

Understanding Society User Support - Support #1408

Family structure

09/11/2020 05:31 PM - Theocharis Kromydas

Status:	Resolved	Start date:	09/11/2020
Priority:	Normal	% Done:	100%
Assignee:			
Category:	Data documentation		
Description Hi there I am trying to construct a family structure variable composed by 4 categories: Single, no children, Coupled no children, Coupled with children and lone parents using the hhtype_dv, nkids_dv marstat and nonepar_dv variables but I think there is some error with some of these variables. When I cross tabulate hhtype_ with nkids_dv I am getting some conflicting results - please see spreadsheet attached. Just to mention here that I have already pooled all 9 waves together and also integrate data from the household to the individual level dataset. Could you please help me with this?			

History

#1 - 09/14/2020 09:41 AM - Rebecca Parsons

- Status changed from New to In Progress

Hi Theocharis,

Thanks for your question. The Understanding Society team is looking into it and we'll get back to you as soon as we can.

Best wishes,
Becky

Understanding Society User Support

#2 - 09/14/2020 11:04 AM - Gundi Knies

- Category changed from Data inconsistency to Data documentation

- Assignee changed from Alita Nandi to Theocharis Kromydas

- Priority changed from High to Normal

- % Done changed from 0 to 50

- Private changed from Yes to No

Hi Theocharis,

the household type variable draws on the number of dependent children in the household (depchl_dv) so includes young people aged 16 to 18 in full-time education and living in a family with his or her parent(s). By contrast, nkids_dv is a simple count of the number of persons aged under 16 in the household; where the age_dv information is missing and the interview outcome variable ivfio indicates that the enumerated member of the household in aged under 16, this information is also used. depchl_dv will not include children under 16 with missing age (but, as indicated before it will include children aged 16-18).

In short, depchl_dv and nkids_dv capture a different universe of young household members and we would expect them to not match up 100 percent. By extension this is true for hhtype_dv.

Best wishes,
Gundi

On behalf of Understanding Society User Support

#3 - 09/14/2020 11:24 AM - Theocharis Kromydas

Hi Gundi

Many thanks for your quick response. So do you recommend me using depchl_dv (I guess you mean ndepchl_dv?) instead of nkids_dv to construct the family structure variable I want? As far as I can understand hhtype_dv has been derived from nkids_dv, so it cannot be combined with ndepchl_dv to construct a new variable. To recap, What I want is to construct a household structure variable that has 4 values: 1)Single no children, 2)Coupled no children, 3)Coupled with children and 4)lone parents. Is this possible using the Understanding Society data? Can it be constructed in a reliable manner?

Many Thanks!

Harry

#4 - 09/14/2020 01:29 PM - Gundi Knies

Hi Harry,

sorry to not have been more precise. No, I am not recommending you to use or not use `depchl_dv`. I was just clarifying that the `hhtype_dv` variable uses `depchl_dv` and that this uses a different definition of 'child' than `nkids_dv`. More specifically, the `hhtype_dv` programme counts the number of household members for whom `depchl_dv` is 1 using the `indall` data file. The variable `ndepchl_dv` which you mention is yet another possible candidate. It is an individual-level variable that counts the number of dependent children a person has in the household (and is similar to `nchild_dv`, which counts the number of own children aged 0-15 in the household).

Which of these variables, if any, is useful for your analysis is not something we can advise on. It really depends on how you define a family unit, is this simply based on co-residence (`hidp`-level) or also on economic dependency (possibly: `buno_dv` level), on biological and or non-biological relationships (no ids readily available)? Are children to be defined on the basis of age and or economic-dependency? Do non-biological relationships count?

Generally, you will have a lot more definition flexibility by drawing on the `egoalt` data file instead of using the existing derived variables in `indall` as most of them are at the `hh`-level, so possibly not the family-level you have in mind. In particular when you add the `age_dv` variable to `egoalt` you can produce individual or household-level counts of number of people in a specific age range fulfilling specific relationship criteria (e.g. biological child aged 0-24); or, if your definition of the family matches that of a benefit unit, you could add the benefit unit number (`buno_dv`) and compute the number of people in a specific age range fulfilling specific relationship criteria (e.g., biological child aged 0-24) in the benefit unit.

Hope this helps,
Gundi

#5 - 09/15/2020 11:15 AM - Theocharis Kromydas

- File *Cross_tab_1.xlsx* added

Dear Gundi

Many thanks for taking the time to look in my query in more detail. Many thanks for your suggestion too. Before I look on `egoalt` datafile let me clarify what I want to do. I would need to create a household structure variable with the aforementioned 4 values (1)Single no children, 2)Coupled no children, 3)Coupled with children and 4)lone parents) where children are defined as those aged 16y that live in the household and are also economically dependent to their parents irrespective on whether they have a biological relationship or not, but I am still confused on which variables I need to use for this. For example after merging household and individual level dataset from wave 9 and cross-tabulating `hhtype_dv` and `ndepchl_dv` I am getting the table I attach. If you look on column C there are plenty of cases in `hhtype_dv` with children under the zero value in `ndepchl_dv`. Does this mean that in these households parents do have children (biological or not) but their children either do not live with them or are not defined as economically dependent?

