



---

**Disability, Care and Participation  
Technical Report: Survey of Carers in Households  
ESRC: Secondary Data Analysis Initiative**

**ESRC 2674 – February 2017**

**Eleni Kampanellou**

---



## Contents

Introduction .....	1
STEP 1: Recoding of socio-economic variables.....	1
Starting sample .....	1
Carers - GHS .....	1
Age and gender of carer .....	1
Marital status of carer .....	2
Ownership of carer .....	2
Economic activity .....	3
Part-time vs full-time .....	4
Education .....	5
First category SCH .....	5
Second category SCH .....	6
Third category SCH.....	6
Fourth category SCH .....	6
Fifth Category SCH .....	6
Health.....	7
LLTI and Limits activities .....	8
STEP 3: Recoding of informal caring variables- carers only.....	8
Defining carers - if carer lives with dependant.....	8
Defining carers - if carer does not live with dependant .....	9
Number of dependants living with carer .....	9
Number of dependants not living with carer.....	10
STEP 4: Recoding of informal caring variables - dependants .....	10
Relationship of dependant to carer .....	11
Demographics of dependant - age and gender .....	12
If dependant lives with carer all the time- only for household members .....	13
If dependant does not live with carer all the time- only for non- household members .....	13
If dependant does not live with carer all the time- only for household members.....	14
Affected.....	14
Typology.....	15
Hours and years spent in caring.....	16
Other non-professional help.....	17
Moretime – Someone else spending more time .....	17

STEP 5: Only main cared for variables .....	18
Main cared for person regular visits (only if relative or in hh) .....	18
Main cared for person activities (only if relative or in hh) .....	19
Main cared for person spends without carer (only if relative or in hh) .....	20
Break of carer (referring only to main cared for person- in and out hh).....	20
Months and years break of carer (referring only to main cared for person) .....	21
STEP 6: Unique identifier main cared for person .....	22
Main dependant/cared for person (referring only to main cared for person).....	22
Sample reduction .....	25
Starting sample .....	25
Acknowledgements.....	26
Appendix SCH: Main and Variables .....	27
SCH- unique identifiers .....	27
SCH- extra variables .....	27

## Introduction

Likewise with GHS the first step is to select the variables of interest and then to begin coding those selected variables. The aim is to create comparable categories with the GHS. No linkage is necessary in this case since the dataset of interests includes all the relevant information. Therefore I use the dataset called carers\_main\_data.dta

## STEP 1: Recoding of socio-economic variables

The recoding process took place on the file of carers\_main\_data.dta in stata12. This includes the chosen and comparable variables found in GHS. In this case there is no need for separation of carers vs non-carers, since the dataset includes only the carers in England. However, it is not a national representative sample therefore the inclusion of weights is necessary. Likewise with the GHS I begin with the recoding of the socio-economic variables.

## Starting sample

Population: 2401 carers.

## Carers - GHS

Definition of carers according to GHS (caredef variable): The GHS definition of Carers excludes those caring as volunteers for a charity or organisation, those caring for someone in an institution, those providing financial support only and those caring for someone with a temporary illness or disability.

So at this stage I need to drop those cases that do not agree with the definition of the GHS. For that reason I use the binary variable caredef and I drop the cases with value no. The sample is now 2199 carers. There are three cases where the carer number is missing. No any other info existing in the subsequent variables. There 3 cases are deleted. So the sample is 2196 cases. Next the recoding begins.

## Age and gender of carer

The first variable to be recoded is age in 10 year intervals due to small numbers. The new variable created is cage with the label Age recoded 10. The variable sex kept its original form: 1=Male 2=Female

Variable: cage Label: Age recoded 10 Missing: 4

Value = 1 Label = 16 - 24

Value = 2 Label = 25 - 34

Value = 3	Label = 35 - 44
Value = 4	Label = 45 - 54
Value = 5	Label = 55 - 64
Value = 6	Label = 65 - 74
Value = 7	Label = 75+
Value = -9	Label = Missing – 4 cases

Original SCH: age

Value = 4	Label = 16 - 24
Value = 5	Label = 25 - 34
Value = 6	Label = 35 - 44
Value = 7	Label = 45 - 54
Value = 8	Label = 55 - 64
Value = 9	Label = 65 - 69
Value = 10	Label = 70 - 74
Value = 11	Label = 75+
Value = 13	Label = Not stated

### Marital status of carer

Variable: cmarstat (original marstat) Label: Marital Status recoded Missing: None.

Value = 1	Label = Single (1)
Value = 2	Label = Married (2 3 4)
Value = 3	Label = Divorced/Separated (5 6)
Value = 4	Label = Widowed (7)

Original SCH: Variable marstat

Value = 1	Label = Single, that is, never married?
Value = 2	Label = Living with someone in this household as a couple
Value = 3	Label = Married and living with your husband/wife?
Value = 4	Label = A civil partner in a legally-recognised Civil Partnership
Value = 5	Label = Married and separated from your husband/wife?
Value = 6	Label = Divorced?
Value = 7	Label = Or widowed?

### Ownership of carer

Variable: cq89 Label: Owner recoded Missing: 2.

