Office for National Statistics

Article

Developing improved estimates of quality adjusted labour inputs using the Annual Survey of Hours and Earnings: a progress report

Improving the precision and granularity of quality adjusted labour inputs (QALI) and multi-factor productivity (MFP)

Contact:
Kris Johannsson
productivity@ons.gsi.gov.uk
+44 (0)1633 455981

Release date:
5 July 2017

Next release:
To be announced

Table of contents

1 . Abstract

This article describes work in progress to improve our estimates of quality adjusted labour inputs (QALI) using data collected in the Annual Survey of Hours and Earnings (ASHE). ASHE provides detailed estimates of the hourly earnings of UK employees, which we plan to use to augment the compilation of QALI indices, which currently rely almost exclusively on the Labour Force Survey (LFS). Because ASHE does not record levels of education this means using information from ASHE and LFS on the occupational classification of workers for the first time.

There are several strands to this work. Firstly, in order to construct a reasonable time series we need to convert historic occupational classifications used in earlier ASHE vintages to the most recent equivalent classification. This process is similar to conversions of industrial classifications, which is a fairly routine occurrence within the Office for National Statistics (ONS). However, conversion of historic occupational classifications is a non-trivial task. Although we have made some progress, more work remains to be done in this area.

Secondly, since ASHE records earnings per paid hour and QALI uses earnings per actual hour worked, we report some exploratory analysis on the relationship between actual and paid hours. This analysis suggests that the relationship between actual and paid hours can be modelled satisfactorily in terms of the characteristics that we use to stratify hourly earnings estimates on ASHE; namely age group, sex, industry and occupation.

Thirdly the article describes a method of benchmarking hourly earnings in the QALI framework to ASHE estimates (the latter adjusted to an actual hours basis). To do this we first need to expand the QALI LFS-based framework to include occupation in addition to the existing age, sex, industry and education dimensions, which leads to a large number of cells with missing pay data, particularly as we also propose expanding the current QALI industry breakdown from 10 to 19 industries. We propose to fill these empty cells using model-based estimates, which capture the relationships between pay and education for each occupation.

Fourthly, ASHE includes some sectoral information, which we have used to re-visit previous work on sectorisation of labour market metrics. In particular ASHE provides an improved source of estimates of non-market sector workers other than those in central and local government, as well as information on the sectoral dimension of second jobs, which is not available on LFS.

Lastly, and not directly related to the use of ASHE, we report some small methodological changes to how QALI deals with LFS respondents who do not report their level of education.

All of the work reported in this article is exploratory. We plan to do more work on converting occupation classifications and on modelling relationships within the LFS microdata and we need to develop the ASHE-LFS benchmarking framework from proof-of-concept to a full operational process. We will report on these planned developments alongside the next QALI release, which is scheduled for October.

As always, your feedback is welcome and can be sent to productivity@ons.gsi.gov.uk or to kris.johannsson@ons.gsi.gov.uk.

2 . Introduction

A Quality Adjusted Labour Index (QALI) augments traditional measures of labour input by taking account of changes in labour composition. As such, it is one measure of the effective supply of labour: weighting changes in the hours worked of relatively high (low) productivity workers more heavily (lightly) to produce an index that reflects both changes in the quantity and quality of the labour supply.

As currently specified, QALI stratifies the employed labour force into 360 segments across four categories: education (six strata), sex (two), age group (three) and industry (10). We collect data from the Labour Force Survey (LFS) on hours worked and hourly earnings of each category in each quarter. These raw estimates are then benchmarked to industry-level estimates of hours worked and labour income. QALI indices are then compiled by weighting (log) changes in hours worked by the income weights implied by the combination of hours worked and average hourly remuneration of each QALI category. Other things equal, a QALI index will increase faster than a simple measure of hours worked when labour composition is shifting towards those categories with relatively higher hourly remuneration, for example, an increasing share of graduates in the employed labour force, or a rising share of labour employed in industries that tend to pay higher wages.

We have published experimental QALI estimates for a number of years. QALI indices are of some interest in their own right, but the principal reason for their compilation by the Office for National Statistics (ONS) is as a set of inputs to our multi-factor productivity (MFP) estimates. In the growth accounting literature, MFP is what is left-over after subtracting contributions to economic growth that can be ascribed to movements in capital services, movements in hours worked and movements in labour composition.

The work reported in this article is motivated by two principal drivers. First, as well as having a larger number of unique respondents than the LFS, the Annual Survey of Hours and Earnings (ASHE) also has the merit of being a survey of businesses about their employees – which is widely thought to avoid some problems of reporting bias and to provide more accurate industry allocation, as well as a lower propensity to round reported hours. A further issue is that LFS collects earnings information only on the first and fifth quarterly wave, resulting in many missing pay estimates in any particular quarterly LFS dataset, which will contain cohorts from all five waves.

Second, utilising a secondary data source provides a route to delivering finer industry granularity. Some earlier work by the growth accounting team suggested that it might be feasible to expand the industry granularity of QALI from the current 10-industry specification. But it is already the case that some QALI cells are very thin (or missing entirely) on the LFS, whereas ASHE is sufficiently large to support a much more detailed granularity.

In the first instance the work reported in this article expands the industry granularity from 10 to 19 industries (all letter level industries in Standard Industrial Classification 2007: SIC 2007 apart from S, T and U which are aggregated). Subject to your feedback we intend to use this breakdown for forthcoming quarterly QALI and MFP estimates. We are planning to develop functionality for a finer industry granularity (around 60 2-digit industries) for QALI and MFP as an annual system.

The layout of the rest of the article is as follows. Section 3 explores issues arising from the use of occupational classification data for the first time. Section 4 reports some work on identifying relationships between actual and paid (or usual) hours worked. Section 5 describes an approach to adjusting LFS hourly pay estimates in terms of QALI categories to align with ASHE estimates adjusted as described in the previous section. This involves expanding the number of pay and hours observations collected from LFS to include occupation groups (as well as finer industry granularity), replacing missing pay observations with estimated equivalents, aligning to the ASHE hourly earnings estimates before re-aggregating back to the original QALI stratification. Initial results suggest that this method generates pay differentials that are similar but not identical to those from LFS alone.

Appendix 1 also uses ASHE data but in the context of sectorisation of the labour market between market and non-market components, and for the purpose of deriving industry level benchmarks for sectoral hours worked and sectoral labour remuneration. Using ASHE for this purpose will have some impacts on market sector QALI that are independent of the use of ASHE component level hourly earnings.

Appendix 2 describes further proposed changes to the QALI methodology that are independent of ASHE, specifically dealing with the treatment of LFS respondents who do not report their level of education.

3 . Working with occupation classifications in LFS and ASHE

To make greater use of Annual Survey of Hours and Earnings (ASHE) data in our Quality Adjusted Labour Index (QALI), it is first necessary to ensure an overlap of the characteristics that we use from the Labour Force Survey (LFS) with those available from ASHE. Our QALI methodology utilises information on education qualifications, along with age, sex and industry of employment. ASHE collects information on age, sex and industry but not on education. The closest alternative to education that is available on ASHE is occupation, which is also available on LFS. As there is a sizeable literature on the relationship between education and occupation, this forms our bridging variable. But before we explore this relationship further – and make the changes outlined previously to both utilise the larger sample size from ASHE and to increase the industry granularity of our QALI estimates – it is first necessary to determine what level of occupational categories to use.

Two considerations guide the choice of occupational grouping. Firstly, a more granular categorisation would ensure that differences in hourly remuneration can be better captured. To the extent that there are notable changes in hours or earnings within an occupational category, these will be averaged away at a higher level of aggregation, but made plain with a more detailed classification. All else equal, a more detailed breakdown is therefore preferred. However, a more detailed classification could result in a large number of cells that are empty or contain few observations, reducing the quality of our estimates. The cell size resulting from a given level of classification is consequently the second consideration.

At the 2-digit level there are 25 different occupation groups (Table 1) and using so many occupational groupings would result in 17,100 QALI categories on our expanded 19-industry granularity. This would result in many categories not having any observations for hourly remuneration and other cells with a small sample of pay observations. Two-digit occupations could be amalgamated; for instance into four skill groups as shown in Table 1. However, these are quite aggregated and are likely to mask significant variation in pay and hours.

Table 1: Comparison of the sub-major groups of Standard Occupational Classification SOC2000 and SOC2010

Skill Level	Sub-major groups of:			
	SOC 2000		SOC 2010	
Level 4	11	Corporate managers	11	Corporate managers and directors
	21	Science and technology professionals	21	Science, research, engineering and technology professionals
	22	Health professionals	22	Health professionals
	23	Teaching and research professionals	23	Teaching and educational professionals
	24	Business and public service professionals	24	Business, media and public service professionals
Level 3	12	Managers and proprietors in agriculture services	12	Other managers and proprietors
	31	Science and technology associate professionals	31	Science, engineering and technology associate professionals
	32	Health and social welfare associate professionals	32	Health and social care associate professionals
	33	Protective service occupations	33	Protective service occupations
	34	Culture, media and sports occupations	34	Culture, media and sports occupations
	35	Business and public service associate professionals	35	Business and public service associate professionals
	51	Skilled agricultural trades	51	Skilled agricultural and related trades
	52	Skilled metal and electrical trades	52	Skilled metal, electrical and electronic trades
	53	Skilled construction and building trades	53	Skilled construction and building trades
	54	Textiles, printing and other skilled trades	54	Textiles, printing and other skilled trades
Level 2	41	Administrative occupations	41	Administrative occupations
	42	Secretarial and related occupations	42	Secretarial and related occupations
	61	Caring personal service occupations	61	Caring personal service occupations
	62	Leisure and other personal service occupations	62	Leisure, travel and related personal service occupations
	71	Sales occupations	71	Sales occupations
	72	Customer service occupations	72	Customer service occupations
	81	Process, plant and machine operatives	81	Process, plant and machine operatives
	82	Transport and mobile machine drivers and operatives	82	Transport and mobile machine drivers and operatives
Level 1	91	Elementary trades, plant and storage related occupations	91	Elementary trades and related occupations
	92	Elementary administration and service occupations	92	Elementary administration and service occupations

Source: Office for National Statistics

Our point of departure is therefore to assess the degree to which there are differences in hourly remuneration between occupation categories and the number of empty pay cells at the level of the nine separate 1-digit Standard Occupational Classification 2010: SOC10 occupation categories shown in Table 2.

Table 2: Standard Occupational Classification (SOC) 1–digit categories

The SOC Hierarchy	
Occupation Group 1	Managers, directors and senior officials
Occupation Group 2	Professional occupations
Occupation Group 3	Associate professional and technical occupations
Occupation Group 4	Administrative and secretarial occupations
Occupation Group 5	Skilled trades occupations
Occupation Group 6	Caring, leisure and other service occupations
Occupation Group 7	Sales and customer service occupations
Occupation Group 8	Process, plant and machine operatives
Occupation Group 9	Elementary occupations

Source: Office for National Statistics

To examine the extent of differences in earnings within skill groups and across 1-digit occupational groups we adopt a regression approach. Regressions on the log of hourly remuneration in ASHE over the period 1997 to 2015 (using a modal mapping of earlier SOC classifications, see the SOC conversion sub-section later in this section) shows that there are quite substantial differences in hourly pay for 1-digit occupation categories that are included in the same skill level, after controlling for other factors likely to affect hourly pay. The regression in Model 1 consists of occupation groups and year. The subsequent models each include additional control variables, so Model 2 adds industry controls, Model 3 adds age group controls to Model 2, and Model 4 adds controls for sex to Model 3 (Table 3).

