

## Understanding Society User Support - Support #2190

### Weights wave 10,12, and 14

01/06/2025 06:06 PM - Abdulaziz Majrashi

<b>Status:</b>	Resolved	<b>Start date:</b>	01/06/2025
<b>Priority:</b>	High	<b>% Done:</b>	100%
<b>Assignee:</b>	Olena Kaminska		
<b>Category:</b>	Weights		

#### Description

Dear Understanding Society Team,

My name is Abdulaziz Majrashi, and I am pursuing a PhD in Economics at Newcastle University. My research primarily focuses on the labour market, specifically examining the impact of work-from-home arrangements on mental health, productivity, and other outcomes in the UK post-COVID-19. For this purpose, I am utilising data from waves 10, 12, and 14 of the Understanding Society dataset, mainly focusing on a restricted sample.

I am encountering some challenges with the application of weights in my analysis. I use the "jbflex7" variable to identify individuals working from home, and after data restrictions, my sample includes about 10,000 observations. However, when applying weights, I noticed a reduction to fewer than 3,000 observations, which is considerably lower than expected.

For my analysis, I have employed a difference-in-differences model using OLS with fixed effects in Stata, using the command: (xtreg Y treated time did i.year [pweight = inding2\_xw], fe vce(cluster pidp))

This command has resulted in issues such as the error "weight must be constant within pidp," collinearity, or insignificant results with the drastically reduced sample size. I have attempted to use various weight variables but continue to face similar issues.

Could you please advise on the appropriate weight(s) to use for this longitudinal analysis involving multiple waves (10,12, and 14)? Should I be integrating multiple weights due to the multiple waves in my study, and if so, how should this be implemented effectively in Stata?

I would greatly appreciate any guidance or example Stata code you could share to help address these issues. Your expertise and insights would be invaluable to my research.

Thank you very much for your time and assistance. Let me know if you need more information

Best regards,

Abdulaziz Majrashi PhD Student

#### History

##### #1 - 01/06/2025 06:08 PM - Abdulaziz Majrashi

Dear Understanding Society Team,

My research primarily focuses on the labour market, specifically examining the impact of work-from-home arrangements on mental health, productivity, and other outcomes in the UK post-COVID-19. For this purpose, I am utilising data from waves 10, 12, and 14 of the Understanding Society dataset, mainly focusing on a restricted sample.

I am encountering some challenges with the application of weights in my analysis. I use the "jbflex7" variable to identify individuals working from home, and after data restrictions, my sample includes about 10,000 observations. However, when applying weights, I noticed a reduction to fewer than 3,000 observations, which is considerably lower than expected.

For my analysis, I have employed a difference-in-differences model using OLS with fixed effects in Stata, using the command: (xtreg Y treated time did i.year [pweight = inding2\_xw], fe vce(cluster pidp))

This command has resulted in issues such as the error "weight must be constant within pidp," collinearity, or insignificant results with the drastically reduced sample size. I have attempted to use various weight variables but continue to face similar issues.

Could you please advise on the appropriate weight(s) to use for this longitudinal analysis involving multiple waves (10,12, and 14)? Should I be integrating multiple weights due to the multiple waves in my study, and if so, how should this be implemented effectively in Stata?

I would greatly appreciate any guidance or example Stata code you could share to help address these issues. Your expertise and insights would be invaluable to my research.

Thank you very much for your time and assistance. Let me know if you need more information.

Best regards,

Abdulaziz Majrashi PhD Student

##### #2 - 01/08/2025 11:06 AM - Understanding Society User Support Team

- Assignee changed from Abdulaziz Majrashi to Olena Kaminska

##### #3 - 01/09/2025 02:45 PM - Olena Kaminska

Abdulaziz,

Thank you for your question. I suggest you start here to read about using our data in multilevel context:

<https://www.understandingsociety.ac.uk/documentation/mainstage/user-guides/main-survey-user-guide/analysis-guidance-for-weights-when-fitting-mu>

If you still have issues with sample size, you may want to create your own tailored weight for your analysis. For this follow our online course:  
<https://www.understandingsociety.ac.uk/help/training/creating-tailored-weights/>

Hope this helps,  
Olena

**#4 - 01/21/2025 12:11 PM - Understanding Society User Support Team**

- *Status changed from New to Feedback*
- *% Done changed from 0 to 80*

**#5 - 02/10/2026 06:06 PM - Understanding Society User Support Team**

- *Status changed from Feedback to Resolved*
- *% Done changed from 80 to 100*
- *Private changed from Yes to No*