Value = 1	Label = Own outright (1)
Value = 2	Label = Mortgage + shared ownership/mortgage (2)
Value = 3	Label = Pay part rent and part mortgage (shared ownership)-(3)
Value = 4	Label = Rent it/rent free (4 5)
Value = 5	Label = Squatting (6)
Value = -9	Label = Missing/Not stated (2)

Original SCH: Variable = q89

Value = 1	Label = Own outright
Value = 2	Label = Buying it with the help of a mortgage or loan
Value = 3	Label = Pay part rent and part mortgage (shared ownership)
Value = 4	Label = Rent it
Value = 5	Label = Rent-free (included relative/friend)
Value = 6	Label = Squatting
Value = 7	Label = Not stated

### Economic activity

Multiple responses of the original binary variables where someone defines himself/herself as employment but also as keeping house or going to school. For those who define themselves as students in employment training but working I have put them in the employed category. For all the other combinations recoding took place as follows:

For example: The student category also includes some who said they are retired/ill. Some employed are also retired/in school/keeping house/other inactivity. Someone defines himself as unemployed and retired/keeping house/or permanently ill. So what I did is the following:

- Category Employed: Only those who defined themselves exclusively employed or they are in training scheme for employment but working as well. It includes also those who are inactive (any type of inactivity) but also are employed. There are 3 cases of females who have given two types of inactivity: keeping house and working voluntarily (one also retired) but also said they are employed. These cases have the serial numbers: 1241110 (1), 1250219 (1) and 1390513 (1).
- Category Unemployed: Includes only the unemployed not any type of inactivity or being employed is included. Someone cannot be unemployed and inactive.
- Category School: Includes only those in education and those in a government scheme/employment training but not working. Not any other type of activity/inactivity.

- Category Illness: Includes only those being sick and those who are sick and unemployed. It does not include those who said they are sick with any other type of inactivity. The same concept is applied for the retired, keeping house and other categories.
- Category Combined inactive: Includes all the people who have indicated multiple types of inactivity more than one.

Variable: cactivity Label: Ownership recoded Missing: None

Value = 1	Label = Employees & Self-Employed (2)
Value = 2	Label = Unemployed (5 6 7)
Value = 3	Label = School or college (1 3)
Value = 4	Label = Inactive due to illness (8)
Value = 5	Label = Retired (9)
Value = 6	Label = Inactive taking care of family (10)
Value = 7	Label = Inactive other (4 11 12)
Value = 8	Label = Combined inactive (1 3 4 11 12 8 9 10)

Original SCH: Variables empstat01-13

Value = 0 Label = No - Value = 1 Label = Yes (Q5)

empstat01: Going to school or college full-time (including while on holiday)

empstat02: In paid employment or self-employed (or sick leave, holiday)

empstat03: On a government scheme for employment training

empstat04: Doing unpaid work for a business that you own, or a relative owns

empstat05: Waiting to take up paid work already obtained

empstat06: Looking for paid work or a government training scheme

empstat07: Intending to look for paid work but prevented by temporary sickness or injury

empstat08: Permanently unable to work because of long-term sickness or disability

empstat09: Retired from paid work

empstat10: Looking after home or family

empstat11: Voluntary work

empstat12: Doing something else

empstat13: Not stated- No missing values.

### Part-time vs full-time

Over 35 hours it is considered full time in SCH. Not applicable refers to unemployed/inactive who do not work. Missing refers to the employed (inactive who work) with no working hours or they said I don't know.

Variable: cftptwk, Variable label = Full vs Part time Missing:



Value = 1	Label = Part-time
Value = 2	Label = Full-time
Value = -8	Label = Not Applicable (refers to unemployed/inactive)
Value = -9	Label = Missing – 10 cases

Original SCH: Variable=ftptwk, Label = Full or Part-time work

Value = 1	Label = Part Time
Value = 2	Label = Full Time
Value = 3	Label = Don't work regular hours (goes to part-time)
Value = 4	Label = Don't know (goes to missing)
Value = -99	Label = Missing

## Education

In GHS education refers only to those who are 16-69. The q93\_ 20 and 21 have the SLC/SCE/SUPE Lower on the second category and Higher q93\_11 in the third level category. The SLC/SCE/SUPE is not divided in categories in GHS. There is one typo where the respondent replied both I don't know and level 4 (level 4 is considered- serial 1181014 (1)). Except the cq93 there is another variable named ccq93, which refers to the GHS version of education. It has an additional N/A category which included the 70+ people.

New variable: cq93, Label=Education recoded, Missing=6

Value = 1	Label = Level 1
Value = 2	Label = Level 2
Value = 3	Label = Level 3
Value = 4	Label = Level 4
Value = 5	Label = Level 5 (other qualifications)
Value = 6	Label = No qualifications
Value = -9	Label = Missing/Don't know-6

## First category SCH

q93_18	Variable label = CSE GRADES 2-5/SCE Ordinary BANDS D-E (Level 1)
q93_19	Variable label = CSE Ungraded: (Level 1)
q93_16	Variable label = GCSE GRADES D-G (Level 1)
q93_14	Variable label = O-level passes taken after 1975 GRADES D-E (Level 1)
q93_27	Variable label = NVQ Level 1/Foundation level GNVQ: (Level 1)

## Second category SCH

q93_12	Variable label = O-level passes taken in 1975 or earlier: (Level 2)
q93_13	Variable label = O-level passes taken after 1975 GRADES A-C (Level 2)
q93_17	Variable label = CSE GRADE 1/SCE BANDS A-C/Standard Grade 1-3
q93_15	Variable label = GCSE GRADES A-C (Level 2)
q93_26	Variable label = NVQ Level 2/Intermediate level GNVQ (Level 2)
q93_20	Variable label = SLC Lower:
q93_21	Variable label = SUPE Lower or Ordinary
q93_22	Variable label = School Certificate or Matric

## Third category SCH

q93_09	Variable label = A-levels/Higher School Certificate: (Level 3)
q93_25	Variable label = NVQ Level 3/Advanced level GNVQ: (Level 3)
q93_10	Variable label = AS level: (Level 3)
q93_11	Variable label = SLC/SCE/SUPE at Higher Grade or Certificate of Sixth Year