Table 3: Modelling pay by occupation

	Model 1	Model 2	Model 3	Model 4
Dependent variable		ln (hourly pay)		
Controls	Year	+ Industry	+ Age Group	+ Sex
Professional occupations	0.000750	0.0116***	0.0344***	0.0477***
	(0.77)	(11.89)	(35.76)	(50.30)
Associate professional and technical occupations	-0.205***	-0.223***	-0.198***	-0.190***
	(-199.84)	(-222.64)	(-201.20)	(-196.36)
Administrative and secretarial occupations	-0.612***	-0.625***	-0.593***	-0.530***
	(-637.25)	(-667.50)	(-643.77)	(-571.77)
Skilled trades occupations	-0.581***	-0.563***	-0.536***	-0.564***
	(-513.72)	(-505.14)	(-490.25)	(-522.26)
Caring, leisure and other service occupations	-0.820***	-0.756***	-0.715***	-0.669***
	(-737.35)	(-660.62)	(-635.24)	(-599.31)
Sales and customer service occupations	-0.916***	-0.831***	-0.772***	-0.716***
	(-875.80)	(-778.02)	(-729.71)	(-678.82)
Process, plant and machine operatives	-0.677***	-0.706***	-0.688***	-0.701***
	(-596.24)	(-620.94)	(-616.45)	(-638.39)
Elementary occupations	-0.929***	-0.864***	-0.831***	-0.815***
	(-937.35)	(-875.04)	(-855.79)	(-851.50)
R^2	0.4758	0.5146	0.534	0.5487
N	3377976	3377976	3377976	3377976

t statistics in parentheses

* p < 0.05, ** p < 0.01, *** p < 0.001

Source: Office for National Statistics

Notes:

1. Estimated coefficients can be interpreted as logs of hourly pay relative to the control group (Managers, directors and senior officials). For instance, Model 4 suggests that workers in Elementary occupations earn around exp(-0.815) = ~0.44 of the hourly pay of the control group.

As expected, higher occupation groups (that is, lower skill groups) tend to receive lower levels of hourly remuneration. The regressions also show that associate professionals and technical occupations receive significantly more pay than skilled trades occupations, despite being in the same skill grouping in Table 1. Thus using 1-digit occupation groups would ensure that differences in labour quality are better captured than would be the case by using skill levels, but are likely to deliver fewer observations based on low cell-counts than a full 2-digit breakdown.

SOC conversion

In order to use ASHE data from 1997 it is necessary to convert earlier Standard Occupational Classification codes (SOC90 and SOC00) into SOC10. There are a number of different methods that can be used to map previous SOC codes to SOC10, most of which depend on correspondence tables that draw on dual coded observations for a limited period, which show how each old classification maps to a new one. For instance, the conversion of SOC90 to SOC00 codes for LFS data was done using correspondence tables produced from dual-coded LFS data from winter 2000 to 2001. The SOC00 to SOC10 conversion uses a correspondence matrix derived from the dual coding of LFS for winter 1996 to 1997, the 2001 Census and the first quarter (January to March) of 2007.

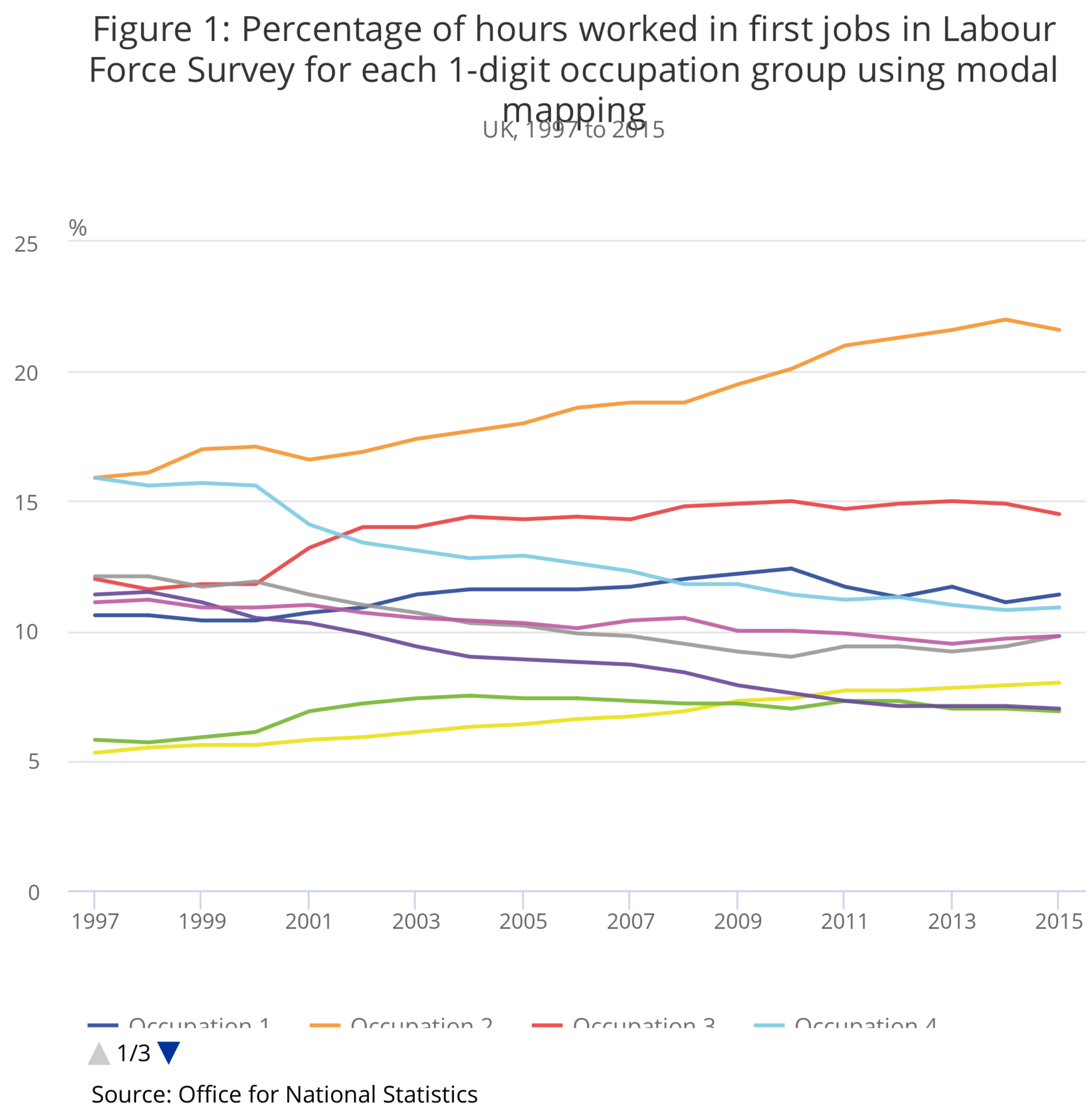
One method of conversion using these data is modal conversion. An example is that for women in SOC90 code 345 (dispensing opticians), the relationship from the correspondence tables is that 75% are coded to SOC00 code 3216 (dispensing opticians) and 25% are coded to SOC00 2214 (ophthalmic opticians). Using a modal conversion all SOC90 code 345 records would be mapped to SOC00 code 3216. A drawback of this method is that, as in this example, correspondence tables generally do not map to a single SOC code.

An alternative method is to use a one-to-many mapping, proportionately splitting existing records and weighting them accordingly. So for the previous example each record for SOC90 code 345 (dispensing opticians) would be split into two; one with SOC00 code 2214 (ophthalmic opticians) with a weight of 0.75 and another into SOC00 code 3216 with a weight of 0.25 (dispensing opticians). This more accurately reflects the relationship of the mapping, but at the cost of significantly increasing the size and complexity of the dataset. This is particularly apparent when converting occupational classifications more than once. For example, where a SOC90 code is converted to 10 different SOC00 codes and each of these is then converted to 10 SOC10 codes, the original SOC90 record will be split into 100 separate records in terms of SOC10, many of which are likely to have negligible weights.

Figure 1 shows the proportion of hours worked in each occupation group in the LFS using a modal mapping and Figure 2 the proportion of hours worked for a proportional mapping. Figure 2 has significantly less variation in the proportion of hours worked in 1-digit occupation categories for changes in SOC code in 2001 and 2011 than in Figure 1.

Figure 1: Percentage of hours worked in first jobs in Labour Force Survey for each 1-digit occupation group using modal mapping

UK, 1997 to 2015



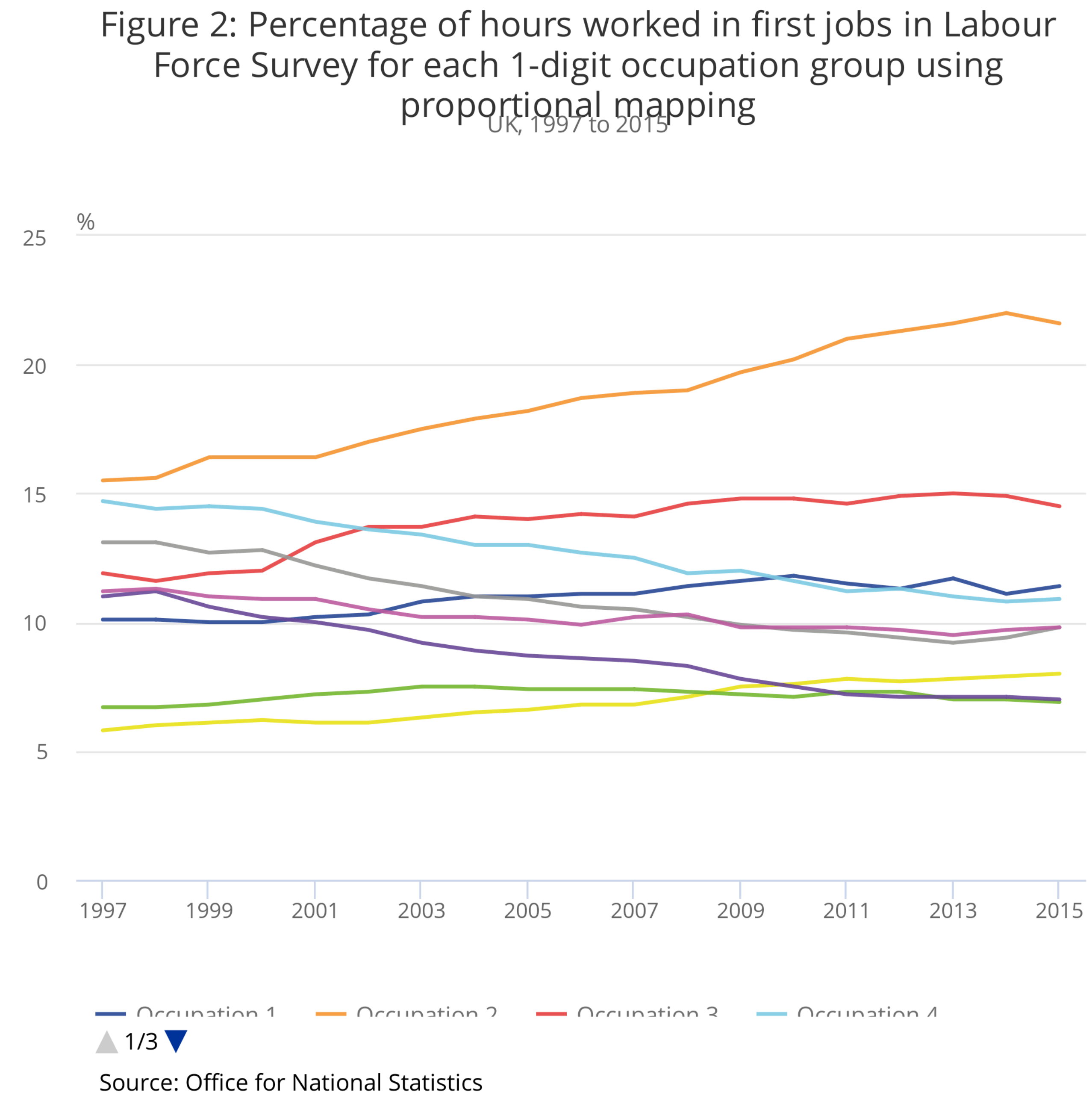
Source: Office for National Statistics

Notes:

1. Occupational groups 1 to 9 as described in Table 2.

Figure 2: Percentage of hours worked in first jobs in Labour Force Survey for each 1-digit occupation group using proportional mapping

UK, 1997 to 2015



Source: Office for National Statistics

Notes:

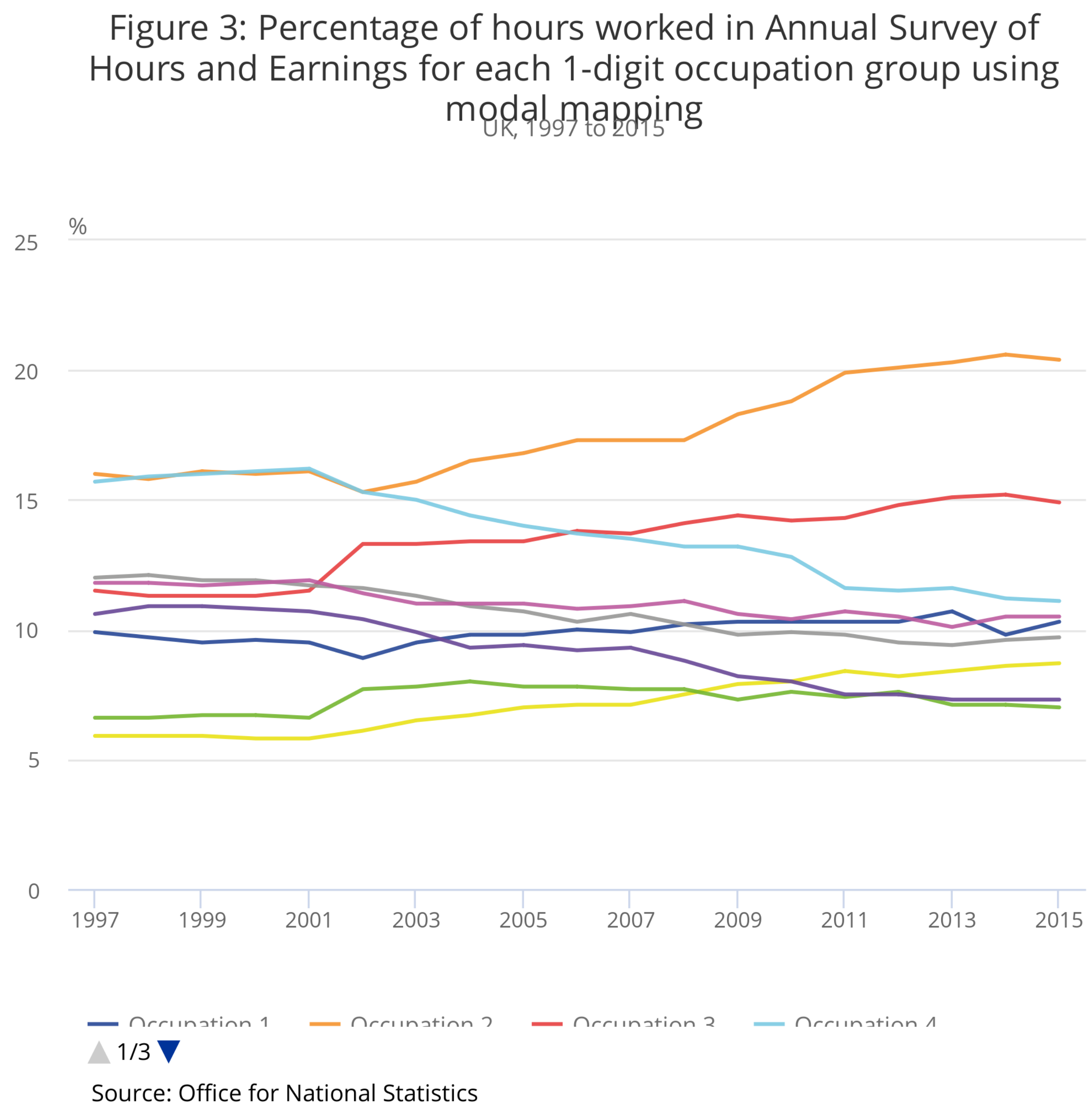
1. Occupational groups 1 to 9 as described in Table 2.

A similar issue arises for the conversion of earlier occupational categories into the most recent version in the ASHE datasets. In order to convert ASHE records from SOC90 to SOC00, a correspondence table was produced by matching records from 2001 coded to SOC90 and records from 2002 coded to SOC00. The records were matched using an ONS serial number and a correspondence table was produced from those that remained in the same job. The correspondence table for the SOC00 to SOC10 conversion was produced using the relationship of occupations for dual-coded ASHE data in 2011.

By mapping occupation using a modal mapping there are some quite large changes in the proportion of hours worked in ASHE for SOC code changes in 2002 and 2011 (Figure 3). There is much less variability in the number of hours worked where there are SOC code changes using a proportional mapping in Figure 4.

Figure 3: Percentage of hours worked in Annual Survey of Hours and Earnings for each 1-digit occupation group using modal mapping

UK, 1997 to 2015



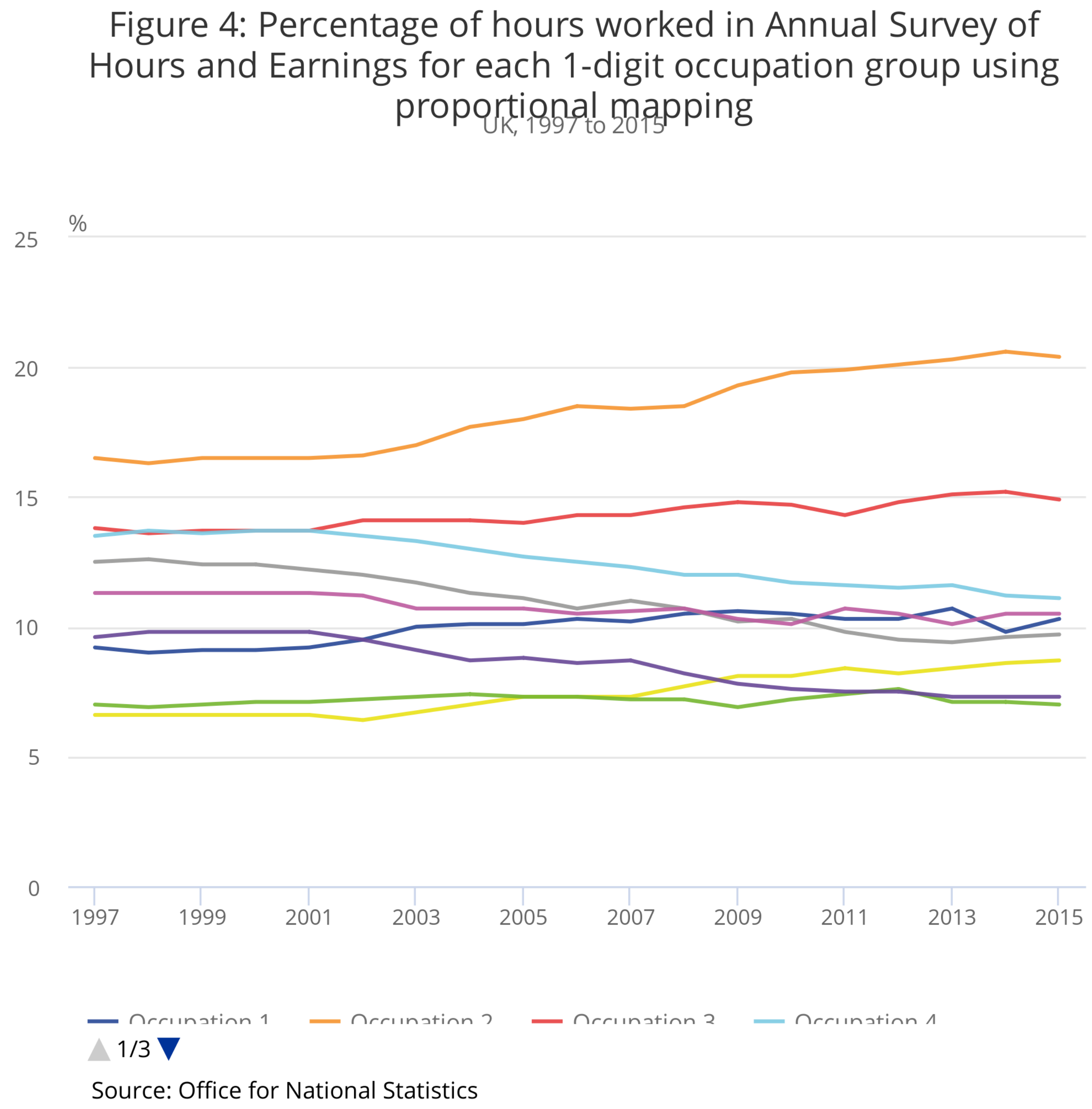
Source: Office for National Statistics

Notes:

1. Occupational groups 1to 9 as described in Table 2

Figure 4: Percentage of hours worked in Annual Survey of Hours and Earnings for each 1-digit occupation group using proportional mapping

UK, 1997 to 2015



Source: Office for National Statistics

Notes:

Occupational groups 1 to 9 as described in Table 2

This work shows that conversion of previous SOC codes using a simple modal mapping can result in unwelcome variability in occupation shares where there are changes in SOC codes. Proportionate mapping improves the time series properties of the data but at the cost of large increases in the size and complexity of the source datasets.

One potential solution to this problem could be to use a probabilistic mapping. In the previous example, this would allocate 75% of records coded to SOC90 345 to SOC00 code 3216, and 25% of records to SOC00 code 2214. An advantage of this approach is that it does not increase the number of records in the dataset, although care needs to be taken to ensure that such a mapping delivers a unique outcome (that is, it always maps individual records to the same destination) and that it takes account of other relevant information, for example, where the destination depends on other record characteristics such as age and sex. We plan to investigate this route, as well as considering using a similar approach for industry codes, where we currently use a modal mapping.

4 . Exploring the relationship between paid and actual hours worked

Our measures of labour productivity and quality adjusted labour inputs (QALI) use measures of actual hours worked, weighted in the case of QALI by estimates of earnings per actual hour worked. The Annual Survey of Hours and Earnings (ASHE) reports paid usual hours, which can vary from actual hours worked for a variety of reasons including holidays, sickness and discretionary leave. To explore the relationship between paid hours and actual hours we have used the Labour Force Survey (LFS), which contains information on both measures.

We have constructed a variable to represent the ratio of actual to paid hours at the individual record level and regressed this variable on the set of categorical variables that we intend to use from ASHE, namely industry, age, sex and occupation. We also included a year variable to capture changes in the relationship between paid and actual hours over time. Note that we would expect the actual:paid hours ratio to vary over the year (for example, because workers are more likely to take leave over the summer months). But because our aim is to adjust paid hours from ASHE, we have annualised the LFS data over 4 quarters on a Quarter 4 (October to December) to Quarter 3 (July to September) basis to align with the ASHE data collection timetable. This approach to adjusting between paid and actual hours differs from the frequency distribution approach used in our new industry by region labour metrics, though both exploit the same statistical properties of the underlying LFS data.

Reflecting issues with converting occupational classifications prior to Standard Occupational Classification 2010: SOC10, regressions are run on pooled LFS data from the first quarter (January to March) of 2011. Results are shown in Table 4.

Table 4: Regression of actual to paid hours ratio on categories of employees in Labour Force Survey

Dependent variable	acthr/usuhr
year	0.000499 ***
	(-230.41)
30 to 49 years	-0.00548***
	(-4.97)
50 to 99 years	-0.0203***
	(-17.34)
Male	0.0479***
	(-49.85)
Professional occupations	-0.0212***
	(-13.04)
Associate professional and technical occupations	-0.0448***
	(-26.70)
Administrative and secretarial occupations	-0.0485***
	(-27.13)
Skilled trades occupations	-0.0747***
	(-40.98)
Caring, leisure and other service occupations	-0.0730***
	(-36.57)
Sales and customer service occupations	-0.0494***
	(-23.88)
Process, plant and machine operatives	-0.0766***
	(-36.11)
Elementary occupations	-0.0714***
	(-38.90)
R^2	0.8364
N	1015447
t statistics in parentheses	
* p < 0.05, ** p < 0.01, *** p < 0.001	
Industry controls all significant at p < 0.001	

