,4	,7	100,0
	Total	1748	65,4	100,0	
Missing	Don't know	2	,1		
	System	923	34,5		
	Total	925	34,6		
Total		2673	100,0		

Variable number 19
Variable name Hinctnt
Variable label Household's total net income, all sources
Variable type Nominal
Values and value labels

Household's total net income, all sources (1 Euro = 8,2 Nkr)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Less than 12 900 Nkr	16	,6	,9	,9
	12 900 - 25 800	13	,5	,8	1,7
	25 800 - 43 000	34	1,3	2,0	3,7
	43 000 - 85 900	42	1,6	2,5	6,2
	85 900 - 128 900	87	3,3	5,2	11,4
	128 900 - 171 800	88	3,3	5,2	16,6
	171 800 - 214 800	103	3,9	6,1	22,7
	214 800 - 257 700	157	5,9	9,3	32,0
	257 700 - 429 500	487	18,2	28,9	60,9
	429 500 - 644 300	386	14,4	22,9	83,8
	644 300 - 859 00	168	6,3	10,0	93,8
	More than 859 00 Nkr	105	3,9	6,2	100,0
Missing	Total	1686	63,1	100,0	
	Refusal	5	,2		
	Don't know	59	2,2		
	System	923	34,5		
Total	Total	987	36,9		
		2673	100,0		



Representativity Indicators for Survey Quality

RISQ data set documentation
Norwegian Survey of Level of Living 2004

Deliverable no. 1.7 (WP 2)

Øyvind Kleven, Li-Chun Zhang
SSB

23 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Levekårsundersøkelsen 2004. In English: Survey of level of living, 2004
Abstract:	The Survey of living conditions has two main purposes. One is to throw light on the main aspects of the living conditions in general and for various groups of people. Another purpose is to monitor development in living conditions, both level and distribution. Over a three-year period the cross-sectional survey of living conditions will cover all main areas of the living conditions. The survey topics change during a three-year cycle. Housing conditions, participation in organisations, leisure activities, offences and fear of crime were topics in 2004.
Topic classification:	Living conditions, Social indicators, Health conditions, Crime and the justice, Dwelling and housing conditions, Working conditions, sickness, absence, Membership in organizations, Sports and outdoor life.
Type of survey	Cross section
Unit of study	Individuals
Target population	All persons of 16 years and older in Norway
Producer:	Statistisk Sentralbyrå (SSB). In English: Statistics Norway
Data file name:	RISQ_dataset_SLL_Norway.sas7bdat
File contents:	Data file contains response indicator, several auxiliary variables from administrative data and 5 survey item from the base questionnaire.
File structure:	Rectangular
Number of cases:	4 837
Number of variables	17
Design weighting:	Equal design weights for all persons
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	Missing items are set at SAS system missing
Data collector:	Statistisk Sentralbyrå (SSB). In English: Statistics Norway
Mode of data collection:	Computer assisted personal interview and computer assisted telephone interview.
Field work period(s):	23.08.04 – 10.01.04.

Sampling procedure: The survey is a two-stage sample, in which the clusters in the first stage are formed by municipalities. From the clusters simple random samples without replacement are drawn consisting of persons.

Interviewers: 130 interviewers, working in 109 interviewer regions. Experience and age of interviewers varies.

Advance information Persons receive an individual pre-notification letter

Call schedules Not applicable

Incentives Non

Refusal conversion Yes

Response rates

Type of response to survey

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	3340	69,1	69,1	69,1
	Refusal	1098	22,7	22,7	91,8
	Unable/language problem	173	3,6	3,6	95,3
	Non contact	218	4,5	4,5	99,8
	Other	8	,2	,2	100,0
	Total	4837	100,0	100,0	

Codeboek of variables

Variable number 1
Variable name Ref_no
Variable label
Variable type Nominal

Variable number 2
Variable name Gender
Variable label Gender of respondent
Variable type Nominal
Values and value labels

Gender of respondent

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	2437	50,4	50,4	50,4
	Female	2400	49,6	49,6	100,0
	Total	4837	100,0	100,0	

Variable number 3
Variable name Age
Variable label Age of respondent
Variable type Interval
Values and value labels

Statistics

Age of respondent

N	Valid	4837
	Missing	0
Median		44,00
Minimum		16
Maximum		86

NB: All persons over 85 have been given the value of 86

Variable number 4

Variable name County

Variable label County - Regional identification

Variable type Nominal

Values and value labels

County - Regional identification

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Østfold	279	5,8	5,8	5,8
Akershus	509	10,5	10,5	16,3
Oslo	577	11,9	11,9	28,2
Hedmark	203	4,2	4,2	32,4
Oppland	184	3,8	3,8	36,2
Buskerud	244	5,0	5,0	41,3
Vestfold	227	4,7	4,7	46,0
Telemark	193	4,0	4,0	49,9
Aust-Agder	98	2,0	2,0	52,0
Vest-Agder	169	3,5	3,5	55,5
Rogaland	400	8,3	8,3	63,7
Hordaland	486	10,0	10,0	73,8
Sogn og Fjordane	110	2,3	2,3	76,1
Møre og Romsdal	241	5,0	5,0	81,0
Sør-Trøndelag	300	6,2	6,2	87,2
Nord-Trøndelag	141	2,9	2,9	90,2
Nordland	244	5,0	5,0	95,2
Troms	161	3,3	3,3	98,5
Finmark	71	1,5	1,5	100,0
Total	4837	100,0	100,0	

Variable number 5
Variable name Municiptyp
Variable label Type of Municipality
Variable type Nominal
Values and value labels

Type of Municipality

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Primary industry municipalities/Mixed aggriculture and manufacturing municipalities/Manufacturing municipalities	907	18,8	18,8	18,8
	Less central, mixed service industry and manufacturing municipalities/Less central service industry municipalities	622	12,9	12,9	31,7
	Central, mixed service industry and manufacturing municipalities/Central service industry municipalities	3293	68,1	68,3	100,0
	Total	4822	99,7	100,0	
Missing	System	15	,3		
Total		4837	100,0		

Variable number 6
Variable name Education
Variable label Level of education
Variable type Ordinal
Values and value labels

Level of education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low	1114	23,0	23,0	23,0
	Middle	2632	54,4	54,4	77,4
	High (Univ)	1091	22,6	22,6	100,0
	Total	4837	100,0	100,0	

