

# Measuring up?

## Exploring data discrepancies in the Labour Force Survey

13 August 2024

Adam Corlett & Hannah Slaughter<sup>1</sup>

Since the onset of the Covid-19 pandemic, significant issues have emerged with the UK's official labour market data. The Office for National Statistics' (ONS) main data source on the labour market, the Labour Force Survey (LFS), estimates that there are over a million fewer workers than the trends seen in other data sources suggest. And understanding the reasons for this discrepancy – and determining which figure is accurate – is of clear importance for public and monetary policy debates.

There are two overarching drivers of this discrepancy: falls in data quality due to lower survey response rates, and issues with the way the data is weighted up to represent the whole population. On the first, there are larger drops in response rates among high-employment groups of renters and outright homeowners compared to their low-employment counterparts, which may indicate that response rates dropped disproportionately for people in work more widely. On the second, adjusting the way the data is weighted to account for the latest population data adds around 340,000 workers to the LFS estimate. Based on alternative data sources, we present a range of plausible scenarios for the employment rate, some of which are above 76 per cent – similar to pre-pandemic levels and quite different from the LFS estimate of 74.5 per cent in Q1 2024. This underscores the importance of the ONS' ongoing work to improve the quality of the LFS and eventually move to an entirely new survey.

### **Data issues are confounding our understanding of the labour market**

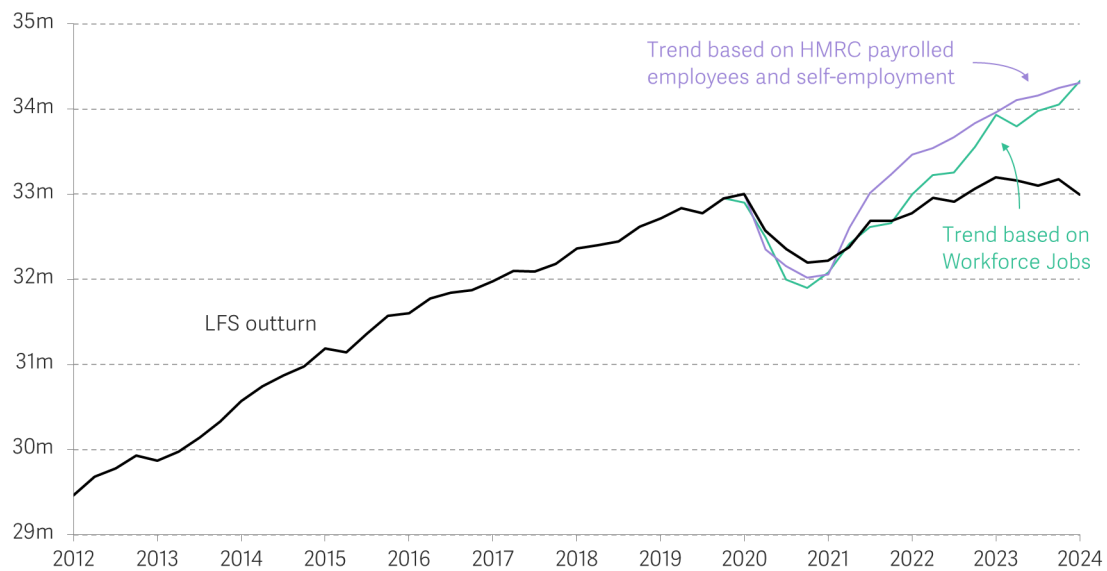
Understanding the labour market is essential to understanding the economy. But recently, this has become much more complicated. The Office for National Statistics (ONS) has faced challenges with its labour market data, leading to the suspension of published estimates based on the Labour Force Survey (LFS) – including the headline employment, unemployment and inactivity rates – between October 2023 and February 2024. Even though the statistics have since been [reintroduced](#) after some methodological updates, concerns over their robustness have resulted in them [losing](#) their 'National Statistics' status.

