

LFS User Workshop



Martina Helme / Simon Woodsford
ONS Social Survey Operations – Labour Force Survey Research Team

20 May 2021

Session 1

Survey history and design

- Background/history
- Methodology
 - (questionnaire content, sampling, imputation, data collection)
- Changes made in response to COVID-19

What is the Labour Force Survey?

- A survey of the employment circumstances of the UK population
- Largest continuous household survey in the UK
- Over 1000 variables (questions and derived variables) covering:



Overall Sample Design



- People living in private households and nurses in NHS accommodation
- Students in halls of residence or boarding schools should be included at their parents address
- The Postcode Address File (PAF) is the main sampling frame
- NHS sampling frame and the Telephone Directory also used
- Other communal establishments are excluded from the sample
- Random sample, stratified by postcode
- Representative of the whole UK population (with a few exceptions)

LFS Sample

- Around 75,000 households sampled each quarter for the LFS and around 40,000 households take part
- This equals around 90,000 individuals every quarter
- Sample intended to be representative of the entire population
- Therefore each case weighted by age, sex and geography
- $65 \text{ million people} \div 90,000 =$ each respondent is equivalent to approximately 800 people
- 75+ households and multiple households will have
- The LFS is not of a sufficient size to allow for analysis at below regional level



The LFS Wave Structure

JM20	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
AJ20	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
JS20	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
OD20	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
JM21	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5

- Respondents are interviewed for five successive waves at three-monthly intervals
- 20% of the sample is replaced every quarter

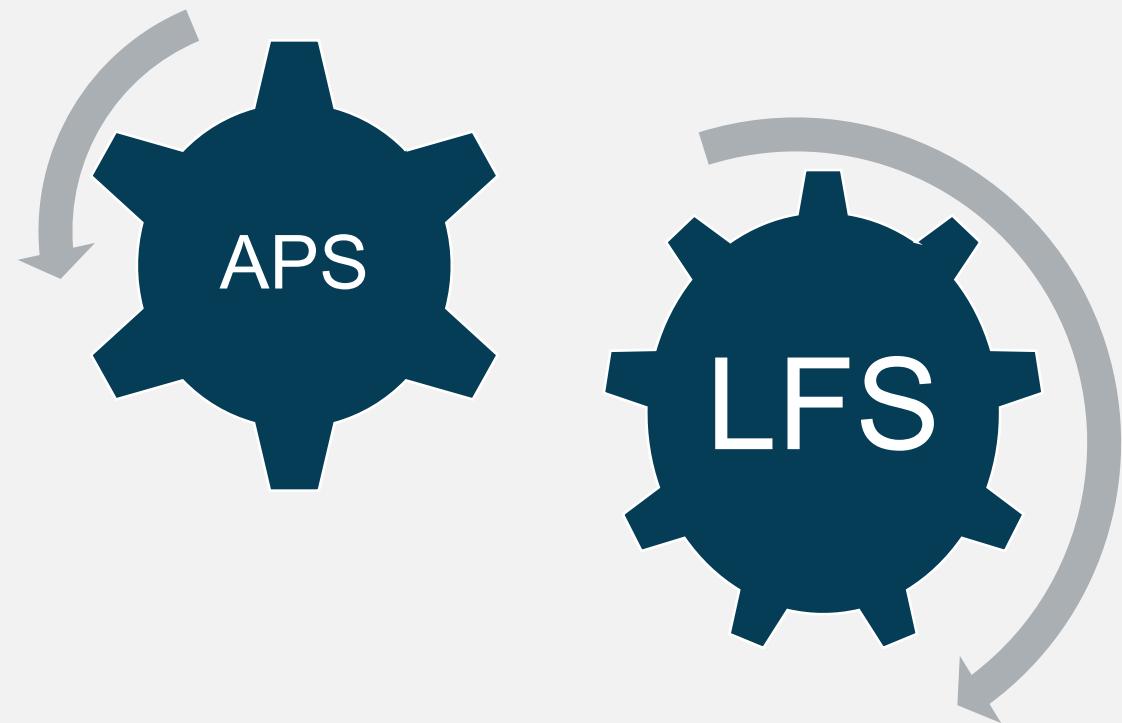


Annual Population Survey (APS) Sample

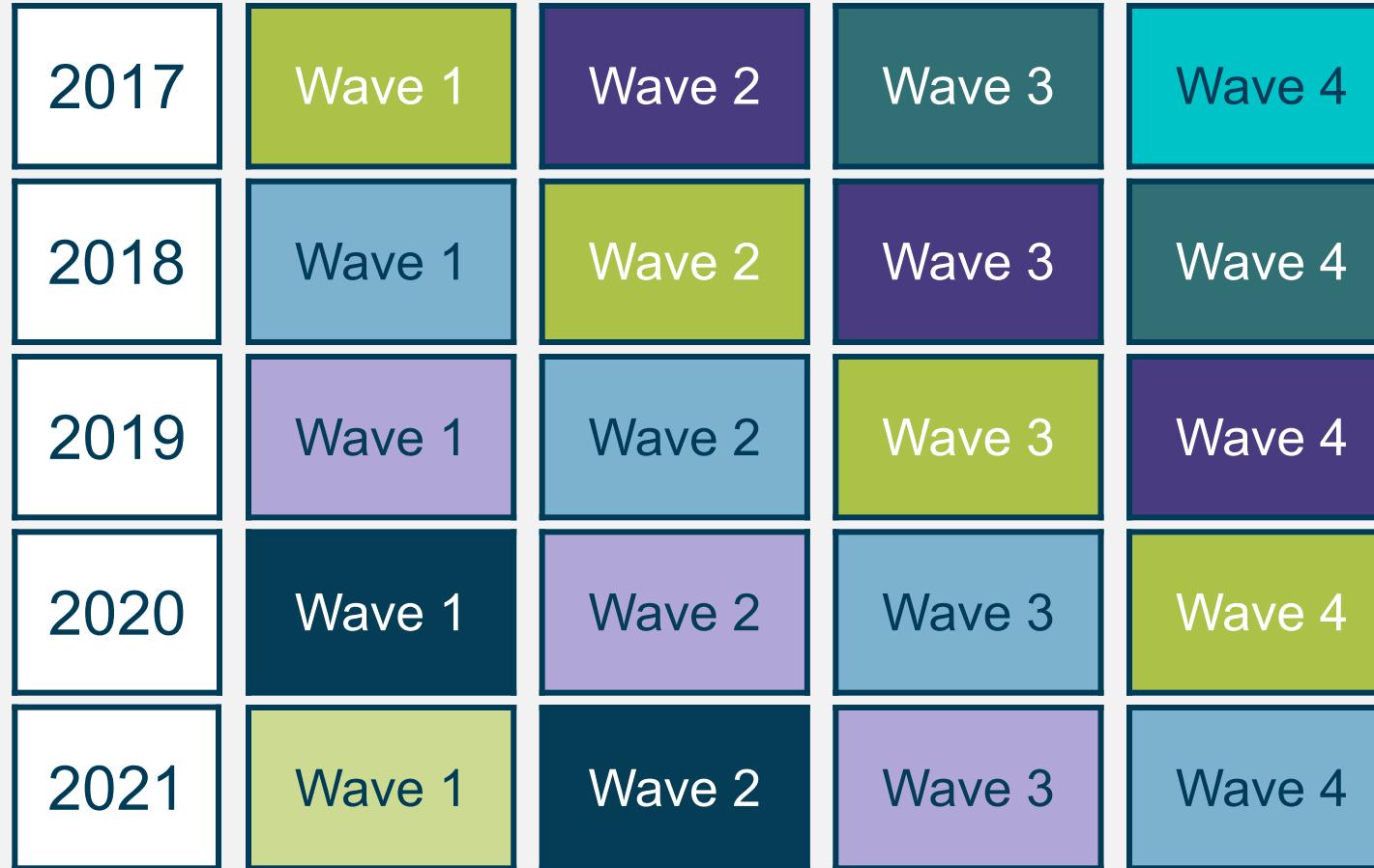
- Main purpose of the APS is that analysis can be conducted at **Local Authority level**
- The ‘boost’ cases ensure a set number of cases for each Local Authority are targeted
- Around **200,000 households** are contacted...
- Around **140,000 households** take part,...
- Equates to around **280,000 individuals**