Variable number 7
Variable name Marstat
Variable label Marital status
Variable type Nominal
Values and value labels

Marital status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not married	1730	35,8	35,8	35,8
	Married	2269	46,9	46,9	82,7
	Widowed	285	5,9	5,9	88,6
	Divorced	553	11,4	11,4	100,0
	Total	4837	100,0	100,0	

Variable number 8
Variable name Norcitizen
Variable label Norwegian citizen or not
Variable type Nominal
Values and value labels

Norwegian citizen or not

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not Norwegian	227	4,7	4,7	4,7
	Norwegian	4610	95,3	95,3	100,0
	Total	4837	100,0	100,0	

Variable number 9
Variable name Mode
Variable label Mode of contact
Variable type Nominal
Values and value labels

Mode of contact

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Telephone	2542	52,6	52,7	52,7
	Personal visit	2203	45,5	45,7	98,4
	Non contact	79	1,6	1,6	100,0

Total	4824	99,7	100,0
Missing System	13	,3	
Total	4837	100,0	

Variable number 10
Variable name Reassigned
Variable label Ordinary or re-assigned case
Variable type Nominal
Values and value labels

Ordinary or re-assigned case

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Ordinary	4399	90,9	91,0	91,0
	Re-assigned	436	9,0	9,0	100,0
	Total	4835	100,0	100,0	
Missing	System	2	,0		
Total		4837	100,0		

Variable number 11
Variable name Int_Data
Variable label Date for when final status was determined
Variable type Interval
Values and value labels

Variable number 12
Variable name Resptype
Variable label Type of response to survey
Variable type Nominal
Values and value labels

Type of response to survey

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent	3340	69,1	69,1	69,1
	Refusal	1098	22,7	22,7	91,8
	Unable/language problem	173	3,6	3,6	95,3
	Non contact	218	4,5	4,5	99,8

Other	8	,2	,2	100,0
Total	4837	100,0	100,0	

Variable number 13
Variable name MNACTIC
Variable label Main activity
Variable type Ordinal
Values and value labels

Main activity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Paid work	2061	42,6	61,7	61,7
	Education (including military service)	412	8,5	12,3	74,0
	Unemployed	77	1,6	2,3	76,3
	Retired	628	13,0	18,8	95,1
	Houswork, looking after children, others	85	1,8	2,5	97,7
	Other	77	1,6	2,3	100,0
	Total	3340	69,1	100,0	
Missing	System	1497	30,9		
Total		4837	100,0		

Variable number 14
Variable name Workhour
Variable label Respondents usually hours of work per week
Variable type Interval
Values and value labels

Statistics

Respondents usually hours of work per week

N	Valid	2276
	Missing	2561
Median		38,00
Minimum		1
Maximum		84

NB: Over 83 has been given the number of 84

Variable number 15
Variable name ORG
Variable label ember of organisation?

Nominal

Percent
16,2
83,8
100,0

16

Weakecon

Financial circumstances made it difficult to pay an unforeseen bill of 5000 nok last 12 months.

Nominal

enforseen b

Percent
76,8
23,2
100,0

17

Unemployed

Have you been unemployed last 3 months?

Nominal

months?

Percent
96,8
3,2
100,0



Representativity Indicators for Survey Quality

RISQ data set documentation
Belgian European Social Survey 2006

Deliverable 1.8 (WP2)

Koen Beullens, Geert Loosveldt
KUL

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title: **European Social Survey (ESS), round 3, Belgium**

NOTE: **Although ESS data are publicly available, this data set contains data from the sample frame that is not disseminated via the internet (ZIP codes). Therefore, these data can only be used among RISQ participants.**

Abstract: The European Social Survey (ESS) is an academically-driven multi-country survey, which has been administered in over 30 countries to date. Its three aims are, firstly - to monitor and interpret changing public attitudes and values within Europe and to investigate how they interact with Europe's changing institutions, secondly - to advance and consolidate improved methods of cross-national survey measurement in Europe and beyond, and thirdly - to develop a series of European social indicators, including attitudinal indicators. In the third round, the survey covers 25 countries and employs the most rigorous methodologies. It is funded via the European Commission's 6th Framework Programme, the European Science Foundation, and national funding bodies in each country. It involves strict random probability sampling, a minimum target response rate of 70% and rigorous translation protocols. The hour-long face-to-face interview includes questions on a variety of core topics repeated from previous rounds of the survey and also two modules developed for Round Three covering personal and social well being and the organisation of the life course in Europe.

Topic classification: Media, social trust, political interest and participation, socio-political orientations, social exclusion, national, ethnic and religious allegiances, timing of key life events and the life course, personal and social well-being and satisfaction with work and life, demographics and socio economics.

Type of survey Cross section

Unit of study Individuals

Target population All persons aged 15 and over resident within private households, regardless of their nationality, citizenship, language or legal status, in Belgium

Producer: Roger Jowell (PI), CCSS, City University, UK together with the Central Co-ordinating Team and the National Coordinator from Belgium:

Geert Loosveldt, Katholieke Universiteit Leuven, Belgium
Marc Jacquemain, University of Liege, Belgium

Data file name: risq_ess3_be.por

File contents:	Data set contains response indicator, contact data and several auxiliary variables on both the individual level as on the level of municipalities where the individual sample unit reside.
File structure:	Rectangular
Number of cases:	3249
Number of variables	194
Design weighting:	Equal design weights for all persons
Adjustment weights	Adjustment weights are not included
Imputation	No missing items are imputed
Missing data:	see description variables
Data collector:	TNS Dimarso
Mode of data collection:	Face to face interviews, CAPI
Field work period(s):	23.10.06 - 19.02.07
Sampling procedure:	Sampling frame:

The basis is the commercial database of 'Orgassim'. Using the National register, Orgassim has developed a database with 'Statistics of inhabitants per building'. With this database it is possible to make an individual database with age, gender and address for each person. The names of the persons are not available in this database. Then, the individual database is linked with another commercial database and 'enriched' with names (65% matches). A person is identified by his or her name or the combination of gender and age. The database is updated annually.

Sampling design:

Stratified two stage probability sampling. The ten provinces and Brussels are used for regional stratification.

Stage 1: The primary sampling units (PSU's) are 'virtual' clusters located in municipalities. This means that the clusters within the municipalities are not further defined regionally. The number of clusters for each province is proportional to the size of the population in each province. For that a list of municipalities with a population distribution (+15) years for each province is used. The number of clusters in a municipality is proportional to the size of its population. The total number of clusters equals 338 (see below).