Source: Office for National Statistics

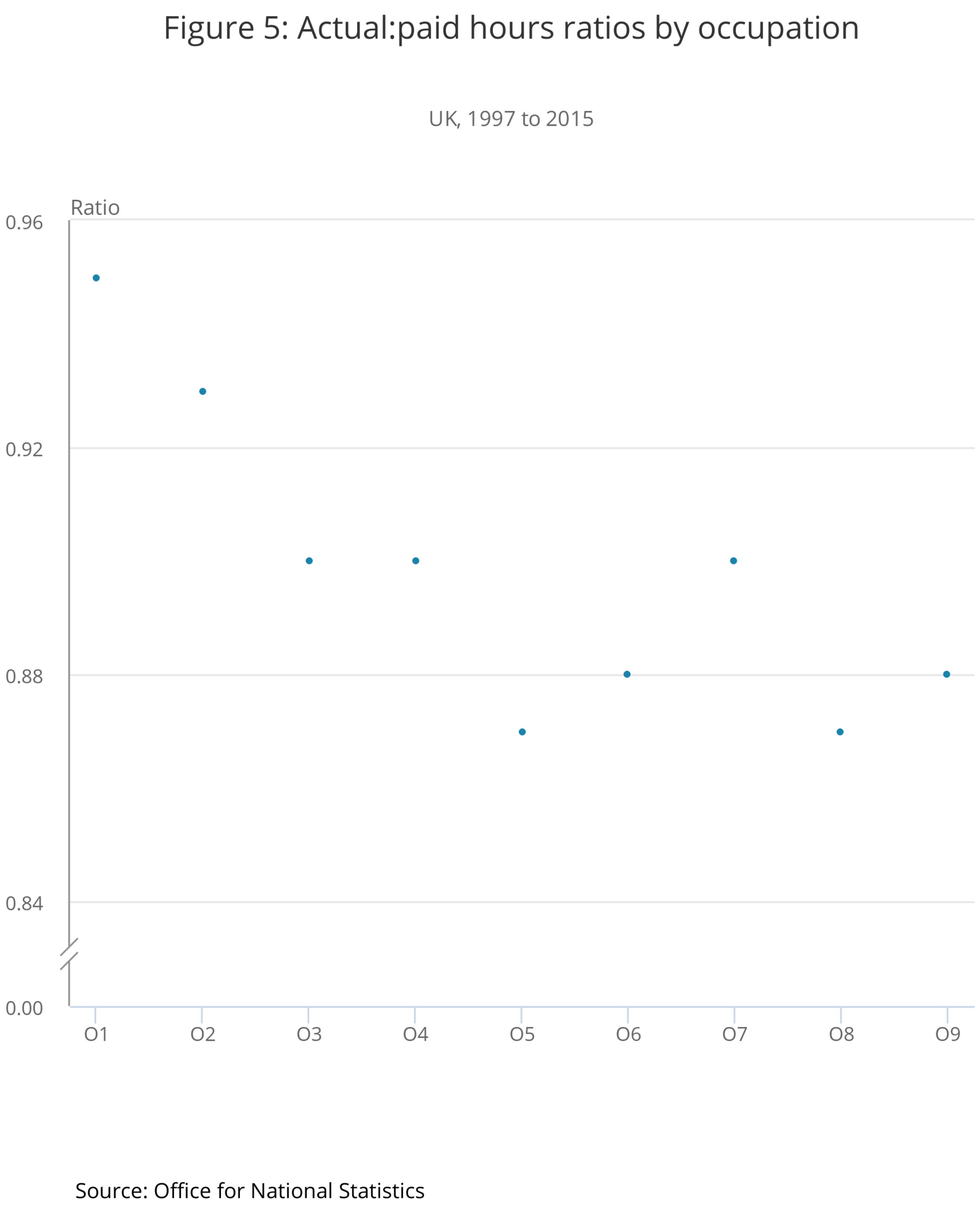
In this set of results, positive coefficients imply higher ratios of actual to paid hours and the other way around. For instance, we find a positive coefficient on the (male) sex dummy variable, indicating that all else equal, males tend to have a higher ratio of actual to usual hours worked than females. In interpreting the negative coefficients on the age and occupation dummy variables it should be borne in mind that these are relative to the first element in each classification. For the age category, for instance, the regression results imply that the actual:paid hours ratio is highest for the youngest age cohort and gets progressively lower for the older age groups. This might be because older workers have accrued more entitlement to paid leave, or perhaps because older workers take more sick leave than younger workers.

Individual industry coefficients are not reported in Table 4 although all are significant at the 0.1% level. Across industries, the actual:paid hours ratio is highest for the first industry in the classification structure (that is, industry A – agriculture, fishing and forestry) and variably lower (by roughly 2 to 10 percentage points) for the remaining industries. Similarly across the occupation groups, the actual:paid hours ratio is lower than the control group (managers, directors and senior officials) for all occupations, by roughly 2 to 8 percentage points.

Figure 5 shows the implied actual:paid hours ratios for each occupation group when the regression coefficients in Table 4 are enumerated for each age, sex, industry and occupation category and averaged over occupation groups. As expected, the ratio is largest for occupation group 1 and ratios are less than one for all occupations.

Figure 5: Actual:paid hours ratios by occupation

UK, 1997 to 2015



Source: Office for National Statistics

Notes:

1. Occupational groups 1 to 9 as described in Table 2; Estimates are unweighted averages from regression results categorised by industry, age group and sex, grouped by occupation and averaged over 1997 to 2015.

Adjustment factors by year, age group, sex, industry and occupation are applied to ASHE hourly earnings per paid hour estimates to derive a set of estimates of hourly earnings per (estimated) actual hour. For example, if the ASHE pay estimate per paid hour for a particular category is £20 per hour and the adjustment factor for this category is 0.95, then the adjusted pay estimate would be 20 divided by 0.95 equals £21.05 per (estimated) actual hour. It is these adjusted earnings estimates that we use in the benchmarking process, described in the following section.

5 . Benchmarking LFS to ASHE

The aim of the benchmarking exercise is to derive a set of hourly earnings by year, age, sex, industry and education, which, when grouped by occupation and weighted by shares of hours worked, are consistent with the adjusted Annual Survey of Hours and Earnings (ASHE)-based earnings estimates described in the previous section. It would of course be simpler to re-parameterise quality adjusted labour inputs (QALI) to replace education with occupation. But we are reluctant to go down this route because research using the Labour Force Survey (LFS) reveals stronger and more consistent relationships between earnings and education than between earnings and occupation. Education also tends to be used in QALI estimates compiled by other organisations in the UK and internationally.

However, since ASHE does not collect any information on education, we need some means of extrapolating ASHE component level earnings estimates across the six educational categories used in QALI. We propose to deal with this issue in three stages.

First, we use pooled LFS quarterly microdata to derive estimates of education pay relatives (that is, the hourly pay of each educational category relative to the average pay of all workers) and hours worked for each educational category by year, age group, sex, industry and occupation. We pool quarterly LFS datasets over 4 quarters centred on the ASHE sampling timetable. But even so, with 19 industries and nine occupations this entails dividing LFS into 6,156 separate cells.

Table 5 shows a breakdown of these 6,156 cells by occupation based on LFS annualised data between 2011 and 2015. The first row reveals large numbers of cells with missing hours worked for each occupation group. For example, 46.3% of cells for occupation group 8 have missing hours over this period. However, the principal problem of disaggregating into smaller categories is not missing hours worked as such, but the instances where hours worked estimates are present and pay observations are missing.

The second row of Table 5 shows that across occupations there are between 10.3% and 23.6% of cells with missing pay estimates and positive estimates of hours worked. These observations reflect the sampling nature of the LFS whereby individual respondents are surveyed up to five times over 5 quarters, but are only asked for pay information on the first and fifth interview.

The third row of Table 5 expresses the cells in the second row as percentages of all hours worked in each occupation. It shows that the problem cells only account for at most 1.8% of total hours worked in an occupation category and only 1.0% of hours worked across all occupations. Intuitively this is because cells in the second row of the table are likely to contain very few individual records and hence account for very small numbers of hours worked. As the number of records in a particular cell increases, so does the probability of observing a pay estimate. As a result using nine 1-digit occupation categories will mean educational pay relatives are ordinarily calculated using LFS pay data and only estimated for a small proportion of hours worked.

Table 5: Labour Force Survey cells with missing hours worked or pay estimates by occupation, 2011 to 2015

1-digit level occupation category	1	2	3	4	5	6	7	8	9	Total
Hours missing	20.0%	22.0%	13.0%	15.3%	33.7%	41.9%	29.9%	46.3%	25.8%	27.6%
Hours present, pay missing	12.3%	11.3%	10.3%	11.5%	16.2%	16.4%	18.2%	23.6%	13.7%	14.3%
as % of hours worked	1.3%	0.4%	0.9%	1.0%	1.2%	1.1%	1.6%	1.8%	1.2%	1.0%

Source: Office for National Statistics

The second stage of the proposed method involves addressing cells in which pay relatives are missing in the LFS data. To resolve this issue, we estimate pay relatives for missing cells using the results from a regression analysis of LFS microdata.

Regression models were estimated separately for each occupation group, as pay premia for higher qualifications are higher for high-skill occupations than for elementary occupations. For instance there are greater pay premiums for higher education levels for professional occupations (occupation group 2) than elementary occupations (group 9). Each regression model fits the log of hourly pay in that occupation from LFS on a set of controls for age group, sex, education and industry. Estimated coefficients on the education controls can then be interpreted as logs of pay of each education category relative to the pay of the no-qualification group.

Table 6: Sample regression results

Dependent variable	ln (hourly pay)			
Occupation group	1	3	4	7
Year	2015	2015	2015	2015
Controls	Age/sex/industry			
GCSEs or equivalent	0.209*	0.154*	0.128**	0.0631
	(2.52)	(1.97)	(2.95)	(1.78)
A – levels or trade apprenticeships	0.280***	0.248**	0.159***	0.179***
	(3.39)	(3.18)	(3.56)	(4.68)
Certificate of Education or equivalent	0.418***	0.287***	0.223***	0.219***
	(4.80)	(3.59)	(4.62)	(4.39)
First Degrees and other degrees	0.490***	0.394***	0.276***	0.261***
	(5.96)	(5.06)	(5.96)	(5.99)
Masters and doctorates	0.644***	0.539***	0.339***	0.278***
	(7.36)	(6.61)	(5.99)	(3.50)
R^2	0.1284	0.1281	0.0746	0.1656
N	4008	5457	5030	3396

t statistics in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Source: Office for National Statistics

Notes:

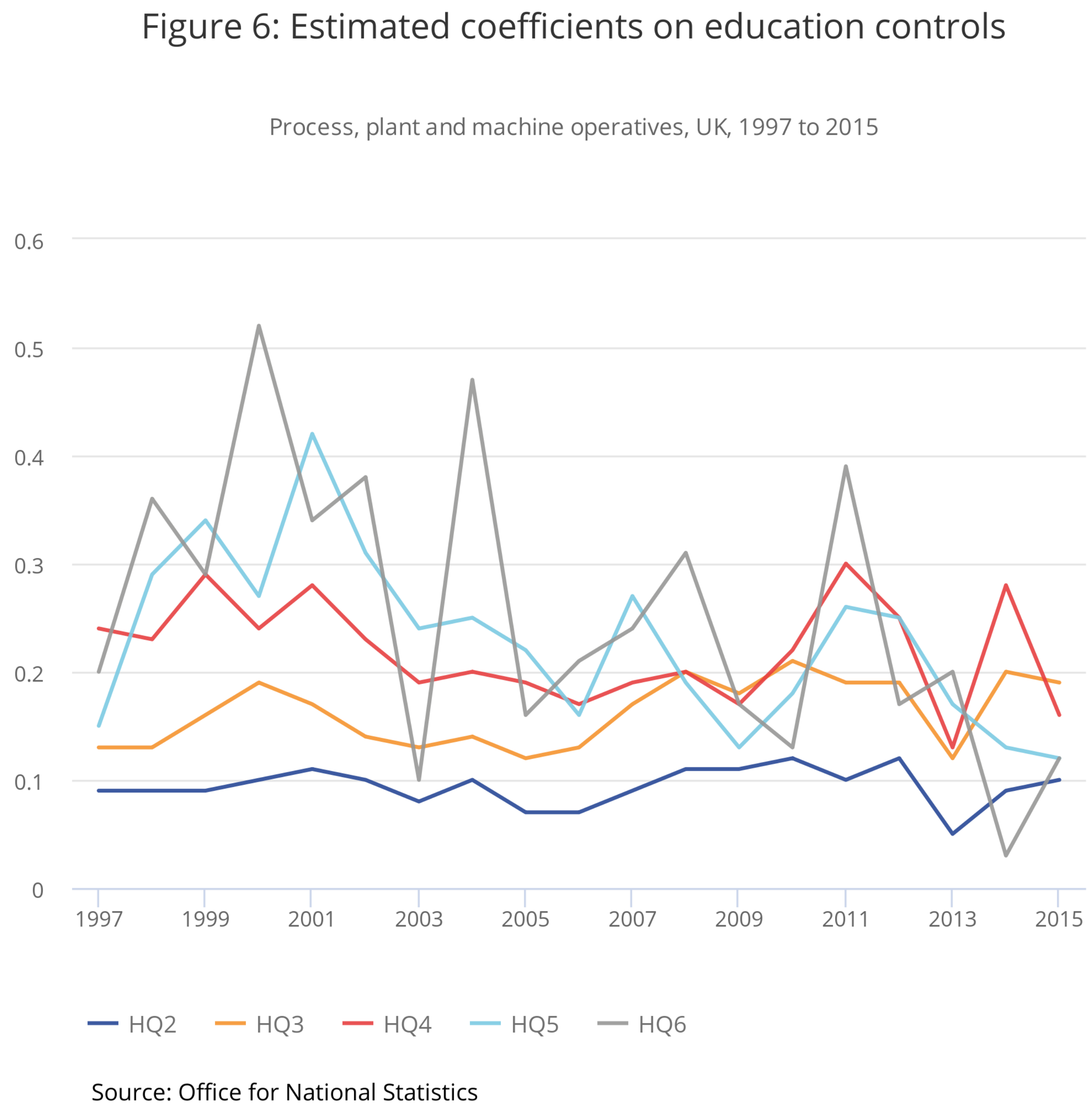
1. Occupation groups are described in Table 2.

