

Article

# 2011 Census and the Labour Force Survey



Contact:  
Emma White  
census.customerservices@ons.  
gsi.gov.uk

Release date:  
11 December 2012

Next release:  
To be announced

## Table of contents

1. [Summary of findings](#)
2. [Introduction](#)
3. [History and purpose of the LFS](#)
4. [The Census](#)
5. [Challenges in making direct comparisons between the 2011 Census and the LFS](#)
6. [The LFS - Census link study](#)
7. [Comparing some key labour market indicators](#)
8. [- Economic activity status](#)
9. [- Economic inactivity](#)
10. [Conclusion](#)
11. [Next steps](#)
12. [More Census analysis](#)
13. [Background notes](#)

# 1. Summary of findings

This paper examines the differences in labour market indicators between the 2011 Census and the LFS and sets out potential reasons why these differences will occur. The LFS is the main source of aggregate labour market statistics in the UK as it is focused primarily on labour market issues designed to elicit responses in line with internationally agreed definitions for economic activity. The Census on the other hand provides an accurate snapshot across a range of household characteristics at a detailed geographical level which is very useful for the analysis of population sub groups. There are various reasons why there will be differences in the outputs from these two sources. These include coverage and data collection issues, proxy responses, coding and classification differences between the Census and the LFS, sampling variability and the self-completion nature of the Census. The main differences observed in this analysis are the following;

- The Census estimates that there are 352,000 more household residents in the labour market, aged 16-64, in England and Wales than the LFS. Note that this excludes all residents of communal establishments.
- The Census estimates a higher number of people who are economically active, mainly those in employment, with a corresponding lower number of people who are economically inactive.
- The Census estimates that there are 602,000 more people who are employed (employment rate of 71.8 per cent) than the LFS (employment rate of 70.8 per cent) while estimating that there are 28,000 fewer unemployed people (unemployment rate of 7.5 per cent) compared to the LFS (unemployment rate of 7.7 per cent).
- Most of the observed differences are for women, with the Census estimating that there are 538,000 more economically active women than the LFS and 272,000 fewer economically inactive women than the LFS.

## 2. Introduction

The labour force survey (LFS) and the census are both rich sources of data in different respects for labour market statistics. This paper explores some conceptual reasons why the outputs from the LFS and the Census might be different. Additionally, it presents some analysis of differences observed between the LFS (February-April, 2011) and the 2011 Census and compares these with outcomes from a similar post-census report carried out for the 2001 Census and published in January 2005.

A paper published with the release of the Census 2011 annual population totals for England and Wales on 16th July 2012 described the differences between the rolled forward population estimates from the 2001 Census and the latest census figures. These rolled forward population estimates are used in the weighting of survey outputs to provide aggregate statistics from the LFS which is the UK's most reliable source of regular labour market and household indicators. Following interest from stakeholders and users, ONS carried out initial ad hoc analysis on the likely effects of these changes in population estimates on the LFS outputs and published it on its [website \(163.8 Kb Pdf\)](#). This also committed to ONS providing more analysis and clarification as more census 2011 data becomes available and released to the public.

### 3. History and purpose of the LFS

The LFS is a highly regarded UK data source for regular household key labour market statistics and provides data at a level of precision not matched by any other regular household survey. It is a household-based survey, for which interviewing is carried out on a continuous cycle. It aims to interview about 60,000 households, comprising 150,000 people, each quarter. The data it provides also includes demographic and household statistics of value to a range of stakeholders and users such as government departments, academics, local government and other institutions. The LFS was initially designed as a biennial survey in 1973, aimed at satisfying the requirements of the Statistical Office of the European Community (Eurostat). From 1984 the LFS was changed to an annual survey and the International Labour Organisation (ILO) definition of unemployment was adopted. Since 1992, the LFS has been further improved with the survey now conducted on a quarterly basis. It has evolved over time to its present role as a major UK National Statistic which provides key macroeconomic indicators every month.

The LFS target population is defined as all persons living in private households in the United Kingdom and also includes people living in some specific communal establishments i.e. nurses and others in NHS accommodation, and students in university halls of residence (sampled at their UK parents home address). The LFS utilises a rotating panel design (see figure 1) with each sample household retained for five consecutive quarters, and a fifth of the sample replaced each quarter. Most households are interviewed face-to-face at Wave 1 (except for households in Scotland north of the Caledonian Canal where all interviews are by telephone) and via telephone during subsequent waves. From January 2011, where a telephone number can be found and matched against an address selected in Wave 1, the household is first approached by telephone. At present about 15% of addresses have their Wave 1 information collected by telephone. Respondents are encouraged to participate through the five waves though some participants cannot be contacted or drop off at wave one (non-response) while others drop off at later waves of the survey (attrition).