## Fourth category SCH

q93_01	Variable label = Degree/degree level qualification (including higher degree)
q93_02	Variable label = Teaching qualification (Level 4)
q93_03	Variable label = Nursing qualifications SRN, SCM, SEN, RGN, RM, RHV, Midwife: q93_04 Variable label = HNC/HND, BEC/TEC Higher, BTEC Higher/SCOTECH Higher
q93_23	Variable label = NVQ Level 5
q93_24	Variable label = NVQ Level 4

## Fifth Category SCH

q93_05	Variable label = ONC/OND/BEC/TEC/BTEC not higher:
q93_06	Variable label = City and Guilds Full Technological Certificate (Level 4)
q93_07	Variable label = City and Guilds Advanced/Final Level (Level 3)
q93_08	Variable label = City and Guilds Craft/Ordinary Level (Level 2)
q93_28	Variable label = Recognised Trade Apprenticeship completed
q93_29	Variable label = Clerical or Commercial Qualification
q93_30	Variable label = HGV Licence:
q93_31	Variable label = NNEB:
q93_32	Variable label = Other (SPECIFY)

Original SCH: Variable=Q93 Label=Highest Educational Qualification

Value = 1      Label = Degree/degree level qualification (including higher degree)

Value = 2	Label = Teaching qualification
Value = 3	Label = Nursing qualifications SRN, SCM, SEN, RGN, RM, RHV, Midwife
Value = 4	Label = HNC/HND, BEC/TEC Higher, BTEC Higher/SCOTTECH Higher
Value = 5	Label = ONC/OND/BEC/TEC/BTEC not higher
Value = 6	Label = City and Guilds Full Technological Certificate
Value = 7	Label = *City and Guilds Advanced/Final Level
Value = 8	Label = **City and Guilds Craft/Ordinary Level
Value = 9	Label = A-levels/Higher School Certificate
Value = 10	Label = AS level
Value = 11	Label = SLC/SCE/SUPE at Higher Grade or Certificate of Sixth Year St
Value = 12	Label = O-level passes taken in 1975 or earlier
Value = 13	Label = *O-level passes taken after 1975 GRADES A-C
Value = 14	Label = **O-level passes taken after 1975 GRADES D-E
Value = 15	Label = GCSE GRADES A-C
Value = 16	Label = GCSE GRADES D-G
Value = 17	Label = CSE GRADE 1/SCE BANDS A-C/Standard Grade LEVEL 1-3
Value = 18	Label = CSE GRADES 2-5/SCE Ordinary BANDS D-E
Value = 19	Label = CSE Ungraded
Value = 20	Label = SLC Lower
Value = 21	Label = SUPE Lower or Ordinary
Value = 22	Label = School Certificate or Matric
Value = 23	Label = NVQ Level 5
Value = 24	Label = NVQ Level 4
Value = 25	Label = NVQ Level 3/Advanced level GNVQ
Value = 26	Label = NVQ Level 2/Intermediate level GNVQ
Value = 27	Label = NVQ Level 1/Foundation level GNVQ
Value = 28	Label = Recognised Trade Apprenticeship completed
Value = 29	Label = Clerical or Commercial Qualification (e.g. typing/book-keepi
Value = 30	Label = HGV Licence
Value = 31	Label = NNEB
Value = 32	Label = Other (SPECIFY)
Value = 33	Label = No qualifications
Value = 34	Label = Don't know

## Health

The health variable has been coded as binary.

New variable: cq94 Label: Health recoded Missing: None.

Value = 1	Label = Good (1 2)
Value = 2	Label = Not good (3 4 5)

Value = -9    Label = Missing (6 7)- No missing values category does not exist

Original SCH: Variable = q94 Variable label = How is your health in general?

Would you say it is...

Value = 1	Label = Very good
Value = 2	Label = Fairly good
Value = 3	Label = Fair
Value = 4	Label = Fairly bad
Value = 5	Label = Very bad
Value = 6	Label = Don't know (does not exist in datasets, missing)
Value = 7	Label = Not stated (likewise)

### **LLTI and Limits activities**

Original variable:q95, Label= LLTI- No need for recoding- No missing values

Value = 1	Label = Yes
Value = 2	Label = No

New variable: cq96, Label= Limits activities recoded (old var. q96)

Value = 1	Label = Yes
Value = 2	Label = No
Value = -8	Label = Not Applicable (no missing values only NA LLTI==2)

At this point I have saved the data as socioeco.dta

## **STEP 3:    Recoding of informal caring variables- carers only**

I start by using the socioeco.dta

### **Defining carers - if carer lives with dependant**

The cdeplivin variable describes whether the dependant lives with the carer. Those who did not say whether they live with the dependant or not or were not certain (value=-99.99), they said later that he/she does not live with them.

Variable: cdeplivin Label: Dep. lives with carer Missing: None

Value = 1      Label = Yes

Value = 2      Label = No

Original SCH: deplivin Label= Dependant living with them

Value = 1      Label = Yes

Value = 2      Label = No

Value = 3      Label = Not sure (goes to No value after crosstab with dlivout)

Value = -9      Label = Not stated (-9 go to No value after crosstab with dlivout)

### **Defining carers - if carer does not live with dependant**

Variable: cdlivout Label= Dep. does not live with carer Missing: None.

