

Creation of household and female weights in PMA2020

PMA2020 collaborates with the National Statistical Organization (NSO) in each country to obtain a representative survey sample. NSO selects enumeration areas (EAs) with probability proportional to size using the master sampling frame stratified by urban-rural areas. PMA2020 receives the sampled EAs as well as their selection probabilities. Each EA is then mapped, and a census list of households is compiled for constructing the HH listing frame for the EA. The selection probability of the household is calculated as the probability of selecting the EA times the probability of selecting the household in the EA. The design-survey weights are then created as the inverse of the HH selection probability. The weights are further adjusted for non-response at the household level within the EA.

With PPS sampling, the selection probability of a unit (i.e., EA) is

$$\pi_i^{pps} = \frac{n * P_i}{\sum_{i \in u} P_i}$$

where n is the sample size of the EA, and P_i is the measure of size (e.g. population size, number of households) in the EA unit i . The denominator P_i is summed over all units in the specific survey domain. The survey domain, for example, will be urban/rural strata by region or province. This selection probability is provided by the NSO.

Next, we calculate the selection probability of the secondary sampling units (SSU) – households. Using Uganda as an example, where 44 households are selected for interview within each EA, the selection probability of the HH in the i -th EA is

$$H_i = \frac{44}{P_i}$$

The design-weight is then

$$w_i = \frac{1}{\pi_i^{pps} * H_i}$$

The design-weight for the i -th EA is adjusted for the (unit) non-response rate (i.e., household non-response rate within an EA) by multiplying the weight w_i by the factor

$$R_i = \frac{1}{1 - f_i}$$

where f_i is the non-response rate.



JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

BILL & MELINDA GATES INSTITUTE *for*
POPULATION *and* REPRODUCTIVE HEALTH

As all females age 15-49 within a household are surveyed, no further selection probability of eligible women is needed to generate the female weight (self-weighted). Instead, the household weight is adjusted for the female non-response at the EA level. To create the final weights, the household and female weights are then each normalized at the national level. In the case of countries where the survey was conducted in a selected administrative area (e.g., Kinshasa, Democratic Republic of Congo), the household and female weights are normalized within the administrative area.