Mitigating the impact of the pandemic on data quality in ONS social surveys, with a specific focus on the Labour Force Survey

Salah Merad and Martina Helme,
Office for National Statistics, UK
salah.merad@ons.gov.uk
martina.helme@ons.gov.uk

International Total Survey Error Workshop (ITSEW) 2021

Outline

- Brief overview of LFS sample design and weighting
- Changes to the LFS since the start of the pandemic
- Impact on LFS response and sample bias
- Adjustments to reduce bias
- Implementing the adjustments: impact on 2020 estimates
- Future developments

Brief overview of LFS sample design and weighting

Office for National Statistics

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

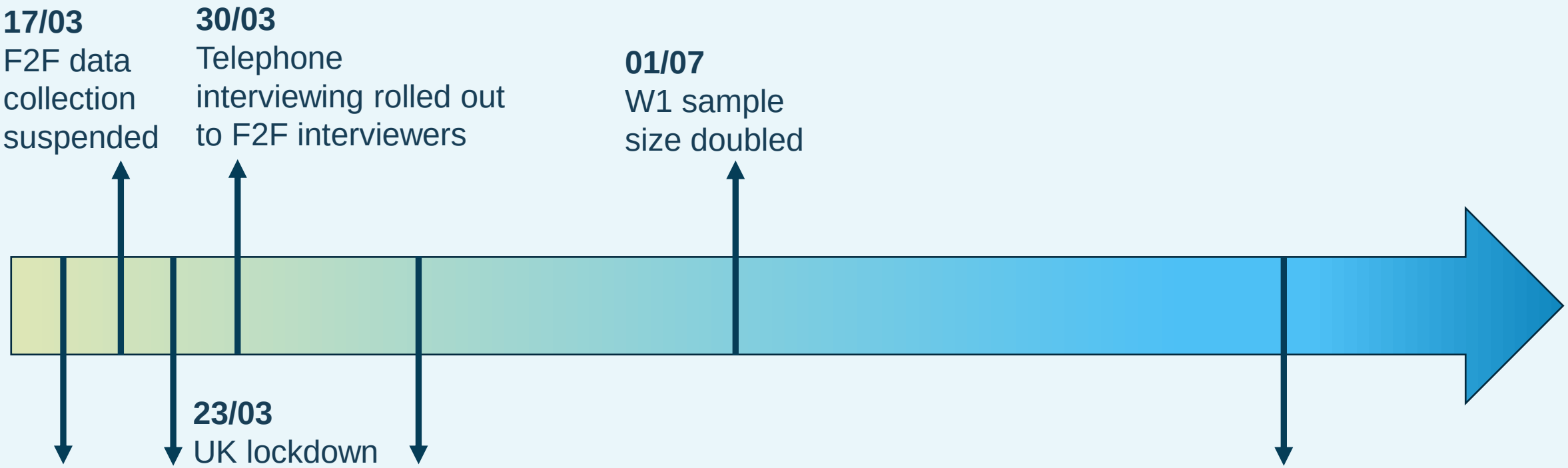
LFS sample design and weighting

- The LFS is a rotating panel probability sample survey (5 waves in each quarter)
- Non-responders in wave k, k=2,…,5, who responded in wave k-1, are imputed
- The data are weighted to the UK resident population by age, sex and geography
  - Students in halls of residence are part of the population but their data are collected from parents
- The population totals for the groups used in weighting are obtained from population projections by single year of age in each local authority that are based on mid-year population estimates

