



```

-----
      name: <unnamed>
      log:  E:\nhats\nhats_code\NHATS data setup\logs\5_nhats_cleaning_21_Mar_2
> 019.smcl
      log type:  text
      opened on:  21 Mar 2019, 10:02:36

1 .
2 .
3 .
4 .
5 . *****
6 .
7 . cd "${work}"
   E:\nhats\data\NHATS working data

8 . *****
9 . //get all 3 waves helper imputed hours into a single dataset
10.
11. use R1_hrs_imputed_added.dta

12.
13. local keepvars spid wave opid op_hrs_mth_i numact yearnotmonth disab ///
   > justone spouse otherrel nonrel regular oneact impute_cat opishelper_allw paid
   > ///
   > op_adl* op_iadl* op_nonadl* son daughter oth_fam rc_staff op_doc_help op_sn o
   > p_age

14.
15. forvalues w=2/$w{
   2. qui append using R`w'_hrs_imputed_added.dta
   3.
16. }

17.
18. tab wave opishelper_allw, missing

      Survey |      opishelper_allw
      wave |      -1      1 |      Total
-----+-----+-----+-----
      1 | 26,500 11,597 | 38,097
      2 | 28,120  9,824 | 37,944
      3 | 27,332  8,413 | 35,745
      4 | 25,931  7,085 | 33,016
      5 | 39,861 12,424 | 52,285
      6 | 40,006 11,168 | 51,174
      7 | 39,533  9,532 | 49,065
-----+-----+-----
    Total | 227,283 70,043 | 297,326

19.
20. // age categories
21. egen op_age=rowmax(dage)

22. label val op_age dage

23. replace op_age=. if inlist(op_age, -8, -7, -1, .)
    (144,342 real changes made, 144,342 to missing)

```

```

24.
25.
26.
27. //identify paid helpers and staff at residence (not helpers)
28. gen paid=0

29. gen rc_staff=0

30. forvalues w=1/$w {
    2.         replace paid=1 if paidh==1 & ((relat>=30 & relat<=40)| ///
>         relat==92) & wave=='w'
    3.         foreach x of varlist hlp *hlp1 *hlp2 *hlpms {
    4.             replace rc_staff=1 if relat==37 & `x'==1 & wave=='w'
    5.         }
    6.     }
(859 real changes made)
(0 real changes made)
(55 real changes made)
(7 real changes made)
(25 real changes made)
(894 real changes made)
(0 real changes made)
(5 real changes made)
(1 real change made)
(25 real changes made)
(369 real changes made)
(0 real changes made)
(3 real changes made)
(0 real changes made)
(27 real changes made)
(406 real changes made)
(0 real changes made)
(2 real changes made)
(0 real changes made)
(31 real changes made)
(1,221 real changes made)
(0 real changes made)
(4 real changes made)
(0 real changes made)
(30 real changes made)
(532 real changes made)
(0 real changes made)
(10 real changes made)
(3 real changes made)
(34 real changes made)
(1,000 real changes made)
(0 real changes made)
(7 real changes made)
(0 real changes made)
(29 real changes made)

31.
32. gen op_sn=0

33. forvalues w=1/$w {
    2.         replace op_sn=socl==1 if wave=='w'
    3.     }
(13,406 real changes made)
(10,743 real changes made)
(9,445 real changes made)
(7,964 real changes made)
(15,335 real changes made)
(13,196 real changes made)
(11,735 real changes made)

```

```

34. label var op_sn "OP in SP's social network"
35.
36. keep `keepvars'
37.
38. replace opishelper_allw=0 if opishelper_allw==-1
    (227,283 real changes made)
39. replace op_hrsnth_i=. if op_hrsnth_i==-1
    (227,283 real changes made, 227,283 to missing)
40. *****
41. //identify primary helper, person with the most hours / month using imputed ho
    > urs
42. sort spid wave opid spouse
43.
44. gen primary_cg=0
45. egen max_hrs_helped=max(op_hrsnth_i) if opishelper_allw==1, by(spid wave)
    (227283 missing values generated)
46. replace primary_cg=1 if op_hrsnth_i==max_hrs_helped & opishelper_allw==1
    (37,829 real changes made)
47. tab primary_cg wave, missing