Value = 1      Label = Yes

Value = 2      Label = No

Original SCH: Variable dlivout

Value = 2      Label = Yes

Value = 3      Label = No

Value = 4      Label = Not sure (goes to No value after crosstab with deplivin)

### **Number of dependants living with carer**

Variable: cdlivno Label= Number of dependants living with career

Value = 1      Label = one

Value = 2      Label = two +

Value = -8      Label = NA- No missing values

Original SCH: Variable DlivNo

Value = 2      Label = one

Value = 3      Label = two (two +)

Value = 4      Label = three (two +)

Value = 5      Label = four (two +)

Value = -99.99 Label = Missing (NA)

## Number of dependants not living with carer

Variable: cdoutno Label= Number of dependants not living with carer. Missing: none. Some overlapping cases with those replying having dependants both in and out of hh (113 cases)

Value = 1	Label = one
Value = 2	Label = two
Value = 3	Label = three +
Value = -8	Label = Not Applicable – No missing values

Original SCH: Variable doutno

Value = 2	Label = one
Value = 3	Label = two
Value = 4	Label = three
Value = 5	Label = four
Value = 6	Label = five
Value = 7	Label = six
Value = -99.99	Label = Missing

I have saved those new variables in the file named dependant.dta.

## STEP 4: Recoding of informal caring variables - dependants

I open the file dependant.dta At first I need to transform the dataset in a long format that contains only the variables about the dependants. Before I do that I create a unique identifier for the carer based on the serial (hh identifier) and person (individual identifier) numbers. This identifier is named idc.

After that I recode the variable dhelp7\_01 that contains only one value about the type of help offered by the carer and it refers to keeping an eye. However this variable name, as it was created initially, indicates personal care and that is wrong because its category is keeping an eye. For example variables dhelp1\_01 to dhelp6\_01 (dependants 1-6) include info about personal care (\_01). The keeping an eye variable is dhelp1\_09 to dhelp6\_09 (\_09), therefore the equivalent variable for the 7th dependant would normally be dhelp7\_09. It should be mentioned that this variable dhelp7\_01 (after the recoding is named dhelp7\_09 referring to keeping an eye) was recoded differently compared to the ones referring to dependants 1-6. For example there are no variables about other types of

help starting from 1-10 (10 different types of help exist in the dataset). Only one stand-alone variable exists for the dependant 7 - dhelp7\_01- which refers to only one case with serial number 1420930 and persno 1.

The next step is to rename all the variables referring to dependants in order to prepare the dataset for the long format. At this point it should be mentioned that the variable dhelp1\_12 refers to the second dependant –personal care. It has not been named properly, therefore it is renamed as dhelp12 (personal care 2<sup>nd</sup> dependant). After that, the dataset is ready for the long format. The original variable dhelp7\_01 and new one dhelp7\_09 are not included in the long format because they don't have the same categories as the other variables about help referring to dependants 1-6. They are stand-alone referring only to the 7<sup>th</sup> dependant. If they are included they are not transformed properly and it creates a mess in the data. However since they are not included in the long format, the software thinks that they are a carer variable therefore it gives value number 1 (6 more times) for the other dependants 1-6. Therefore in total there are seven values instead of one. This can be fixed easily since I know which case has the value 1/keeping an eye.

In order to do that I find the case 1420930 and I enter number 0 to the dependants 1-6 and at the same time I find the variable dhelp9 which refers to the keeping an eye help and I enter value 1 for the 7<sup>th</sup> dependant and 0 for the other types of help referring to the 7<sup>th</sup> dependant. This way the 7<sup>th</sup> dependant is integrated within the proper variable keeping an eye and there is no need for additional one. The original stand-alone and its recoded form are not needed any more.

The only last thing is that because the variable dhelp7\_09 (original dhelp7\_01) is not integrated in the long format, the software cannot produce dhelp1-dhelp10 variables for the 7<sup>th</sup> dependant. These variables referring to other types of help do not exist in the original dataset). Therefore the 7<sup>th</sup> dependant in the whole dataset has dots (.) since any info does not exist. These dots are replaced by -99.99 following the same pattern of the SCH (for other variables and dependants) when info does not exist. When I return back to the wide format this problem is solved because all the types of help for all dependants have the same structure.

### **Relationship of dependant to carer**

Using the same dataset dependant in a long format the coding begins. The number of dependants is 15372 (2196\*7). In this case I delete value 9, which is client of voluntary organization. It is better to do delete those cases in a long format because in a wide format not only the dependant who is in a voluntary organization is deleted but the whole wide row, meaning all dependants. One example is the case with the serial number 1330805, person 1 who has multiple dependants (7) and six of them are in voluntary organization and

one of them is a child. After the deletion of those six dependants I replace the values of the totaldep1=7 to the value 1, indicating one dependant, a child. Otherwise there will be confusion when I create the main cared for variable.

But after transforming all the relevant variables in a long format, now the sample is 15366 cases. Other category of the original variables goes to friend/neighbour. This other category value 10 does not exist in GHS. Also the grand-parent category goes to other relative.

Because not everyone has 7 dependants, the cases, where the relationship does not have any value, are deleted from the sample in order to get proper frequencies. The original variable of relationships in the wide format has not missing values. Now the sample is 2651 dependants

Variables: cdrelinf Label: Relationship recoded- Missing none.

Value = 1	Label = Spouse
Value = 2	Label = Own/adopted/step child/
Value = 3	Label = Parent
Value = 4	Label = Parent-in-law
Value = 5	Label = Other relative
Value = 6	Label = Friend or neighbour & other

Original SCH: Variable drelinf1- drelinf7 (for each dependant) (Q.19).