Changes to the LFS data collection since the start of the pandemic

Office for National Statistics

Timeline of events



17/03
F2F data collection suspended

30/03
Telephone interviewing rolled out to F2F interviewers

01/07
W1 sample size doubled

09/03
Public awareness starts affecting survey participation

23/03
UK lockdown begins

20/04
Online portal for telephone details

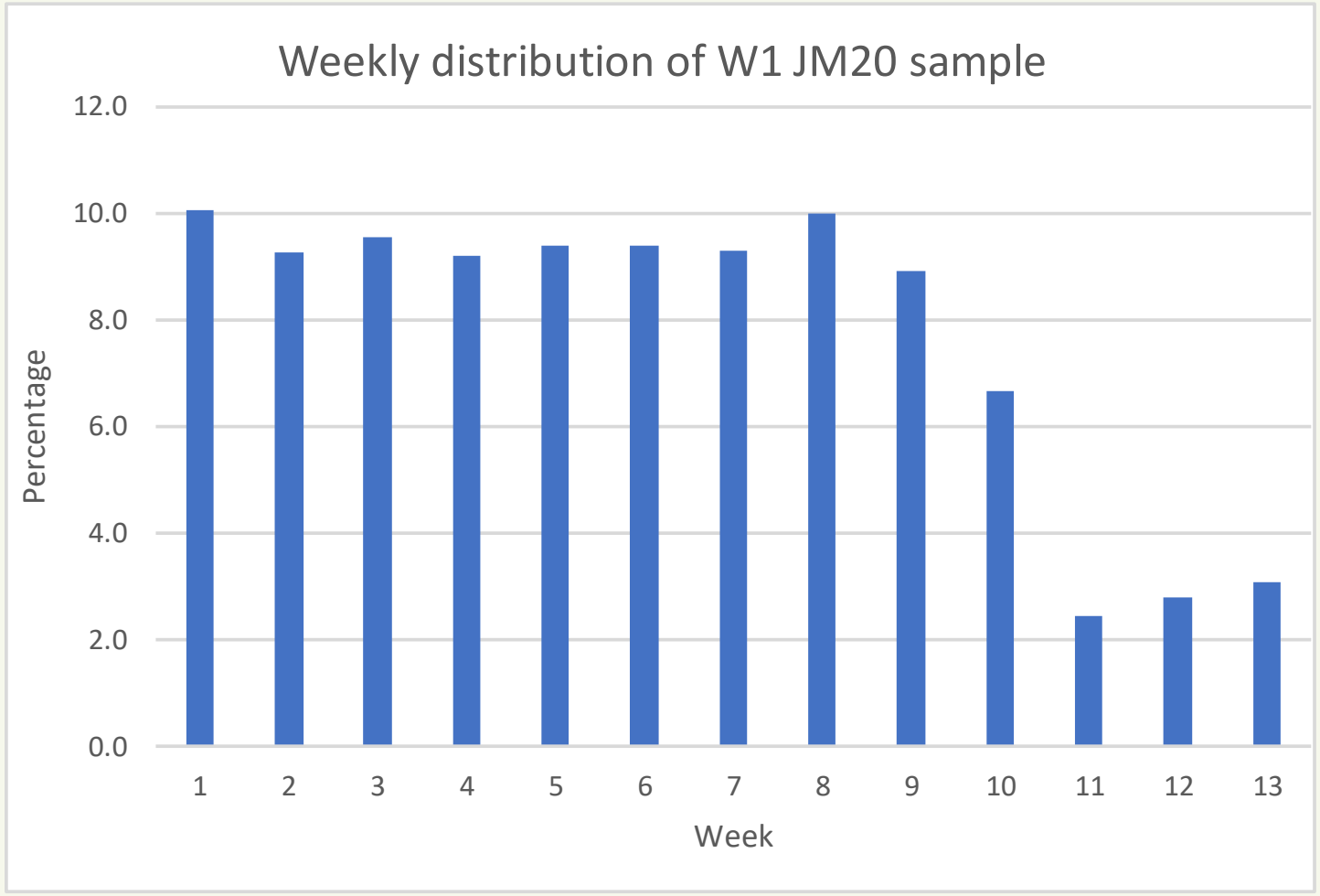
01/04