What is the connection between LFS and APS?



The LFS Boost Wave Structure



- England, Scotland and Wales all fund Local Labour Force Survey interviews (LLFS, Boost or Enhancement cases)
- Respondents are interviewed for four successive waves at 12-monthly intervals
- 25% of the sample is replaced every quarter



The APS Wave Structure

The APS is made up of LFS W1 and W5, and the LFS Boost W1-4

2017	LFS W1	LFS W5	Boost W1	Boost W2	Boost W3	Boost W4
2018	LFS W1	LFS W5	Boost W1	Boost W2	Boost W3	Boost W4
2019	LFS W1	LFS W5	Boost W1	Boost W2	Boost W3	Boost W4
2020	LFS W1	LFS W5	Boost W1	Boost W2	Boost W3	Boost W4
2021	LFS W1	LFS W5	Boost W1	Boost W2	Boost W3	Boost W4



Imputation

Rolled forward imputation

- A successful interview at one wave followed by a non-response at the next
- Rolled forward cases appear as IOUTCOME = 6 on LFS and APS 'person' and 'household' micro data
- Rolled forward imputation only occurs once and only for LFS main cases

Donor imputation (employment focused)

- Non-response at wave 1 or a second consecutive wave of non-response
- Non-responder cases appear as IOUTCOME = 3 only on LFS and APS 'household' micro data
- Donor imputation occurs for LFS and APS cases



Data collection (pre-COVID)

- CAPI (Computer Assisted Personal Interviewing)

Wave 1

Face-to-Face

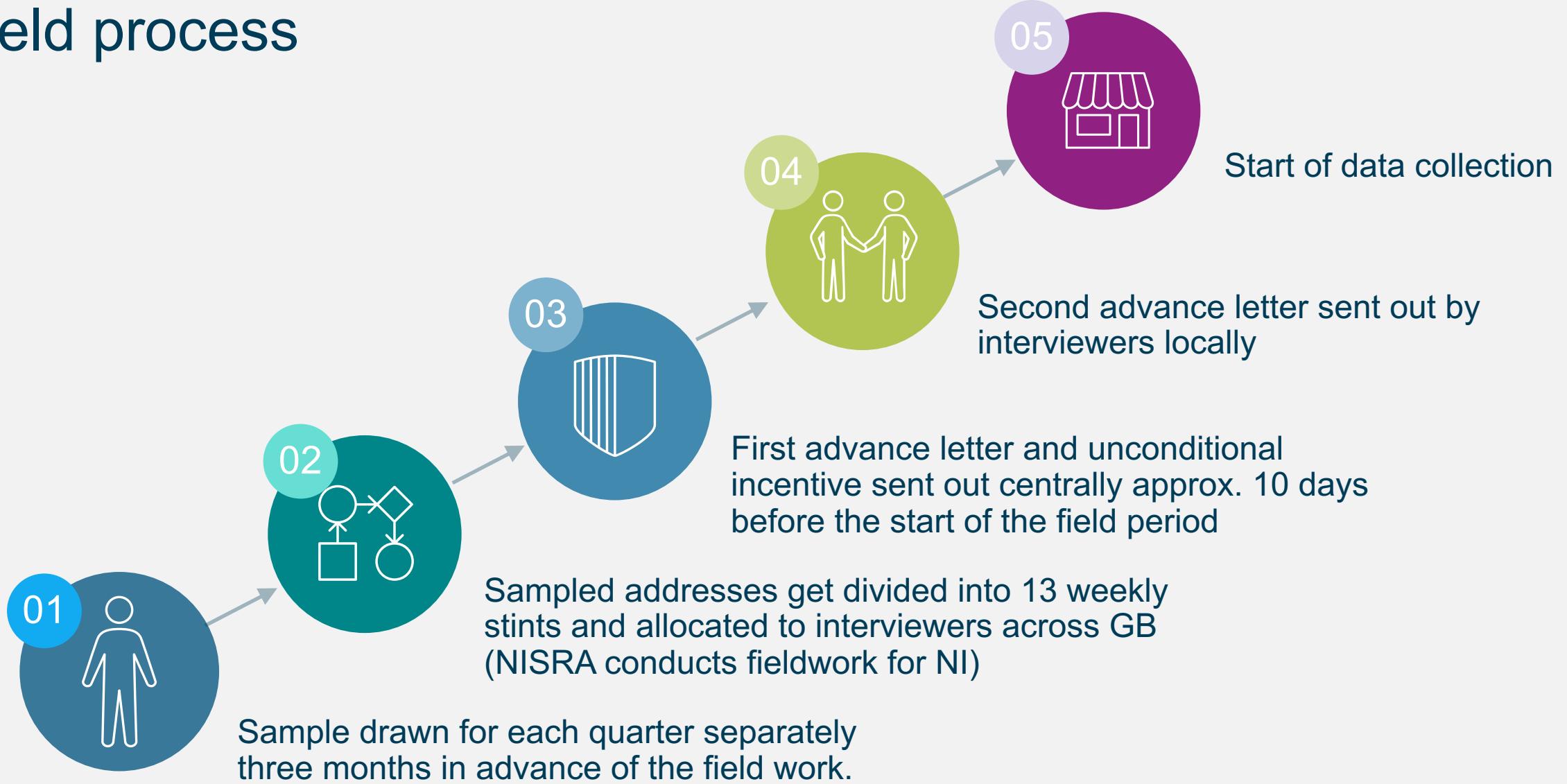
Waves 2-5

Telephone

- Scotland north of Caledonian Canal all done by telephone unit
- Approximately 600 field interviewers and 200 telephone interviewers
- Respondents can answer for others in the household = **Proxy Interviews**
- Around 1/3 of LFS data collected in this way



Field process



Variable frequency

- Certain variables are wave specific
- Others are quarter specific

SIDFTFQN
Please use showcard
Which of the options on this card best describes how you think of yourself? Please just read out the number next to the description.

