Subject: Re: Why I am getting different total observations when using iweight for tabulating a variable
Posted by sujata on Sat, 01 Apr 2023 06:23:27 GMT
View Forum Message <> Reply to Message

Dear Tom,
Thank you very much for looking into this.
I understand that this is outside the forum's scope, and I really appreciate that you spared some time for this.
However, I wanted to clarify further my understanding of how to treat the weights in my analysis. I want to ensure that I use them correctly and get accurate results.

Firstly, As per your suggestion, I normalized the data so that the mean of the shweight_PR in the PR file is equal to 1000000.

```
gen unwtd=1000000
total unwtd shweight
matrix B=e(b)
matrix list B
scalar sfactor=B[1,1]/B[1,2]
scalar list sfactor
gen shweight_PR=round(sfactor*shweight)
```

After that, I generated wgt_shweight_PR= shweight_PR/1000000. The mean of wgt_shweight_PR is 1.

```
svyset [pw= wgt_shweight_PR ], psu( hv021) strata( hv022)

sum wgt_shweight_PR

Variable              Obs     Mean     Std. dev.    Min      Max

wgt_shweig~R      52,682    1     .6350303    .05442   4.638086

egen raw_rank_CE=rank(sv271s), unique
sort raw_rank_CE
qui sum shweight_PR
gen wi = shweight_PR /r(sum)
gen cusum = sum(wi)
gen wj= cusum[_n-1]
replace wj=0 if wj==.
gen rank_CE=wj+0.5*wi

sum rank_CE

Variable    Obs     Mean       Std. dev.     Min         Max

rank_CE 52,682  .4857322    .2892787     6.00e-06 .9999868
```

I am getting the same mean (0.4857) with wgt_shweight_PR as well.

Is it the right way to use weights?

Thank you.