In this spotlight, we explore the considerable uncertainty around labour market data, focusing on people in work – where, as Figure 1 shows, LFS estimates are out of step with those from other datasets. The LFS estimated that, in Q1 2024, there were just 43,000 more people in work than before the Covid-19 pandemic. But other data sources suggest that there are many more people (over a million) in work now than in 2019. For example, if LFS estimates of the number of workers had tracked the trends in the Workforce Jobs data since 2019 (which is mostly based on a survey of businesses)<sup>2</sup> – as they have tended to do [in the past](#) – then they would have suggested 1.4 million more people in work than pre-pandemic (the green line in Figure 1).<sup>3</sup>

Administrative data shows a similar story to the Workforce Jobs numbers. HM Revenue & Customs (HMRC) publishes [real-time estimates](#) of the number of payrolled employees, and these have risen far faster than LFS employee numbers (up by 4.8 per cent and 3.3 per cent, respectively, between Q4 2019 and Q1 2024). And for the self-employed, HMRC data on the number of people with self-employment income indicates that the number of self-employed people was stable between 2019-20 and 2022-23, whereas the LFS shows a large drop.<sup>4</sup> Adding these together (and assuming zero change in self-employment in 2023-24) gives the purple trend in Figure 1.<sup>5</sup>

Figure 1 **The LFS is underestimating the number of workers relative to the trends in other data sources**

Number of people in employment, outturn and scenarios following trends in other data sources: UK



Notes: The latest data point is Q1 2024.

Source: RF analysis of ONS, Labour market statistics; ONS, Workforce jobs; ONS/HMRC, Earnings and employment from Pay As You Earn Real Time Information, seasonally adjusted; Freedom of Information request to HMRC, FOI2024/46873, May 2024.

The ONS is, very sensibly, working to [address such data issues](#), including working towards replacing the LFS with the new Transformed LFS (TLFS) – but this is an ongoing process, with no set date for its implementation. In the meantime, the sheer *scale* of the discrepancies has significant real-world implications. The Government has [ambitious plans](#) to boost employment in the context of a post-pandemic [rise in economic inactivity](#) driven by long-term sickness. But policy debates and priorities would likely look very different if there were actually over a million more people in work than both current estimates and pre-pandemic levels suggest (notwithstanding that other data sources, such as [claims for disability and sickness benefits](#), confirm that there is a real trend in ill-health). In addition, labour market statistics have knock-on impacts on other economic data, such as productivity (which, all else equal, would look even weaker if the employment level is indeed higher than currently estimated).

So, to help add to economists' and policy makers' understanding of labour market statistics, the rest of this spotlight explores two overarching drivers of the ongoing LFS issues: first, falling survey response rates and potential related bias, and second, the weighting of responses to match population totals.

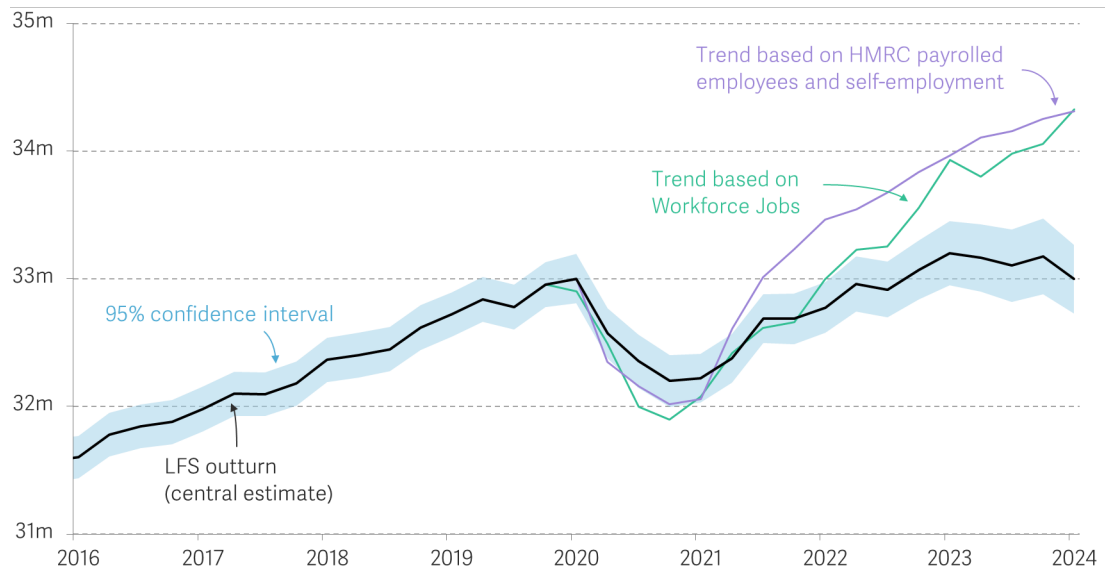
### **Falling LFS response rates have made estimates more uncertain, but any disproportionate fall among workers would be particularly worrying**