Stage 2: In each of the 338 clusters, 9 persons are

selected for the gross sample by simple random sampling. This means that the number of contacted persons in each municipality equals the number of clusters in the municipality times 9.

Interviewers:

Interviewer selection

Total number of interviewers:	118
Number of experienced interviewers:	118
Numbers of inexperienced interviewers:	0

Briefing of interviewers

Number of interviewers who received ESS specific personal briefing:	118
Total length of ESS specific personal briefing(s) per interviewer:	½ day or less
Training in refusal conversion:	Yes
Written ESS specific instructions:	Yes

Employment status of interviewers

Free-lance interviewers:	Yes
Employees of the survey organisation:	No
Other:	No

Payments of interviewers

Hourly rate:	No
Per completed interview:	Yes
Assignment fee (set fee for working on a set of sample units):	No
A regular fixed salary:	No
Bonus arrangement:	No
Other:	No

Advance information

Use of advance letter:	Yes
Use of brochure:	Yes

Call schedules

First contact by:	Visit
Number of minimum required visits per respondent:	4
Number of visits (per respondent) required to be on a weekend:	1
Number of visits (per respondent) required to be in the evening:	1

Incentives

Respondent incentive:	No
Unconditional monetary incentives, paid before the interview:	No
Conditional monetary incentives, upon completion of the interview:	No
Unconditional non-monetary incentives, paid before the interview:	No
Conditional non-monetary incentives. upon completion of the interview:	No

Refusal conversion

Yes. 720 units that were either 'non-contact' or 'refusal' were re-issued and re-contacted again. This was necessary to achieve the minimum effective size (N=1800). Other interviewers than the original ones were used for these activities.

Response rates

resp_code	Frequency	Percent	Cumulative Frequency	Cumulative Percent
ineligible	322	9.91	322	9.91
noncontact	80	2.46	402	12.37
other nonresponse	302	9.30	704	21.67
refusal	727	22.38	1431	44.04
respondent	1818	55.96	3249	100.00

Codebook of variables

Variable number	1
Variable name	NIS
Source	Sample frame
Level of observation	Individual - municipality
Variable label	nis code (~ZIP code)
Variable type	Nominal
Values and value labels	221 values ranging from 11001 up to 93022
Missing code	No missing values
Variable number	2
Variable name	IDNO
Source	Sample frame, Contact data and Main data
Level of observation	individual
Variable label	Respondents identification number
Variable type	Nominal
Values and value labels	3249 values ranging from 10101 up to 91509
Missing code	No missing values
Variable number	3-22
Variable name	DATEV1 DATE2-DATE20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	Date of the i th visit
Variable type	
Values and value labels	1-31
Missing code	'99' or `.`
Variable number	23-42
Variable name	MONV1-MONV20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	Month of the i th visit
Variable type	
Values and value labels	1 (= January) - 12 (= December)

Missing code	'99' or `.`
Variable number	43-62
Variable name	DAYV1-DAYV20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	DAY of the i th visit
Variable type	
Values and value labels	1 (= Monday) – 7 (= Sunday)
Missing code	'9' or `.`
Variable number	63-82
Variable name	HOURV1-HOURV20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	HOUR of the i th visit
Variable type	
Values and value labels	0-23
Missing code	'99' or `.`
Variable number	83-102
Variable name	MINV1-MINV20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	MINUTE of the i th visit
Variable type	
Values and value labels	0-59
Missing code	'99' or `.`
Variable number	103-122
Variable name	MODEVB1-MODEVB20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	CONTACT MODE of the i th visit -
Variable type	NOMINAL
Values and value labels	1: personal face-to-face 2: telephone 3: personal but only by intercom 4: information through survey organization 5: other

Missing code	'9' or \.'
Variable number	123-142
Variable name	RESULB1-RESULB20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	RESULT of the i th visit
Variable type	NOMINAL
Values and value labels	1: Completed Interview with respondent 2: Partial interview 3: Contact with someone, Target Respondent not yet selected 4: Contact with Target Respondent, but no interview 5: Contact with somebody other than Target Respondent 6: No contact at all 7: Address is not valid (unoccupied, demolished, institution,...) 8: other information about sample unit
Missing code	'9' or \.'
Variable number	143-162
Variable name	OUTNIA1-OUTNIA20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	OUTCOME of the i th visit WHEN THERE WAS NO INTERVIEW
Variable type	NOMINAL
Values and value labels	01= appointment 02: refusal of respondent 03: Refusal by someone else, on behalf of the respondent 04: Household Refusal: Refusal, don't know if target respondent 05: Respondent not available/away 06: Respondent mentally or physically not able 07: Respondent deceased 08: Respondent moved out of country or destination unknown 09: Respondent moved, still in country 10: Language barrier of Respondent 11: Other
Missing code	'9' or \.'

Variable number	163
Variable name	TYPE
Source	Contact data (interviewers' impression)
Level of observation	Individual
Variable label	Type of house respondent lives in
Variable type	NOMINAL
Values and value labels	01: Farm 02: Single unit: Detached house 03: Single unit: Semi-detached house 04: Single unit: Terraced house 05: Only housing unit in building with other purpose 06: Multi-unit house, flat 07: Multi-unit: Student apartments, rooms 08: Multi-unit: Sheltered/retirement housing 09: House -trailer or boat 10: Other 88: Don't know
Missing code	'99' or \.'

Variable number	164
Variable name	PHYS (interviewers' impression)
Source	Contact data
Level of observation	Individual
Variable label	Physical state of the buildings and dwellings in the area
Variable type	ORDINAL
Values and value labels	1: In a very good state 2: In a god state 3: In a very satisfactory state 4: Bad state 5: Very bad state
Missing code	'9' or \.'

Variable number	165
Variable name	LITTER (interviewers' impression)
Source	Contact data
Level of observation	Individual
Variable label	How common is litter or rubbish lying around the immediate area
Variable type	ORDINAL
Values and value labels	1: Very common 2: Fairly common 3: Not very common 4: Not at all common
Missing code	'9' or '.'

Variable number	166
Variable name	VANDA (interviewers' impression)
Source	Contact data
Level of observation	Individual
Variable label	How common is vandalism, graffiti or deliberate damage to property
Variable type	ORDINAL
Values and value labels	1: Very common 2: Fairly common 3: Not very common 4: Not at all common
Missing code	'9' or '.'