Value = 1	Label = Spouse/partner (new value 1)
Value = 2	Label = Own/adopted/step child (new value 2)
Value = 3	Label = Foster child (new value 2)
Value = 4	Label = Parent (new value 3)
Value = 5	Label = Parent-in-law (new value 4)
Value = 6	Label = Grand-parent (new value 5)
Value = 7	Label = Other relative (new value 5)
Value = 8	Label = Friend or neighbour (new value 6)
Value = 9	Label = Client of voluntary organisation (deleted)
Value = 10	Label = Other (new value 6)

## Demographics of dependant - age and gender

Variables: cdage.

Original SCH: dsex (1-7 for SCH), dage (1-7 for SCH).

No need for recoding variable dsex.



Variable: cdage Label: Dependant's age recoded

Value = 1	Label = 0-15
Value = 2	Label = 16 - 24
Value = 3	Label = 25 - 34
Value = 4	Label = 35 - 44
Value = 5	Label = 45 - 54
Value = 6	Label = 55 - 64
Value = 7	Label = 65 - 74
Value = 8	Label = 75 +
Value = -9	Label = Missing (19)-Coming from the don't know original category

### **If dependant lives with carer all the time- only for household members**

Variable: clivsew Label: If dep does live with carer- hh members

Value = 1	Label = Lives here all the time
Value = 2	Label = Sometimes lives somewhere else
Value = -8	Label = Not applicable (refers to those out of hh)-1450
Value = -9	Label = Missing - 1

Original SCH: Variable livsew1-livsew7 (only for household members) Q.26.

Value = 1	Label = Lives here all the time
Value = 2	Label = Sometimes lives somewhere else
Value = -99	Label = NA/Missing

### **If dependant does not live with carer all the time- only for non- household members**

Yes applies to the dependants living either in another hospital/residential or nursing home or no if he/she does not live in those. If no the interviewer continues up to variable moretime, otherwise it goes to next dependant or to next chapter if no dependant.

Questions: May I ask is the dependant living in hospital/residential or nursing home?

Variable: clivinst Label: If dep does not live with carer- non hh members.

Value = 1	Label = Yes
Value = 2	Label = No
Value = -8	Label = Not applicable (refers to those in hh)
Value = -9	Label = Missing-1

Original SCH: Variable livinst1- livinst7 (only for non-household members) Yes NO.

Q.25 *Does the cared for person usually live in a hospital, residential or nursing home?*

Value = 1	Label = Yes
Value = 2	Label = No
Value = -9	Label = Not stated/ Missing

### **If dependant does not live with carer all the time- only for household members**

New variable: cwhereelse

Value = 1	Label = In another private household
Value = 2	Label = Care home or hospital
Value = 3	Label = Other
Value = -8	Label = Not applicable (refers to non-hh & hh value 1 here all time)
Value = -9	Label = Missing - 1

Original SCH: Variable = whereelse1- whereelse7 (for every dependant household member)  
(Q.27)

Value = 1	Label = In another private household (new value 1)
Value = 2	Label = Boarding school/residential college (value 3)
Value = 3	Label = Residential home (new value 2)
Value = 4	Label = Nursing or care home (new value 2)
Value = 5	Label = Hospital (new value 2)
Value = 6	Label = Other institution (value 3)
Value = 7	Label = Other (value 3)
Value = -9	Label = Not stated/ Missing

### **Affected**

New variable: caffected- Refers to those in hh or out of hh if only they do not live in institution/residential home etc.

Value = 1	Label = Physically
Value = 2	Label = Mentally
Value = 3	Label = Both
Value = 4	Label = Old age
Value = 5	Label = Other
Value = -9	Label = NA – 40 (out of hh where dep livinst=yes)

Original SCH: Variable = affected1- affected7 (Q.29)

Value = 1	Label = Physically
Value = 2	Label = Mentally
Value = 3	Label = Both physically and mentally
Value = 4	Label = Old age
Value = 5	Label = Other
Value = -9	Label = Not stated/Missing

## Typology

SCH has an additional variable/category dealing with services. This does not exist in GHS. A typo exists dhelp1\_12 is the personal help of second dependant. It has been renamed as dhelp2\_01. Refers to those in hh or out of hh if only they do not live in institution/residential home etc. and those who have permanent illness.

New variables: typo

Value = 1	Label = Personal and physical caring (with or without other types)
Value = 2	Label = Personal but not physical (with or without other types)
Value = 3	Label = Physical but not personal (with or without other types)
Value = 4	Label = Other Practical help (with other types except personal or physical)
Value = 5	Label = Practical help only
Value = 6	Label = Other help (all other combinations)
Value = -8	Label = NA – 62 (22 of those refers to those with non-permanent illness, 40 to those living in institution)

**Note: Person cared for has a permanent illness disability only in SCH**

Variable = dhelp1_01-Personal care	Value = 0	Label = No	Value = 1	Label = Yes
Variable = dhelp1_02-Physical help				
Variable = dhelp1_03-Dealing with services				
Variable = dhelp1_04-Paperwork or financial matters				
Variable = dhelp1_05-Other practical help				
Variable = dhelp1_06- Keeping him/her company				
Variable = dhelp1_07-Taking him/her out				
Variable = dhelp1_08-Giving medicines				
Variable = dhelp1_09-Keeping an eye on him/her to see he/she is all right				
Variable = dhelp1_10-Other things				

## Hours and years spent in caring

Refers to those in hh or out of hh if only they do not live in institution/residential home etc.

New variable: chelphrs Label: Help in hours recoded

Value = 1	Label = 0-19 hours per week
Value = 2	Label = 20 + hours per week
Value = 3	Label = other
Value = -8	Label= NA-40 (refers to those in institution)

I have also created another variable that divides between less than 35 or more.