Many thanks

#6 - 09/15/2020 01:02 PM - Gundi Knies

Dear Harry,

as I said, `ndepchl_dv` is an individual-level variable: how many dependent children does this enumerated household member have in this household. By contrast, `hhtype_dv` is a household level variable. Individuals may very well have no dependent child of their own but live with a dependent child in the household.

To decide which of these variables, if any, is useful I would suggest you run through a couple of example households and think about what exactly it is that you want to capture in your family classification. For example:

1. a three person three generations household. There is a parent (aged 49, so not counted as a pensioner) and their child (aged 15), who is the parent of a newborn (aged 0). The mother of the newborn has `ndepchl_dv`=1, the newborn and his/her grandparent have `ndepchl_dv`=0. `depchl_dv`=1 only for the newborn. All have `nkids_dv`=2 (because there are two persons aged 0-15 in the household). The grandparent and parent of the newborn have `nchild_dv`=1 as both are responsible for a child aged 0-15 in the household. Both the grandparent and the parent of the newborn are single (`couple_dv`=0).
2. a four-person family - father, mother and two children aged 16 and 17; one child is no longer a dependent child but the other one is. The father and mother will both have `ndepchl_dv`=1, both children will have `ndepchl_dv`=0; the dependent child will have `depchl_dv`=1, the child that does no longer count as a dependent child has `depchl_dv`=0, as do the father and mother. `nkids_dv` and `nchild_dv` will be 0 for all household members as both children are no longer aged 0-15.
3. a five person household two siblings in their late 20s, one in a couple with a child aged 2 and the other a single with a dependent child aged 3. For the couple and single parent `ndepchl_dv`=1; `nchild_dv`=2 for all five household members, `depchl_dv`=1 for the two children, and 0 for the three adults.
4. ...
5. ...

How do you want your classification to deal with these cases? Is [3] a couple or a single parent household or is it two families? Is [2] a 1-child couple family or is it a two-child couple family? Is [1] a single-parent family with two children? Or two one parent families with a child each? If [1] is two families then you have the issue that the 15-year old is part of two different families?

How much detail can you afford to loose in your specific analysis? Perhaps you can lump all non-standard living arrangements in a miscellaneous category 5 "Other family type?", for example, or you may need to work at the family level instead of at the household level..

Hope this helps,
Gundi

#7 - 09/15/2020 04:42 PM - Theocharis Kromydas

Hi Gundi

Thanks again for your very informative reply. I would need to test some of these classification using the data I am currently working on and then get back to you.

Many thanks for your valuable help.

Best,
Harry

#8 - 09/18/2020 03:30 PM - Theocharis Kromydas

Hi Gundi again

I read your last email carefully and it seems to me that the main problem I have is with definitions. My main aim is to classify individuals in household types based on the presence of dependent children or not. However, I am not sure on the definition of dependent child used by Understanding Society. The hhtype_dv variable says that is using LFS definition but I searched on-line and could not find consistent answers. However, because of the specific requirement of the study I am working on at the moment I would need to use the definition of the dependent child from the Family Resources Survey that defines Dependent child as an individual aged under 16 or aged between 16 and 19 and:

- Not married nor in a Civil Partnership nor living with a partner; and
- Living with parents; and
- In full-time non-advanced education or in unwaged government training.

Could you please let me know if this is the definition used by Understanding Society as well to define dependent children?

If yes then I can work it out how to recode the hhtype_dv variable and apply it to individuals.

Many thanks,
Harry

#9 - 09/21/2020 03:29 PM - Alita Nandi

- Assignee changed from Theocharis Kromydas to Gundi Knies

#10 - 09/25/2020 10:03 AM - Gundi Knies

Hi Harry,
the definition of the dependent child variable seems to have changed over time and we are currently reviewing the code frame for our depchl_dv variable in light of changes in the definition. Any changes will be indicated in the next release of the data and online documentation.

Currently, depchl_dv in Understanding Society is mainly based on age_dv. those who are aged 0-15 (or with missing age_dv but an interview outcome code marking them out as a child aged under 16) are considered dependent children. Children aged 16-19 are considered in full term education based on jbstat not being 1 or 2 and employ not being 1; mnspro/fnspro are used to define whether 16-19 year olds children live without a parent; resp16_dv is used to define whether 16-19 year-olds have a child.

In the meantime, to match the specific definition you are after, you might want to derive the measure according to what you think best represents the FRS definition.

Best wishes,
Gundi

#11 - 09/25/2020 11:06 AM - Gundi Knies

- Assignee changed from Gundi Knies to Theocharis Kromydas

#12 - 09/29/2020 06:43 PM - Theocharis Kromydas

Thanks for letting me know Gundi!

Harry

#13 - 10/22/2020 10:14 AM - Alita Nandi

- Status changed from In Progress to Resolved

- % Done changed from 50 to 90

#14 - 10/13/2021 11:14 AM - Understanding Society User Support Team

- Assignee deleted (Theocharis Kromydas)

- % Done changed from 90 to 100

Files

Cross_tab.xlsx	10.2 KB	09/11/2020	Theocharis Kromydas
Cross_tab_1.xlsx	10 KB	09/15/2020	Theocharis Kromydas