Enter a numeric value between -9 and 99

ONS ✓ GOV

Applies to respondents who face interview
IF AGE > 15 and TELEPH:

MATLVE
...and was your maternity leave (in the week ending Sunday the ...) ...
Code one only

1 a legal entitlement (statutory maternity leave/absence),
2 or leave your employer has allowed (contractual maternity leave)?
3 don't know

ONS ✓ GOV ✓ SPL ✓ EUL ✓

Applies if on maternity leave
IF YLESS6=3 AND SEX=2

PERSONAL ONLY

EQ
W1F
IHS

UK
AJ
MAIN

Data quality at point of collection (1)

We provide showcards for questions such as nationality, ethnicity and sex ID



SID1
{For 1 1 1, please use showcard 1}

Which of the options on this card best describes how you think of yourself?
Please just read out the number next to the description.

The numbers on each card are different for each person



Office for National Statistics

Data quality at point of collection (2) – Soft checks

Forms Answer Navigate Options Help Show Watch Window

LFS0321 Household_Information Person[1] Person[2] Household_Reference_Person Benefit_unit Information_about_next_interview Timing_administration_and_household_outcome

MR JOE BLOGGS (1)
What period did this cover?

1. One week
 2. Two weeks
 3. Three weeks
 4. Four weeks
 5. Calendar month/five week month
 7. Two months
 8. Eight times a year
 9. Nine times a year
 10. Ten times a year
 13. Three months

26. Six months
 52. One year

Active Signal

MR JOE BLOGGS (1)
Gross weekly pay = £70000.00. This seems high.
Please check the amount and period covered with the respondent.
Press 'S' to suppress this message.

Gross99 70000 Net99
GrsExp
GrsPrd 1 NetPrd
BandG BandN
UsGrs99 UsNet99
UsuGPay UsuNPay
IncChk

Questions involved

QIndiv[1].QRouting.QEarning.GrsPrd	Value	One week
QIndiv[1].QRouting.QEarning.Gross99	70000	

Suppress Close Goto

Data quality at point of collection (3) – Hard checks

Forms Answer Navigate Options Help Show Watch Window

LFS0321 Household_Information Person[1] Person[2] Household_Reference_Person Benefit_unit Information_about_next_interview Timing_administration_and_household_outcome

Age at end of reference week should be
31
Check with respondent and enter agreed age.

① If D.O.B. not known record respondent's (or your) estimate of age.
② Enter code 99 if aged 99 or more

Enter a numeric value between 0 and 99

	PerNo	RelTxt	Title	FstNme
Line[1]	1	person1	Mr	Joe
Line[2]	2			
Line[3]				
Line[4]				
Line[5]				
Line[6]				
Line[7]				
Line[8]				

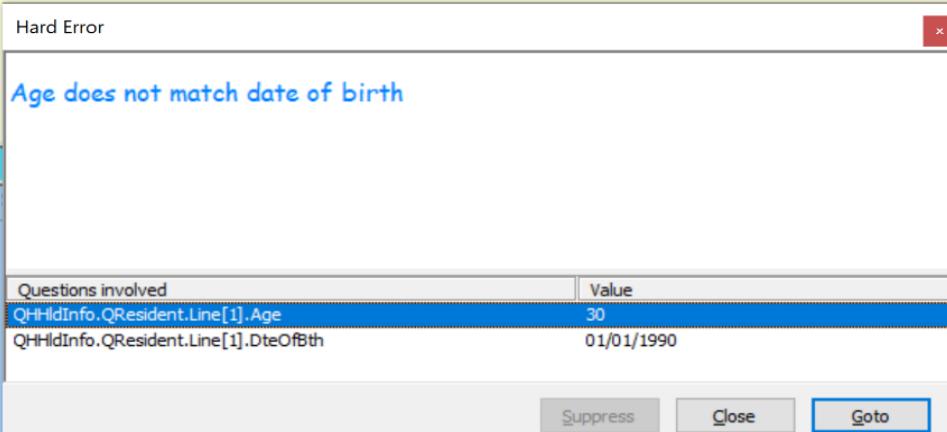
Hard Error

Age does not match date of birth

Questions involved	Value
QHhdInfo.QResident.Line[1].Age	30
QHhdInfo.QResident.Line[1].DteOfBth	01/01/1990