New variable: helphrs35 Label: Help 35 hours division

Value = 1	Label = 0-34 hours per week
Value = 2	Label = 34 + hours per week
Value = 3	Label = Varies - 20 hours or more
Value = 4	Label = other
Value = - 8	Label = NA-40 (refers to those in institution)

Original SCH: Variable = helphrs\_1- helphrs\_7 (Q.34)

Value = 1	Label = 0-9 hours per week (new value 1)
Value = 2	Label = 10-19 hours per week (new value 1)
Value = 3	Label = 20-34 hours per week (new value 2)
Value = 4	Label = 35-49 hours per week (new value 2)
Value = 5	Label = 50-99 hours per week (new value 2)
Value = 6	Label = 100 or more hours per week (new value 2)
Value = 7	Label = Varies - Under 20 hours (new value 1)
Value = 8	Label = Varies - 20 hours or more (new value 2)
Value = 9	Label = Other (new value 3)

Refers to those in hh or out of hh if only they do not live in institution/residential home etc.

New variable chelpyrs

Value = 1	Label = Less than 1 year
Value = 2	Label = 1 year, less than 3 years
Value = 3	Label = 3 years, less than 5 years
Value = 4	Label = 5 years, less than 10 years
Value = 5	Label = 10 years, less than 15 years

Value = 6	Label = 15 years, less than 20 years
Value = 7	Label = 20 years or more
Value = -8	Label = NA-40 (refers to those in institution)

Original SCH: Variable = helpyrs\_1- helpyrs\_7 (Q.35) (for everyone)

Value = 1	Label = Less than 6 months
Value = 2	Label = 6 months, less than 1 year
Value = 3	Label = 1 year, less than 3 years
Value = 4	Label = 3 years, less than 5 years
Value = 5	Label = 5 years, less than 10 years
Value = 6	Label = 10 years, less than 15 years
Value = 7	Label = 15 years, less than 20 years
Value = 8	Label = 20 years or more

### Other non-professional help

Refers to those in hh or out of hh if only they do not live in institution/residential home etc. In GHS it includes also paid helpers in this variable.

New variable: cothhelp

Value = 1	Label = Yes
Value = 2	Label = No
Value = 3	Label = Other person(s) spend equal time
Value = 4	Label = Don't know
Value = -8	Label = NA-40 (refers to those in institution)

Original SCH: Variable= othhelp1- othhelp7 (Q. 36)

Value = 1	Label = Yes
Value = 2	Label = No
Value = 3	Label = Other person(s) spend equal time
Value = 4	Label = Don't know
Value = -9	Label = Not stated

### Moretime – Someone else spending more time

It refers only to those who replied yes in othhelp-previous variable.

New variable cmoretime

Value = 1	Label = Yes
Value = 2	Label = No
Value = 3	Label = Other person(s) spend equal time
Value = 4	Label = Don't know
Value = -8	Label = NA (those with values 2-4 othhelp and those in institution)

SCH: Variable = moretime\_1- moretime\_7 (non-professional in and out household)

Value = 1	Label = Yes
Value = 2	Label = No
Value = 3	Label = Other person(s) spend equal time
Value = 4	Label = Don't know

All the recoded variables are saved with the name depenlongallcoded.dta dataset.

## **STEP 5: Only main cared for variables**

In SCH form now will include only the main cared for person and some variables refer to the main cared for person in hh or who is a relative. In order to identify the main cared for person, I need to transform the dataset in a wide format. Further explanation at the end of the report. However because I have used a more broad threshold of hours there are some differences between my definition of main dependant and the SCH (25 in total). So the NA applicable category of the subsequent variables refer to those who are friends/neighbours/ other and do not live in the same house.

### **Main cared for person regular visits (only if relative or in hh)**

Refers only to the main cared for person and only to those who are relatives or living in hh. There is one case with idc=1462 (aka serial=1290849, 1 and depno=2) who is a friend living out and carer replied to question about visits. That is not applicable.

New variable visits

Value = 1	Label = Doctor & Nurse (only those visits)
Value = 2	Label = Health visitor
Value = 3	Label = Social worker
Value = 4	Label = Home help & Meals on wheels (only those)
Value = 5	Label = Voluntary worker
Value = 6	Label = Other (& warden/chiropracist)
Value = 7	Label = More than one for the above
Value = 8	Label = I do not know/no visits

Value = -8      Label = Not applicable (refers to friends/neighbours/other not in hh)  
Value = -9      Label = Missing – 1

SCH: Variable dvis- **only for the main cared for person (Q.40)**

**Note: Only if relative or in same household**

dvis\_01- Doctor    Value = 0 Label = No - Value = 1 Label = Yes  
dvis\_02- Nurse  
dvis\_03- Health visitor  
dvis\_04- Social worker/care manager  
dvis\_05- Home help  
dvis\_06- Meals on wheels  
dvis\_07- Voluntary worker  
dvis\_08- Occupational therapist ( new value 5)  
dvis\_09- Educational Professional ( new value 5)  
dvis\_10- Specialist/nursing/palliative care ( new value 5)  
dvis\_11- Community mental health services ( new value 5)  
dvis\_12- Other professional visitor ( new value 5)  
dvis\_13- No one  
dvis\_14- I don't know

SCH: Variable dvis\_13 Label: No visits Values =0 No 1=Yes

### **Main cared for person activities (only if relative or in hh)**

New variable: activities (there are in some cases multiple responses)

Value = 1      Label = Work  
Value = 2      Label = Adult training centre or day centre  
Value = 3      Label = School or college  
Value = 4      Label = Day hospital  
Value = 5      Label = Club centre  
Value = 6      Label = Any of the above  
Value = 7      Label = None of these  
Value = -8      Label = Not applicable (refers to friends/neighbours/other not in hh)  
Value = -9      Label = Missing-1