It is a known issue that LFS response rates have been [steadily falling](#) for several years, with a particularly dramatic drop since the onset of the Covid-19 pandemic. Over the past decade, the LFS sample size has roughly halved, from 99,300 to 50,800, as the survey's response rate has plummeted from 48 per cent in 2014 to 39 per cent in 2019 and to just 17 per cent at the start of 2024.

This falling response rate has had two important impacts. First, as Figure 2 shows, the smaller sample has inevitably increased the level of uncertainty around the estimates. Between Q4 2019 and Q1 2024, the 95 per cent confidence interval around the employment level [widened](#) from a margin of 177,000 in either direction to 268,000. But while this increased uncertainty suggests we should be more cautious in interpreting headline LFS estimates, it is small relative to the gap with other data sources and cannot be the sole cause of the discrepancies shown in Figure 1.

Figure 2 **Sampling variability in the LFS has been growing as the number of responses declines**

Number of people in employment, central estimate, 95 per cent confidence interval and scenarios following trends in other data sources: UK



Notes: The central estimate comes from the published outturn. Sampling variability data does not get revised when the underlying data is revised, so here we apply the sampling variation as originally published to the most recent outturn data. The latest data point is Q1 2024.

Source: RF analysis of ONS, Labour market statistics; ONS, Workforce jobs; ONS/HMRC, Earnings and employment from Pay As You Earn Real Time Information, seasonally adjusted; Freedom of Information request to HMRC, FOI2024/46873, May 2024; ONS, Labour Force Survey sampling variability.

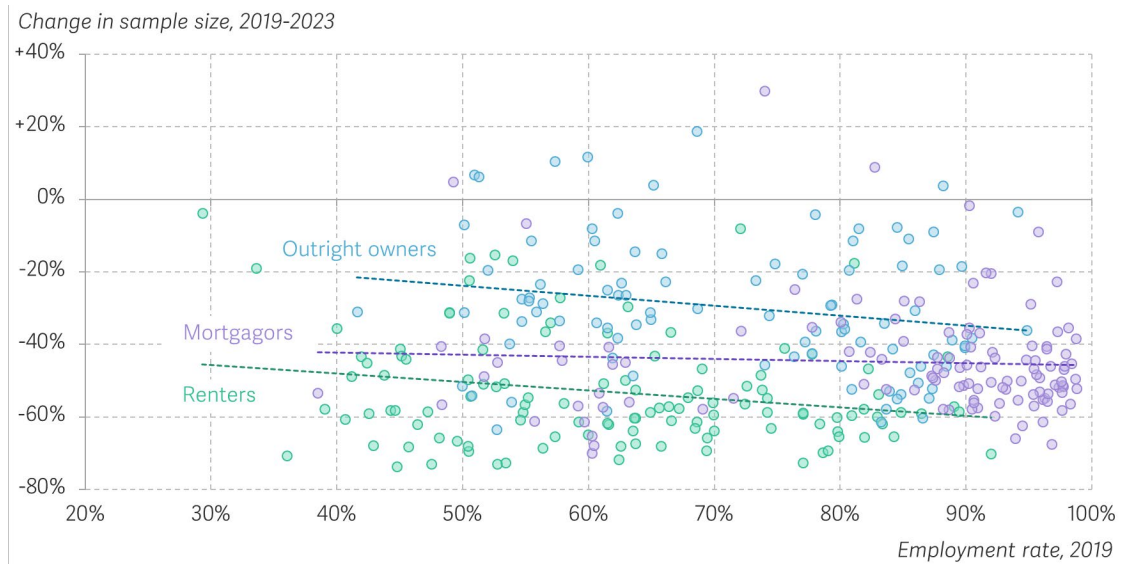
Second, and of greater concern, is the possibility that the decline in responses may not have been evenly distributed across different groups. This would make the survey less representative of the population, and – if this were not corrected for fully when the ONS produces sample weights – potentially bias the estimates. In particular, sampling bias could explain the employment gap in Figure 1 if the LFS sample was becoming more skewed towards low-employment groups in ways that the ONS’ [reweighting process](#) does not adjust for.