Roll-out of ‘Knock-to-Nudge’ alongside W1 sample reduction

What is ‘Knock to Nudge (KtN)’?

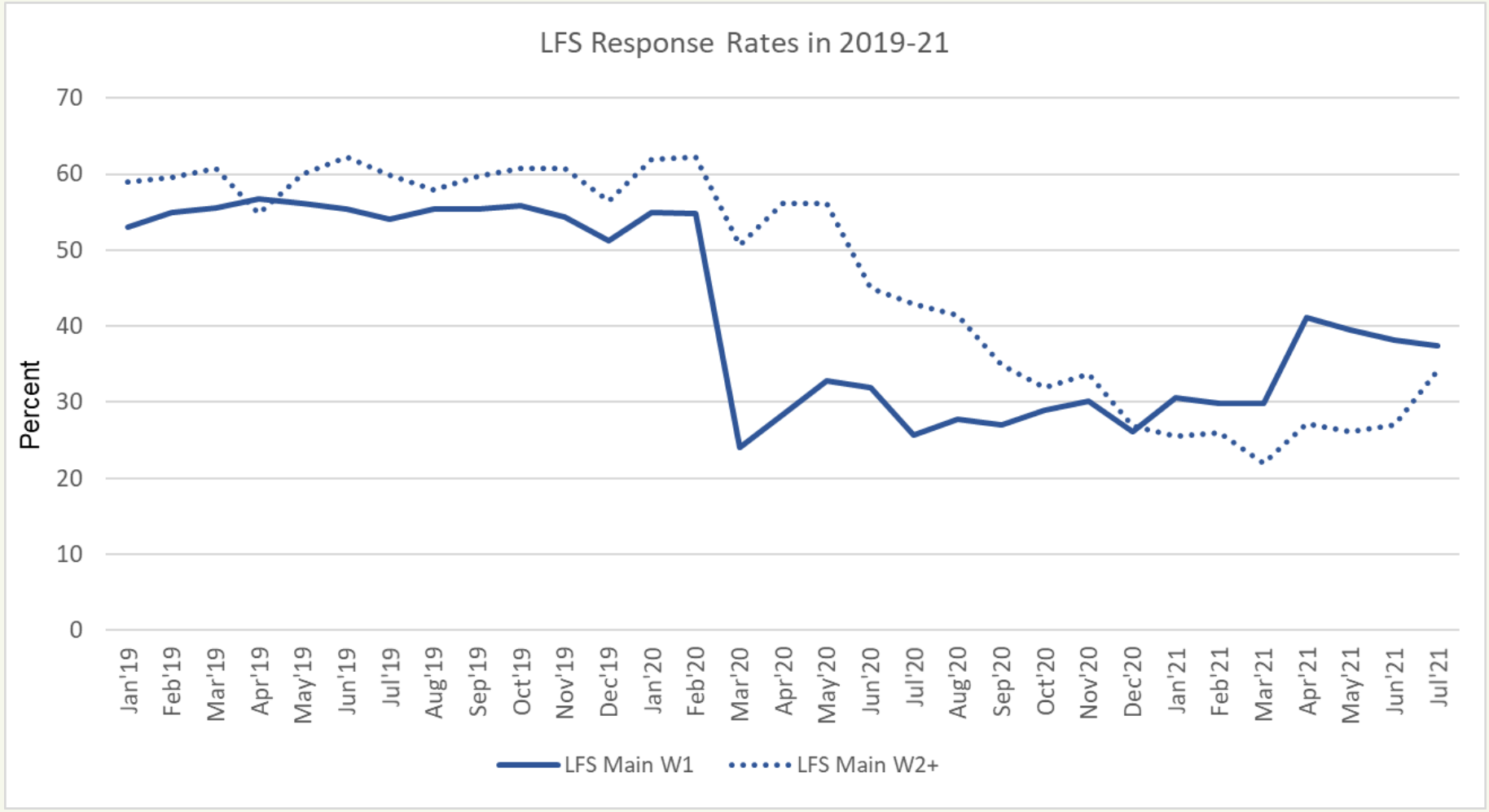
- Cases without phone contact details are assigned to an F2F interviewer
- Interviewer will make 3 call attempts to obtain phone contact details at the doorstep followed by a telephone interview if successful
- KtN is currently conducted on wave 1 cases, but only on wave 2-5 where phone contact details were incorrect or supplied late.