Table 6 shows sample regression results for four occupation groups in 2015, with coefficients on levels of education representing the estimated contributions relative to those with no qualifications. Predictably those with higher qualifications in each occupation group receive greater remuneration than those with lower levels of education. And the relationship between increased pay for higher educational qualifications is stronger for lower occupational groups (that is, the more skilled occupations).

In order to capture changes in education pay relatives over time we run annual regressions for each occupational group. However, in some cases the resulting regression coefficients on education can be quite volatile. This is mainly as a result of small sample sizes for particular categories; where sample sizes are larger, the estimated coefficients tend to be more stable over time. Figure 6 plots the coefficients on the education controls from each annual regression in occupation group 8 (process, plant and machine operatives). It suggests that for this occupation group, those with postgraduate degrees (highest qualification 6 (HQ6)) are paid less than those with just A-levels (HQ3) or GCSEs (HQ2) for some years, but earn significant premiums in other years.

Figure 6: Estimated coefficients on education controls

Process, plant and machine operatives, UK, 1997 to 2015



Source: Office for National Statistics

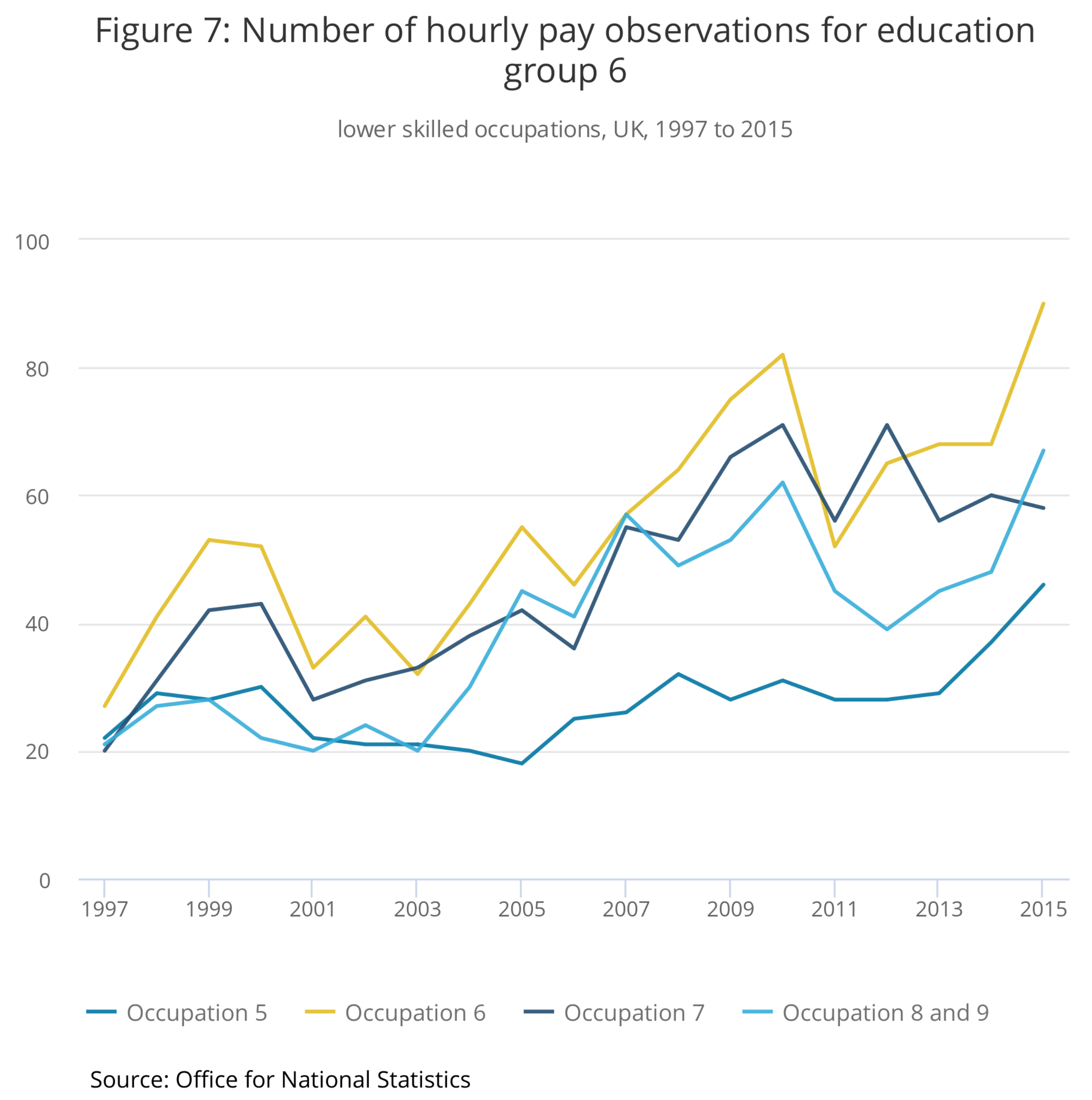
Notes:

1. Coefficients reflect log differences in hourly pay relative to the HQ1 no-qualifications group
   HQ2: GCSEs or equivalents
   HQ3: A-levels or trade apprenticeships
   HQ4: Certificates of education or equivalent
   HQ5: First and other degrees
   HQ6: Masters and doctorates

Further investigation reveals that the volatility of some coefficients shown in Figure 6 is associated with thin cell sizes for certain combinations of occupation and education. Figure 7 illustrates that there are a lack of pay records for occupation groups 5 to 9 (that is, the lower skilled occupations) for those in the highest education group (HQ6). As a result annual regression results are likely to be unreliable due to the small sample sizes. Generally speaking there is a diagonal relationship between occupation and education – workers in highly skilled occupations tending to be more highly educated and the other way around, with falling cell counts as we move away from the diagonal. Education categories HQ2 (GCSEs) and HQ3 (A-levels) generally have a minimum of 500 pay observations annually, but each of the other education groups have combinations with occupation groups where a shortage of observations can give rise to parameter volatility.

Figure 7: Number of hourly pay observations for education group 6

lower skilled occupations, UK, 1997 to 2015



Source: Office for National Statistics

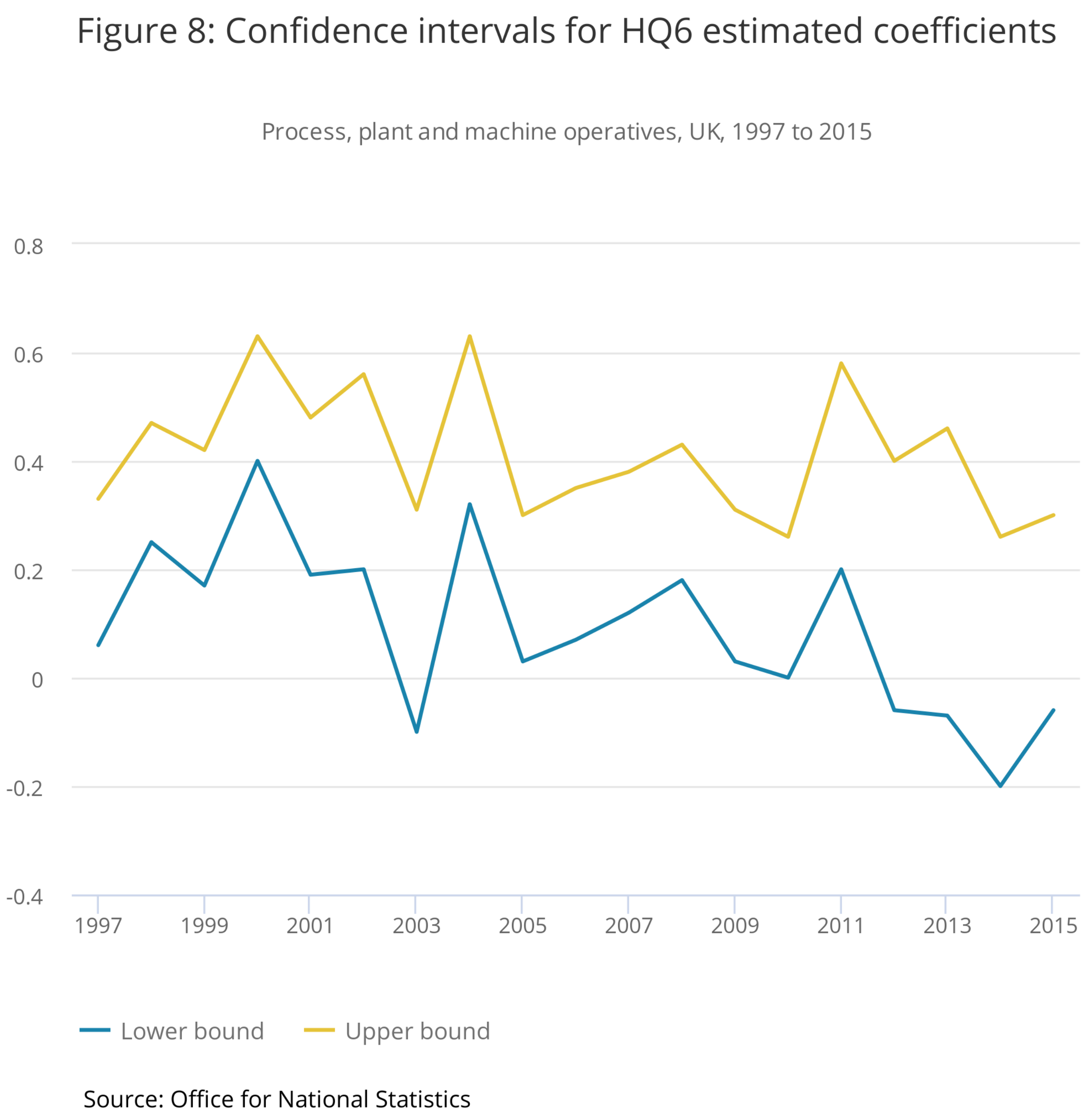
Notes:

1. Occupation groups are as described in Table 2.

Figure 8 shows 95% confidence intervals on estimated coefficients of those with postgraduate degrees in occupation group 8 (process, plant and machine operatives) over workers with no qualifications. Such large confidence intervals demonstrate the uncertainty of pay premia for combinations of education and occupation groups with few records.

