

---

Subject: Weight to use for regression that has children and mother level variables  
Posted by [hbsheldon](#) on Mon, 27 Jun 2022 23:07:29 GMT

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

---

We are intending to run a child-level regression that also uses variables from the mother. For example, our dependent variable is the weight-for-age z-score and the independent variables are a combination of child level variables (ex: was the child breastfed) and mother level variables, such as her responses to questions on decision-making involvement.

Of course, mothers have different numbers of children. The guidance seems to be that it is ok to use the women's weight, but we have doubts about this because, for example, all women in a cluster share the same weight even if they have different numbers of children. Does this woman's weight account for the number of children a woman has in any way? If not, would a child-level regression with a mother's characteristics and survey weight succumb to the problem of "double-counting" (since some mothers have 10 kids and some have only 1, for example)? If yes, what is the best way to adjust for this?

---

---

Subject: Re: Weight to use for regression that has children and mother level variables

Posted by [Janet-DHS](#) on Wed, 06 Jul 2022 13:38:21 GMT

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

---

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

We would normally do the kind of analysis you describe with the KR file, which has one record for each child born in the past five years. Virtually all of the mother's information is on the child's record. For some outcomes, the response is only obtained for the youngest child (the most recent birth). This introduces some bias, if your interest is in ALL children, mainly because the youngest child is biased toward low-fertility women (see <https://www.dhsprogram.com/pubs/pdf/MR14/MR14.pdf>). If you were to weight up those children, in proportion to the number of children under five, you would not be reducing that bias--you would be magnifying it. I recommend that you not do any such weighting, but just describe the selectivity in the data.

Other variables, such as the height-for-age (HAZ) score, do not have this kind of selectivity. Our practice would be to use all the children (in the KR file, since you want to relate to the mother; the Z scores are also in the PR file for all children in the household) and just use v005. The mother's weight is assigned to the child.

You may also be thinking of maternal clustering. Children of the same mother tend to have similar outcomes because of omitted variables. Some people have gone to a lot of trouble to adjust for maternal clustering. However, it's hard to apply a multi-level model to such data because there are so few level-1 units (children) per level-2 unit (mothers)

---