Impact on response and sample bias

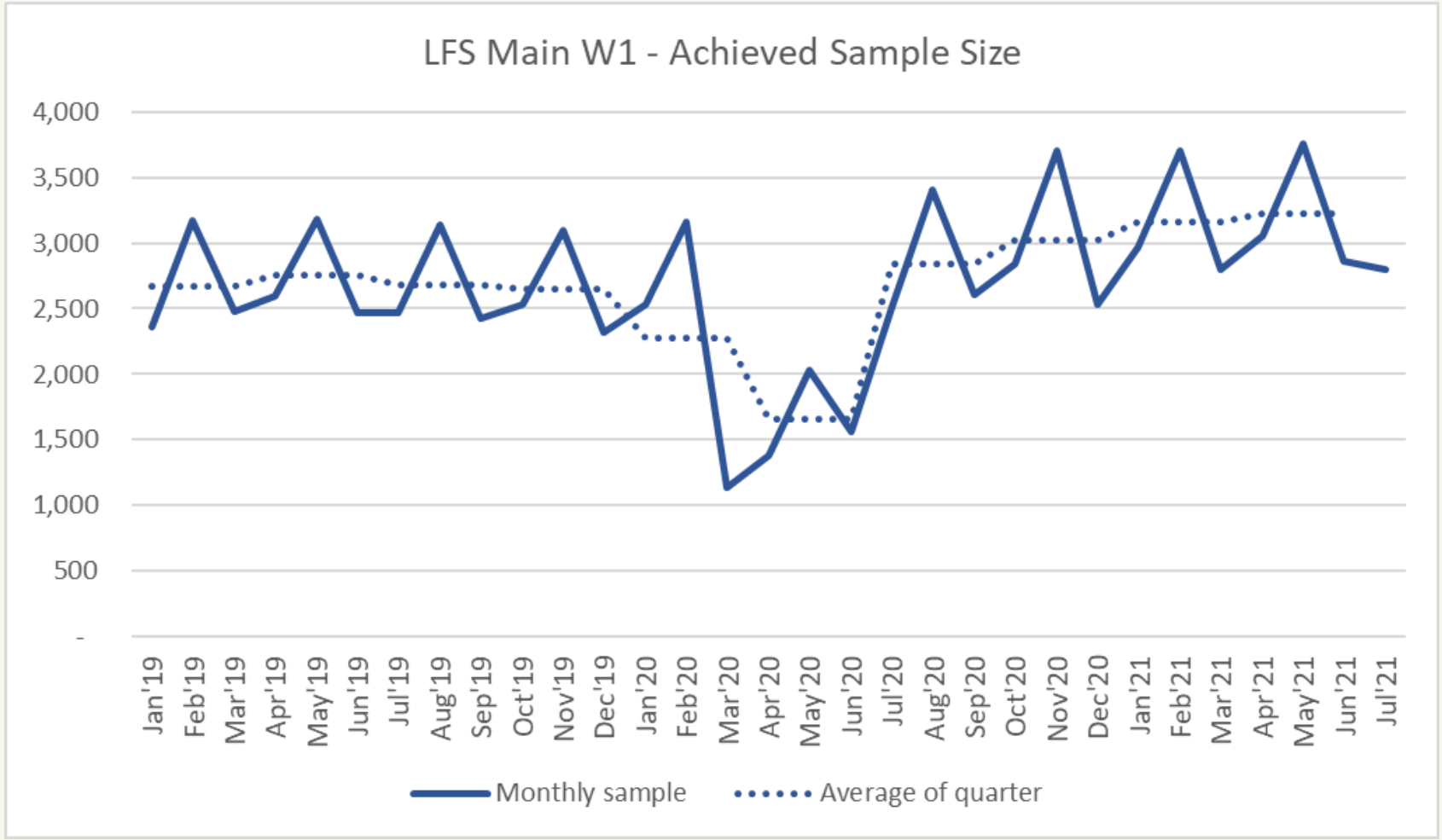
Impact of weekly distribution of Wave 1



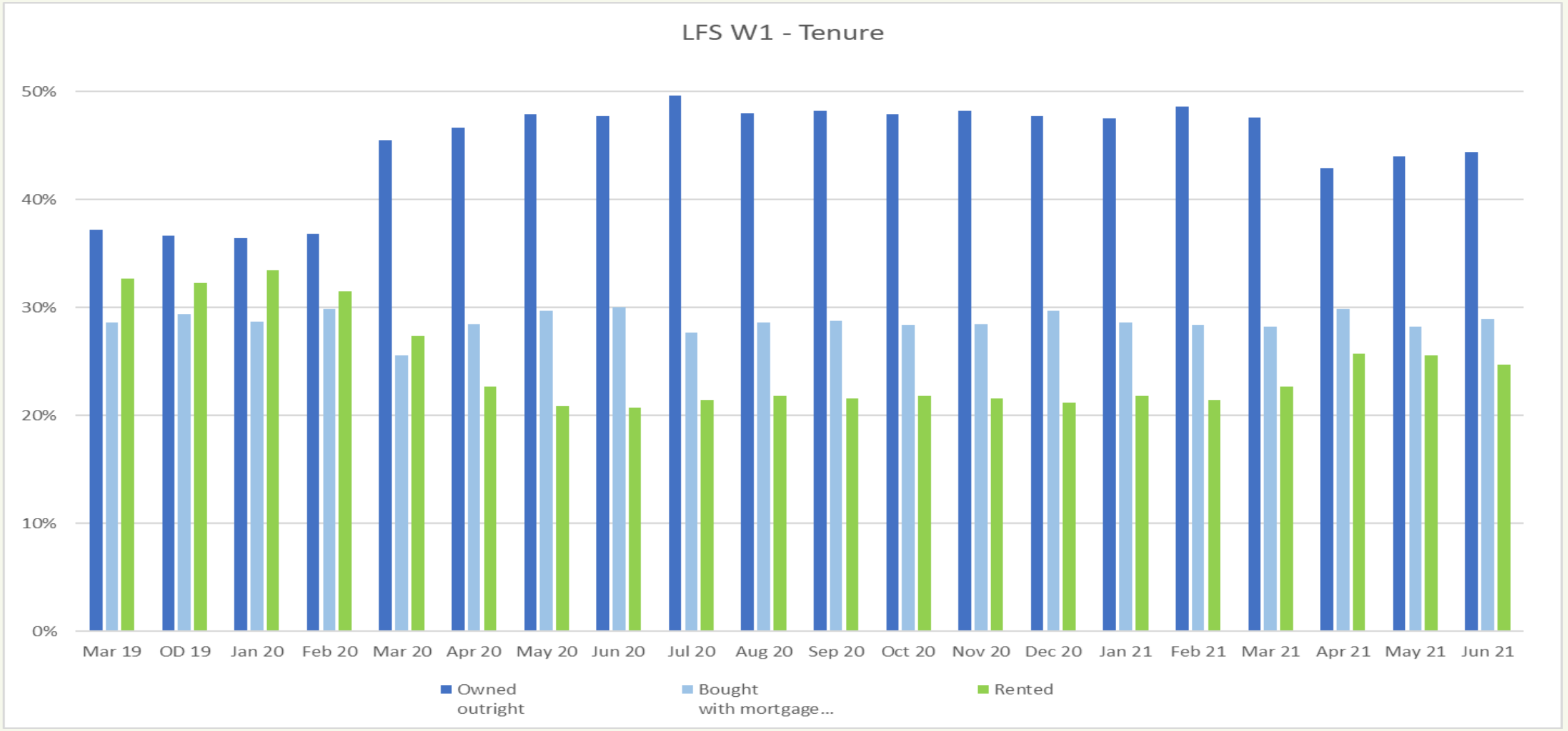
Impact of COVID-19 on LFS Response Rates