MarChk Liv12W HRPId Estimate

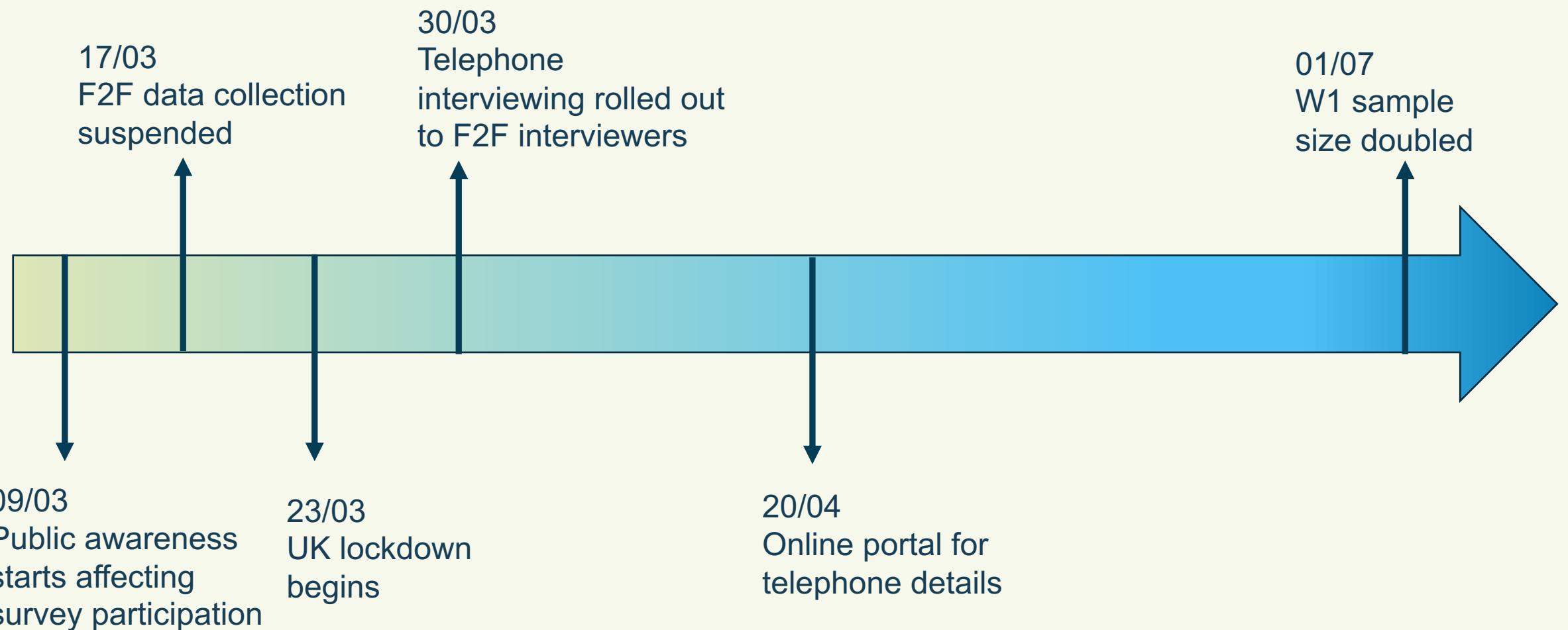
Suppress Close Goto



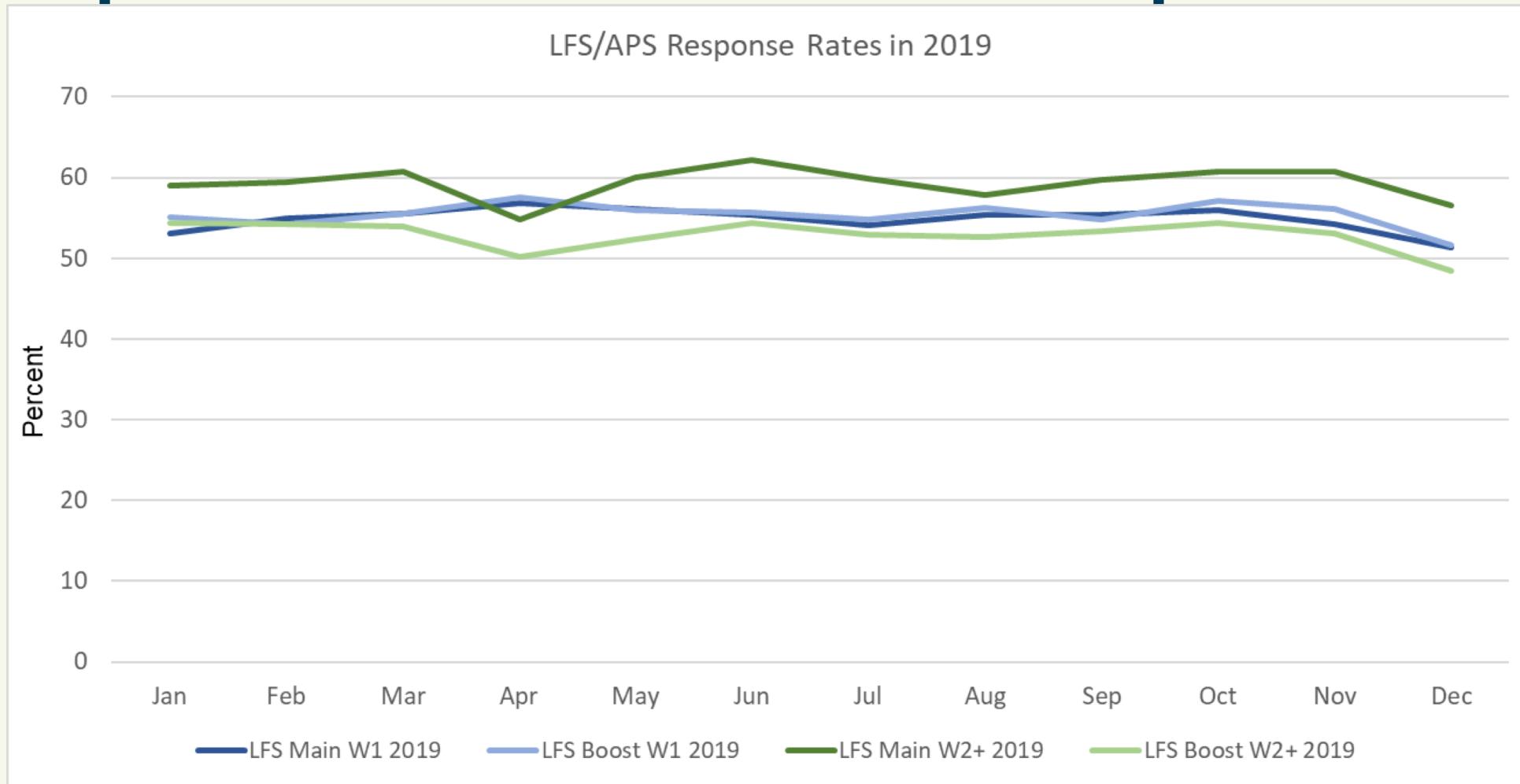
Changes to the LFS/APS since the start of the pandemic



