
Subject: Why can't individual WI data and values be released
Posted by [Marco Stamazza](#) on Wed, 19 Oct 2022 08:36:13 GMT
[View Forum Message](#) <> [Reply to Message](#)

I see that many of us are trying to reproduce the wealth indices from past surveys. It is painful and time consuming.

I don't understand why the values (and possibly the recoded data) can't be made available. It would save us a lot of time and likely mistakes.

Is there a way for asking for these files?

Subject: Re: Why can't individual WI data and values be released
Posted by [Janet-DHS](#) on Mon, 24 Oct 2022 17:09:23 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member Tom Pullum:

The wealth quintiles are constructed from the continuous wealth index, and both are in the data files. In the HR and PR files, the quintiles are hv270 and the continuous index is hv271. In the IR, BR, and KR files, those variables are v190 and v191. In the MR file they are mv190 and mv191. The continuous index is constructed from household-level characteristics that are in the HR and PR files. The survey-specific instructions are on the website. What else are you looking for?

Subject: Re: Why can't individual WI data and values be released
Posted by [Marco Stamazza](#) on Wed, 26 Oct 2022 10:39:46 GMT
[View Forum Message](#) <> [Reply to Message](#)

Great, thanks, I found the values.

Am I right thinking that 270/1 are for the wi computed on all data and 270a/1a for the wi computed using regression on urban and rural?

One more question. Is there a file with a list of variables to use (I know I'll have to add some specific ones). I am doing it by hand and it is taking days. I don't use SPSS, so can't run the macros.

Many thanks,

Subject: Re: Why can't individual WI data and values be released
Posted by [Janet-DHS](#) on Fri, 28 Oct 2022 16:57:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member Tom Pullum:

The wealth index is based on household-level variables and is the same for every person in the same household. Your use of the word "individual" is a little misleading.

The continuous index, hv271, is used to construct the quintiles, hv270 for all households, hv270a for urban households, and hv270b for rural households. The difference between the three versions of hv270 is just in the specification of the 4 cutpoints that divide the continuous index into 5 pieces. The bottom quintile, for example, applies to the households that have the lowest values of hv271, up the scale to include the bottom 1/5 of weighted de jure household members. For hv270a the application is to urban households only, and for hv270b to rural households only. There is no use of regression. Stata includes a "pctile" command to find the cutpoints.

The variables that are used to construct hv271, with principal components analysis, vary somewhat from one survey to another but are specified on the website for each survey. Not sure why you would want to reconstruct the wealth index. Yes, it's a lot of work.

Subject: Re: Why can't individual WI data and values be released

Posted by [Marco Stamazza](#) on Sat, 29 Oct 2022 01:56:36 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello, thank you very much for your and Tom's prompt help!!

I am sorry I used the term individual, I saw that the women survey "IR" data is labeled as individual recode.

I am investigating some properties of the WIs and I need to reproduce them. It is a lot of work but I also find a lot of help and explanation, including the rdhs R package. I am impressed!!!

Subject: Re: Why can't individual WI data and values be released

Posted by [Janet-DHS](#) on Wed, 16 Nov 2022 14:15:13 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member Tom Pullum:

I just checked, and it appears that this survey was typical of DHS surveys. A variable hv117 in the PR file is "eligibility for the women's interview". It is 1 if the woman is de facto (i.e. available for the interview) and age 15-49. I constructed a check for the coding of hv117

```
gen hv117_test=0
replace hv117_test=1 if hv103==1 & hv104==2 & hv105>=15 & hv105<=49
```

```
tab hv117_test hv117,m
```

The "test" variable matches hv117 perfectly. There don't seem to be any other criteria. Some of the eligible women will not actually appear in the IR file, usually because they refused or they were not (after all) available to meet with an interviewer. But there doesn't seem to be any systematic reason for eligible women to be missing from the IR file. Hope this answers your

question. If not, let me know.