Figure 8: Confidence intervals for HQ6 estimated coefficients

Process, plant and machine operatives, UK, 1997 to 2015

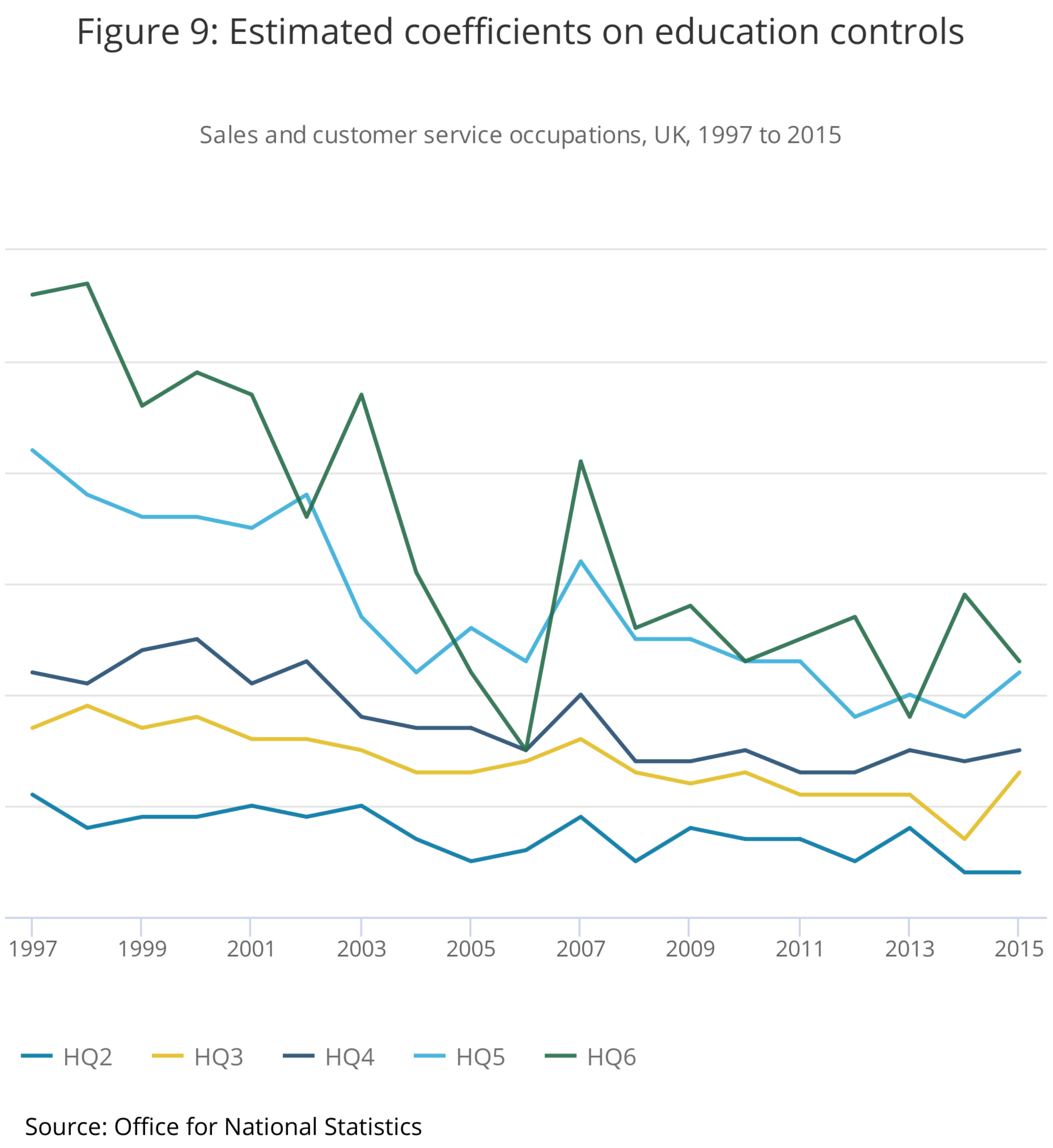


Source: Office for National Statistics

A potential solution to the problem of small sample sizes for combinations of education and occupation is to pool LFS micro-data over the entire time period from 1997 to 2015 and include a year variable in the regression specification. This would ensure that there are large enough sample sizes, but relies on there being either no trend in pay premia or a stable trend over in pay premia over time. For example, figure 9 shows that pay premia for sales and customer service occupations (occupation group 7) have fallen over time for each education group relative to workers with no qualifications.

Figure 9: Estimated coefficients on education controls

Sales and customer service occupations, UK, 1997 to 2015



Source: Office for National Statistics

Notes:

1. Education groups as described in Figure 6.

We plan to do some further work on annual versus panel regressions. However, it is worth re-stating that the share of hours worked where pay estimates are missing on LFS is typically very small, so the impact of alternative approaches to estimating pay relatives for these cells is limited.

Using the annual regression coefficients described in this section it is comparatively straightforward to populate a complete set of pay estimates for each of the 6,156 cells categorised by age group, sex, industry occupation and education. Aggregating across the six education categories we can then compute pay relatives (that is, the pay of each education category relative to the average pay of all education categories) for each age group, sex, industry, occupation and year.

Our approach is then simply to use the LFS pay relative where LFS pay data exist and the estimated pay relative otherwise. The final step is to convert from relatives to pay levels so as to hit the ASHE benchmark, taking account of the distribution of hours worked from LFS. This is illustrated in Table 7, which provides a stylised example of the proposed method to benchmark to ASHE for an example age, sex, industry and occupation category where there are empty cells for some education groups in the LFS data.

Table 7: Stylised example

		Hourly pay	hours worked (%)	Pay relatives		Adjusted Hourly pay	
				from LFS	estimated	unbenched	benched
LFS	No qualifications	.	2%		0.712	£13.53	£13.92
	GCSEs or equivalent	£14.00	4%	0.811		£15.40	£15.84
	A-Levels or trade apprenticeships	.	5%		0.782	£14.87	£15.29
	Certificate of education or equivalent	.	10%		0.892	£16.95	£17.43
	First degree or other degrees	£16.50	42%	0.955		£18.15	£18.66
	Masters and doctorates	£18.50	37%	1.071		£20.35	£20.93
	Weighted Average	£17.27	100%			£18.48	£19.00
ASHE		£19.00					

Source: Office for National Statistics

Estimated pay relatives for the missing observations are computed using regression coefficients. Unbenched estimates in the penultimate column are then computed as actual or estimated pay relatives multiplied by the ASHE benchmark and these estimates are re-scaled in the final column to take account of the distribution of hours worked.

Note that this process can result in cases as in this stylised example where the adjusted pay levels are not strictly monotonic with respect to education. However, as noted previously, in practice estimated pay accounts for only a tiny percentage of hours worked. Moreover at the detailed age, sex, industry or occupation category level there are examples within the LFS data of non-monotonic pay rates.

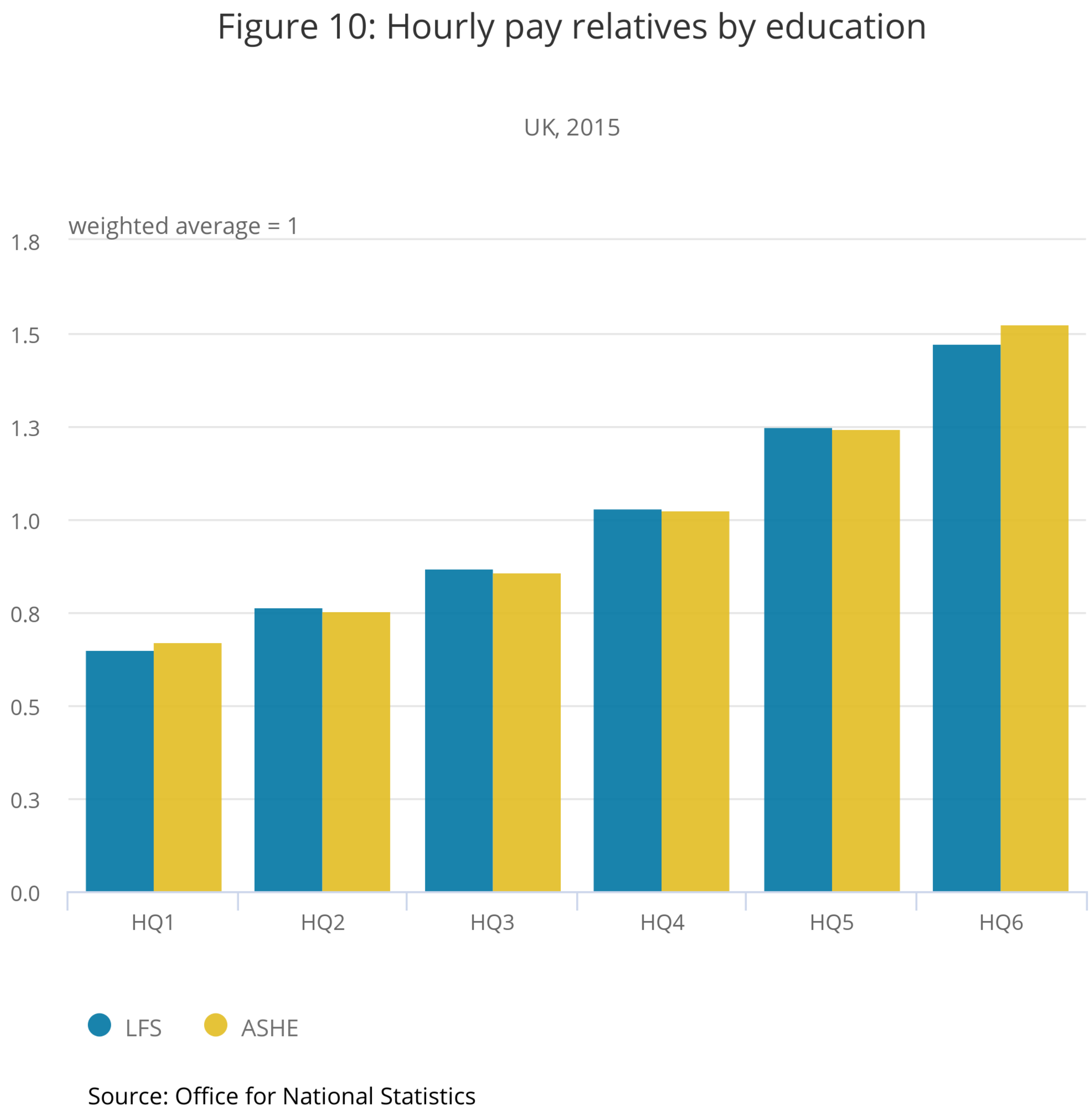
Provisional results

Here we focus on results across some of the QALI categories, in terms of differences in pay relatives between those taken purely from the LFS microdata and the results from the benchmarking exercise described earlier in this section. Further development, such as converting earlier occupational classifications, may lead to some changes in these results. You should also note that for comparison with ASHE, LFS quarterly datasets have been grouped on a Quarter 4 (October to December) to Quarter 3 (July to September) basis, rather than the calendar years reported in our QALI releases.

We do not present results by industry for two reasons. First, these results have been produced at 19-industry level rather than the 10-industry level used in QALI releases. Second, we intend to continue to benchmark industry level hours worked and aggregate labour remuneration (that is, the sum of employee and self-employed labour remuneration) to a set of top-down industry estimates derived from the income side of the national accounts. This means that, for industries unaffected by increased granularity such as manufacturing and construction, using ASHE pay information will have no effect on the aggregate industry pay weights. Pay relatives shown in this section are calculated before application of these industry-level constraints.

Figure 10: Hourly pay relatives by education

UK, 2015



Source: Office for National Statistics

Notes:

1. Education groups as described in Figure 6. LFS data are pooled over Q4 2014 to Q3 2015 and are relative to the weighted average of all LFS employees (=1). ASHE estimates are LFS education by occupation components benchmarked to ASHE hourly earnings adjusted for actual hours worked.

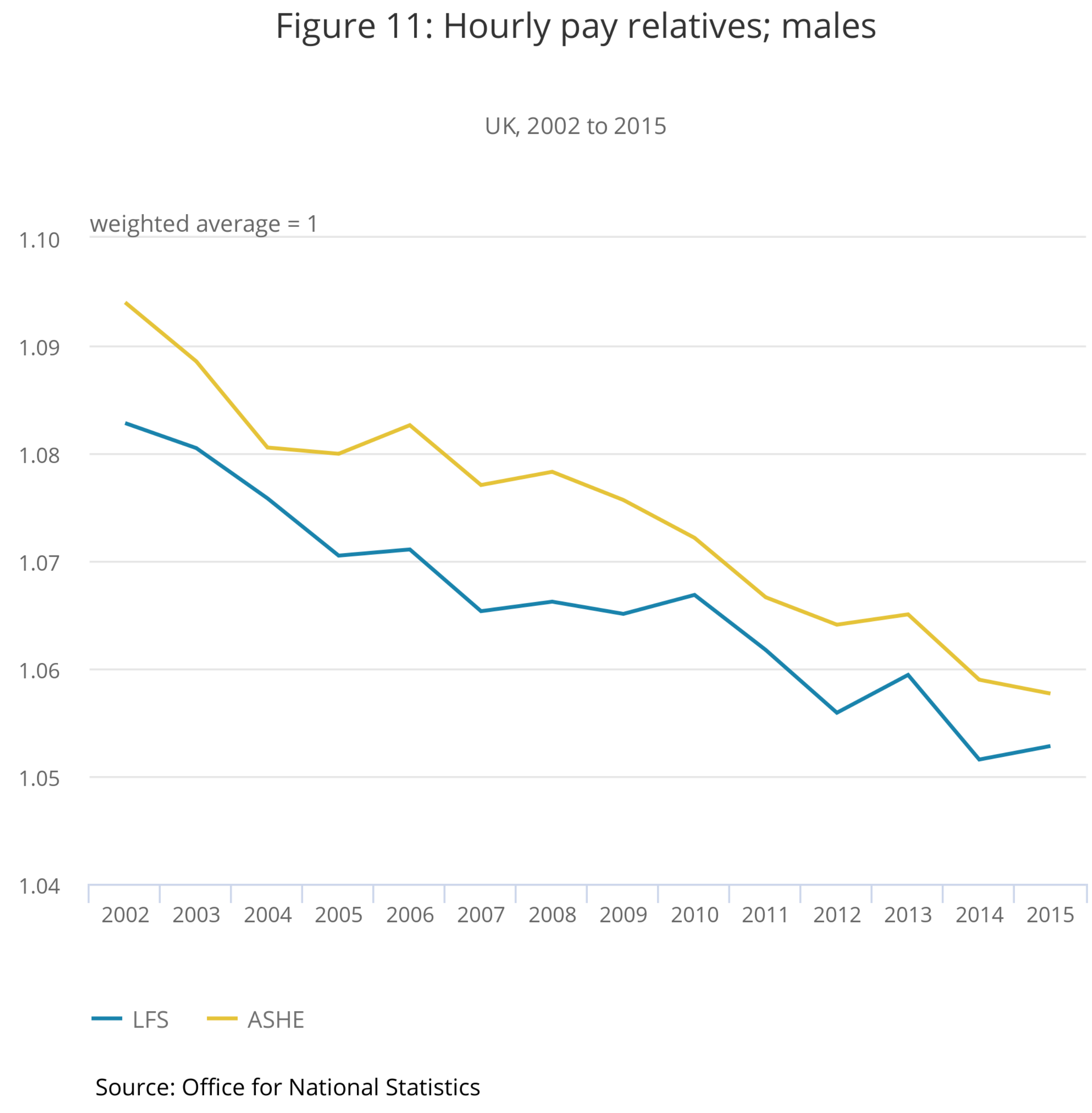
Pay relatives for several education categories are remarkably similar between the raw LFS data and the results implied by benchmarking to component level ASHE estimates. ASHE estimated pay relatives are higher than LFS for education category 1 (no qualifications) and education category 6 (masters and doctorates) and are a little lower than LFS for the intermediate categories (Figure 10).