Variable number	167
Variable name	AGE
Source	Sample frame
Level of observation	Individual
Variable label	age of the sample unit
Variable type	METRIC
Values and value labels	
Missing code	NONE

Variable number	168
Variable name	MALE
Source	Sample frame
Level of observation	Individual
Variable label	Respondent is a man (=yes)
Variable type	NOMINAL
Values and value labels	0: female 1: male
Missing code	NONE

Variable number	169
Variable name	RESP_CODE
Source	Calculated from contact data
Level of observation	Individual
Variable label	Final response code
Variable type	NOMINAL
Values and value labels	'Respondent' 'refusal' (by respondent or proxy) 'ineligible' (deceased, R moved, house not built or demolished, business or institution, house not occupied) 'other nonresponse' (partial/invalid interview, broken appointment, R not available, away, mentally or physically unable, language barrier, address not tracable, not attempted) 'noncontact'
Missing code	NONE

Variable number	170
Variable name	LATITUDE
Source	External source (Statistics Belgium http://www.statbel.fgov.be/) Linked to NIS CODE
Level of observation	municipality
Variable label	Latitude of municipality of sample unit
Variable type	METRIC
Values and value labels	Ranging from 49.5833 to 51.45
Missing code	NONE

Variable number	171
Variable name	LONGITUDE
Source	External source (Statistics Belgium http://www.statbel.fgov.be/) Linked to NIS CODE
Level of observation	municipality
Variable label	Longitude of municipality of sample unit
Variable type	METRIC
Values and value labels	Ranging from 2.5833 to 5.95
Missing code	NONE
Variable number	172
Variable name	POP_DENSITY
Source	External source (Statistics Belgium http://www.statbel.fgov.be/) Linked to NIS CODE
Level of observation	municipality
Variable label	Population density of municipality of sample unit (inhabitants/km ²)
Variable type	METRIC
Values and value labels	Ranging from 29 to 17907
Missing code	NONE
Variable number	173
Variable name	PERC_FOREIGN
Source	External source (Statistics Belgium http://www.statbel.fgov.be/) Linked to NIS CODE
Level of observation	municipality
Variable label	Percentage if foreigners in municipality of sample unit
Variable type	METRIC
Values and value labels	Ranging from 0.0058 to 0.4106
Missing code	NONE

Variable number	174
Variable name	INCOME
Source	External source (Statistics Belgium http://www.statbel.fgov.be/) Linked to NIS CODE
Level of observation	municipality
Variable label	average annual per capita income in municipality of sample unit (euro)
Variable type	METRIC
Values and value labels	Ranging from 8836€ to 18497€
Missing code	NONE

Variable number	175-194
Variable name	INT1-INT20
Source	Calculated from contact data (totcint1-totcont3 and intnum1- intnum3)
Level of observation	Contact level (within individual)
Variable label	interviewer who conducted i th visit
Variable type	nominal
Values and value labels	Ranging from 1 to 117
Missing code	\.'

Variable number	195-214
Variable name	YEARV1-YEARV20
Source	Calculated from contact data (monv1-mon20)
Level of observation	Contact level (within individual)
Variable label	Year of i th visit
Variable type	
Values and value labels	2006 or 2007
Missing code	\.'



RISQ data set documentation
Flemish Housing survey 2005

Deliverable 1.9 (WP2)

Koen Beullens, Geert Loosveldt
KUL

18 July, 2008



The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.

RISQ SURVEY DATA FILE DOCUMENTATION

Title: **Flemish Housing Survey**

NOTE: **Data can only be used among RISQ participants.**

Abstract: The Flemish Housing Survey is conducted by the 'Research Network on Sustainable Housing Policy' commissioned by the Ministry of the Flemish Community, Housing Policy Department. The Flemish Housing study consists of two parts. The second part is a face-to-face survey. This part focuses on the screening of profiles, expectation and needs of the Flemings as housing consumers. Independently from the face-to-face survey, inspectors were used to register auxiliary data on the exterior characteristics of the private residences of each of the sample units. This is the first part of the project. Of the 8400 dwellings that were technically inspected (first part), the occupants of 7770 units were subsequently selected for a face-to-face survey. Representativity of the face-to-face part is of main interest with regard to the RISQ project.

Topic classification: Housing quality, as perceived by the occupants.

Type of survey Cross section

Unit of study Houses

Target population Reference persons of the families occupying houses in the Flemish region.

Producer: The Flemish Housing Survey is conducted by the 'Research Network on Sustainable Housing Policy' commissioned by the Ministry of the Flemish Community, Housing Policy Department. It is coordinated by the centre of Survey Methodology of the University of Leuven.

Data file name: risq_housing_survey.por

File contents: Data set contains response indicator, contact data and several auxiliary variables attributable to both the reference persons as well as their houses.

File structure: Rectangular

Number of cases: 8400

Number of variables 106

Design weighting: The steering committee has decided that the number of sample units per district should be at least 270 and should not exceed 600. This decision disturbs the representativity equilibrium particularly with regard to

the greatest and the smallest districts. This disproportionality can be corrected by using weight variables.

Adjustment weights	dweight_frame corrects the data frame (n=8400) towards the population while dweight_gross corrects the units selected for face-to-face survey (n=7770).
Imputation	No missing items are imputed
Missing data:	see description variables
Data collector:	Dimarso-NID, coordinated by Centre for Survey Methodology (K.U.Leuven)
Mode of data collection:	Face to face interviews, CAPI
Field work period(s):	04.05 - 02.06
Sampling procedure:	Sampling for technical scrutiny of houses.

8400 persons (reference persons – head of family) have been drawn from the national register in only 201 of the 308 Flemish municipalities. The information includes name, address, date of birth, gender and identification number. Per district (Flanders count 23 districts) a pre-specified number of clusters had to be drawn under PPS_WR-conditions (probability proportional to size with replacement). Each cluster referred to a Flemish municipality.

Sampling for face-to-face survey.

Only 7770 units of the 8400 were selected for face-to-face survey, due to financial and organizational reasons. The probability that a cluster was sampled for the face-to-face round was disproportional to the over-or underrepresentation with regard to the districts. This makes the variance in dweight_gross smaller than in dweight_frame.

