Thank you for sharing the do file. It's very helpful. If at all possible for you can you share the merge_children_mothers_fathers.dta file. I want to compare my merged data file with it if I have got it right or made some mistakes in the process.

Please clarify one thing for me. In your example "effect of maternal and paternal education on child survival" you are looking at the effects of mother and father in separate regressions, right? Not taking them together in one regression? because if we do then we will be left with near about those children observations only for whom we have both parent's info who are alive and live with them. so roughly 35-40k observations.

Also, I had one more query. I just need father's education and mother's education to create my variable of interest and other controls could be mother, child, HH characteristics. So can I use the info about the child's father line no. in PR file and from there merge into KR file?

I tried this as well and now I have father's education variable for approximately 1,73, 000 children (in KR file out of 2,32,920).

Can I now use these 1,73,000 observations as my sample size and apply women's weights and do my analysis? Will that be correct?