#### Figure 1: The rotating panel design of the LFS

Figure 1 - Infographic

Key indicators of a survey's quality include its sampling errors, response rates, the consistency of its questions and the coherence of its modules. An area of concern to ONS and the users of LFS data is the observed decline in response rates, which has been seen across most social surveys over time. This is the case not just in the UK but across most of the world for non statutory surveys. At its peak in the early 1990s, the LFS response rate reached 84 per cent. This has declined over time, dropping to 74 per cent in the late 1990s and more recently dipping below 60%. Weighting the data is one adjustment made to improve the quality of outputs. The intention is to adjust for some of the systematic imbalances in the sample that could arise from differential non-response.

### 4. The Census

A census has been carried out every ten years since 1881 (apart from in 1941 as a result of the Second World War), with the most recent census carried out in March 2011. The 2011 Census had a household response rate of 95 per cent, with a person response rate of 94 per cent. This is similar to what was achieved in the 2001 Census though the variability in response rate across local authorities has been much reduced in 2011.

A census is designed to provide the most accurate snapshot of a population at a point in time. The 2011 census date was March 27th 2011. Once the estimates from the Census became available, they formed the benchmark from which to develop rolled forward population estimates. The 2011 Census collected data on living arrangements and personal characteristics of the usual resident population. A section was dedicated to the economic activity of the household. The census included all residents of communal establishments and included a special exercise to try to count homeless people. It is designed as a self-completion questionnaire and is therefore prone to some respondent error as there is no interviewer present to provide clarifications or resolve any queries that might arise. As the Census collects data from the entire population, it is ideally suited for providing information at more detailed levels of geography and across small sub-groups.

## **5. Challenges in making direct comparisons between the 2011 Census and the LFS**

This paper describes some of the reasons why differences may occur between outputs from these two key data sources and attempts to address a number of them. Some of the key differences are described below:

### **Mode of data collection**

The way that the data are collected in the LFS and the census is likely to lead to differences between estimates from the two sources. Census respondents do not have an interviewer to help explain and clarify concepts, whereas LFS respondents do, with these interviewers probing inconsistencies and checking for understanding. The LFS also uses computer-assisted interviewing, which should lead to better quality data than from a self-completion questionnaire owing to automatic routing and immediate identification of inconsistencies in responses. Computer-aided interviewing also allows a much greater amount of information to be collected than from a paper questionnaire that has a limited amount of space for answering questions.

The census asks summary questions which are interpreted by the respondent, whereas the LFS asks a detailed set of questions. Therefore, the census measures some respondent defined concept, and the LFS measures the ILO standard concept. The census is also constrained by space and scanning requirements compared with the LFS, and this reduces the amount of questioning that can take place. This can lead to response ordering effects.

### **Context effects**

Given that the LFS is a survey that concentrates on labour market issues, this could lead respondents to give more consideration to labour market questions asked on the LFS than on a multi-purpose questionnaire like the census. This should lead to less non-sampling variability in the LFS compared with the census.

### **Coverage of the census**

Estimates obtained from the census are based on data from around 98 per cent of the population, with the main coverage difference being communal establishments, and supports analysis of very small subgroups. Although the LFS is based on data from a sample of the population, results are weighted to the total household population. However, because the sample size is relatively small estimates for subgroups with small populations typically have large sampling variances, which is not the case with the census. Additionally, the LFS collects data across a quarter, so that estimates are the average over the quarter, while the Census collects the data on a single day, census day.

### **Processing – coding and classification**

The way that variables such as industry and occupation are coded differs between the LFS and the census. In the LFS, the data are coded by interviewers using either a computer-assisted system or a manual system assisted by lookups. For the census, automatic coding methods are used, assisted by expert coding using computer-aided technology where necessary. Decision rules on the census are also not always consistent with the ILO definition as it applies on the LFS. Respondents on government work schemes are identified through detailed questions on the LFS and coded appropriately as employed, unemployed or inactive but are automatically coded as employed on the census.

## Sample and population totals

The census aims to cover all the resident population within the UK. The questionnaire collects information on all people who are normally resident within a particular household or communal establishment and some homeless people are also approached. The target population of the LFS on the other hand covers all people in private households, along with residents of two particular types of communal establishments; student halls of residence and NHS accommodation. Residents of other communal establishments are not covered e.g. nursing homes, boarding houses, and prisons. The LFS is therefore currently weighted to the adjusted mid-year population estimates, which are an estimate of the household population, rather than to the standard resident population estimates published by ONS from the Census. The difference in the population estimates used accounts for the difference of around 470,000 between the population bases in the census and the LFS in March 2011.