Impact of COVID-19 on LFS achieved sample size

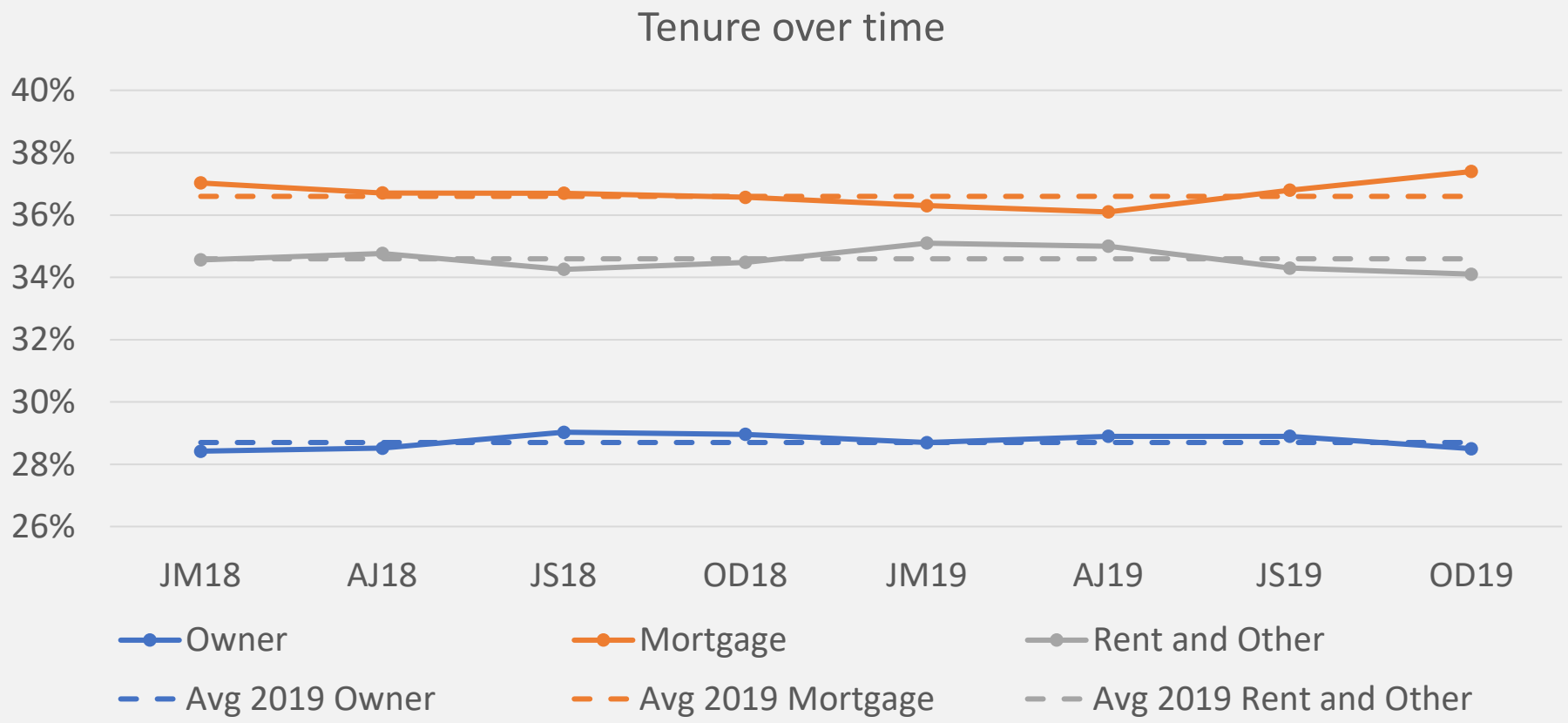


Impact of COVID-19 on profile of LFS W1 responses



Adjustment to reduce bias

Weighted distribution of household tenure

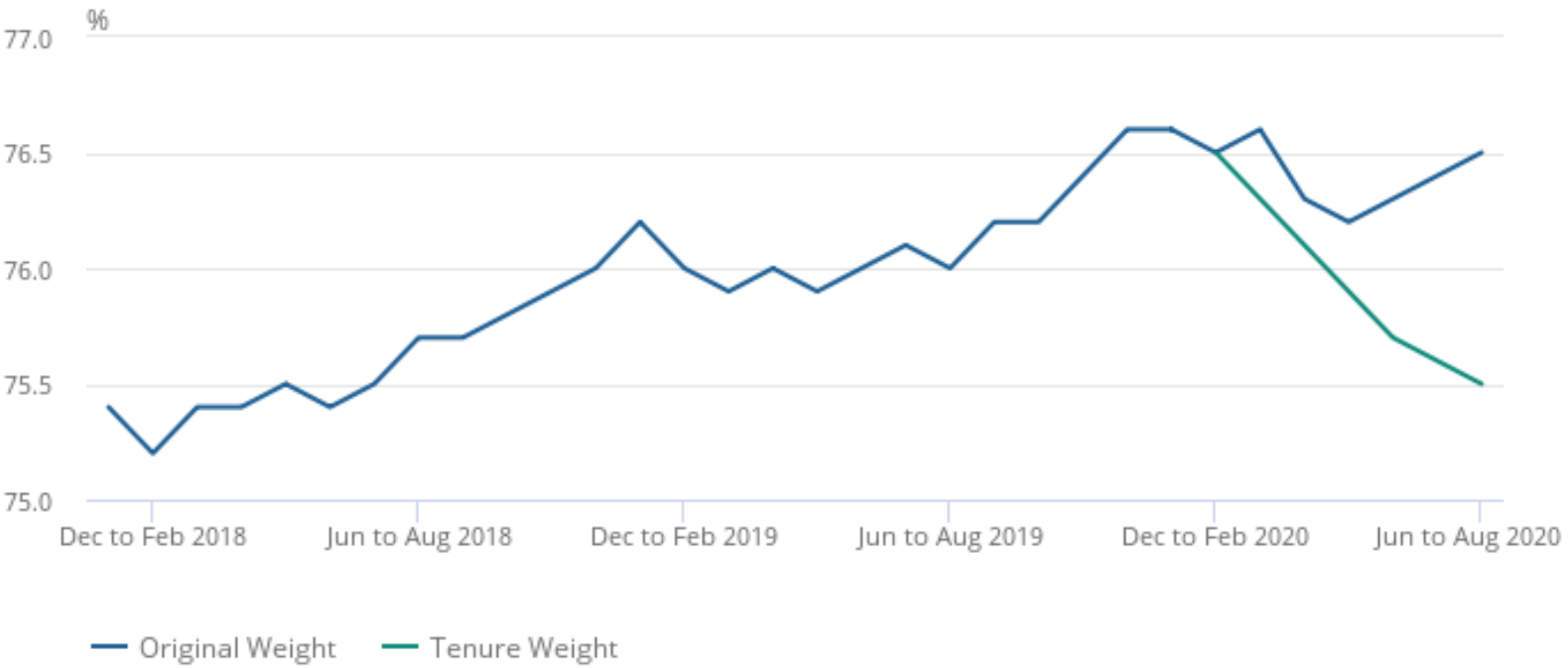


Adjustments to weighting

- There is no explicit non-response adjustment in current weighting method
- We introduced additional calibration constraints
  - Uniform weekly distribution of weighted sample by country
  - Distribution of tenure set to the average of 2019 distribution
  - Assumes minor structural change to the population
  - Assumes available population projections to be fairly accurate