```

primary_cg	Survey wave					Total
	1	2	3	4	5	
0	31,264	32,449	31,325	29,322	45,616	259,497
1	6,833	5,495	4,420	3,694	6,669	37,829
Total	38,097	37,944	35,745	33,016	52,285	297,326

primary_cg	Survey wave		Total
	6	7	
0	45,423	44,098	259,497
1	5,751	4,967	37,829
Total	51,174	49,065	297,326

```

48.
49. //check that each sp has 1 primary cg identified, doesn't matter!
50. egen num_primary_cg=total(primary_cg), by(spid wave)
51. tab num_primary_cg wave, missing

```

num_prim y_cg	Survey wave					Total
	1	2	3	4	5	
0	4,738	6,594	6,729	6,210	8,886	52,232
1	31,436	29,780	27,771	25,671	41,756	234,234
2	1,575	1,370	1,116	1,027	1,461	9,653
3	258	169	119	78	159	956
4	61	23	10	22	14	174
5	12	0	0	8	0	43
6	17	0	0	0	0	17
7	0	8	0	0	0	8
9	0	0	0	0	9	9
Total	38,097	37,944	35,745	33,016	52,285	297,326

num_primary_cg	Survey wave		Total
	6	7	
0	9,284	9,791	52,232
1	40,198	37,622	234,234
2	1,574	1,530	9,653
3	93	80	956
4	12	32	174
5	13	10	43
6	0	0	17
7	0	0	8
9	0	0	9
Total	51,174	49,065	297,326

```

52.
53. //assign relationship to primary helper
54. //if same number of hours for two different helpers, then assign to spouse fir
   > st
55. gen prim_helper_cat=1 if primary_cg==1 & spouse==1
   (280,849 missing values generated)

56. replace prim_helper_cat=2 if primary_cg==1 & otherrel==1 & prim_helper_cat==.
   (15,620 real changes made)

57. replace prim_helper_cat=3 if prim_helper_cat==.
   (265,229 real changes made)

58. la def helpcat 1"Spouse" 2"Other relative" 3"Other, not relative"

59. la val prim_helper_cat helpcat

60. tab prim_helper_cat wave, missing

```

prim_helper_cat	Survey wave				Total
	1	2	3	4	
Spouse	3,122	2,386	1,900	1,502	16,477
Other relative	2,845	2,278	1,812	1,578	15,620
Other, not relative	32,130	33,280	32,033	29,936	265,229
Total	38,097	37,944	35,745	33,016	297,326

prim_helper_cat	Survey wave			Total
	5	6	7	
Spouse	2,991	2,449	2,127	16,477
Other relative	2,668	2,368	2,071	15,620
Other, not relative	46,626	46,357	44,867	265,229
Total	52,285	51,174	49,065	297,326

```

61.
62. //get categorical relationship
63. gen op_relationship_cat=6

64. local i=1

```

```

65. foreach x in spouse daughter son oth_fam paid {
    2.         replace op_relationship_cat=`i' if `x'==1
    3.         local i=`i'+1
    4.     }
    (20,932 real changes made)
    (70,229 real changes made)
    (68,581 real changes made)
    (62,715 real changes made)
    (5,281 real changes made)

66. label define op_relationship_cat 1 "Spouse/Partner" 2 "Daughter (incl. step-/-
> in-law)" ///
> 3 "Son (incl. step-/-in-law)" 4 "Other family" 5 "Paid CG" 6 "Unpaid non-famil
> y CG"

67. label values op_relationship_cat op_relationship_cat

68.
69. //get total number hours/month help received and number of helpers per spid/wa
> ve
70. egen tot_hrsmonth_help_i=total(op_hrsmonth_i), by(spid wave)

71. egen n_helpers=count(op_hrsmonth_i), by(spid wave)

72. gen tot_hrswk_help_i=tot_hrsmonth_help_i/(52/12)

73. //family caregivers
74. gen op_family=op_relationship_cat<=4 if opishelper==1
    (227,283 missing values generated)

75. label var op_family "OP is family helper"

76. egen n_family_helpers=total(op_family),by(spid wave)

77. label var n_family_helpers "Number family helpers"

78. gen ind_family_helper=n_family_helpers>0

79. label var ind_family_helper "Indicatory any family helpers"

80. //paid caregivers
81. egen n_paid_helpers=total(paid), by(spid wave)

82. gen ind_paid_helper=n_paid_helpers>0

83. label var n_paid_helpers "Number paid helpers"

84. label var ind_paid_helper "Indicator any paid helpers"

85. egen ind_rc_staff=max(rc_staff), by(spid wave)

86. label var ind_rc_staff "Indicator Facility Staff in OP file"

87. gen num_helpers_cat=0

88. replace num_helpers_cat=1 if n_helpers>0 & n_helpers<.
    (245,094 real changes made)

89. replace num_helpers_cat=2 if n_helpers>3 & n_helpers<.
    (32,590 real changes made)