## 6. The LFS - Census link study

ONS methodologists are currently carrying out work that will help understand the differences between the census and the LFS and improve labour market estimates. The primary focus of this work is to examine non response in the LFS to ascertain whether the lower response rates have introduced a bias. If this is found to be the case, a bias adjustment will be introduced for LFS weighting. It will also compare responses to both surveys to see if there is evidence of a response bias where questions are being answered differently.

### Non- response bias

This occurs in the LFS when some sampled households cannot be contacted or drop out after wave one of the survey. If these non-responders are different from those who have responded there is possibility of some bias being introduced into the aggregate statistics. At present, the LFS applies population based weighting (by age, sex and region) to allow for differential non-response but does not include a bias adjustment for any other factors. The LFS weighting procedure uses a calibration technique to ensure that the weights adjust for proportionate imbalances between the sample respondents and the population according to their age, sex and region. To the extent that non-response is associated with age, sex and region, the adjustment will remove non-response bias. However, it is possible that some other characteristics are associated with non-response, that the calibration has not accounted for, and these characteristics may be influential on labour market outcomes, which could result in bias.

ONS is currently carrying out work that will help understand the differences between characteristics of respondents and non-respondents to the census and a range of social surveys including the LFS. Differences correlated with the age, sex and region of the respondent are already accounted for by the calibration of LFS responses to population totals. The census non-response link study provides an opportunity to examine whether propensity to respond is additionally correlated with other variables, and if there is such a correlation, then ONS will consider whether an adjustment to the LFS weighting procedure is appropriate.

### Response bias

This is possible where certain aspects of the LFS survey and census administration could engender some variations. Some form of response variation is likely to occur where respondents on the census (which is mostly administered as a self-completion questionnaire) respond differently to a similar question on the LFS (which is administered by an interviewer either face-to-face in the first wave or by telephone in subsequent waves). A perfect match between responses obtained for labour market questions from both sources would generate a table similar to Table A below.

### Table A

Table A - Table

However, where some variation exists then Table B below is more likely. Additionally, where proxy responses occur there may also be some scope for inaccuracies and error in what data are collected.

## Table B

Table B - Table

The ONS LFS - Census link study will analyse these differences and try to evaluate the level and impact this has on published labour market statistics. It cannot be assessed at this stage but is part of methodological work that will build up a better understanding of the different response patterns.

## 7. Comparing some key labour market indicators

The measurement of labour market status is not as straightforward an activity as it may sound. There a number of questions used in the process and some coding priorities are applied when multiple classifications are possible. For example, all respondents on the census who indicate they are on a government work scheme are coded as employed whereas the LFS asks detailed questions on the specific scheme the respondent is on in order to establish if they are to be classified as employed, unemployed or inactive. For the Feb – Apr 2011 LFS, around a quarter of those who indicated they were on a government scheme were coded as either unemployed or inactive.

Over the past 10 years there has been a concerted effort to align the census questions and response categories to accepted ILO definitions and standards similar to those used in the LFS. Hence, questions on current economic status are more consistent with the LFS and thus the pattern of differences observed in 2001 would no longer apply. The differences are more likely to be attributed to changes in decision rules on the census over the 10 year period for some questions, such as reasons for inactivity and differences in terms of how these are coded on the LFS. These changes and the differences are detailed below.

This section focuses on comparisons of economic activity, reasons for inactivity, and regional distribution of economic activity. It is important to note that the published census tables are for all usual residents and include those in communal establishments while the LFS only includes a small proportion of those in communal establishments (students in student's halls of residence sampled at their UK parent's residence and nurses in NHS accommodation).

To assess differences between the LFS and Census figures, a comparison is made summarising economic activity status by sex across England and Wales. The comparisons use data from the 2011 Census and figures from the February – April 2011 quarterly LFS which spans the period when the census occurred. This specific dataset is the closest in time to Census day. The 2011 Census outputs for economic activity are published for those aged 16 – 74 years while the LFS outputs on Labour market activity are published for 16-64 and 16+. We have created tables for those aged 16-64 on both the LFS and the Census while removing those in communal establishments on the census and the LFS in order to compare the LFS and the census on a like for like basis.