[We have previously shown that](#) changes in the sample make-up since the start of the Covid-19 pandemic have been similar for both low- and high-employment groups overall.<sup>6</sup> But Figure 3 shows that when we also break down the sample by housing tenure, groups with higher employment rates (defined by age, sex and region/nation) have dropped out of the sample at higher rates than their low-employment counterparts. Among renters, for example, sample sizes have fallen more for high-employment renters than for low-employment renters. The same is true for outright owners, although the relationship is less clear for mortgagors.<sup>7</sup> This may be an indication that response rates dropped disproportionately for people in work (separately from the disproportionate drop in responses among renters that quickly became apparent during the pandemic) – including (crucially) *within* the groups that the LFS weighting process accounts for.

Figure 3

### Falls in the LFS sample have been bigger for high-employment groups of renters and homeowners than their low-employment counterparts

Employment rate in 2019 (horizontal axis) and the change in the Labour Force Survey sample size between 2019 and 2023 (vertical axis) among 16-64-year-olds, by selected characteristics: UK



Notes: Each data point represents a combination of housing tenure, ten-year age band, sex and region/nation. The change in the proportion of the sample is calculated on unweighted data, and the employment rate is calculated on weighted data.

Source: RF analysis of ONS, Labour Force Survey.

Even a mild downwards slope in this chart may represent an important bias. Based on Figure 3, we estimate that if sample size falls had been equal across all groups within each tenure – such as if all renter groups saw falls in the sample size at the same rate as the average across all renters – then the estimated employment rate would have been 0.7 percentage points higher in 2023 than the published figure. This is equivalent to 290,000 more 16-64-year-olds in work.<sup>8</sup> Additional regression analysis shows that, for the groups in Figure 3, a 1 percentage point increase in a group's 2019 employment rate is associated with a 0.17 percentage point drop in the sample size. This suggests that if we compare all those in work (who have a 100 per cent employment rate) to the rest (whose employment rate is 0 per cent), in-work respondents could have seen a 17-percentage-point higher fall in their response rate than those who are not in work. The ONS' weighting process (which accounts for age, sex, region and housing tenure) should correct for this particular issue, but it could indicate that people in work are being under-represented relative to people not in work more widely.

