Understanding Society User Support - Support #1172

seeming errors in fnspid and mnspid variables

03/26/2019 05:19 PM - Katy Morris

Status:	Resolved	**Start date:**	03/26/2019
Priority:	Normal	**% Done:**	70%
Assignee:	Katy Morris		
Category:			

Description

Hello,

I am trying to match parents and children in UKHLS (using via pidp, fnspid and mnspid variables) and obtain a peculiar error when doing so in Stata.

Having merged all waves, I'm using the egen command in order to pick up any parents who are not identified in the wave the respondent enters UKHLS, as follows:

```
egen fpidp=max(fnspid), by (pidp)
format fpidp %12.0g
```

Unfortunately this generates incorrect data, for example pipd 29925, whose father is officially fnspid 614482685 somehow becomes 614482688 using the above syntax. This is one of many, many errors that I can see. Having tried to find ways around this, it turns out that *clonevar* clones the variable correctly but is vulnerable to the same error when *egen* is subsequently used. A simple gen fpidp= fnspid produces the same (mostly wrong) values as the *egen* procedure outlined above.

I find this very puzzling - is it a Stata error? Grateful for all and any advice on how to proceed.

Katy

History

#1 - 03/27/2019 10:44 AM - Alita Nandi

- Status changed from New to In Progress

- % Done changed from 0 to 70

USE

egen long fpidp=max(fnspid), by (pidp)

INSTEAD OF

egen fpidp=max(fnspid), by (pidp)

Best wishes,
Alita

#2 - 03/27/2019 11:08 AM - Alita Nandi

- Assignee set to Katy Morris

#3 - 03/27/2019 11:08 AM - Alita Nandi

- Private changed from Yes to No

#4 - 03/27/2019 01:24 PM - Katy Morris

Alita Nandi wrote:

> USE
>
> egen long fpidp=max(fnspid), by (pidp)
>
> INSTEAD OF

egen fpidp=max(fnspid), by (pidp)

Best wishes,
Alita

Problem solved - thank you!

#5 - 03/27/2019 02:29 PM - Alita Nandi

- Status changed from In Progress to Resolved