```

```

90. replace num_helpers_cat=3 if n_helpers>=7 & n_helpers<.
    (2,745 real changes made)

91. label define num_helpers_cat 0 "No helpers" 1 "1-3 helpers" 2 "4-6 helpers" //
    > /
    > 3 "7+ helpers"

92. label values num_helpers_cat num_helpers_cat

93.
94. gen phrs=op_hrsmonth_i if paid==1
    (292,045 missing values generated)

95. egen tot_hrsmonth_paid_i=total(phrs), by(sp1d wave)

96. gen tot_hrswk_paid_i=tot_hrsmonth_paid_i/(52/12)

97. label var tot_hrsmonth_paid_i "Hours/month help recived, paid helpers, imputed"
    > "

98. label var tot_hrswk_paid_i "Hours/week help received, paid helpers, imputed"

99. gen op_hrswk_i=op_hrsmonth_i/(52/12)
    (227,283 missing values generated)

100 label var op_hrswk_i "OP hours/week help SP"

101
102 la var tot_hrsmonth_help_i "Hours / month help received, all helpers, imputed"

103 sum tot_hrsmonth_help_i, detail

      Hours / month help received, all helpers, imputed
-----
      Percentiles      Smallest
1%              0              0
5%              0              0
10%             0              0      Obs          297,326
25%             4              0      Sum of Wgt.    297,326

50%            27.5
75%             90      Largest      Mean          93.41678
90%            242.48      2661.06      Std. Dev.    185.5261
95%             480      2661.06      Variance     34419.93
99%            834.1727      2661.06      Skewness     3.869629
                                Kurtosis     23.51947

104 la var tot_hrswk_help_i "Hours / week help received, all helpers, imputed"

105
106 la var n_helpers "Number helpers reported by SP"

107 tab n_helpers wave, missing

```

Number helpers reported by SP	Survey wave					Total
	1	2	3	4	5	
0	4,738	6,594	6,729	6,210	8,886	52,232
1	15,424	13,403	11,036	10,098	17,418	97,790
2	9,665	9,295	9,107	8,489	13,100	74,830
3	4,659	4,930	4,928	4,405	6,895	39,884
4	2,044	2,071	2,131	2,165	3,400	18,128
5	892	1,026	1,091	977	1,462	8,559
6	356	372	466	342	470	3,158
7	203	97	148	121	335	1,492
8	73	64	44	64	125	568
9	20	59	65	41	126	412
10	10	22	0	48	40	120
11	0	11	0	41	16	85
12	13	0	0	0	12	37

13	0	0	0	15	0	31
Total	38,097	37,944	35,745	33,016	52,285	297,326

Number helpers reported by SP	Survey wave		Total
	6	7	
0	9,284	9,791	52,232
1	15,620	14,791	97,790
2	12,869	12,305	74,830
3	7,244	6,823	39,884
4	3,310	3,007	18,128
5	1,694	1,417	8,559
6	620	532	3,158
7	375	213	1,492
8	87	111	568
9	71	30	412
10	0	0	120
11	0	17	85
12	0	12	37
13	0	16	31
Total	51,174	49,065	297,326

108 la var prim\_helper\_cat "Primary helper relationship, missing if no helper"

109 tab prim\_helper\_cat wave, missing

Primary helper relationship, missing if no helper	Survey wave				Total
	1	2	3	4	
Spouse	3,122	2,386	1,900	1,502	16,477
Other relative	2,845	2,278	1,812	1,578	15,620
Other, not relative	32,130	33,280	32,033	29,936	265,229
Total	38,097	37,944	35,745	33,016	297,326