### LFS statistics are calculated based on outdated population estimates

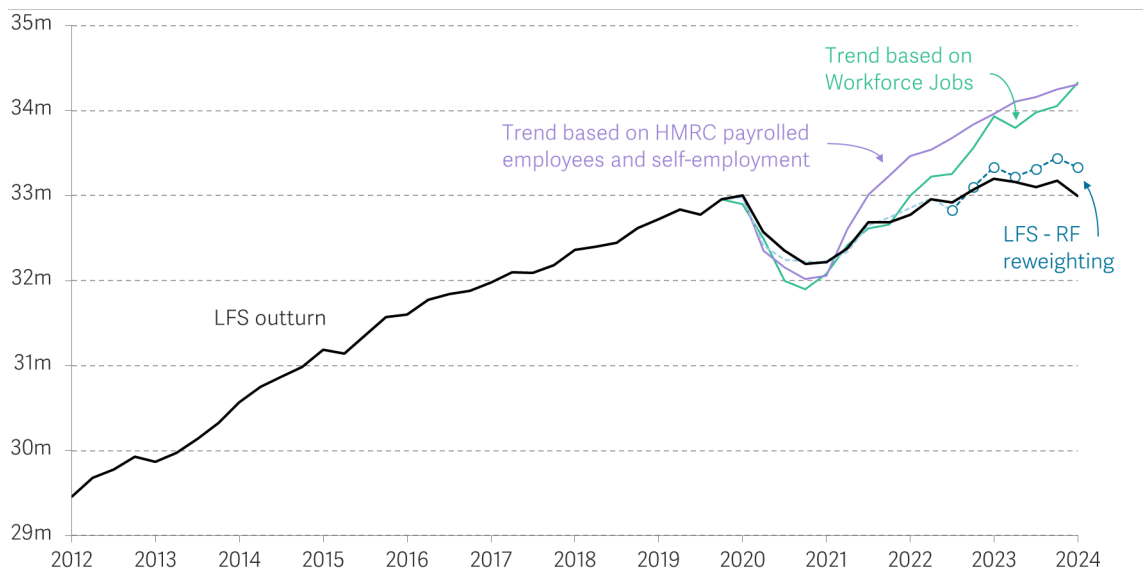
The second possible explanation for the difference lies in how the ONS [weights](#) the raw responses from the LFS. It does this both to make its sample more representative of the population, in terms of factors like age and sex, and to gross up its estimate to the total population.

The ONS continually updates its weighting as new data on the population becomes available. Most recently it [reweighted](#) the LFS earlier this year to incorporate the latest population estimates, increasing its assumption about the size of the population by 740,000 people in September-November 2023 due to higher-than-expected migration levels. But as the Bank of England [has pointed out](#) and the ONS itself [has acknowledged](#), this reweighting does not account for [the latest population projections](#), which further revised up the size of the population from 2023 onwards following a continued migration surge. This means that recent LFS estimates are still out of step with official population figures.

In Figure 4, we show the impact of bringing the LFS in line with the most recent population levels by age and sex, as an illustrative approximation of the impact of any eventual ONS reweighting.<sup>9</sup> Doing so adds 340,000 people to the employment estimate (equivalent to around a quarter of the gap between the LFS and Workforce Jobs in Q1 2024).

Figure 4 **Reweighting the LFS to account for a growing population closes some of the discrepancy with other data sources**

Number of people in employment, outturn and scenarios following trends in other data sources: UK



