Following is a response from Senior DHS Specialist, Tom Pullum:

I believe there is a problem with your approach, because it does not take into account that a cluster's probability of selection is not a constant within each stratum, but is proportional to the size of the cluster. Without knowing the size of the cluster, you cannot calculate its probability of selection.

We at DHS hope to prepare guidelines for dealing with the requirement for separate weights for multi-level models. I don't think the results will be very sensitive to the allocation (as your comparison suggests) but we do need something better than just wij and 1. Your approach is worth pursuing but I think the best that can possibly come out of it will be a better approximation, because we simply don't know the size of the cluster.