Primary helper relationship, missing if no helper	Survey wave			Total
	5	6	7	
Spouse	2,991	2,449	2,127	16,477
Other relative	2,668	2,368	2,071	15,620
Other, not relative	46,626	46,357	44,867	265,229
Total	52,285	51,174	49,065	297,326

110

111 la var tot\_hrsmonth\_help\_i "Hours / month help received, all helpers, imputed"

112 sum tot\_hrsmonth\_help\_i, detail

Hours / month help received, all helpers, imputed				
Percentiles	Smallest			
1%	0	0		
5%	0	0		
10%	0	0	Obs	297,326
25%	4	0	Sum of Wgt.	297,326

50%	27.5		Mean	93.41678
		Largest	Std. Dev.	185.5261
75%	90	2661.06		
90%	242.48	2661.06	Variance	34419.93
95%	480	2661.06	Skewness	3.869629
99%	834.1727	2661.06	Kurtosis	23.51947

113 la var tot\_hrswk\_help\_i "Hours / week help received, all helpers, imputed"

114

115 la var n\_helpers "Number helpers reported by SP"

116 tab n\_helpers wave, missing

Number helpers reported by SP	Survey wave					Total
	1	2	3	4	5	
0	4,738	6,594	6,729	6,210	8,886	52,232
1	15,424	13,403	11,036	10,098	17,418	97,790
2	9,665	9,295	9,107	8,489	13,100	74,830
3	4,659	4,930	4,928	4,405	6,895	39,884
4	2,044	2,071	2,131	2,165	3,400	18,128
5	892	1,026	1,091	977	1,462	8,559
6	356	372	466	342	470	3,158
7	203	97	148	121	335	1,492
8	73	64	44	64	125	568
9	20	59	65	41	126	412
10	10	22	0	48	40	120
11	0	11	0	41	16	85
12	13	0	0	0	12	37
13	0	0	0	15	0	31
Total	38,097	37,944	35,745	33,016	52,285	297,326

Number helpers reported by SP	Survey wave		Total
	6	7	
0	9,284	9,791	52,232
1	15,620	14,791	97,790
2	12,869	12,305	74,830
3	7,244	6,823	39,884
4	3,310	3,007	18,128
5	1,694	1,417	8,559
6	620	532	3,158
7	375	213	1,492
8	87	111	568
9	71	30	412
10	0	0	120
11	0	17	85
12	0	12	37
13	0	16	31
Total	51,174	49,065	297,326



```
117 la var prim_helper_cat "Primary helper relationship, missing if no helper"
```

```
118 tab prim_helper_cat wave, missing
```

Primary helper relationship, missing if no helper	1	2	3	4	Total
Spouse	3,122	2,386	1,900	1,502	16,477
Other relative	2,845	2,278	1,812	1,578	15,620
Other, not relative	32,130	33,280	32,033	29,936	265,229
Total	38,097	37,944	35,745	33,016	297,326

Primary helper relationship, missing if no helper	5	6	7	Total
Spouse	2,991	2,449	2,127	16,477
Other relative	2,668	2,368	2,071	15,620
Other, not relative	46,626	46,357	44,867	265,229
Total	52,285	51,174	49,065	297,326

```
119
```

```
120
```

```
121 //number in social network
```

```
122 by spid wave, sort: egen n_sn=total(op_sn)
```

```
123 replace n_sn=0 if missing(n_sn)
```

```
(0 real changes made)
```

```
124 label var n_sn "Number in social network"
```

```
125
```

```
126 //save as op file before dropping down to one obs per spid
```

```
127 saveold "E:/nhats/data/NHATS cleaned/op_round_1_$w.dta", replace version(12)
```

```
(saving in Stata 12 format, which can be read by Stata 11 or 12)
```

```
file E:/nhats/data/NHATS cleaned/op_round_1_7.dta saved
```

```
128
```

```
129 //keep the entry for the primary caregiver only, this will then be the
```

```
130 //information merged into the SP dataset
```

```
131 keep if primary_cg==1
```

```
(259,497 observations deleted)
```

```
132
```

```
133 //sort and keep only first primary cg, again preference given if spouse
```

```
134 gsort spid wave -spouse otherrel opid
```

```
135 duplicates drop spid wave, force
```

```
Duplicates in terms of spid wave
```

```
(1,662 observations deleted)
```

```

136
137 //now one obs per sp - wave
138
139 sum tot_hrsmonth_help_i, detail