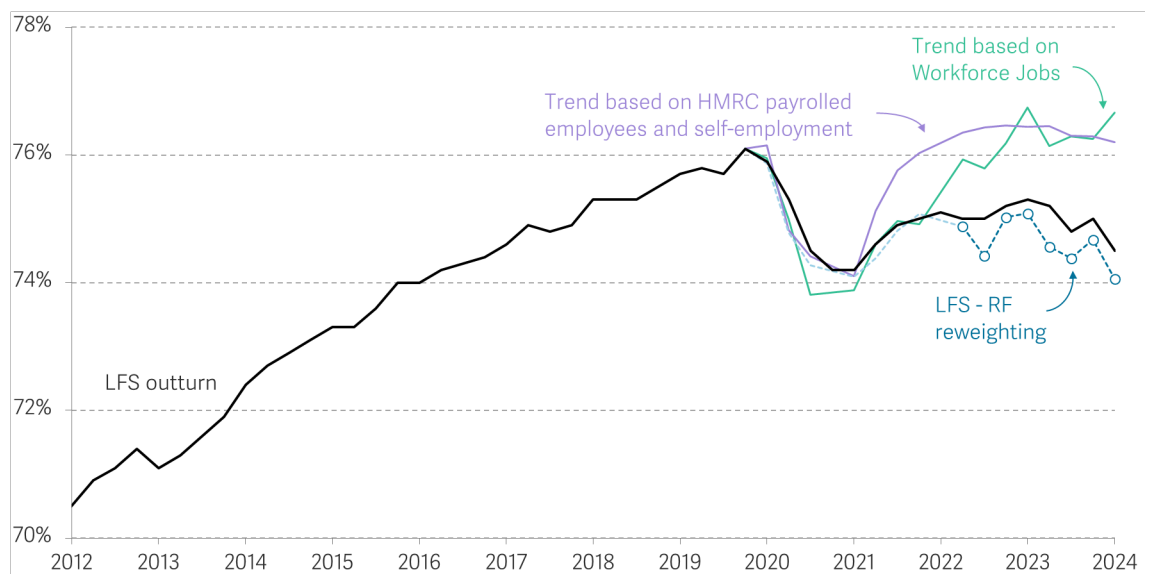
Notes: The LFS outturn shows the published labour market statistics. The reweighted LFS estimates use the reweight2 package in Stata to calibrate LFS microdata to headline population estimates and projections by sex and five-year age bands. The latest data point is Q1 2024.

Source: RF analysis of ONS, Labour market statistics; ONS, Workforce jobs; ONS/HMRC, Earnings and employment from Pay As You Earn Real Time Information, seasonally adjusted; Freedom of Information request to HMRC, FOI2024/46873, May 2024; ONS, Labour Force Survey; ONS, Population estimates: Detailed time series 2011 to 2022; ONS, 2021-based interim national population.

Although this reweighting exercise increases the estimated number of people in employment, it also increases the estimates of those *not* in employment. Overall, it reduces the estimated employment rate (the blue line in Figure 5), lowering it by 0.4 percentage points in Q1 2024 (to 74.1 per cent in our reweighted scenario, compared to a 74.5 per cent outturn). So, although the exact impact of any future reweighting by ONS may differ from ours (and it will certainly be more sophisticated), it appears that adjusting for new population estimates will do little to close the gap in the estimated employment rate with other datasets – and may even push in the opposite direction.

Finally, the green and purple lines in Figure 5 show the implied estimates of the employment rate based on trends in Workforce Jobs and administrative data (as shown in Figure 1) and the latest population statistics.<sup>10</sup> Both of these suggest that the employment rate could be much higher than the LFS estimates – above 76 per cent, similar to pre-pandemic levels and close to record highs.<sup>11</sup>

Figure 5 **Estimates of the employment rate are not affected by accounting for recent population growth, but are still out of step with other data sources**  
16-64 employment rate, outturn and scenarios following trends in other data sources: UK



Notes: The LFS outturn shows the published labour market statistics. The reweighted LFS estimates use the reweight2 package in Stata to calibrate LFS microdata to headline population estimates and projections by sex and five-year age bands. The latest data point is Q1 2024. The trend based on HMRC payrolled employees and self-employment uses payroll data for 18-64-year-olds only (the closest age match to 16-64 available in the HMRC data) and adjusts the self-employment figures for the share of self-employed workers in the LFS who are not aged 16-64 as a better proxy than the data for all age groups for trends among the 16-64 age group.

Source: RF analysis of ONS, Labour market statistics; ONS, Workforce jobs; ONS/HMRC, Earnings and employment from Pay As You Earn Real Time Information, seasonally adjusted; Freedom of Information request to HMRC, FOI2024/46873, May 2024; ONS, Labour Force Survey; ONS, Population estimates: Detailed time series 2011 to 2022; ONS, 2021-based interim national population.