Interviewers (face-to-face survey):

Interviewer selection

Total number of interviewers:	194
Number of experienced interviewers (at least one year)	194
Experienced interviewers in scientific research (no commercial surveys)	50%
Introductory training for inexperienced interviewers	yes

Briefing of interviewers

Number of interviewers who received briefing: 192
 Training in refusal conversion: no

Control survey by telephone

Number of interviewers: 5

Advance information

Use of advance letter: Yes
 Use of brochure: Yes

Call schedules

First contact by: Visit
 Number of minimum required visits per respondent: 4
 Number of visits (per respondent) required to be on a 1
 weekend:
 Number of visits (per respondent) required to be in 1
 the evening:

Refusal conversion

Yes, only when a refusal was given by telephone.

Replacement of respondents:

Only within the family when the original reference person was not capable of participating or had moved. This happened in only 429 cases.

New interviewer:

In 64 cases, the original interviewer has been replaced by a new one. Unfortunately, the contact efforts of the first interviewer as well as the reason for the replacement have not been documented in the contact sheet data.

Response rates (unweighted)

resp_code	Frequency	Percent	Cumulative Frequency	Cumulative Percent
ineligible	551	7.09	551	7.09
noncontact	652	8.39	1203	15.48
other nonresponse	59	0.76	1262	16.24
refusal	1292	16.63	2554	32.87
respondent	5216	67.13	7770	100.00

Frequency Missing = 630 (not selected for face-to-face survey)

Codebook of variables

Variable number	1
Variable name	IDNO
Source	Contact data and Main data
Level of observation	individual
Variable label	Respondents identification number
Variable type	Nominal
Values and value labels	8400 values ranging from 110101 up to 531915
Missing code	No missing values

Variable number	2
Variable name	IDNO_NEW
Source	Contact data and Main data
Level of observation	individual
Variable label	Respondents identification number (is different from IDNO if there was a replacement of the reference person (n=429))
Variable type	Nominal
Values and value labels	8400 values ranging from 110101 up to 531915
Missing code	No missing values

Variable number	3
Variable name	AGECLASS
Source	Contact data
Level of observation	individual
Variable label	ageclass
Variable type	Ordinal
Values and value labels	1 = 18-20 years 2 = 21-25 years 3 = 26-30 years 4 = 31-35 years 5 = 36-40 years 6 = 41-45 years 7 = 46-50 years 8 = 51-55 years 9 = 56-60 years 10 = 61-65 years 11 = 66-70 years 12 = 71-75 years 13 = 76-80 years 14 = 81-85 years 15 = 86-90 years 16 = 91-95 years 17 = 96-100 years 18 = >100 years
Missing code	No missing values

Variable number	4
Variable name	GENDER
Source	Contact data
Level of observation	individual
Variable label	Gender
Variable type	Nominal
Values and value labels	1 = Male 2 = Female
Missing code	No missing values

Variable number	5
Variable name	DISTRICT
Source	Contact data
Level of observation	individual
Variable label	District, subset of province
Variable type	Nominal
Values and value labels	11 = Antwerpen 12 = Mechelen 13 = Turnhout 21 = Halle-Vilvoorde 22 = Leuven 31 = Brugge 32 = Diksmuide 33 = Ieper 34 = Kortrijk 35 = Oostende 36 = Roeselaere 37 = Tielt 38 = Veurne 41 = Aalst 42 = Dendermonde 43 = Eeklo 44 = Gent 45 = Oudenaarde 46 = Sint - Niklaas 51 = Hasselt 52 = Maaseik 53 = Tongeren
Missing code	No missing values

Variable number	6
Variable name	PROVINCE
Source	Contact data
Level of observation	individual
Variable label	Province
Variable type	Nominal
Values and value labels	1 = Antwerpen 2 = Vlaams - Brabant 3 = West - Vlaanderen 4 = Oost - Vlaanderen 5 = Limburg
Missing code	No missing values

Variable number	7
Variable name	RESP_CODE
Source	Contact data
Level of observation	individual
Variable label	Final response code (short)
Variable type	Nominal
Values and value labels	'respondent' 'refusal' (explicit or excuses) 'noncontact' (including R on journey during fieldwork) 'ineligible' (language barrier, illness, disabled, demented, deceased, address not occupied or not traceable) 'other nonresponse' (interviews not validated)
Missing code	'.' (630)

Variable number	8
Variable name	GROSS
Source	Contact data
Level of observation	individual
Variable label	Selected for survey
Variable type	Nominal
Values and value labels	1 = selected to cooperate (7770) 0 = not selected to cooperate (630)
Missing code	No missing values

Variable number	9
Variable name	SWITCH_INT
Source	Contact data
Level of observation	individual
Variable label	initial interviewer has been replaced
Variable type	Nominal
Values and value labels	1 = yes (64) 0 = no (7706)
Missing code	'.' (630)

Variable number	10
Variable name	INTNUM
Source	Contact data
Level of observation	individual
Variable label	Number of interviewer (in case the initial interviewer has been replaced, the number of the second interviewer is given)
Variable type	Nominal
Values and value labels	187 interviewer, number ranging from 886 to 93373
Missing code	`.`

Variable number	11
Variable name	SWITCH_RESP
Source	Contact data
Level of observation	individual
Variable label	respondent has been replaced (by other family member)
Variable type	Nominal
Values and value labels	1 = yes (7341) 0 = no (429)
Missing code	`.`

Variable number	12-21
Variable name	DAYV1-DAYV10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	DAY of the i th visit
Variable type	
Values and value labels	1 (= Monday) – 7 (= Sunday)
Missing code	`.`

Variable number	22-31
Variable name	MONV1-MONV10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	Month of the i th visit
Variable type	
Values and value labels	1 (= January) - 12 (= December)
Missing code	\.'

Variable number	32-41
Variable name	DATEV1 DATE2-DATE20
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	Date of the i th visit
Variable type	
Values and value labels	1-31
Missing code	\.'

Variable number	42-51
Variable name	YEARV1-YEARV10
Source	Calculated from contact data (monv1-monv10)
Level of observation	Contact level (within individual)
Variable label	Year of i th visit
Variable type	
Values and value labels	2005 or 2006
Missing code	\.'

Variable number	52-61
Variable name	HOURV1-HOURV10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	HOUR of the i th visit
Variable type	
Values and value labels	0-23
Missing code	\.'