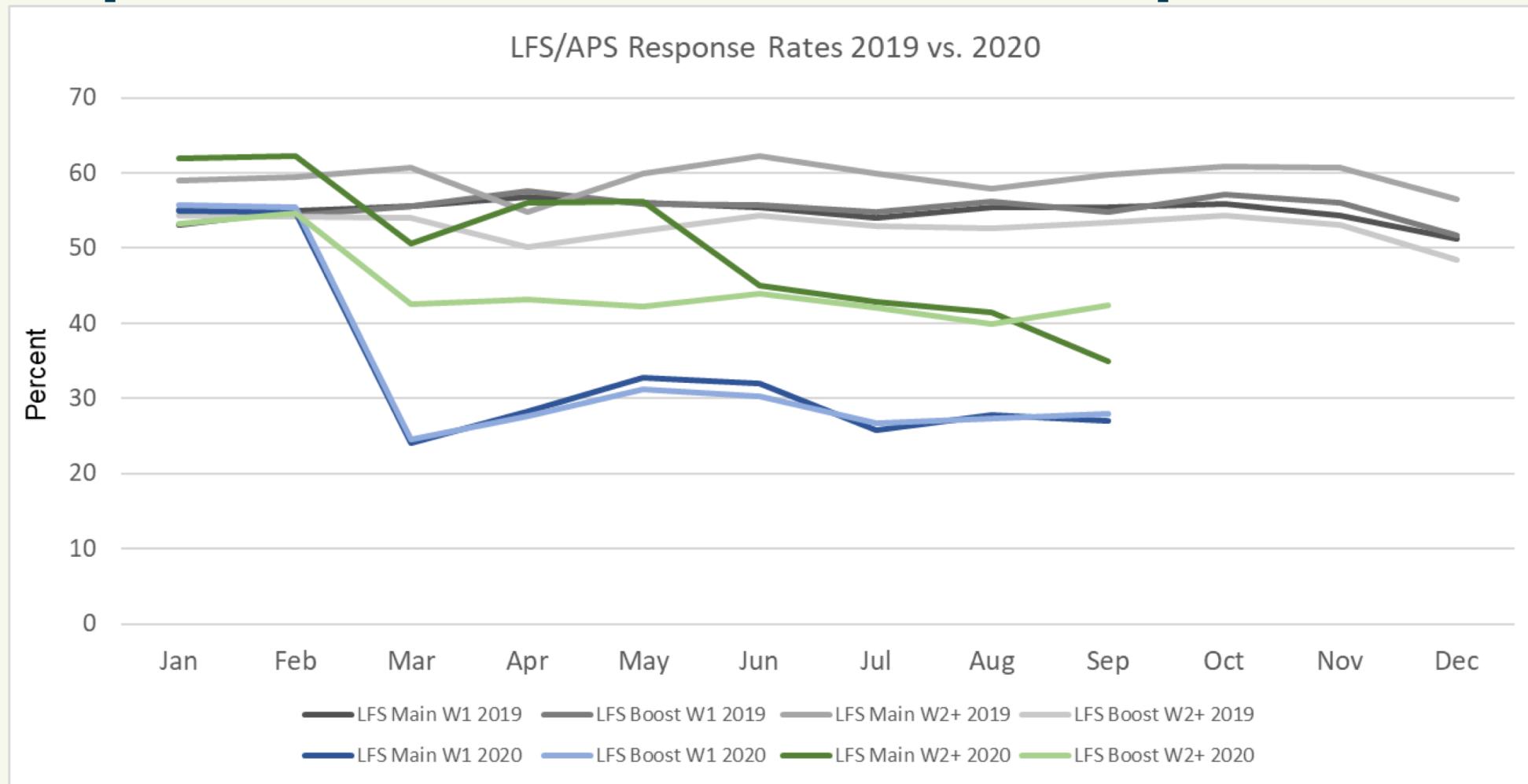
Timeline of events



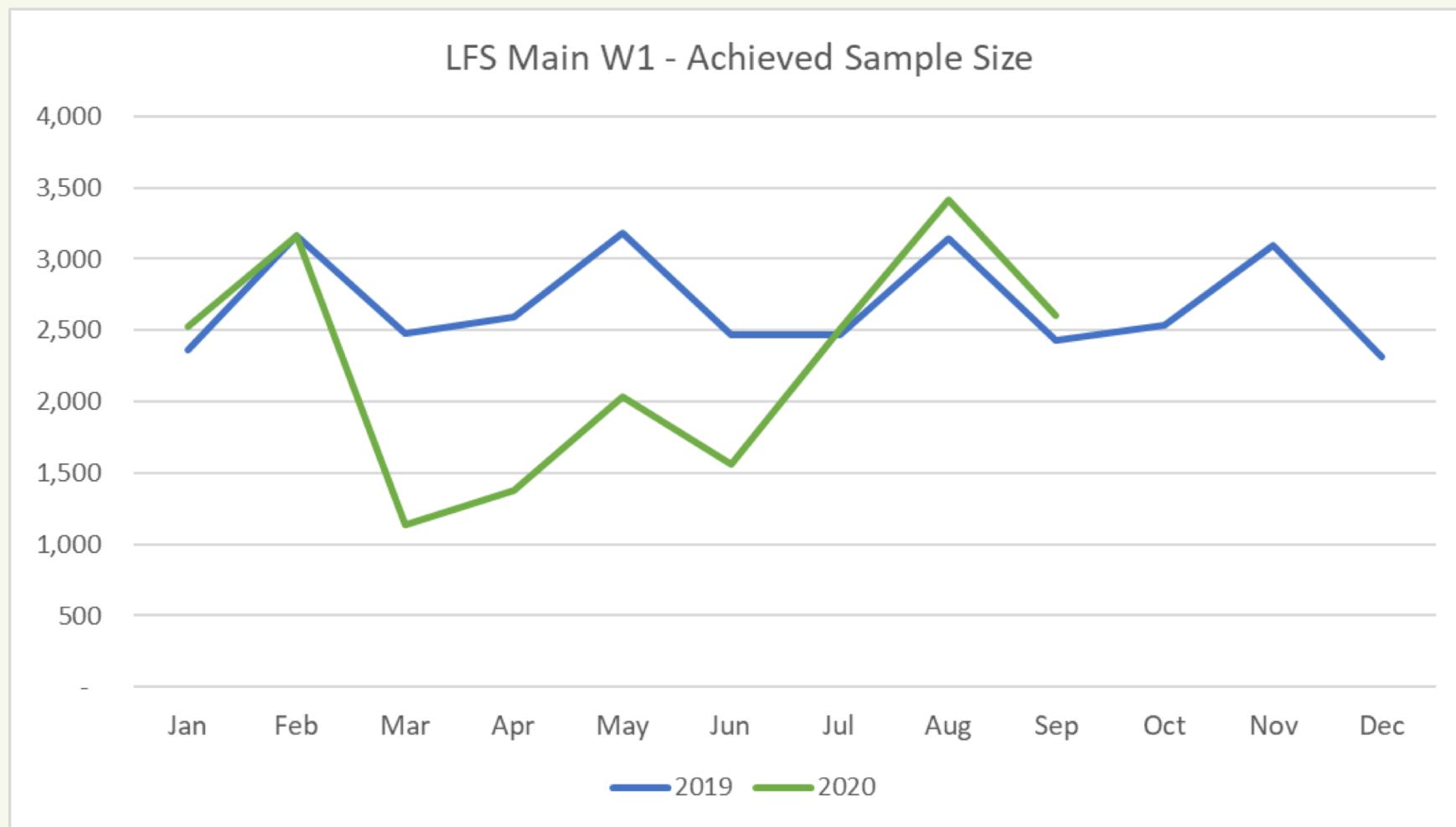
Impact of COVID-19 on LFS Response Rates