```

Hours / month help received, all helpers, imputed					
Percentiles		Smallest			
1%	0	0			
5%	2	0			
10%	4	0	Obs	36,167	
25%	12.67796	0	Sum of Wgt.	36,167	
50%	36.17251		Mean	108.6939	
		Largest	Std. Dev.	192.8727	
75%	103.92	2290.57			
90%	273	2537	Variance	37199.88	
95%	621.97	2581.12	Skewness	3.511812	
99%	836	2661.06	Kurtosis	19.63268	

```

140 tab n_helpers wave, missing

```

Number helpers reported by SP	Survey wave					Total
	1	2	3	4	5	
1	3,356	2,526	1,856	1,529	2,982	16,670
2	1,817	1,528	1,326	1,115	1,931	10,927
3	764	733	627	516	889	5,104
4	302	266	245	227	379	2,062
5	123	113	116	94	146	876
6	39	38	47	33	44	301
7	22	10	15	10	27	130
8	8	5	4	5	10	46
9	2	5	5	3	9	32
10	1	2	0	3	3	9
11	0	1	0	2	1	5
12	1	0	0	0	1	3
13	0	0	0	1	0	2
Total	6,435	5,227	4,241	3,538	6,422	36,167

Number helpers reported by SP	Survey wave		Total
	6	7	
1	2,374	2,047	16,670
2	1,718	1,492	10,927
3	841	734	5,104
4	347	296	2,062
5	161	123	876
6	55	45	301
7	30	16	130
8	6	8	46
9	6	2	32
10	0	0	9
11	0	1	5
12	0	1	3
13	0	1	2
Total	5,538	4,766	36,167

141 tab prim\_helper\_cat wave, missing

Primary helper   relationship,   missing if no   helper	1	2	3	4	Total
Spouse	3,122	2,386	1,900	1,502	16,477
Other relative	2,498	2,049	1,649	1,449	14,195
Other, not relative	815	792	692	587	5,495
Total	6,435	5,227	4,241	3,538	36,167

Primary helper   relationship,   missing if no   helper	5	6	7	Total
Spouse	2,991	2,449	2,127	16,477
Other relative	2,468	2,185	1,897	14,195
Other, not relative	963	904	742	5,495
Total	6,422	5,538	4,766	36,167

142 \*\*\*\*\*

143 //merge this helper data with the main SP dataset

144 save helper\_by\_sp.dta, replace  
file helper\_by\_sp.dta saved

145

146 use round\_1\_\$w\_cleanv3, clear

147

148 merge 1:1 spid wave using helper\_by\_sp.dta, keepusing(opid n\_helpers n\_fam ind  
> \_fam ///  
> \_tot\_hrsmonth\_\* prim\_helper\_cat ind\_pa n\_pai tot\_hrswk\* num\_helpers\_cat ind\_rc  
> \_\*)

Result	# of obs.	
not matched	11,611	
from master	11,611	(_merge==1)
from using	0	(_merge==2)
matched	36,167	(_merge==3)

149

150 drop \_merge

151

152 foreach x in n\_helpers ind\_pa n\_paid num\_helpers\_cat(  
    2. replace `x'=0 if `x'==.  
    3. )  
(11,611 real changes made)  
(11,611 real changes made)  
(11,611 real changes made)  
(11,611 real changes made)

```

153
154 gen ind_no_helpers=.
    (47,778 missing values generated)

155 replace ind_no_helpers=1 if n_helpers==0
    (11,611 real changes made)

156 replace ind_no_helpers=0 if n_helpers>0
    (36,167 real changes made)

157 la var ind_no_helpers "Indicator no Helpers reported"

158 tab ind_no_helpers wave, missing

```

Indicator   no Helpers   reported	1	2	Survey wave 3	4	5	Total
0	6,435	5,227	4,241	3,538	6,422	36,167
1	1,810	1,848	1,558	1,199	1,912	11,611
Total	8,245	7,075	5,799	4,737	8,334	47,778