The time series properties of the LFS and ASHE estimates are broadly similar for all education categories.

Benchmarking to ASHE component level pay estimates results in small increases in the pay relatives for males, compared with the raw LFS estimates, although the downward trend is similar and the gap between the two series has narrowed a little over time (Figure 11). The corollary is that ASHE pay relatives for females are a little lower than the LFS equivalents.

Figure 11: Hourly pay relatives; males

UK, 2002 to 2015



Source: Office for National Statistics

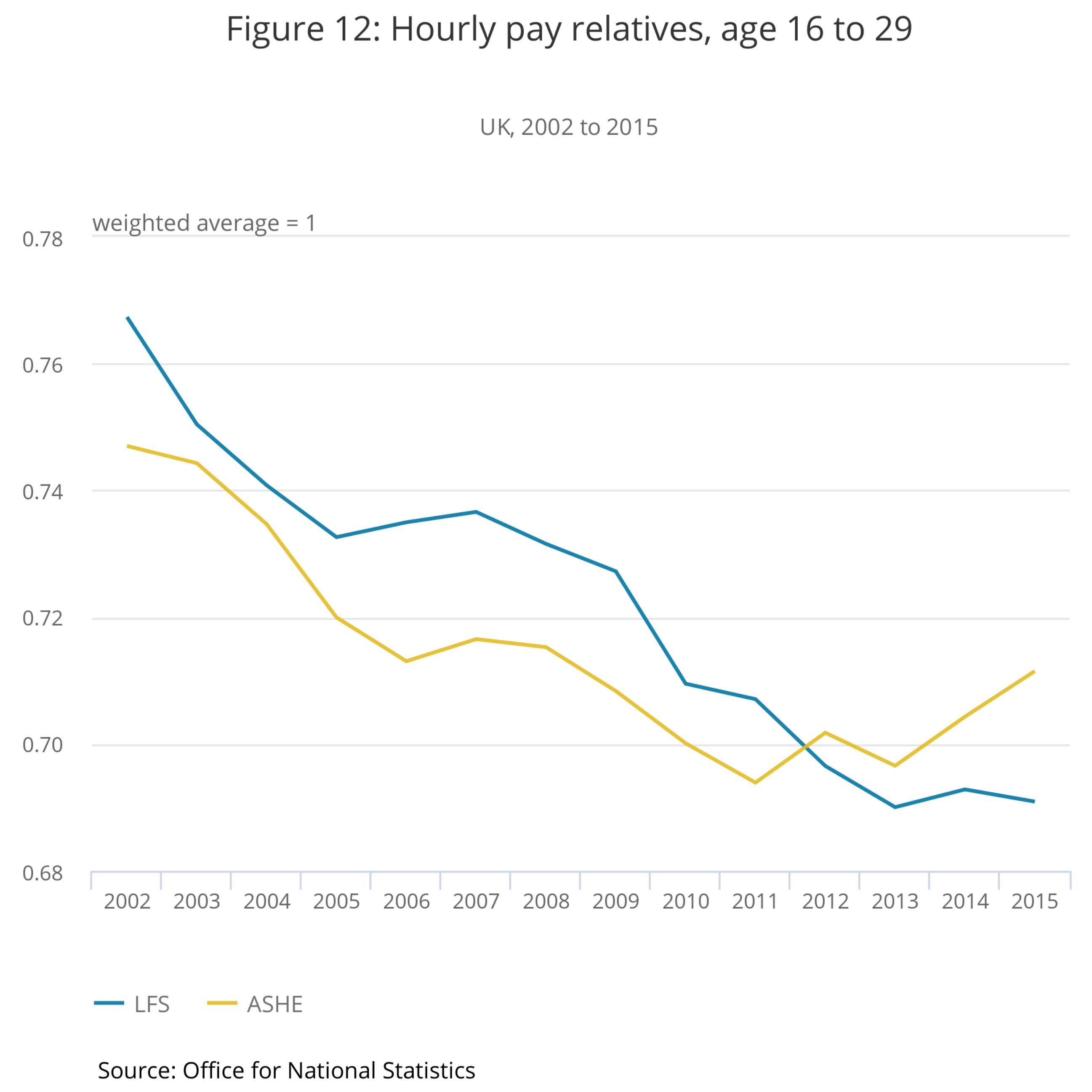
Notes:

1. Labour Force Survey data are relative to the weighted average of all LFS employees in each year.

ASHE estimates for the relative pay of the 16 to 29 age group are lower than LFS estimates up to 2011. However, ASHE shows an increase in relative pay for this cohort in recent years, in contrast to a broadly flat LFS profile and in contrast to the trend decline in relative pay for this group up to 2011 (Figure 12).

Figure 12: Hourly pay relatives, age 16 to 29

UK, 2002 to 2015



Source: Office for National Statistics

Notes:

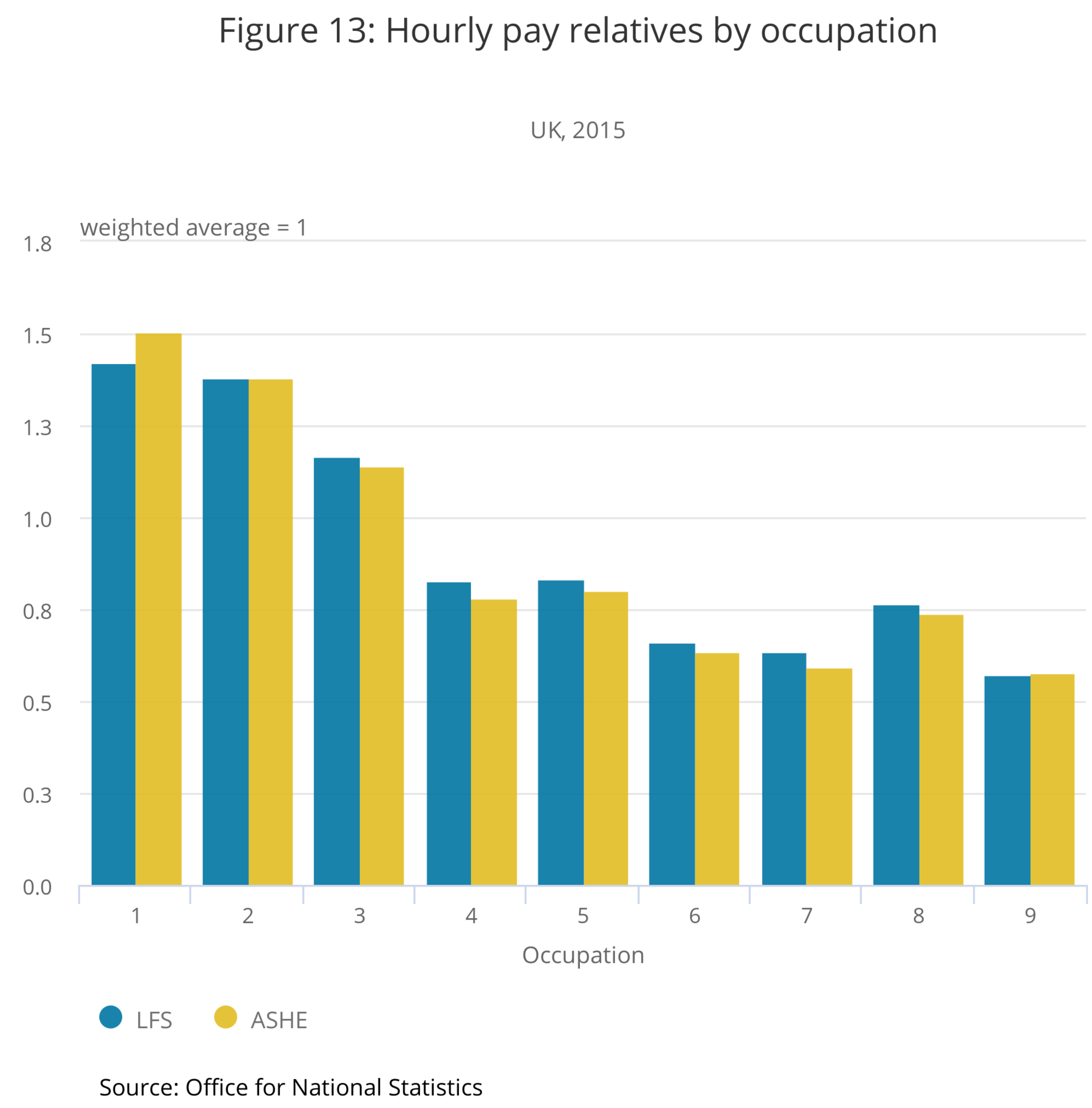
1. Labour Force Survey data are relative to the weighted average of all LFS employees in each year.

According to LFS, the pay premia of the 30 to 49 age cohort have been fairly stable at about 11 percentage points above the average of all employees over the period 2002 to 2015. The average premia over the whole period according to ASHE is virtually identical, although ASHE shows a more distinct trend over the period, with an average pay premia of around 12 percentage points over 2002 to 2008 and an average of around 10 percentage points since then.

Figure 13 shows a snapshot of pay relatives by occupation in 2015. Unlike education, pay is not monotonic in terms of the standard occupation taxonomy – the relative pay of occupation group 5 (skilled trades) is slightly higher than group 4 (administrative and secretarial) and relative pay of occupation group 8 (process, plant and machine operatives) is above that of groups 6 (caring, leisure and other services) and 7 (sales and customer service occupations). ASHE-based estimates are above LFS estimates for occupations 1 (managers, directors and senior officials), 2 (professional occupations) and 9 (elementary occupations) and below LFS estimates in the remaining occupations. The largest difference between the two sources is occupation 1, where ASHE estimates of relative pay were some 10 percentage points higher than LFS in 2015.

Figure 13: Hourly pay relatives by occupation

UK, 2015



Source: Office for National Statistics

Notes:

1. Occupation categories range from 1 (managers, directors, senior officials) to 9 (elementary occupations). A full list is shown in Table 2. LFS data are pooled over Q4 2014 to Q3 2015 and are relative to the weighted average of all LFS employees (=1). ASHE estimates are LFS education by occupation components benchmarked to ASHE hourly earnings adjusted for actual hours worked.

6 . Appendix 1: Re-visiting sectorisation using ASHE

A previous article described development of labour market metrics for the market sector. Estimates of market sector hours worked and labour remuneration at a 10-industry component level were used for the first time to derive component level market sector quality adjusted labour inputs (QALI) estimates used in our multi-factor productivity estimates estimates published on 5 April 2017. This methodology relied heavily on a sector marker in the Labour Force Survey (LFS) in identifying workers employed in the general government and non-profit institutions serving households (NPISH) institutional sectors.