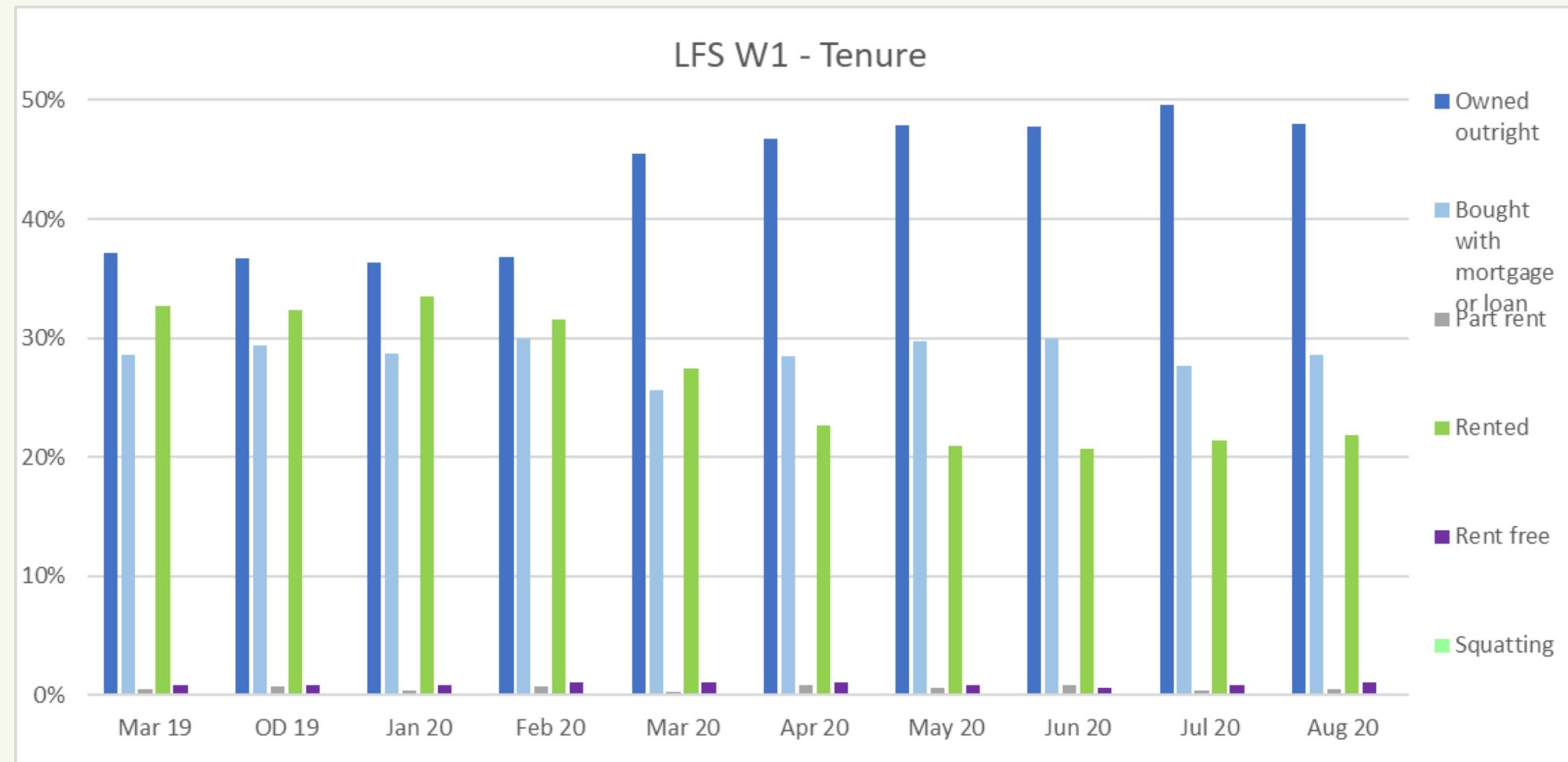
Impact of COVID-19 on LFS Response Rates



Impact of COVID-19 on LFS achieved sample size



Impact of COVID-19 on LFS estimates



Impact of COVID-19 on LFS estimates

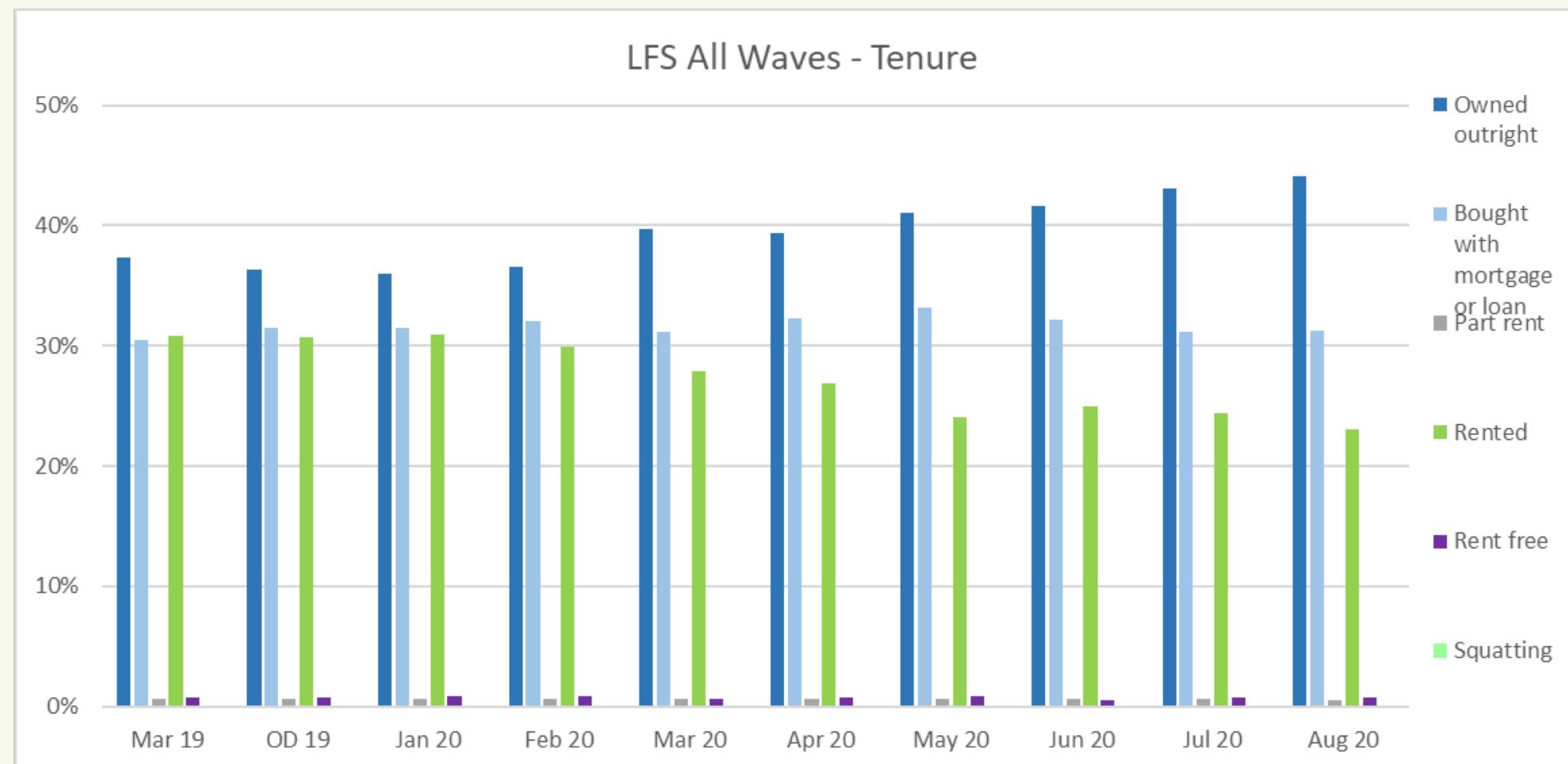


Figure 1: In June to August 2020, the adjusted employment rate was 1.0 percentage points lower than under usual weights

