

Subject: FW: JOINING LINK Introduction to the Family Resources Survey workshop

From: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Date: 11/11/2022, 14:55

To: Pierre Walthery <pierre.walthery@manchester.ac.uk>

Hi Pierre,

Hope second day workshop went well yesterday.

I am forwarding some resent correspondence with the FRS team at the DWP about accounting for survey design. It came following a question from one of the workshop attendees.

Have a good weekend.

Jen

From: Owen Justyna DIGITAL GROUP Sheffield Kings Court <JUSTYNA.OWEN@DWP.GOV.UK>

Sent: 10 November 2022 09:08

To: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>; Brandon-Bravo Alexander DIGITAL GROUP Manchester Corporate HUB <Alexander.Brandon-Bravo@dwp.gov.uk>

Cc: Britton Anna DIGITAL GROUP Longbenton BPV <Anna.Britton@dwp.gov.uk>

Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Hi Jen,

With thanks to Alex, please see response to the original query below:

Hi Jen/Tammy,

Thank you for getting in touch with your query, Tammy, and thank you for the response you already provided Jen. Everything you mentioned in your original response sounded right on the money to me and covers the majority of it, but I will elaborate where I can.

Indeed, the FRS sampling process involves a combination of stratification, systematic sampling, and clustering, all of which have an effect on how generally representative of the population the sample can be assumed to be. For instance, clustering reduces the general representativeness of any given sample, as if we were to repeat the sampling process, the random selection at a cluster level would produce a new estimate likely to be more different to the first than we would expect in a simple random sampling process. This lowers our confidence in the estimates from any given sample and widens our CIs. The weights supplied that you refer to primarily enhances the accuracy of estimates reported by the survey, but if you want to compute robust measures of the certainty on these estimates, the complex survey design should be accounted for.

On a technical note, the sampling process is not the same across all parts of the UK, for instance there is no clustering in Northern Ireland and the stratification regime is based on different factors, and so the amount of households per stratum is very different to that of Great Britain. In Northern Ireland, the process of systematic sampling is in effect, whereas this is not considered to be present in Great Britain. Guidance on the sampling process for the most recent FRS publication can be found here: [Family Resources Survey: background information and methodology - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/publications/family-resources-survey-background-information-and-methodology). When accounting for complex survey design in the calculation of CIs,

care must be taken to account for these distinct sampling processes.

In the Safe Room version of the dataset, clustering information for Great Britain is encoded within the SerNum identifying each household. However in the EUL version of the dataset the SerNum is replaced by a different household identifier as part of the EUL dataset anonymisation schedule. This is because original SerNum is classed as a disclosive variable, which can potentially be used to identify a household. Anonymised SerNum does not contain any information relating to stratification and clustering and is assigned to each household randomly. Some guidance for calculating CIs from FRS data both with and without this information can be found here: [Uncertainty in Family Resources Survey-based analysis - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/publications/uncertainty-in-family-resources-survey-based-analysis). This source also describes how to account for stratification by grouping adjacent clusters into pairs.

In the absence of complex survey design information, the main advice is to calculate standard error assuming a simple random sample, and then use published design factors (DEFTs) to broaden or tighten the result, by using design factors that concern an estimate as similar as possible to the one you have calculated. Advice on this can be found generally using the paper Jen referenced ([ONS methodology working paper series no. 9 - Guide to calculating standard errors for ONS Social Surveys - Office for National Statistics](https://www.ons.gov.uk/methodology/workingpapers/onsmethodologyworkingpapers/onsmethodologyworkingpapersseriesno9)) and the detailed advice published by the ONS on the Labour Force Survey can also be very useful ([Labour Force Survey - user guidance - Office for National Statistics \(ons.gov.uk\)](https://www.ons.gov.uk/peoplepopulationandcommunity/employmentandlabourmarket/employmentandlabourmarketintheuk/labourforcesurvey)).

The re-sampling process you referred to earlier is likely to be in reference to a bootstrapping or jack-knife process that can be used to produce CIs for any measure calculated from the survey. With the Safe Room dataset it would be entirely possible to call a process that incorporates both bootstrapping as well as complex survey design as support for this is certainly available in SAS and I imagine other statistical software packages. The advice in [Uncertainty in Family Resources Survey-based analysis - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/publications/uncertainty-in-family-resources-survey-based-analysis) actually falls short of such a method, describing first a process that takes into account complex survey design (PROC SURVEYFREQ etc.) but is inappropriate for non-linear measures, and then a bootstrapping process which does not take into account complex survey design. Nevertheless, comparing these methods could be instructive. This publication also includes code examples that are all in SAS, so unfortunately you may find them of reduced utility if you are working in STATA. I have no experience myself of STATA so cannot advise any further on code.