Original SCH: Variable dgo1-dgo6 (**only for the main cared for person in home or relative**)  
(Q.43)

dgo1 - Work Value = 0 Label = No - Value = 1 Label = Yes  
dgo2 - Day Centre

- dgo3 - School or college
- dgo4 - Day hospital (maybe new value 5)
- dgo5 - Social club
- dgo6 - None of these

### **Main cared for person spends without carer (only if relative or in hh)**

A new categorical variable: hoursout (**only for the main cared for person in home or relative**)

Value = 1	Label = 0-5
Value = 2	Label = 1-5
Value = 3	Label = 6-11
Value = 4	Label = 12-17
Value = 5	Label = 18-23
Value = 6	Label = 24+
Value = -8	Label = Not applicable (Friend/neighbours/other not in hh + value 7 of variable activities-none of these)
Value = -9	Label = Missing-1 activities no hours, cannot give average

Original SCH: Variable dgohrs (Q.44) (**only for the main cared for person in home or relative**)

Value = 1	Label = None (zero)
Value = 2	Label = 1-5
Value = 3	Label = 6-11
Value = 4	Label = 12-17
Value = 5	Label = 18-23
Value = 6	Label = 24+
Value = 7	Label = Cannot give average (Goes to missing)
Value = 8	Label = Not stated

### **Break of carer (referring only to main cared for person- in and out hh)**

Note: In GHS it is only for those caring for 20 or more. More precisely it includes all the categories with values 5-8 and 10 of the original variable helphrs. All the rest are not applicable. I used also the variable resp2dys because the variable break includes only those who said yes, need someone else in variable resp2dys. Test the cbreak2 if I crosstab cbreak2 chelphrs2 and cbreak2 resp2dys.

New variable: cbreak, Missing values: none



Value = 1      Label = Yes  
Value = 2      Label = No  
Value = -8     Label = NA (refers to those with no in resp2dys)

New variable: cbreak2, Missing values: none for over 20 hours

Value = 1      Label = Yes  
Value = 2      Label = No  
Value = -8     Label = NA (refers to those with less than 20 hours and other/na  
category of variable chelphrs2 and with no in resp2dys)  
Value = -9     Label = Missing (refers to those with no hours and yes in resp2dys)

Original SCH: Variable = break

Value = 1      Label = Yes  
Value = 2      Label = No  
Value = -99.99 Label = Not stated

### **Months and years break of carer (referring only to main cared for person)**

**Note:** *This section refers to those caring for 20 hours or more in the GHS and the main cared for people in SCH. Make this comparable*

This is a comparable variable that takes into account more than 20 hours of caring and for those only who had a break

New variable: cmonbreak cyearbreak

Value = 1      Label = 1 month  
Value = 2      Label = 2 months  
Value = 3      Label = 3 months  
Value = 4      Label = 4 months  
Value = 5      Label = 5 months  
Value = 6      Label = 6 months  
Value = 7      Label = 7 months  
Value = 8      Label = 8 months  
Value = 9      Label = 9 months  
Value = 10     Label = 10 months  
Value = 11     Label = 11 months  
Value = -8     Label = NA (<20 hours, said no break or said break >11 months)  
Value = -9     Label = Not stated/None

Value = 1	Label = 1 year
Value = 2	Label = 2 year
Value = 3	Label = 3 year
Value = 4	Label = 4 year
Value = 5	Label = 5 year
Value = 6	Label = 6 year
Value = 7	Label = 7 years and more
Value = -8	Label = NA (<20 hours, said no break or said break <1 year)
Value = -9	Label = Not stated/None

## **STEP 6: Unique identifier main cared for person**

### **Main dependant/cared for person (referring only to main cared for person)**

In order to construct this variable multiple steps have taken place. I have used the file `depenlongallcoded.dta`

#### **First step: Unique identifiers for dependant and carer**

Create a unique id for carer named: `idc` based on `hserno` and `person`.

Create a dependant id name `depno`.

#### **Second step: Create of a variable where dependant lives-cnow**

Value = 1	Label = Deps in and out HH
Value = 2	Label = Deps in HH only
Value = 3	Label = Deps out HH only

#### **Third step: Creating of variable - number and location of dependants**

This variable describes the number of dependants whether they live in and/out of hh. For that I used the variables `depno` (total number of dependants) and `cnow` (if current carer)

Variable `deplocus`

Value = 1	Label = Only one (in or out)
Value = 2	Label = 2 and more in hh
Value = 3	Label = 2 and more out hh
Value = 4	Label = 2 and more in and out

#### Fourth step: Recoding of variable nhours – more and less than 20

Variable chelphrs2 Label: Help in hours second recoded

Value = 1	Label = 0-19 hours
Value = 2	Label = 20 +
Value = -8	Label = NA (in institution and other category)

#### Fifth step: Creating variable– Location number of dependants and hours

This variable describes the number of dependants whether they live in and/out of hh and number of hours. This variable is based on the previous variable deplocus and the variable chelphrs2. Save the dataset as mainlong.dta – folder named coded.

Variable: hourslocus

Value = 1	Label = Main only one
Value = 2	Label = Two + in hh <20
Value = 3	Label = Two + in hh >20
Value = 4	Label = Two + out hh <20
Value = 5	Label = Two + out hh >20
Value = 6	Label = Two + in and out hh <20
Value = 7	Label = Two + in and out hh >20
Value = -9	Label = Missing -57

#### Sixth step: Transform the dataset in wide format and binary variables

Using the mainlong.dta file the next step is to transform the dataset into a wide format in order to identify the main cared for person. For that reason I created additional binary variables based on the categories of the above variable hourslocus. In a wide format, where all dependants are in a row, there are 2196 carers.