UK employment rates (aged 16 to 64 years), not seasonally adjusted, between November to January 2018 and June to August 2020



Source: Office for National Statistics – Labour Force Survey



Figure 2: In June to August 2020, the adjusted unemployment rate was 0.2 percentage points higher than under usual weights

UK unemployment rates (aged 16 years and over), not seasonally adjusted, between November to January 2018 and June to August 2020

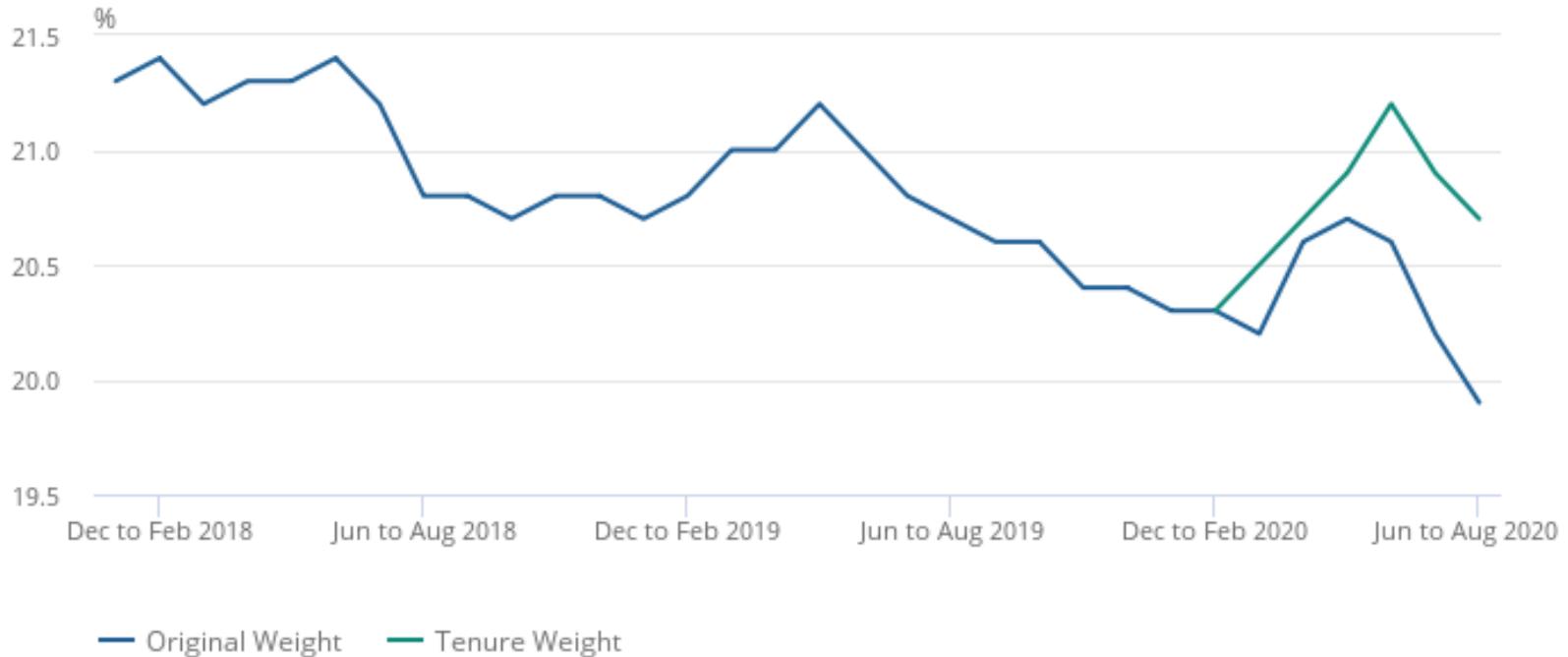


Source: Office for National Statistics – Labour Force Survey



Figure 3: In June to August 2020, the adjusted economic inactivity rate was 0.8 percentage points higher than under usual weights

UK economic inactivity rate (all people aged 16 to 64 years), not seasonally adjusted, between November to January 2018 and June to August 2020



Source: Office for National Statistics – Labour Force Survey



Questionnaire up-dates

- **Sickness absence**
- Reasons respondent was **away from work or worked fewer hours** in the reference week
- Reasons respondent **worked more hours than usual** in the reference week
- Reasons why the respondent was **paid less than usual**
- **State benefits** (furlough, SEISS)
- **Home/flexible working**



Survey changes since AJ20 as a response to COVID-19

- Collection of email address for each household (Main W1-4, Boost W1-3) to allow respondents to be contacted via email in case the postal system is disrupted.
- Telematching of entire W1 sample
- Telephone interviews for entire sample
- Online portal to collect telephone numbers of respondents
- LFS W1 sample size increase by 100% for JS20 and OD20, and by 60% from AJ21 onwards (inc. Knock to Nudge)
- Adjusted weight calibration using tenure to address bias in the estimates with a further weight adjustment to be applied later this year.



What the future holds

- Reweighting exercise on going this year.
 - ✓ Weight adjustment based on comparison with [RTI data](#)
 - ✓ Detailed methods paper to be published in a few weeks
- Currently working on a roadmap back to in-house/F2F interviewing (trial planned for later this year)
- Ongoing development of the Labour Market Survey
 - ✓ LMS together with a combination of admin data will replace the LFS
 - ✓ On-line first
 - ✓ Mixed mode data collection
 - ✓ Latest results from the 2019 mixed-mode test published on ONS website

More information on the LMS

- [LMS Technical Report](#)
- [LMS Characteristics Report](#)
- [LMS Comparative Estimates Report](#)

Session 2

Considerations when using the data

- Datasets - scope/limitations
- What to know about weights
- What to know about variables
- Data linking/pooling

Dataset overview

- **LFS person level** data - quarterly
- **LFS household level** data - now all quarters
- **LFS longitudinal** data – 2 quarters and 5 quarters
- **APS person level** data - quarterly, covering 12 months
- **APS household level** data – annually
- **APS 3 year pooled** data – annually

- Government level files
- Public files (now only end user licence)

- Data accessible from UK Data Archive/SRS



LFS Person level data

- Purpose** → Personal level analysis at a national or regional level
- Frequency** → Quarterly – official released around 6 weeks after the end of data collection period, available through UKDS/SRS several weeks later
- Time periods** → Normally released for calendar quarters only (e.g. Jan to March)
- Availability** → AJ92 to AJ01 (2007 weight)
JS01 to AJ11 (2014 weight)
JS11 onwards (2018 weight)
- Weights** → PWT (person analysis) and PIWT (earnings analysis)
- Limitations** → Not for analysis below regional level
No earnings information for self employed
Some variables are specific to certain quarters
Not for Well-Being, Sexual Orientation, Family/Household analysis