## 8. - Economic activity status

In order to examine the differences in economic activity status, we have created Table 1 below which shows economic activity for those aged 16-64 for the 2011 Census and the February – April 2011 LFS. In general the LFS shows lower levels of employment than the census. There is little difference in the levels of those who are unemployed, with the LFS estimating that there are 28,000 more people (1.3 per cent difference) than on the census. The largest relative difference is for unemployed students (-16.9 per cent) where the census estimates that there are 50,000 more full-time students who are unemployed than are estimated on the LFS. Note though that this group is relatively small and subject to larger variation than some of the other groups. The LFS shows higher levels of the economically inactive than the census. The LFS estimates that there are 574,000 fewer people that are economically active compared to the census (-2.1 per cent difference) of which there are 602,000 (-2.4 per cent) fewer people employed on the LFS than on the census, while estimating that there are 222,000 more people (2.8 per cent difference) who are economically inactive.

In summary, the census estimates a higher employment rate of 71.8 per cent when compared to the LFS rate of 70.8 per cent, while estimating a lower unemployment rate<sup>1</sup> of 7.5 per cent when compared with the 7.7 per cent unemployment rate estimated on the LFS. Most of the difference is explained by the difference observed for women.

Several factors could have contributed to the differences observed between the LFS and the census. Firstly, the differences in the nature of the reference period for the census and the LFS can introduce some level of inconsistency in responses. The census asks for economic activity for the week just before the census on the 27th of March 2011. The LFS used in this analysis is for the period February – April 2011 with data collected throughout this period and respondents are advised to confirm if they had any paid work in the 7 days ending on the Sunday before the interview. Thus, on the LFS, 'last week' referred to many different weeks. Therefore, we can expect some of the differences to be true differences. This is because as the reference period shifted, some peoples' labour force status could have also changed. Previous studies have shown that some level of inconsistency may be introduced through this effect at difference measurement periods.

### Table 1: People aged 16-64 by economic activity status and sex

#### England and Wales, 2011

Table 1 - Table

The LFS estimates for women who are economically active in England could be lower than the true population value because the LFS is currently weighted to mid-year population totals that have been estimated using births, deaths and migration since the 2001 Census. The current mid-year population totals are lower than the 2011 Census estimates and the current plan is to re-weight the LFS to totals based on the 2011 Census for release of the revised estimates in early 2014.

In addition, the response rates of the LFS have decreased over the past 10 years by approximately 15 per cent. There is therefore a possibility that certain labour market groups (e.g. working women) are under-represented in the sample. This is being investigated in a census to LFS address linking study. The study will establish if certain population groups, who have different employment status, have a greater or lesser propensity to respond to the LFS. Following this study, an evaluation for non-response weighting on the LFS to adjust for any discovered bias will be undertaken. ONS carried out a similar non-response link study in 2001 to quality assure the LFS; and on the basis of the evidence found in that study decided that it was not necessary to introduce explicit non-response weights into the LFS. The 2011 Census non-response link study is one of a number of work-packages occurring as a part of the Social Survey non-response programme, which has been on-going since 2008.

On the other hand, the Census estimates for women who are economically active could be higher than the true population value. It is a self completion questionnaire while the LFS is collected by personal interview. The LFS leads interviewees through a range of questions to determine economic status but the census provides a summary list of economic statuses from which the respondent needs to select an option. Both allow proxy interviews where a member of the household completes the questionnaire for another member but for the LFS they will be supported by a trained interviewer. It is therefore probable that for some economic groups on the Census there is a propensity for a responder to have their economic status under an ILO definition mis-reported. In addition, the LFS is designed to collect employment data while the census is primarily designed to measure the population. The impact of these differences on the census estimates may become more apparent once more of the census data is open for analysis.

Finally, the LFS is a sample and is therefore subject to sampling error. It is also conducted over a 13 week period unlike the Census that is based on a single time point of one day.

It is probable that a combination of the above factors could account for the differences between the estimates from the two collection instruments. Further analysis over the coming months as information on the census link study and cross tabulations on the census becomes available, will help to increase the understanding of these differences.

Tables 2, 3 and 4 are produced on the same basis as Table 1 and show economic activity for the regions in England and Wales by gender. The tables are split into those employed, unemployed and inactive. The sections on employment and unemployment exclude students in full-time education to ensure comparability with census outputs. This is because at this time the census has not produced economic activity estimates including full-time students. As the majority of full-time students are inactive, the section on inactivity includes them.

As observed earlier, the LFS estimates that there are fewer people employed than on the census. The differences are larger in some parts of the country than others and for employment, the largest differences, both in levels and relative percentage, are seen for London (-169,000 and -4.6 per cent), West Midlands (-107,000 and -4.5 per cent) and the North West (-94, 000 and -3.1 per cent). Most of this difference can be attributed to women in employment across England with particular emphasis on London, the West Midlands and the North West. In terms of regional difference in unemployment, the LFS generally estimates slightly higher unemployment levels than the census. The most significant being a difference of 25.1 per cent (31,000) in the South West and 10.6 per cent (20,000) in Yorkshire and the Humber.

