Understanding Society User Support - Support #2119

Svy Commands and Fixed Effects Regressions

06/06/2024 12:29 PM - Emma Kemp

Status:	Feedback	**Start date:**	06/06/2024
Priority:	High	**% Done:**	50%
Assignee:	Understanding Society User Support Team		
Category:	Weights		

Description

Hi there,

I'm using UKHLS panel data for my MSc Behavioural Science dissertation in which I am trying to explore the impact of perceived neighbourhood social cohesion (NSC_index) on life satisfaction (lfsato). However, I can't work out how to run a fixed effects regression accounting for the complex survey design of UKHLS..

UKHLS recommends using the svy suite of commands so I have set up my do-file as follows:

```
// DECLARE COMPLEX SURVEY DESIGN
use UKHLS_long_acfil_cleaned_usable.dta
```

- set correct weights
  svyset, clear
  svyset l_psu [pweight = l_indscus_lw], strata(l_strata) singleunit(scaled)

My first question is: Have I done this correctly? Should l_psu be pidp instead given that is the smallest unit I am looking at? and is single unit (scaled) correct?

Then, I declare the panel data set up:

```
// DECLARE PANEL DATA SET UP
//Use xtset command to tell stata that this data has a panel structure - pidp being the unique identifier and wave being the time variable
sort pidp wave
xtset pidp wave
```

I am now trying to run fixed effects regressions to work out whether a change in perceived neighbourhood social cohesion leads to a change in life satisfaction however, the command I would normally use for fixed effects regressions (xtreg) is not compatible with svy. Does anyone know of a command that could do this?

I have since come up with the following options:

```
//OPTIONS TO ACCOUNT FOR COMPLEX DESIGN/WEIGHTS
```

1. svy: reg lfsato NSC_index_nm i.wave (this leads to really high estimates as it doesn't account for individual fixed effects)
2. svy: reg lfsato NSC_index_nm i.wave, absorb (pidp)
3. xtreg lfsato NSC_index_nm i.wave [pweight=l_indscus_lw], fe vce(cluster pidp)
4. areg lfsato NSC_index_nm i.wave [pweight=l_indscus_lw], absorb(pidp) cluster(pidp)
5. reghdfe lfsato NSC_index_nm i.wave [pweight=l_indscus_lw], absorb(pidp) vce(cluster pidp strata) //chatGPT told me to add strata and then this command should mimic the syvset command?

In summary, my key questions are:

1. When I am declaring the complex survey design - have I done this correctly? Should l_psu be pidp instead? and is single unit (scaled) correct?
2. What syntax do I use to run a fixed effects regression that accounts for the complex survey design of UKHLS

Thank you in advance for any advice you can provide.

Emma

History

#1 - 06/07/2024 03:01 PM - Understanding Society User Support Team

- Category set to Weights

- Status changed from New to In Progress

- Assignee changed from Understanding Society User Support Team to Olena Kaminska

#2 - 06/18/2024 12:10 PM - Understanding Society User Support Team

- File MLM weights advice 20240121.pdf added

- Status changed from In Progress to Feedback

- Assignee changed from Olena Kaminska to Understanding Society User Support Team

- % Done changed from 0 to 50

- Private changed from Yes to No

Hello Emma,

Given that two-level models, where the higher level corresponds to clusters in the sample design, are the only models supported by developed theory, our weighting team has produced guidance on addressing complex survey design for random effects in a multilevel model.

I’m attaching a short PDF document with the guidance mentioned.

I hope this information is helpful.

Best wishes,
Roberto Cavazos
Understanding Society User Support Team

Files

MLM weights advice 20240121.pdf	184 KB	06/18/2024	Understanding Society User Support Team