

Hi,

I have a couple of questions related to examples 11.5 and 11.6 of the manual.

(1) Example 11.5:

You mention on page 348 that all variables part of the NAMES list are used to impute data (on the variables listed under IMPUTED). Is there any way NOT to use some variables that are part of the list in the imputation ?

(2) Example 11.6: Will this be the same in Plausible values imputations ? Will all the variables listed in the NAMES list be used in generating the plausible values or only those included in the MODEL section ? If yes, is there a way again to not use some variables ?

(3) It is possible to generate multiple imputation data sets (5-10-20-etc.) including imputed values for missing data on observed variables and plausible values in the same data sets.

(4) How can we include additional variables in the saved multiple imputation and/or plausible values data set (lets say the z variables from example 11.5)? I do not necessarily want to impute these data or to use them in the imputation algorithm, just to have them saved in the created data sets so as to be able to use them in subsequent analysis. Will the simple AUXILIARY function (without e-m-r) work ?

Thank you !

I am glad you asked so that this can be clarified.

(1) and (4):

UG ex 11.5 is not as clear as it could be on this point. A user would typically work with not only the NAMES and IMPUTE lists of variables, but also a USEVARIABLES list and an AUXILIARY list. The NAMES list simply reads the variables in the original data set. The USEVARIABLES list is a smaller subset of variables from the NAMES list, just as in an ordinary analysis. The USEVARIABLES list variables are the variables used to create the imputations. In UG ex11.5, the USEVARIABLES list is absent and therefore defaults to the NAMES list. Typically you also want to save into the imputed data set other variables that are not to be used in the imputation and to do that you put those variables on the AUXILIARY list.

(2) Same thing.

(3) The SAVE = data set contains what you are asking for. The PLAUSIBLE = data set gives summary statistics for plausible values.