One final wrinkle however is that even were you to compute such detailed CIs, these CIs may still be an overestimate because the effect of weighting (variably known as grossing, control or calibration weighting, or post-stratification) on CIs has not yet been properly taken into account. To do so requires access to the weighting algorithm and control totals used by the FRS team. CIs and DEFTs published this year will hopefully take this final factor into account and therefore may appear somewhat tighter than any calculated by yourself.

That all sounds like a lot, but it very much depends on how versatile and robust you need your method to be. With access to the Safe Room dataset and armed with something like the SAS procedure PROC SURVEYFREQ, it is very possible to produce decently robust CIs that take into account complex survey design very quickly. With access only to the EUL dataset, the only way to incorporate complex survey design

into any confidence intervals stated is to use published DEFTs, but the versatility of this method is vastly reduced.

Please let me know if anything there requires further clarification and in the meantime I hope that was helpful.

Kind regards,

Justyna

Justyna Owen | Family Resources Survey | Surveys Branch | Data as Statistics |
Department for Work and Pensions | Kings Court | 80 Hanover Way | Sheffield | S3 7UF | MS
Teams call: +44 300 087 2782 | www.dwp.gov.uk

E-M-A-I-L-B-L-O-C-K



We care



We deliver



We adapt



We work together



We value everybody

From: Owen Justyna DIGITAL GROUP Sheffield Kings Court

Sent: 03 November 2022 16:33

To: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>; Brandon-Bravo Alexander DIGITAL GROUP Manchester Corporate HUB <Alexander.Brandon-Bravo@dwp.gov.uk>

Cc: Britton Anna DIGITAL GROUP Longbenton BPV <Anna.Britton@dwp.gov.uk>

Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Hi Jen,

I am forwarding your query to Alex.

Alex is FRS resident expert on standard errors and hopefully will be able to help 😊

Kind regards,

Justyna

Justyna Owen | Family Resources Survey | Surveys Branch | Data as Statistics |
Department for Work and Pensions | Kings Court | 80 Hanover Way | Sheffield | S3 7UF | MS
Teams call: +44 300 087 2782 | www.dwp.gov.uk

E-M-A-I-L-B-L-O-C-K



We care



We deliver



We adapt



We work together



We value everybody

From: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Sent: 02 November 2022 13:23

To: Britton Anna DIGITAL GROUP Longbenton BPV <Anna.Britton@dwp.gov.uk>; Owen Justyna DIGITAL GROUP Sheffield Kings Court <JUSTYNA.OWEN@DWP.GOV.UK>

Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Hi Anna,

Thank you for your offer, if you do know someone who might be able to offer some further insight that would be great.

Best wishes,
Jen

From: Britton Anna DIGITAL GROUP Longbenton BPV <Anna.Britton@dwp.gov.uk>

Sent: 18 October 2022 09:49

To: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>; Owen Justyna DIGITAL GROUP Sheffield Kings Court <JUSTYNA.OWEN@DWP.GOV.UK>

Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Morning Jen,

I must admit that I don't know much about standard errors with the FRS. However, I'd be happy to forward this on to another member of the team who knows more on the subject, if that would help?

Kind regards,

Anna

Anna Britton | Statistical Officer | Family Resources Survey | Surveys Branch | Data as Statistics | Department for Work and Pensions | Benton Park View | Longbenton | Newcastle Upon Tyne | NE98 1YX | www.dwp.gov.uk

E-M-A-I-L-B-L-O-C-K



We care



We deliver



We adapt



We work together



We value everybody

From: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Sent: 12 October 2022 13:09

To: Owen Justyna DIGITAL GROUP Sheffield Kings Court <JUSTYNA.OWEN@DWP.GOV.UK>; Britton Anna DIGITAL GROUP Longbenton BPV <Anna.Britton@dwp.gov.uk>

Subject: FW: JOINING LINK Introduction to the Family Resources Survey workshop

Hello Justyna and Anna,

I hope you are both well. I am following up a query I received following the workshop. It is about the calculation of correct standard errors with the FRS (in the absence of survey design variables). It would be really useful to have your perspective on the query. You can see the full correspondence below. The question the possibility of including pseudo strata/PSUs is a request we have had from users of other social surveys such as the Wealth and Assets survey and the LFS.

Best wishes,
Jen

From: Campbell I, T <T.Campbell1@lse.ac.uk>

Sent: 11 October 2022 10:10

To: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Hi Jen,

Thank you for getting back to me – this is really helpful.

So if I understand correctly, the FRS could easily be analysed using svy in Stata to produce accurate CIs etc – but the issue is that the cluster and strata variables are only available / can only be used in the saferoom.

I'm interested to know why these variables are seen as disclosive? Do you know, by any chance? I'm wondering particularly because they are easily available for other surveys like the MCS.

I was also wondering whether there would be any possibility of non- or less disclosive versions of the cluster and strata variables being made available for the EUL version? For example, in the MCS, the strata variables are named so they give information about the area – but this does not need to be the case in order for them to work with svy commands – they just need to distinguish the strata from one another. If there is a similar issue with the FRS, couldn't the strata / cluster variables be de-identified?

