Subject: Re: merging HIV results and couples data

Posted by user-rhs on Tue, 12 Apr 2016 17:00:35 GMT

View Forum Message <> Reply to Message

That's because the IR file only contains information on women. If the women's dataset is the one in memory ("master," the dataset that you have opened), as it is in Trevor's code, and the HIV dataset is one you're merging it with ("using," the dataset that comes after "using" in the -merge- command), then, based on the coding for the resultant _merge variable (which, by the way, merge results are shown immediately after a merge):

numeric code	equivalent word (results)	description
1	master	observation appeared in master only
2	using	observation appeared in using only
3	match	observation appeared in both
4	match_update	observation appeared in both, missing values updated
5	match_conflict	observation appeared in both, conflicting nonmissing values

Source: -help merge-

You can see that _merge==2 means that the unique identifier existed in the using/HIV data only, and not the women's/master dataset.

Stata keeps all observations, regardless of matching status, unless you specify to keep just the ones in the master dataset or in the using dataset, i.e., supposing you have IR dataset in memory:

```
merge 1:1 v001 v002 v003 using "HIV.dta", assert(match master) /*Keeps only those who are in the IR dataset, regardless of matching status*/
merge 1:1 v001 v002 v003 using "HIV.dta",assert(match using) /*Keeps only those in the HIV dataset, regardless of matching status and therefore sex of the respondent*/
```

hth,
rhs