

Understanding Society User Support - Support #1089

Does the data collection procedure of US fit the rule of randomization?

11/07/2018 03:11 PM - Jing Shen

Status:	Resolved	Start date:	11/07/2018
Priority:	Normal	% Done:	100%
Assignee:	Jing Shen		
Category:	Survey design		
Description			
I'm conducting a study about the impact of the Brexit Referendum on life satisfaction using a difference-in-difference approach, which requires a random assignment between the control and treated groups. My question is: can we say respondents were not selectively interviewed before and after the Referendum (June 23, 2016); namely, the chance for one to be interviewed before and after a certain date is random? Can we make such an assertion?			

History

#1 - 11/07/2018 05:13 PM - Stephanie Auty

- Status changed from New to In Progress
- % Done changed from 0 to 10
- Private changed from Yes to No

Many thanks for your enquiry. The Understanding Society team is looking into it and we will get back to you as soon as we can.

Best wishes,
Stephanie Auty - Understanding Society User Support Officer

#2 - 11/07/2018 09:19 PM - Peter Lynn

- Assignee set to Jing Shen
- % Done changed from 10 to 50

Not really. The chance to be assigned to a particular monthly sample can be considered random, but within each monthly sample people interviewed early in the month (easy-to-contact and co-operative) are systematically different from those interviewed later. So, you'd probably be safe to compare, say, the Jan-Apr samples with the July-Dec samples. But for the May and June samples you would need additional controls for selection into before/after.

#3 - 11/08/2018 11:00 AM - Peter Lynn

- % Done changed from 50 to 60

Follow-up question from Jing (by email):

You mentioned both May and June samples would not be random. I understand the June sample would be problematic, but I'm puzzled why the May sample should also be excluded. What makes the May sample different from, say, the April sample, since respondents were all interviewed before June, in which the Referendum took place?

Answer from me:

May sample members are not all interviewed in May. A non-negligible minority were interviewed after referendum day:

. ta g_istrdatm if g_month==17

Interview	start date			
(month)	Freq.	Percent	Cum.	
5	786	53.76	53.76	
6	443	30.30	84.06	
7	140	9.58	93.64	
8	45	3.08	96.72	
9	23	1.57	98.29	
10	25	1.71	100.00	
Total	1,462	100.00		

. ta h_istrdatm if h_month==5

Interview			
start date			
(month)	Freq.	Percent	Cum.
missing	1	0.05	0.05
5	635	32.95	33.00
6	907	47.07	80.07
7	246	12.77	92.84
8	67	3.48	96.32
9	48	2.49	98.81
10	23	1.19	100.00
Total	1,927	100.00	

#4 - 12/21/2018 10:30 PM - Stephanie Auty

- Status changed from In Progress to Feedback

#5 - 04/11/2019 12:53 PM - Jing Shen

Thank you very much. This is very helpful!

#6 - 05/23/2019 02:57 PM - Jing Shen

Can I follow up this issue with some new concerns? 1. Would sampling strata (Strata) and primary sampling unit (psu) have any influence on the date, on which an individual would be interviewed? 2. Early on, you mentioned that easy-to-contact and cooperative households/individuals would be interviewed first, and difficult cases were dealt with later. Then, my question is who were those easy-to-contact respondents and who were difficult? Would it be possible that people who work full-time, or who live in remote/rural areas, or who don't speak English are more difficult cases? In short, I'm wondering if there is any explanation/mechanism underlying the non-random distribution of interview dates within a month. Thank you very much in advance.

#7 - 05/24/2019 12:47 PM - Stephanie Auty

- % Done changed from 60 to 70

Dear Jing Shen,

To answer your first question, no, the sampling strata and psu do not have any influence on the date on which an individual would be interviewed.

Regarding "easy-to-contact and cooperative households/individuals would be interviewed first, and difficult cases were dealt with later" – this isn't a pre-determined state, that is, interviewers don't deal with 'easy' cases first and leave the 'difficult' cases until later, they try to contact all the sample at around the same time, and those who are easier to contact are more likely to be contacted early. Those who are more difficult to contact are more likely to have missed interviews in previous waves as they may miss appointments, be difficult to get hold of, or reluctant. It takes interviewers more time to chase and interview these participants.

There is a large literature that looks at who are the easy-to-contact and difficult respondents, and this is also something that you could explore empirically with the data.

Best wishes,
Stephanie

#8 - 08/10/2022 11:20 AM - Understanding Society User Support Team

- Category set to Survey design

- Status changed from Feedback to Resolved

- % Done changed from 70 to 100