Variable number	62-71
Variable name	MINV1-MINV10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	MINUTE of the i th visit
Variable type	
Values and value labels	0-59
Missing code	\.'
Variable number	72-81
Variable name	MODEVB1-MODEVB10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	Contact mode of the i th visit -
Variable type	NOMINAL
Values and value labels	1: telephone 2: face-to-face 3: sample unit contacted DIMARSO NID 4: sample unit contacted Centre for Survey Methodology
Missing code	\.'
Variable number	82-91
Variable name	WHOV1-WHOV10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	contact with whom at i th visit
Variable type	NOMINAL
Values and value labels	1: initial reference person 2: replacement of initial reference person 3: another person of family 4: head of family new family 5: replacement of head of new family 6: another person of new family
Missing code	\.'

Variable number	92-101
Variable name	RESULB1-RESULB10
Source	Contact data
Level of observation	Contact level (within individual)
Variable label	Result of the i th visit
Variable type	NOMINAL
Values and value labels	1: Completed Interview with respondent 2: appointment 3: French-speaking 4: language barrier 5: illness, disables, demented 6: R at home, but no occasion 7: on holiday, business trip 8: deceased 9: explicit refusal 10: refusal, excuses 11: at home, but no appointment possible 12: other (in case there was contact with R) 13: not at home 14: home but nobody answered 15: address not traceable 16: address not occupied 17: other (in case of no contact) 97: not validated interview (too short) 98: not validated interview (bad quality) 99: system error
Missing code	\.

Variable number	102
Variable name	DWEIGHT_FRAME
Source	Population data
Level of observation	district
Variable label	design weights applicable to sample frame (n=8400)
Variable type	metric
Values and value labels	From 0.23 to 2.33
Missing code	none

Variable number	103
Variable name	DWEIGHT_GROSS
Source	Population data
Level of observation	district
Variable label	design weights applicable to gross sample (n=7770)
Variable type	metric
Values and value labels	From 0.28 to 2.22
Missing code	none
Variable number	104
Variable name	LITTER (interviewers' impression)
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	How common is litter or rubbish lying around the immediate area
Variable type	metric
Values and value labels	From 0 (no litter or rubbish) to 2.56 (much litter or rubbish)
Missing code	`.' (159)
Variable number	105
Variable name	SCORE_HOUSE
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Objectified score of house quality according to expert report
Variable type	metric
Values and value labels	From 1 (house in excellent state) to 4.84 (house in very bad state)
Missing code	`.' (179)

Variable number	106
Variable name	TYPE
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Type of house respondent lives in
Variable type	NOMINAL
Values and value labels	1: single unit house 2: multi-unit house (apartment)
Missing code	`.' (178)

Variable number	107
Variable name	CONSTR_YR
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Year of construction of the house
Variable type	Ordinal
Values and value labels	1: <1919 2: 1919-1945 3: 1946-1960 4: 1961-1970 5: >1970
Missing code	`.' (166)

Variable number	108
Variable name	FLOORS
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Number of floors of building
Variable type	Count
Values and value labels	1: 1 2: 2 3: 3 4: 4 5: 5 6: 6 7: 7 8: 8 9: 9 10: 10 or >10
Missing code	`.' (174)

Variable number	109
Variable name	FRONT_WIDTH
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Housefront width
Variable type	ordinal
Values and value labels	1: <4 m 2: 4-6 m 3: >6 m
Missing code	`.' (173)

Variable number	110
Variable name	GARAGE
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Presence of garage/drive
Variable type	binary
Values and value labels	0: no 1: yes
Missing code	`.' (206)

Variable number	111
Variable name	AREA
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Housing density of area
Variable type	nominal
Values and value labels	1: exclusively houses 2: mixed use of building (e.g. industry) 3: rural area
Missing code	`.' (160)

Variable number	112
Variable name	GREEN
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	Presence of green elements / plants in neighbourhood
Variable type	Ordinal / metric
Values and value labels	1: none ... 6: very much
Missing code	`.' (160)

Variable number	113
Variable name	UNOCCUP_NEGL
Source	Technical scrutiny by experts
Level of observation	Individual
Variable label	presence of unoccupied or neglected houses in neighbourhood
Variable type	Ordinal / metric
Values and value labels	1: none 2: presence of unoccupied <u>or</u> neglected houses in neighbourhood 3: presence of unoccupied <u>and</u> neglected houses in neighbourhood
Missing code	`.' (162)

Variable number	114
Variable name	RESP_CODE2
Source	Contact data
Level of observation	individual
Variable label	Final response code (long)
Variable type	Nominal
Values and value labels	1: Completed Interview with respondent 3: French-speaking 4: language barrier 5: illness, disables, demented 6: R at home, but no occasion 7: on holiday, business trip 8: deceased 9: explicit refusal 10: refusal, excuses 11: at home, but no appointment possible 12: other (in case there was contact with R) 13: not at home 14: home but nobody answered 15: address not traceable 16: address not occupied 17: other (in case of no contact) 97: not validated interview (too short) 98: not validated interview (bad quality) 99: system error
Missing code	`.` (630)



Representativity Indicators for Survey Quality

RISQ data set documentation
Slovenian ICT survey 2007

Deliverable 1.10 (WP2)

Katja Rutar
SURS

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Usage of information-communication technologies (ICT) in enterprises
Abstract:	The survey gives the information whether the enterprises use computers, the internet, electronic commerce and other ICTs. The data are EU comparable. Data was collected in the first quarter of the year 2006 via posted paper questionnaire. Enterprises had the possibility to receive and send the questionnaire via e-mail. That option was used by 10% of the enterprises that cooperated in the survey. Enterprises are legally obliged to answer to the survey but no sanctions are done if they do not respond. It is recommended that a person who is familiar with information technologies answers the questionnaire.
Topic classification:	Information and communication technologies, information society, enterprises, business statistics
Type of survey	Cross section
Unit of study	Enterprise
Target population	Enterprises registered on the territory of the Republic of Slovenia performing selected NACE classification activities with 5 or more employees.
Producer:	Statistical office of the Republic of Slovenia
Data file name:	ICT-ENTERP-2007
File contents:	Number of employees, NACE activity group, revenue, selected key ICT usage variables and response status.
File structure:	Rectangular
Number of cases:	1.998
Number of variables	9
Design weighting:	Design weighting is done but no separated design weights are included in the final data file.
Adjustment weights	Final weight which include design weight, nonresponse correction and callibration to the number of persons employed.
Imputation	Not done on selected variables
Missing data:	Missing items are set at SAS missing values.
Data collector:	Statistical office of the Republic of Slovenia
Mode of data collection:	post or e-mail questionnaires