Again the possible reasons provided above apply to these differences, especially the self- completion effects of the census. This would also reflect both the different economic conditions of the regions and the how these questions are understood.

## **Table 2: People aged 16-64 by economic activity status and region**

### **England and Wales, 2011**

Table 2 - Table

## **Table 3: Males aged 16-64 by economic activity status and region**

### **England and Wales, 2011**

Table 3 - Table

## **Table 4: Females aged 16-64 by economic activity status and region**

## England and Wales, 2011

Table 4 - Table

### Notes for - Economic activity status

1. Unemployment rate is calculated by dividing the number of those aged 16-64 who are unemployed by the sum of those economically active i.e. sum of those employed and unemployed.

## 9. - Economic inactivity

Table 5 below summaries the reasons for inactivity reported on the 2011 Census and the LFS. Overall, the LFS reports more people economically inactive than on the census. The LFS estimates that there are 222,000 more people that are inactive than on the Census. The detailed picture is much more varied. The Census reports that there are 417,000 more people whose reason for inactivity is their retired status and 51,000 respondents who indicate they are inactive students than on the LFS. Conversely, the LFS reports more people who give their reason for inactivity as 'looking after the family or home' (305,000), 'permanently sick/ disabled' (251,000) and 'other' which includes those who do not want a job, temporarily sick, or are waiting for the outcome of a job application (134,000). The split by gender shows most of the difference can be attributed to women in the population. There are more women 'looking after the family or home' on the LFS than on the Census (298,000) and 106,000 more who are permanently sick/disabled. There are more men on the LFS than the Census reporting their reason for inactivity as 'permanently sick/disabled' (145,000).

On the breakdown of economic inactivity into the sub groups, the differences are likely to be caused because the Census allows the responder to identify with more than one sub group while the LFS does not. The reasons for inactivity selected by the census are then subject to some decision rules to determine what category they would default to. In the 2011 Census, where more than one reason is given and 'student' is not selected, then the response defaults to 'retired'. This might explain the higher number of those estimated as retired on the census when compared to the LFS. In 2001, the decision rule was such that the higher priority was given to those who selected permanently sick/ disabled and this explains the different profile of responses at that time.

Crucially, there is a fundamental difference between the questions on reasons for inactivity on the LFS and the census. On the census, the question asks, "\*\*\*last week were you.....retired etc ?\*\*" On the LFS, the reasons for inactivity is based on the respondents response to this question, "\*\*\*What was the main reason you did not look for work in the last four weeks?\*\*\*" On the LFS caring for family members is more likely to be regarded as a barrier to job search activity than being retired. As such there is good reason to suggest that these two sources of reasons for inactivity are difficult to compare.

This will lead to differences between the LFS and census estimates between the sub groups but not in the overall totals of economically inactive.

### Table 5: Economically inactive people aged 16-64 by sex and reason for inactivity

## England and Wales, 2011

Table 5 - Table

## 10. Conclusion

Comparisons of the census and LFS are difficult because their main purpose, coverage and mode of collection all vary. These differences will explain why observed estimates from the LFS may not be exactly the same as those from the census. The most authoritative and up to date estimates of labour market status including employment and unemployment are the labour market statistics that ONS publishes monthly. The census is valuable in providing a detailed picture at the time of the census of the economically active population. Differences between the census and labour market surveys are not just a UK phenomenon and have been observed in other countries for similar reasons. The observed differences in this paper tend to impact more on women than men and this could be down to a variety of the different factors explained which are difficult to pinpoint until further work can be done.

## 11. Next steps

Further work will continue to look at the observed differences when more information from the census is made available, in particular looking at sub groups within the population such as age and country of birth. ONS is also conducting an LFS – Census link study which will shed more light on any potential non-response bias that may explain some of the observed differences. This will consider whether the propensity to respond is correlated with factors not currently weighted for and if so, ONS will consider whether an adjustment to the LFS weighting procedure is appropriate. ONS is planning to publish the results of this further work in the summer of 2013. A re-weighting of the LFS is planned when the population estimates derived from the 2011 Census will replace population estimates rolled forward from the 2001 Census.

## 12. More Census analysis

[Census Analysis landing page](#)

## 13. Background notes

1. Details of the policy governing the release of new data are available by visiting [www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html](http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html) or from the Media Relations Office email: [media.relations@ons.gsi.gov.uk](mailto:media.relations@ons.gsi.gov.uk)

These National Statistics are produced to high professional standards and released according to the arrangements approved by the UK Statistics Authority.