Variable single: 1=yes, 0=no	Label=main
Variable inless20: 1=yes, 0=no	Label=Two + in hh <20
Variable inmore20: 1=yes, 0=no	Label=Two + in hh >20
Variable outless20: 1=yes, 0=no	Label=Two + out hh <20
Variable outmore20: 1=yes, 0=no	Label=Two + out hh >20
Variable inoutless20: 1=yes, 0=no	Label= Two + in and out hh <20
Variable inoutmore20: 1=yes, 0=no	Label=Two + in and out >20

### Seventh step: In a wide format a new variable with restrictions

Staying in the wide format the next step is to create a new variable based on the restrictions of the SCH definition of cared for person. These restrictions agree with the GHS definition.

The following restrictions are:

1. If one dependant keep him/her as main whether in/out of hh.
2. If more than one keep the one with more hours (in hh or out hh or both in/out dep).
3. If more than one with same hours keep the one in hh (carers having both in/out dependant).
4. If more than one with same hours and in hh keep the first one.
5. If more than one and all live out of hh keep the one with more hours.
6. If more than one and all live out of hh and have same hours keep the first one.

This leads to the creation of a new variable called main in order to identify those different conditions. The above binary ones were used for this. The values are similar with the variable hourslocus by adding 4 additional categories: Those refer to the cases where there are >20 and <20 hours in and out, only in, only out and a final one where hours are unknown and dependants are out. Those extra values except the final one with missing hours refer to the condition number 2 where if there are more than one dependant keep the one with more hours. This file is saved as mainforwide.dta

Variable main Label = Identify conditions

Value = 1	Label = Main only one
Value = 2	Label = Two + in and out hh <20
Value = 3	Label = Two + in and out hh >20
Value = 4	Label = In and out >20 & <20
Value = 5	Label = Two + in hh <20
Value = 6	Label = Two + in hh >20
Value = 7	Label = In >20 & <20
Value = 8	Label = Two + out hh <20
Value = 9	Label = Two + out hh >20
Value = 10	Label = Out >20 & <20
Value = 11	Label = Out unknown hours

### Eighth step: In a long format final main cared for variable

Using the file mainforwide.dta I reshape the dataset into a long format in order avoid having numbers (for every dependant) attached to every variable. Since not everyone has seven

dependants I delete the dots using the complete variable cdrelinf. There are 2651 dependants. A final variable is created based on the previous variable main as well as the variables totaldep1 (number of dependants), nlive sew (where dependant in hh lives) and chelphrs2 (hours of help). Every category corresponds to the restrictions of the SCH survey. The variable is called final. These changes are saved in the finalSCH.dta

Variable final Label = Main cared for person

Value = 1	Label = Main only one
Value = 2	Label = Keep in hh Two + in and out hh <20
Value = 3	Label = Keep in hh Two + in and out hh >20
Value = 4	Label = More Hours- In and out >20 & <20
Value = 5	Label = First Two + in hh <20
Value = 6	Label = First Two + in hh >20
Value = 7	Label = More Hours In >20 & <20
Value = 8	Label = First Two + out hh <20
Value = 9	Label = First Two + out hh >20
Value = 10	Label = More Hours Out >20 & <20
Value = 11	Label = Out unknown hours-2 cases

In the last category because the hours are unknown for every cared for person (for those who have more than one cared for person) and I deleted those like I have done with the GHS (2 cases). There are some missing hours for those who have only one cared for person. These cases are kept. There is also one dependant with temporal illness. The idc number is 2149. This one is deleted, so there are 2193 carers. Sometimes the unknown hours exist in the second dependant than in the first so the mdeprno variable was useful in order to select the dependant with no missing hours as main. Because I have used a more broad range of hours there are 25 cases that do not agree with the main cared for person variable in SCH. Need to mention that all the variables related to the main dependant exist in the level of carer not the level of dependant.

## Sample reduction

### Starting sample

Population: 2401- carers.

Only carers by GHS: 2199.

Care with no missing number: 2196.

Delete cases with missing hours in final variable: 2194.

Delete a case with a dependant with temporal illness: 2193.

## Acknowledgements

Data for analysis of the SCH were from:

GfK NOP, Information Centre for Health and Social Care. (2011). *Survey of Carers in Households, 2009-2010*. [data collection]. UK Data Service. SN: 6768, <http://doi.org/10.5255/UKDA-SN-6768-1>

We acknowledge with thanks the originators of the data sets and those who made them available for secondary analysis via the UK Data Service: [www.ukdataservice.ac.uk](http://www.ukdataservice.ac.uk)

## Appendix SCH: Main and Variables

### SCH- unique identifiers

SCH: Variable=serial, Label = Serial- Household identifier

SCH: Variable=persno, Label= Person number- Person identifier (1,2,3,4)

SCH: Variable=DpersN (but not appearing in dataset only in documentation)

SCH: Variable=mdepno Label = Main Cared For Person (only for the main cared for)

SCH: Variable=caredef Label=Carer-GHS Definition

SCH: Variable=weight Label = Weight

### SCH- extra variables

SCH: Total Number of Dependants: Variable totdep1

SCH: Total number of people in the household: Variable hh\_total

SCH dillness1\_01- dillness1\_11 dillness1- dillness7

SCH dillchk1- dillchk7(temporary vs permanent illness).

SCH dfinchk\_1- dfinchk\_7 (in the case of temporal illness financial help or not)