Following is a response from DHS staff member, Tom Pullum:

You have raised an interesting issue. I did not expect to see variation in d005 with clusters. It took a while for me to understand it, but it makes sense.

The DV module is only administered to one woman in each household. If more than one woman in the household listing (the PR file) is eligible for the women's interview (the larger interview, not specifically the DV module), then one woman is selected at random using a "Kish grid'. That is, before the interview with the women in the household actually begins, the eligible respondents (women with hv117=1) are listed and one of them is selected at random.

With this sampling scheme, if there are two eligible women, the weight for the one who is selected is approximately doubled. If there are three eligible women, the weight for the one who is selected is approximately tripled, and so on. I say "approximately" because there is an adjustment for nonresponse.

Below I will paste a short Stata program that calculates the number of eligible women from the PR file (I call it nelig), merges that onto the IR file, and then calculates the standard deviation of d005 within clusters and also within values of nelig. It shows that d005 IS constant within clusters, if you take account of nelig. You will also see this if you list v001 v002 v003 v005 d005 nelig within some representative clusters. (You will see that v005 is constant within clusters, regardless of nelig.) I use the Kenya 2022 survey for an illustration.

This means that the variation you are finding in d005 within clusters has nothing to do with svyset. You can get the cluster-level proportions just with "proportion ipv [fweight=d005], over(v001)", as you did. fweight is ok because the factor of 1000000 is in both the numerator and the denominator of the proportion and cancels out.

I recommend that you use the usual svyset command, with d005 in place of v005. This is our standard recommendation for analyses of the DV variables.
* The sampling weight for the DV respondents is constant within combinations of

*   clusters and the number of women who are eligible

* nelig is the number of eligible women in the hh.

* it is the number of women in the household with hv117=1e

* Specify a workspace

```
cd e:\DHS\DHS_data\scratch


* Find the number of eligible women usig the PR file

use "...KEPR8CFL.DTA" , clear

keep if hv117==1

collapse (sum) hv117, by(hv001 hv002)

rename hv001 v001

rename hv002 v002

rename hv117 nelig

label variable nelig "Number of women in hh eligible for DV module"

save temp.dta, replace


* Open the IR file, reduce to the women selected for DV, and add nelig to each woman

use "...KEIR8CFL.DTA" , clear

keep if d005<.

merge 1:1 v001 v002 using temp.dta

tab _merge

keep if _merge==3

drop _merge


* Show that there is d005 is constant within cluster, taking nelig into account

collapse (sd) v005 d005, by(v001 nelig)

summarize
```