Indicator   no Helpers   reported	Survey wave 6	7	Total
0	5,538	4,766	36,167
1	1,738	1,546	11,611
Total	7,276	6,312	47,778

```

159
160 la var tot_hrsmonth_help_i "Hours / month help received, all helpers, imputed"

161 sum tot_hrsmonth_help_i, detail

```

Hours / month help received, all helpers, imputed				
Percentiles		Smallest		
1%	0	0		
5%	2	0		
10%	4	0	Obs	36,167
25%	12.67796	0	Sum of Wgt.	36,167
50%	36.17251		Mean	108.6939
		Largest	Std. Dev.	192.8727
75%	103.92	2290.57		
90%	273	2537	Variance	37199.88
95%	621.97	2581.12	Skewness	3.511812
99%	836	2661.06	Kurtosis	19.63268

```

162 la var tot_hrswk_help_i "Hours / week help received, all helpers, imputed"

163
164 la var n_helpers "Number helpers reported by SP"

165 tab n_helpers wave, missing

```

Number helpers reported by SP	Survey wave					Total
	1	2	3	4	5	
0	1,810	1,848	1,558	1,199	1,912	11,611
1	3,356	2,526	1,856	1,529	2,982	16,670
2	1,817	1,528	1,326	1,115	1,931	10,927
3	764	733	627	516	889	5,104
4	302	266	245	227	379	2,062
5	123	113	116	94	146	876
6	39	38	47	33	44	301
7	22	10	15	10	27	130
8	8	5	4	5	10	46
9	2	5	5	3	9	32
10	1	2	0	3	3	9
11	0	1	0	2	1	5
12	1	0	0	0	1	3
13	0	0	0	1	0	2
Total	8,245	7,075	5,799	4,737	8,334	47,778

Number helpers reported by SP	Survey wave		Total
	6	7	
0	1,738	1,546	11,611
1	2,374	2,047	16,670
2	1,718	1,492	10,927
3	841	734	5,104
4	347	296	2,062
5	161	123	876
6	55	45	301
7	30	16	130
8	6	8	46
9	6	2	32
10	0	0	9
11	0	1	5
12	0	1	3
13	0	1	2
Total	7,276	6,312	47,778

166 la var prim\_helper\_cat "Primary helper relationship, missing if no helper"

167 tab prim\_helper\_cat wave, missing

Primary helper relationship, missing if no helper	Survey wave				Total
	1	2	3	4	
Spouse	3,122	2,386	1,900	1,502	16,477
Other relative	2,498	2,049	1,649	1,449	14,195
Other, not relative	815	792	692	587	5,495
.	1,810	1,848	1,558	1,199	11,611
Total	8,245	7,075	5,799	4,737	47,778

Primary helper   relationship,   missing if no   helper	Survey wave			Total
	5	6	7	
Spouse	2,991	2,449	2,127	16,477
Other relative	2,468	2,185	1,897	14,195
Other, not relative	963	904	742	5,495
.	1,912	1,738	1,546	11,611
Total	8,334	7,276	6,312	47,778

```

168
169 rename opid prim_helper_opid

170 deststring prim_helper_opid, replace
    prim_helper_opid: all characters numeric; replaced as byte
    (11611 missing values generated)

171 la var prim_helper_opid "Primary helper OPID"

172
173 *****
174 //save the dataset with helper information added
175
176 save round_1_$w_clean_helper_added.dta, replace
    file round_1_.dta saved

177
178 saveold round_1_$w_clean_helper_added_old.dta, replace
    (saving in Stata 13 format)
    (FYI, saveold has options version(12) and version(11) that write files in
    older Stata formats)
    note: variable label "R2 F SS DEAF PRIOR OR CURRENT ROUND " contains unicode
    and thus may not display well in Stata 13.
    file round_1_.dta saved

179
180 *****
181 log close
    name: <unnamed>
    log: E:\nhats\nhats_code\NHATS data setup\logs\5_nhats_cleaning_21_Mar_2
> 019.smcl
    log type: text
    closed on: 21 Mar 2019, 10:02:56
-----

```