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

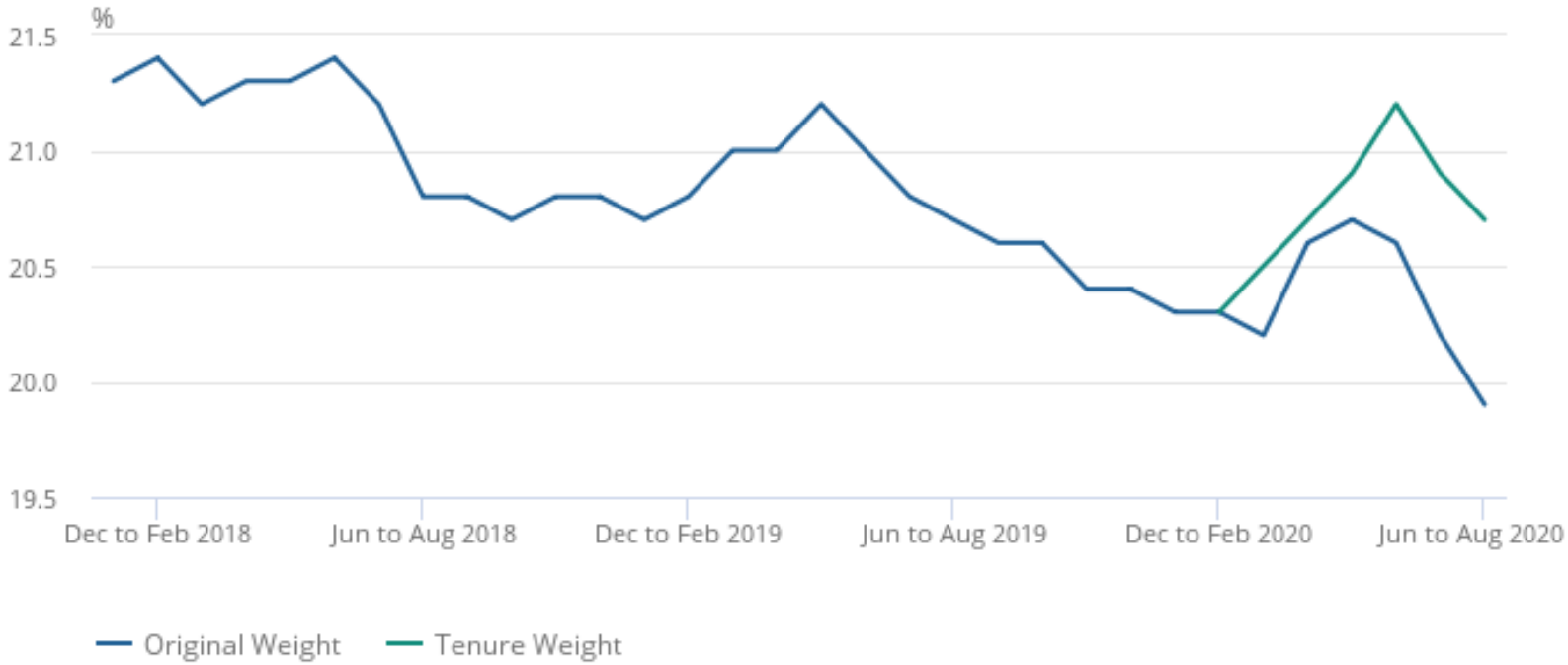
Figure 2: In June to August 2020, the adjusted unemployment rate was 0.2 percentage points higher than under usual weights