Field work period(s):	Data collection was done in the first quarter of the year 2007.
Sampling procedure:	Stratified systematic simple random sampling. Strata are defined according to the classes by the number of persons employed and sector of activity.
Interviewers:	Not applicable
Advance information	Information letter is part of the questionnaire.
Call schedules	Three reminder letters are sent.
Incentives	No
Refusal conversion	Only with reminder letters
Response rates	

	Frequency	Percent	Valid Percent	Cumulative Percent
Response	1751	87,6%	87,6%	87,6%
Uneligible	10	0,5%	0,5%	88,1%
Nonresponse	237	11,9%	11,9%	100,0%
Total	1998	100,0%	100,0%	

Codebook of variables

Variable number	1
Variable name	Number_employees
Variable label	Number of employees in the enterprise
Variable type	Ordinal
Values and value labels	From 5 to 11.321

Variable number	2
Variable name	NACE
Variable label	NACE activity group of enterprise
Variable type	Nominal
Values and value labels	

	Frequency	Percent	Valid Percent	Cumulative Percent
Manufacturing activities DA-DE	234	11,7%	11,7%	11,7%

Manufacturing activities DF-DH	64	3,2%	3,2%	14,9%
Manufacturing activitiied DI-DJ	180	9,0%	9,0%	23,9%
Manufacturing activities DK-DN	211	10,6%	10,6%	34,5%
Construction	338	16,9%	16,9%	51,4%
Sale, repair of motorvehicles	94	4,7%	4,7%	56,1%
Wholesale trade, commission trade	225	11,3%	11,3%	67,4%
Retail trade, repair of household goods	163	8,2%	8,2%	75,5%
Hotels and restaurants	42	2,1%	2,1%	77,6%
Transport and storage	111	5,6%	5,6%	83,2%
Post and telecommunications	32	1,6%	1,6%	84,8%
Computer and related activities	47	2,4%	2,4%	87,1%
Real estate, renting	234	11,7%	11,7%	98,8%
Motion picture, video, radio and TV	23	1,2%	1,2%	100,0%
Total	1998	100,0%	100,0%	

Variable number 3
Variable name Revenue
Variable label Revenue of the enterprise in the year 2006
Variable type Ordinal
Values and value labels From 0 to 405 mio SIT

Variable number 4
Variable name LAN
Variable label Availability of LAN in the enterprise
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	1277	75,5%	75,5%	75,5%
No	415	24,5%	24,5%	100,0%
Total	1692	100,0%	100,0%	

Variable number 5
Variable name Intranet
Variable label Availability of intranet in the enterprise

Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	495	29,3%	29,3%	29,3%
No	1197	70,7%	70,7%	100,0%
Total	1692	100,0%	100,0%	

Variable number	6
Variable name	Rate_computer_use
Variable label	rate of employed using computer connected to internet at least once a week
Variable type	Ordinal
Values and value labels	From 0 to 100 %

Variable number	7
Variable name	Internet
Variable label	Does the enterpriese have an internet page?
Variable type	Nominal
Values and value labels	

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	1065	64,0%	64,0%	64,0%
No	599	36,0%	36,0%	100,0%
Total	1664	100,0%	100,0%	

Variable number	8
Variable name	IT-Experts
Variable label	Does the enterpriese employ IT experts?
Variable type	Nominal
Values and value labels	

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	430	25,4%	25,4%	25,4%
No	1262	74,6%	74,6%	100,0%
Total	1692	100.0%	100.0%	



Representativity Indicators for Survey Quality

RISQ data set documentation
Slovenian LFS survey 2007 – quarter 4

Deliverable 1.11 (WP2)

Katja Rutar
SURS

18 July, 2008



*The RISQ Project is financed by the 7th Framework Programme (FP7) of the European Union.
Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

RISQ SURVEY DATA FILE DOCUMENTATION

Title:	Labour Force Survey (LFS)
Abstract:	The Labour force survey collects employment relevant data and core demographic data. The questionnaire is ILO and EU harmonized. All members of selected household are interviewed. First time households are interviewed CAPI, in repeated waves CATI is used.
Topic classification:	Labour force survey, Labor market
Type of survey	Continuous rotating panel survey. Each household is interviewed five times according to the rotation pattern 3-1-2 (households are interviewed for three consecutive quarters, excluded for one quarter, and included for another two consecutive quarters).
Unit of study	Individual
Target population	All persons of 15 years and older living in Slovenia
Producer:	Statistical office of the Republic of Slovenia
Data file name:	LFS2007Q4
File contents:	Core demographic questions and key employment variables.
File structure:	Rectangular
Number of cases:	14690 respondents 15 years and older living in 5679 private households, 740 nonresponding households with data from previous interviews (waves), and 681 nonresponding households with only territorial data (nuts3 region and type of settlement)
Number of variables	14 selected variables
Design weighting:	Design weighting is done but no separated design weights are included in the final data file.
Adjustment weights	Final weight which include design weight, nonresponse correction and post-stratification to population totals.
Imputation	Not done on selected variables
Missing data:	No missing data allowed by core demographic questions and key employment variables.
Data collector:	Statistical office of the Republic of Slovenia
Mode of data collection:	CAPI, CATI
Field work period(s):	Quarterly samples. Data is collected during all weeks of the fourth quarter of the year 2007.

Sampling procedure:	Stratified systematic simple random sampling. Strata are defined with size and type of settlement.
Interviewers:	30 field interviewers and 12 phone interviewers.
Advance information	Advance letter to selected households
Call schedules	Blaise procedures are used
Incentives	No
Refusal conversion	No explicit refusal conversion procedures are applied

Response rates – household level

	Frequency	Percent	Valid Percent	Cumulative Percent
Respondents	5679	80,0%	80,0%	80,0%
Uneligible	72	1,0%	1,0%	81,0%
Refusals + absent	888	12,5%	12,5%	93,5%
Absent	211	3,0%	3,0%	96,5%
Unable	86	1,2%	1,2%	97,7%
Language problems	2	0,0%	0,0%	97,7%
Non-contact + uneligible	129	1,8%	1,8%	99,5%
Other	33	0,5%	0,5%	100,0%
Total	7100	100,0%	100,0%	