We do not claim any definitive answers here. Separate [modelling by the Bank of England](#) suggests that the employment rate may be higher than indicated by the LFS, and unemployment and economic inactivity lower. Our work points tentatively in the same direction. But the key point is that the range of trends in Figure 5 is an indication of just how uncertain we should be about current labour market statistics.

So, it is welcome that the ONS is taking measures to both boost the size of the LFS sample and conduct a further reweighting exercise later this year, and the long-awaited move to the TLFS should mark an improvement in the way we track our jobs market. In the meantime, the ONS has emphasised the importance of considering a range of data sources, and administrative data is playing a growing role.<sup>12</sup> But at least until the TLFS is implemented, the LFS will remain the main source of official labour market data – with major consequences for policy decisions, as well as knock-on impacts on other statistics – and so it is crucial for the ONS to improve the survey data as much as possible and to present a more definitive story about what has happened to employment over the past few years.

---

<sup>1</sup> Thanks to Nye Cominetti, Charlie McCurdy and Greg Thwaites for advice and discussions. However, any errors remain the authors' own.

<sup>2</sup> In the Workforce Jobs data, employee jobs estimates come from a survey of businesses (private sector employee jobs) and a quarterly census of the public sector (public sector employee jobs). Self-employment figures come from the LFS. Further information on the underlying data is available [on the ONS website](#).

<sup>3</sup> A version of Figure 1 previously appeared in: N Cominetti, [Flying blind?: The case of the missing employment data](#), Resolution Foundation, October 2023.

<sup>4</sup> In theory, definitional differences may play a role in explaining differing trends. For example, job growth may have been concentrated among students or other groups who do not class themselves as employees or self-employed in the LFS; or there may be an increased overlap between employees and the self-employed, with admin-based sources double-counting people. In practice, however, we are sceptical that factors like these can explain a significant share of the divergence between data sources.

<sup>5</sup> Self-employment numbers come from an FOI request to HMRC and are based on the number of people with positive sole trader / partner income. [Published data](#) also fails to replicate the LFS' self-employment fall.

<sup>6</sup> This remains the case if we update the analysis using more recent data.

<sup>7</sup> This finding is consistent with the overall lack of relationship between sample size changes and employment rates among age-sex-region-tenure groups, because the larger sample size falls among high-employment renters and outright owners are offset by the larger-than-average falls among renters as a whole, who generally have lower employment rates than those in other tenures.

<sup>8</sup> To calculate this effect, we apply the average sample size change (in percentage terms) for each tenure to calculate what their sample size would have been if sample sizes had fallen evenly within each tenure. We then calculate the employment rate under each scenario for renters, mortgagors and outright owners by taking the weighted average of these modelled employment rates in each subgroup. Finally, we average the impact across these three tenures, weighted by their share of the 16-64-year-old population.

<sup>9</sup> We do this using the [reweight2](#) package in Stata, taking the ONS' existing person weights as our starting point. Note that this exercise does not perfectly replicate the ONS' LFS weighting process, which accounts for a fuller set of characteristics than we are able to (for example, detailed region) and is not designed to match the full population (for example, it [excludes](#) those living in most communal establishments).



---

10 The trend based on HMRC payrolled employees and self-employment in Figure 5 uses payroll data for 18-64-year-olds only (the closest age match to 16-64 available in the HMRC data) as a better proxy than the data for all age groups for trends among 16-64-year-olds.

11 It is worth noting that another possibility is that the new population projections are still underestimating recent growth, for example if migration has again surpassed the ONS' estimates. This would lower our employment rate estimates using this method. However, reconciling the blue and purple lines would require a 3 per cent upwards revision of the working-age population – which would itself be a very significant change in UK statistics.

12 Linking ONS survey data with HMRC administrative data could be a promising way for the ONS to reconcile the two and should be explored.