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

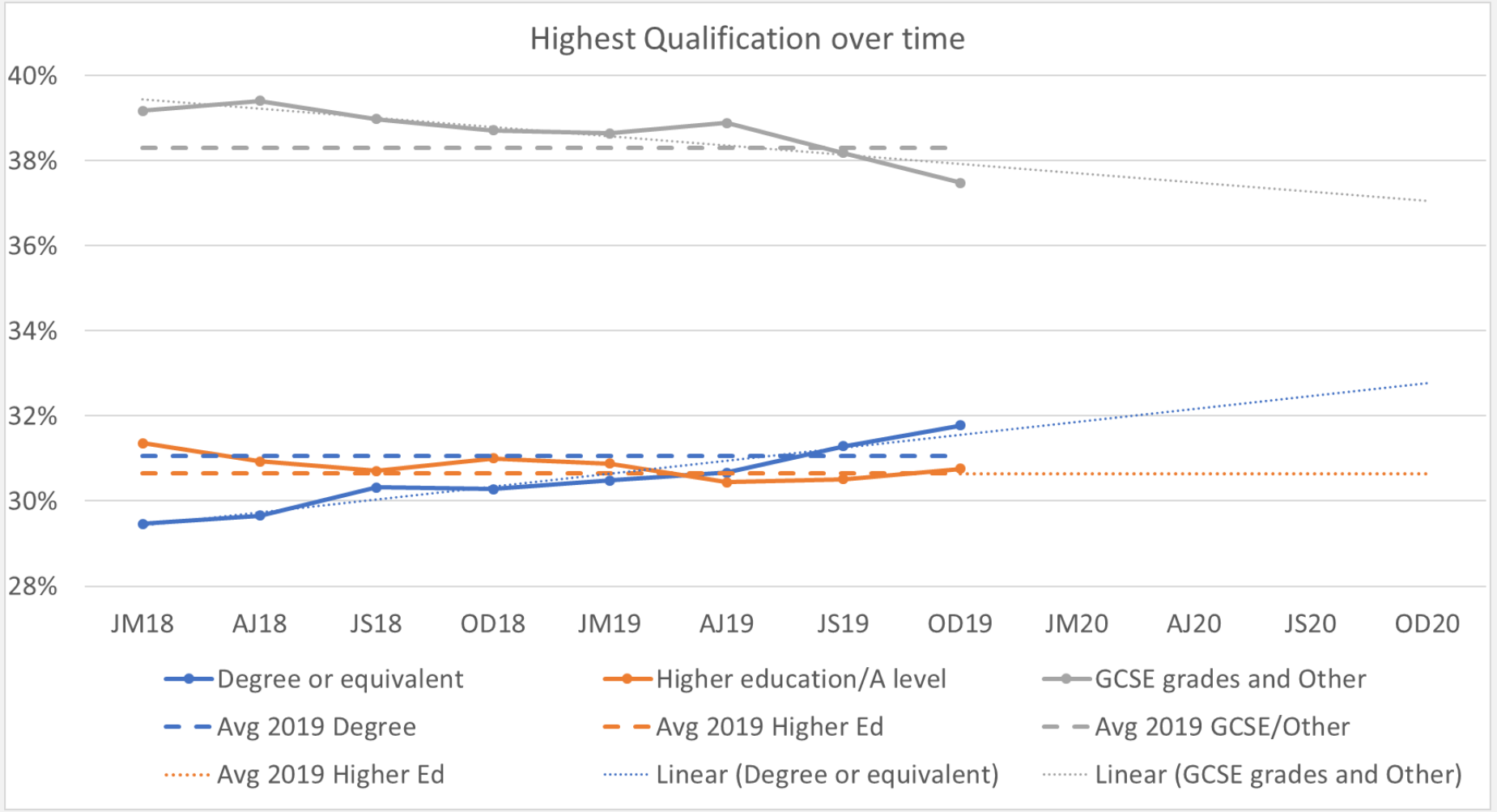
Figure 3: In June to August 2020, the adjusted economic inactivity rate was 0.8 percentage points higher than under usual weights

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

Considering other socio-economic variables: Highest qualification



Impact of pandemic on LFS estimates by country of birth

- Published estimates of employment by country of birth before recent reweighting

	Levels			Year-on-year changes		
Quarter	Total[1]	UK	Total Non UK	Total[1]	UK	Total Non UK
Jan-Mar 2019	32,641	26,838	5,796			
Apr-Jun 2019	32,752	26,967	5,777			
Jul-Sep 2019	32,802	27,028	5,770			
Oct-Dec 2019	32,983	27,171	5,809			
Jan-Mar 2020†	32,998	27,097	5,893	357	258	97
Apr-Jun 2020†	32,592	27,004	5,574	-160	38	-203
Jul-Sep 2020†	32,515	27,328	5,176	-288	299	-594
Oct-Dec 2020†	32,441	27,422	5,014	-542	251	-795

- Estimates by country of birth look implausible from April 2020 onwards

Impact of pandemic on non-UK born population

Using all cases

Quarter	2019	2020	Change
JM	9,436,495	9,246,686	-189,809
AJ	9,413,650	8,968,317	-445,333
JS	9,248,509	8,354,743	-893,766

Excluding imputed cases and wave 1 cases after mid-March 20 (75+ only)

Quarter	2019	2020	Change	Waves used
JM	8,778,739	8,997,224	218,485	All waves
AJ	8,941,893	8,052,690	-889,203	Waves 2-5
JS	8,453,521	7,860,823	-592,698	Waves 3-5
OD	8,651,552	7,931,272	-720,280	Waves 4-5

Attrition analysis

- Attrition rates decreased in UK born after pandemic but increased in non-UK born in all age groups
- EU born dropped out more out of sample than non-EU born

Age group	Country of birth	Jan-Mar 19 - Apr-Jun 19	Jan-Mar 20 - Apr-Jun 20	Difference
18-24	UK	49.7	47.6	-2.1
	Non-UK	56.7	65.6	8.9
18-24 excluding students	UK	51.0	46.7	-4.3
	Non-UK	57.0	65.5	8.5
25-74	UK	30.4	27.4	-3.0
	Non-UK	41.0	41.6	0.6
25-44	UK	42.5	38.4	-4.1
	Non-UK	45.8	45.8	0.0
45-74	UK	24.6	22.0	-2.6
	Non-UK	34.8	36.8	2.0
Under 75	UK	34.2	31.0	-3.2
	Non-UK	41.6	43.3	1.7

Comparing LFS and RTI based estimates of growth (from HMRC, an administrative source)