Codebook of variables

Variable number	1
Variable name	Gender
Variable label	gender for respondents
Variable type	Nominal
Values and value labels	

	Frequency	Percent	Valid Percent	Cumulative Percent
Male	7167	48,8%	48,8%	48,8%
Female	7523	51,2%	51,2%	100,0%
Total	14690	100,0%	100,0%	

Variable number	2
Variable name	Age
Variable label	age for respondents

Variable type	Ordinal
Values and value labels	From 15 to 98 years
Variable number	3
Variable name	Education
Variable label	education level for respondents
Variable type	Ordinal
Values and value labels	

	Frequency	Percent	Valid Percent	Cumulative Percent
Less than 4 years	146	1,0%	1,0%	1,0%
From 4 to 7 years of primary school	489	3,3%	3,3%	4,3%
Primary school	3386	23,0%	23,0%	27,4%
Lower vocational school	3635	24,7%	24,7%	52,1%
Vocational secondary school	3689	25,1%	25,1%	77,2%
General secondary school	960	6,5%	6,5%	83,8%
Post secondary vocational education	763	5,2%	5,2%	89,0%
Higher profesional education	418	2,8%	2,8%	91,8%
University degree	1032	7,0%	7,0%	98,8%
Post-graduate education	172	1,2%	1,2%	100,0%
Total	14690	100,0%	100,0%	

Variable number	4
Variable name	Citizenship
Variable label	citizenship for respondents
Variable type	Nominal
Values and value labels	

	Frequency	Percent	Valid Percent	Cumulative Percent
Slovene	14579	99,2%	99,2%	99,2%
Other (detailed citizenship available)	111	0,8%	0,8%	100,0%
Total	14690	100,0%	100,0%	

Variable number	5
Variable name	HHMembers
Variable label	number of members in the responding households

Variable type

Ordinal

Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
1	762	13,4%	13,4%	13,4%
2	1700	29,9%	29,9%	43,4%
3	1359	23,9%	23,9%	67,3%
4	1245	21,9%	21,9%	89,2%
5 +	613	10,8%	10,8%	100,0%
Total	5679	100,0%	100,0%	

Variable number

6

Variable name

Wave

Variable label

interview sequence for respondents

Variable type

Ordinal

Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
1	3924	26,7%	26,7%	26,7%
2	3130	21,3%	21,3%	48,0%
3	2696	18,4%	18,4%	66,4%
4	2564	17,5%	17,5%	83,8%
5 +	2376	16,2%	16,2%	100,0%
Total	14690	100,0%	100,0%	

Variable number

7

Variable name

Nuts3

Variable label

nuts3 region for respondents

Variable type

Nominal

Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Pomurska	992	6,8%	6,8%	6,8%
Podravska	2622	17,8%	17,8%	24,6%
Koroška	503	3,4%	3,4%	28,0%
Savinjska	1794	12,2%	12,2%	40,2%
Zasavska	355	2,4%	2,4%	42,7%
Spodnjeposavska	467	3,2%	3,2%	45,8%
Jugovzhodna Slovenija	944	6,4%	6,4%	52,3%
Osrednjeslovenska	3508	23,9%	23,9%	76,1%
Gorenjska	1630	11,1%	11,1%	87,2%
Notranjsko-kraška	373	2,5%	2,5%	89,8%

Goriška	920	6,3%	6,3%	96,0%
Obalno-kraška	582	4,0%	4,0%	100,0%
Total	14690	100,0%	100,0%	

Variable number 8
Variable name Settlement_Type
Variable label type of settlement for respondents
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Rural settlement with up to 2000 inhabitants	4259	29,0%	29,0%	29,0%
Urban settlement with up to 2000 inhabitants	3692	25,1%	25,1%	54,1%
Settlements with 2000 - 10000 inhabitants	2344	16,0%	16,0%	70,1%
Settlements with over 10000 inhabitants	1893	12,9%	12,9%	83,0%
Maribor	777	5,3%	5,3%	88,3%
Ljubljana	1725	11,7%	11,7%	100,0%
Total	14690	100,0%	100,0%	

Variable number 10
Variable name Formal_Education_Participation
Variable label participation in formal education in last 4 weeks for respondents
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	2311	15,7%	15,7%	15,7%
No	12379	84,3%	84,3%	100,0%
Total	14690	100,0%	100,0%	

Variable number 11
Variable name Nonformal_Education_Participation
Variable label participation in nonformal education in last 4 weeks for respondents
Variable type Nominal

Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	1459	9,9%	9,9%	9,9%
No	13231	90,1%	90,1%	100,0%
Total	14690	100,0%	100,0%	

Variable number 12
Variable name Official_Employment_Status
Variable label official employment status for respondents
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Employed, selfemployed	7099	48,3%	48,3%	48,3%
Unemployed	1019	6,9%	6,9%	55,3%
Student	1780	12,1%	12,1%	67,4%
Houswife	286	1,9%	1,9%	69,3%
Retired	4371	29,8%	29,8%	99,1%
Unable to work	109	0,7%	0,7%	99,8%
Others	26	0,2%	0,2%	100,0%
Total	14690	100,0%	100,0%	

Variable number 13
Variable name ILO_Employment_Status
Variable label ILO employment status for respondents
Variable type Nominal
Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Unemployed	410	2,8%	2,8%	2,8%
Employed	6743	45,9%	45,9%	48,7%
Selfemployed	889	6,1%	6,1%	54,7%
Unpaid family workers	447	3,0%	3,0%	57,8%
Student	1344	9,1%	9,1%	66,9%
Houswife	194	1,3%	1,3%	68,3%
Retired	3997	27,2%	27,2%	95,5%
Unable to work	105	0,7%	0,7%	96,2%
Other inactive	561	3,8%	3,8%	100,0%
Total	14690	100,0%	100,0%	

Variable number 13

Variable name Activity_Sector

Variable label sector of activity for employed, selfemployed and unpaid family workers

Variable type Nominal

Values and value labels

	Frequency	Percent	Valid Percent	Cumulative Percent
Agriculture	901	11,2%	11,2%	11,2%
Industry	2831	35,0%	35,0%	46,2%
Services	72	0,9%	0,9%	47,1%
Unknown	4275	52,9%	52,9%	100,0%
Total	8079	100,0%	100,0%	

Variable number 14

Variable name Hours_worked

Variable label hours usually worked per week for employed, selfemployed and unpaid family workers

Variable type Ordinal

Values and value labels From 1 to 100