LFS longitudinal data for 2 quarters

- Purpose** → Conducting analysis on flows over 3 months
- Frequency** → Quarterly – official released around 6 weeks after the end of data collection period, available through UKDS/SRS several weeks later
- Time periods** → Two calendar quarters only (e.g. JM to AJ)
- Availability** → Q197-Q297 to Q306 –Q406 (2007 weight)
Q107-Q207 to Q211-Q311 (2014 weight)
Q311-Q411 onwards (2018 weight)
- Weights** → LGWT
- Limitations** → Only available for person level analysis
Only includes respondents who responded in both quarters
Only includes variables asked in both quarters
Doesn't contain all variables available on LFS person datasets

LFS longitudinal data for 5 quarters

- Purpose** → Conducting analysis on flows over 12 months
- Frequency** → Quarterly – official released around 6 weeks after the end of data collection period, available through UKDS/SRS several weeks later
- Time periods** → Five calendar quarters only (e.g. JM to AJ)
- Availability** → Q1'05-Q1'06 to Q2'10-Q2'11 (2010 weight)
Q3'10-Q3'11 to Q2'11 to Q2'12 (2014 weight)
Q2'12 –Q3'13 onwards (2018 weight)
- Weights** → LGWT
- Limitations** → Only available for person level analysis
Only includes respondents who responded in all 5 quarters
As a result sample size drastically reduced
Doesn't contain all variables available on LFS person datasets

APS Person level data

- Purpose** → Personal level analysis below regional level or for breakdowns on certain characteristics
- Frequency** → Quarterly – official released around 11 weeks after the end of data collection period, available through UKDS/SRS several weeks later
- Time periods** → Rolling 12 months (e.g. January to December, April to March)
- Availability** → JD04 to JD11 (2014 weight)
A11M12 onwards (2018 weight)
- Weights** → PWTA (A05M onwards, before then PWAPS, PWLFS)
PIWTA (for earnings analysis) – JD only from 2012-17, all periods from 2018 onwards
SIDWT (for Sexual Orientation analysis) – 2011 onwards
NPWT (for Personal Well-being analysis) – 2011 onwards
- Limitations** → Some LFS quarterly specific variables not available on APS

APS longitudinal data for 2 years

- Purpose** → Personal level analysis on flows below regional level over 12 months
- Frequency** → Annual
- Time periods** → January to December for two consecutive years
- Availability** → JD12 - JD13 onwards (2018 weight)
- Weights** → LGWT
- Limitations** → Only at person not household level
Only for a sub-set of variables
Only includes respondents who have responded in both years

APS 3-year pooled person level data

- Purpose** → Personal level analysis at lower level geographies and for certain topics where achieved sample is much smaller
- Frequency** → annual – normally released around July every year
- Time periods** → January to December covering 3 years
- Availability** → JD13 to JD15 (2016 weight)
JD14 - JD16 to JD15 - JD17 (2017 weight)
JD16 - JD18 to JD17 to JD19 (2018 weight)
- Weights** → PWTA**C, SIDWT**C and NPWT**C
- Limitations** → Point in time estimates and not to be used for any time series analysis

LFS Household level data

- Purpose** → household/family level analysis at a national or regional level
- Frequency** → Quarterly - Officially released around 8 weeks after the end of data collection period, available through UKDS/SRS several weeks later
- Time periods** → Pre 2004 – Apr-Jun periods only;
Between 2004 and 2011 Apr-Jun and Oct-Dec only;
Quarterly from AJ12 onwards
- Availability** → AJ96 to AJ01 (2007 weight)
AJ02 to AJ11 (2014 weight)
OD11 onwards (2018 weight)
- Weights** → PHHWT
- Limitations** → Not suitable for household earning analysis as no earnings household weight

APS Household level data

Purpose → Household and Family level analysis below regional level or for breakdowns on certain characteristics

Frequency → Annually - released around July every year

Time periods → January to December

Availability → JD04 to JD05 (2010 weight)
JD06 to JD11 (2014 weight)
JD12 onwards (2018 weight)

Weights → PHWTA

Limitations → Some LFS variables are not available on APS datasets, particularly LFS quarterly specific variables.
Not suitable for household earning analysis as no earnings household weight



What to know about LFS/APS weights

- LFS/APS collects data on a representative sample of UK population
- Grossing weights are created to estimate for the rest of the UK population we don't sample and the survey design
- Weighting groupings are by age, sex and geography
- Each individual represents around 800 people of similar characteristics (LFS person weights)
- Sum of the weights make up the eligible population in the UK
- Critical to make sure you use the appropriate weight for the analysis you are doing

What to know about variables

- Datasets contain questionnaire variables; derived variables; geography variables; system variables, all can be used for analysis
- Details of all variables can be found in UG2, UG3 and UG4
- Not all variables will be on all datasets
 - Well-Being; Sexid; Smoking; Veterans; General Health; Quality of work – APS only
 - Family and Household variables – Household datasets only
 - Quarterly specific variables – LFS datasets only
 - Reduced variables list on Longitudinal datasets
- Understand the variables you are using (e.g. periods; type; routing)
- Datasets will include ‘missing values’
 - -9 means not routed to that specific question
 - -8 means respondent refused to answer question or answered don’t know

Data linking/pooling

- It isn't advised to carry out any dataset linking or pooling yourself
- Pitfalls of doing this include no suitable weights and duplicate cases
- The LFS and APS datasets available should allow you to carry out all permitted analysis
- Data linking procedure links cases across waves using
 - ✓ A unique case identifier (not available via UKDS)
 - ✓ Sex
 - ✓ DoB
- Name and address information not available on any LFS/APS datasets
- Pseudo identifiers available on LFS UKDS datasets

Which dataset to use when?

- Analysis on sexual orientation → APS person (SIDWT)
 - Analysis on wellbeing → APS person (NPWT)
 - Analysis on earnings below reg. level → APS Person (PIWTA)
 - Analysis on other topics below reg. level → APS Person (PWTB)
-
- Analysis on earnings at or above reg. level → LFS Person (PIWT)
 - Analysis on other topics at or above reg. level → LFS Person (PWT)
-
- Longitudinal analysis below regional level → APS Person 2yr (LGWT)
 - Longitudinal analysis at or above regional level → LFS 2Q or 5Q (LGWT)
-
- Analysis at hhold/family level below reg. level → APS Hhold (PHHWTB)
 - Analysis at hhold/family level at or above req. level → LFS Hhold (PHHWT)

How to get started...

- Always refer to LFS/APS user guides first:
 - UG1 – LFS (and some APS) Methodology
 - UG2 to UG4 – Questionnaire and Derived Variables
 - UG5 – Classifications (SIC/SOC)
 - UG 6 - APS
 - UG8 – Household and family variables
 - UG9 – Eurostat
 - UG10 – Advice on which dataset to use
 - UG11 – Longitudinal
- Key things depending on your analysis:
 - Use the correct dataset
 - Use the correct weight
 - Use the correct variables

Q & A