Thanks again for your help,

Tammy

From: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>
Sent: 10 October 2022 16:48
To: Campbell1,T <T.Campbell1@lse.ac.uk>
Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Dear Tammy,

Thank you for your message.

Without further information, I am not sure about the sampling and re-sampling your colleague has mentioned. However, one of the main issues about CIs with the FRS is that the survey uses a complex survey design (there is clustering and stratification). If you calculate CIs without taking the survey design into account your CIs will not be correct – typically we expect the standard errors to be underestimated and as a result the CIs will be narrower than they should be.

There are methods to account for the survey design when calculating standard errors in most statistical packages (Some of these techniques may be the sampling and re-sampling method you mention). In each case, you will need to have the survey designs variables (Strat, PSU) in the dataset. The problem with the FRS, is that these variables are not included in the standard EUL version of the data – you would need access to the Safe Room version. You therefore have some choice to make around what to do. In some cases, you can use something called design factors to approximate the standard errors. The ONS have written a [Guide to calculating standard errors for ONS Social Surveys](#), which summarises the issues.

I hope this helps in some case, though the answer is not straightforward. I am happy to discuss further and to go the DWP for their response.

Best wishes,
Jen

Sent: 10 October 2022 11:45

To: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Subject: RE: JOINING LINK Introduction to the Family Resources Survey workshop

Dear Jen,

Thanks for the training last week.

I was wondering if it would be possible to re-send me the link to the materials that was shared on the day, as I seem not to have saved them?

I also have a follow-up question I hope you / the other presenters could help with –

In the examples and the descriptions of weights, analyses, etc, there wasn't any mention of anything complicated that needs to be done to produce CIs when using the FRS. However, colleagues I've spoken to after the training session have mentioned that their understanding is that there is a very complicated process of sampling and re-sampling that needs to be undertaken for any / all analyses and estimates.

I'd really like to clarify whether this is necessary, or whether just using the weights supplied (and not going below the regional level) is sufficient for accurate analyses. I'm interested in using the survey, but need properly to scope the time needed for a bid – and my colleagues advising that using the FRS is very complicated and time-consuming will make a big difference to this. If it's not necessarily, in fact, that would be great! I hope this makes sense? – any advice from you / the presenters would be really helpful.

Thanks,

Tammy

From: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Sent: 30 September 2022 11:13

To: Jennifer Buckley <jennifer.buckley@manchester.ac.uk>

Subject: JOINING LINK Introduction to the Family Resources Survey workshop

Dear all,

Thank you for booking a place on the Introduction to the Family Resources Survey workshop, next Tuesday (4 October).

The meeting will run in Microsoft teams, [Click here to join the meeting](#)

Meeting ID: 354 671 628 500

Passcode: txg28R

[Download Teams](#) | [Join on the web](#)

After the workshop, please take a few minutes to complete our [Workshop feedback form](#). Your feedback will help inform further events.

You do not need to do anything prior to the workshop. However, the information section below includes some optional preparation that may help you get the most from the workshop.

Information for workshop participants

- We will be using Mentimeter to ask questions in this workshop so you may benefit from having a second device ready for this such as a smartphone.
- The practical exercises include the option of inspecting the datafiles; to take part in this element, you can download the files from the UK Data Service catalogue prior to the workshop (see attached for instructions). However, alternative activities will be available including accessing data during the practical session.
- To access the Family Resources Survey, you need to be registered with the UK Data Service. Though you do not need to do this to join the workshop, you can visit our website for [information about how to register](#). Non-academic users may find this information especially useful as they may need to request a user name and password.

We look forward to you joining us next week.

Best wishes,

Jen Buckley (UK Data Service), Justyna Owen (DWP) and Anna Britton (DWP)

Dr Jennifer Buckley
UK Data Service
Cathie Marsh Institute, University of Manchester

jennifer.buckley@manchester.ac.uk

This document is strictly confidential and is intended only for use by the addressee. If you are not the intended recipient any disclosure, copying, distribution or other action taken in reliance of the information contained in this email is strictly prohibited.

Any views expressed by the sender of this message are not necessarily those of the Department for Work and Pensions.

If you have received this transmission in error please tell us and then permanently delete what you have received.

This email was scanned for viruses by the Department for Work and Pensions antivirus services and was found to be virus free.

Please note: Incoming and outgoing email messages are routinely monitored for compliance with our Email Policy.

This document is strictly confidential and is intended only for use by the addressee. If you are not the intended recipient any disclosure, copying, distribution or other action taken in reliance of the information contained in this email is strictly prohibited.

Any views expressed by the sender of this message are not necessarily those of the Department for Work and Pensions.

If you have received this transmission in error please tell us and then permanently delete what you have received.

This email was scanned for viruses by the Department for Work and Pensions antivirus services and was found to be virus free.

Please note: Incoming and outgoing email messages are routinely monitored for compliance with our Email Policy.