	RTI data		LFS data	
	Year-on-year percentage change in total employees		Year-on-year percentage change in total employment	
Period	EU	Non-EU	EU	Non-EU
Oct-Dec 2018 to Oct-Dec 2019	2.6	8.4	1.6	3.8
Mar-Jan 2019 to Mar-Jan 2020	0.5	7.1	-0.6	4.0
Apr-Jun 2019 to Apr-Jun 2020	-2.7	4.3	-9.1	1.7
Jul-Sep 2019 to Jul-Sep 2020	-4.6	2.4	-16.2	-4.8

- LFS-based growth rates for non-UK decreased much more between Oct-Dec 19 and in Jul-Sep 20 than RTI-based growth rates
- It's very likely that LFS estimates suffer from bias stemming from differential non-response between UK and non-UK born that the additional tenure constraint hasn't reduced sufficiently
- We need to reduce this bias by adding a constraint on the structure of the population by country of birth

How can we use RTI data to estimate EU and non-EU born subpopulations?

- We want to estimate year-on-year population growth in each rolling quarter in 2020 onwards
- Very few observations from RTI are available to fit a statistical model

- Use simple assumptions to derive an expression of change in population growth rates in terms of change in RTI-based employee growth rates
    - Oct-Dec 2019 (OD19) is the base period: Population growth is know

- Assumptions
    - change in the population growth rate of the non-UK sub-populations is in the same direction as the change in their RTI employee growth rate
    - the magnitude of change in population growth rate does not exceed that of change in RTI employee growth rate

- We consider July-September 2020 (JS20)

Method

- Let $\theta_{RTI,JS20,EU}$ denote the RTI employee total growth rate of the EU born population between JS20 and JS19
- Adjust RTI year-on-year percentage changes in non-UK subpopulations by differencing UK nationals' rates

$$\theta_{RTI,JS20,EU,adj} = \theta_{RTI,JS20,EU} - \theta_{RTI,JS20,UK}$$

$$\theta_{RTI,JS20,non-EU,adj} = \theta_{RTI,JS20,non-EU} - \theta_{RTI,JS20,UK}$$

- Let $\gamma_{OD19,EU}$ denote the population growth rate of the EU born population between OD18 and OD19
- Population growth rate for EU born since pandemic, $\gamma_{JS20,EU,adj}$ for JS20 quarter, satisfies

$$\boldsymbol{\gamma_{JS20,EU,adj} - \gamma_{OD19,EU} \cong b\left(\theta_{RTI,JS20,EU,adj} - \theta_{RTI,OD19,EU,adj}\right)}$$

Where $0 < b \leq 1$

- b is set to ½ to minimise mean prediction error

Evaluation using past RTI data and long-term international migration estimates (LTIM)

Period	Country of birth	LTIM (000s)	RTI-based estimates (000s)
October - December 2016	EU	133	170
	Non-EU	179	156
October - December 2017	EU	99	92
	Non-EU	199	174
October - December 2018	EU	75	59
	Non-EU	205	198
October - December 2019	EU	50	56
	Non-EU	282	329
Mean percentage deviation from LTIM			0.0
Mean percentage absolute deviation from LTIM			14.3
Mean absolute deviation from LTIM (in 1,000s)			21

Impact on the population size by country of birth

LFS year-on-year change in population levels by country of birth – Figures in 1000s

Period	Labour Force Survey			RTI-based method		
	UK	Non-UK	All	UK	Non-UK	All
January - March 2020	575	-192	383	96	302	398
April - June 2020	773	-394	379	34	265	299
July - September 2020	1,253	-880	373	16	214	230
October - December 2020	1,569	-1,203	366	-2	112	110

Implementing the adjustments: impact on 2020 estimates

Office for National Statistics

Impact on economic activity estimates

- A non-response adjustment was applied to cases that joined the sample after the change in data collection using area level data from the census

Quarter	Economic activity (ILODEFR)	Labour Force Survey Levels (000s)	Labour Force Survey Rates (%)	RTI-based method Levels (000s)	RTI-based method Rates (%)	Differences Levels (000s)	Differences Rates (%)
January - March 2020	Employed	31,601	76.3	31,589	76.3	-12	0.0
	Unemployed	1,337	4.1	1,331	4.0	-6	-0.1
	Inactive	8,457	20.4	8,484	20.5	27	0.1
April - June 2020	Employed	31,416	75.8	31,261	75.6	-154	-0.2
	Unemployed	1,363	4.2	1,332	4.1	-31	-0.1
	Inactive	8,641	20.9	8,777	21.2	135	0.3
July - September 2020	Employed	31,186	75.3	31,006	75.0	-180	-0.3
	Unemployed	1,593	4.9	1,711	5.2	118	0.3
	Inactive	8,662	20.9	8,635	20.9	-28	0.0
October - December 2020	Employed	31,082	75.0	30,874	74.8	-207	-0.2
	Unemployed	1,714	5.2	1,750	5.4	36	0.2
	Inactive	8,664	20.9	8,676	21.0	12	0.1

Future developments

- Consideration of further roll-out of KtN and reintroduction of face-to-face interviewing
- Continuing our work to transform population and migration statistics using admin data
- As further sources for estimates of the population become available (e.g. Census 2021), performance of LFS reweighting model will be assessed and further reweighting may take place if needed

Thank you

ONS Methodology contact:

Salah.Merad@ons.gov.uk

ONS LFS Survey Research contact:

Martina.Helme@ons.gov.uk