There have been three developments since our previous article. First, a review of the mapping between the LFS marker and the national accounts revealed an inconsistency in that the national accounts currently treat universities as outside the market sector whereas we had allocated LFS respondents flagged as working in universities to the market sector (that is, not to the NPISH sector).

Second, the development of experimental QALI estimates for the non-market sector uncovered a few cases where the implied non-market benchmarks for labour remuneration became negative (that is, the market sector estimates were greater than the totals for that particular industry and time period). Further investigation revealed that this was due to the method used to benchmark overall market sector labour remuneration to a top-down estimate derived from our sector and financial accounts.

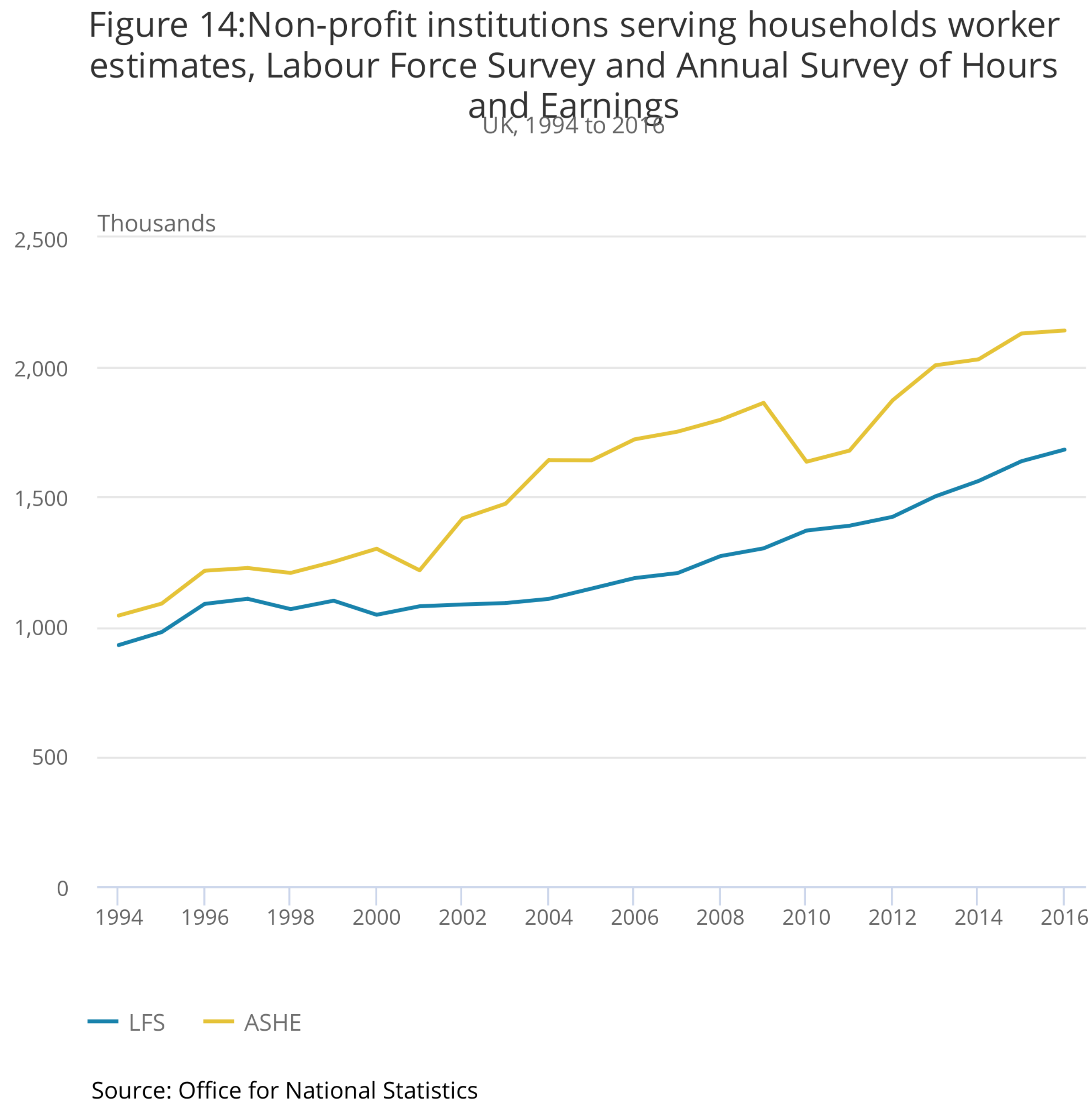
Third, the Annual Survey of Hours and Earnings (ASHE) provides an additional source of sectoral information, including more robust information than LFS on the numbers of NPISH workers and their distribution across industries and information on the sectoral distribution of second jobs, which is missing on LFS.

Workers

Re-classifying university workers to the non-market sector increases the NPISH category derived from LFS by about 700,000 workers. Moreover, ASHE estimates for employees in “non-profit or mutual association” workplaces are systematically higher than their LFS equivalents after this re-classification (Figure 14). The ASHE time series shows more pronounced downturns in 2001 and 2010, and the ASHE series has grown faster than the LFS series.

Figure 14:Non-profit institutions serving households worker estimates, Labour Force Survey and Annual Survey of Hours and Earnings

UK, 1994 to 2016



Source: Office for National Statistics

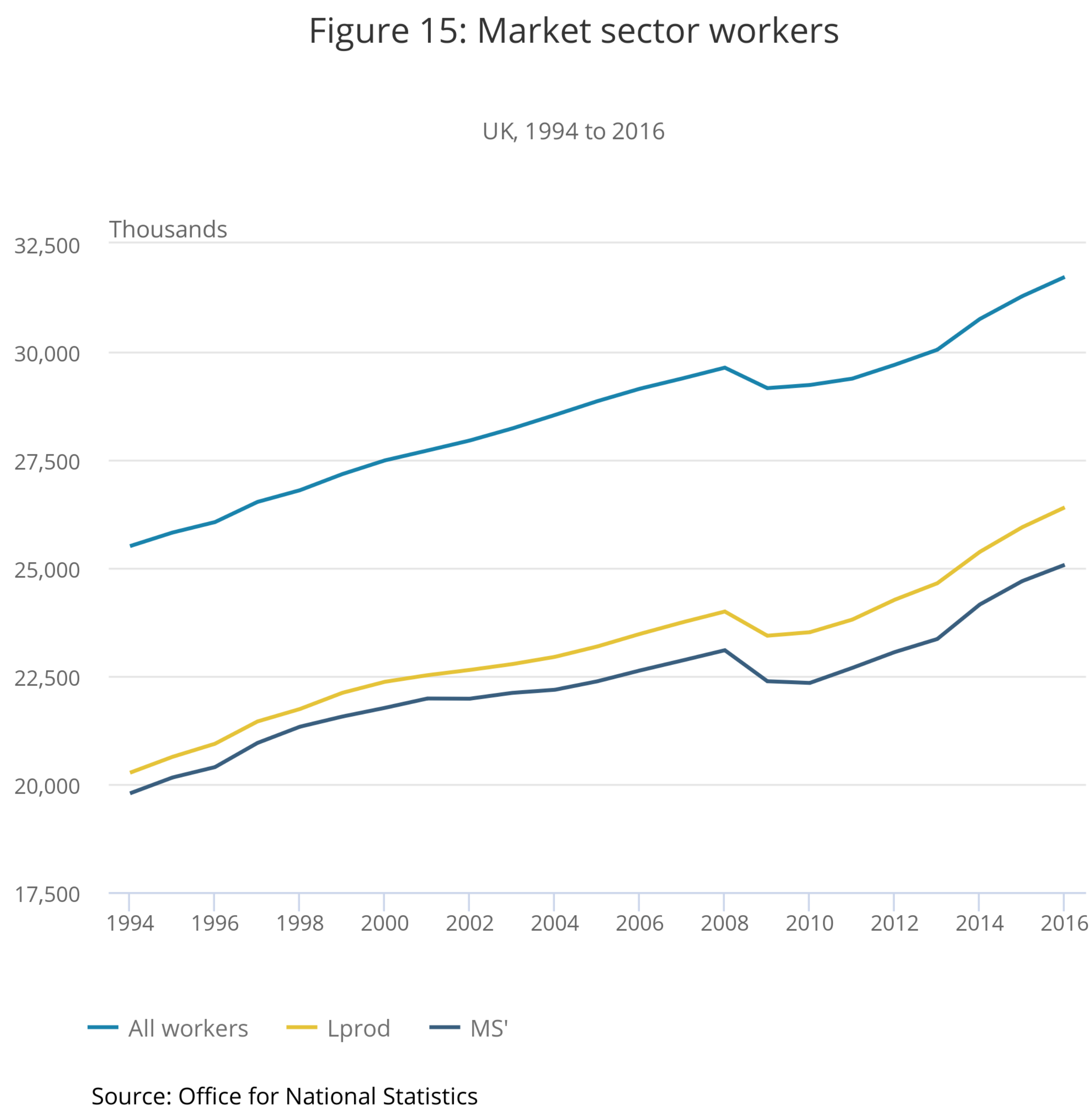
Notes:

1. Labour Force Survey estimates are annual averages centered on end-March. ASHE estimates prior to 1997 are backcasts using LFS growth rates.

We propose using ASHE as our main source for NPISH workers for the reasons noted previously, taking the ASHE estimate as an annual benchmarking and overlaying a quarterly profile from LFS. We propose to continue to use estimates derived from our survey of public sector employment as the source for general government workers by industry. This change has the effect of raising estimates of non-market sector workers, therefore reducing estimates of market sector workers and widening the gap between market sector worker estimates using this methodology and estimates of market sector workers used in our labour productivity system (Figure 15).

Figure 15: Market sector workers

UK, 1994 to 2016



Source: Office for National Statistics

Notes:

1. LPROD: Labour Productivity release

2. MS': This publication

Hours worked

We are proposing to make two small changes to the previous methodology, which applies average hours taken from LFS to the headcount measure of non-market sector workers to derive estimates of market sector hours worked by residual. First, we propose to use information from ASHE on hours worked in second jobs by sector and industry to fine-tune our adjustments for hours worked in second jobs. The previous method only used industry-level second jobs information from LFS as the LFS does not collect any sectoral information on second jobs.

Readers will recall that ASHE collects information on paid hours rather than actual hours worked; section 4 of this article discusses this issue in much more detail. Currently our adjustments between paid and actual hours do not separately distinguish hours worked in second jobs. We intend to review this issue in the future. In the present context, we assume that the uplift for hours worked in second jobs in terms of ratios of actual hours in first and second jobs on LFS can be proxied by the ratio of paid hours in first and second jobs for the equivalent category of worker on ASHE.

Second, we propose to apply an adjustment such that the sum of market sector and non-market sector hours worked in each industry is always equal to the estimate of total hours worked in that industry as in the labour productivity system. This method jointly benchmarks the two sectoral components, whereas the previous method computed estimates for non-market sector hours worked and then calculated market sector hours worked as the residual. A consequence of this change is that the small differences between industry level estimates of hours worked between the sectoral decompositions and the labour productivity system noted in our previous article (Figure 3) disappear.

Labour remuneration

We propose to make a number of changes to the previous methodology in order to remove anomalies, improve consistency with other published estimates and utilise additional information from ASHE.

The basic approach is to take ASHE information on labour remuneration (pay per paid hour plus employer pension contributions) by industry, converted from annual to quarterly frequency using LFS trajectories and multiplied by hours worked estimates derived as previously. These unbenched estimates are replaced by published compensation of employment (COE) series where these are available and used to split aggregated industry-level COE down to the 19-industry level in the remaining cases. For example, published COE estimates are available only for industries G, H and I combined, so we split these estimates into G, H and I separately using the unbenched industry shares.

For the hybrid, part market sector, part non-market sector industries, sectorisation proceeds by apportioning COE according to the unbenched market:non-market shares derived analogously to the combined industry-level unbenched estimates. This method yields market:non-market COE estimates that reflect the shares of hours worked and the ASHE-based differences in hourly remuneration between the market and non-market cohorts.

It should be noted that this approach generates differences between the sum of industry-level market sector components of COE and the top-down estimate of market sector COE derived from the sector and financial accounts and currently used in the compilation of market sector unit labour costs (Figure 16).

Figure 16: Market sector compensation of employment, shares of total

UK, 1994 to 2016

Source: Office for National Statistics

Notes:

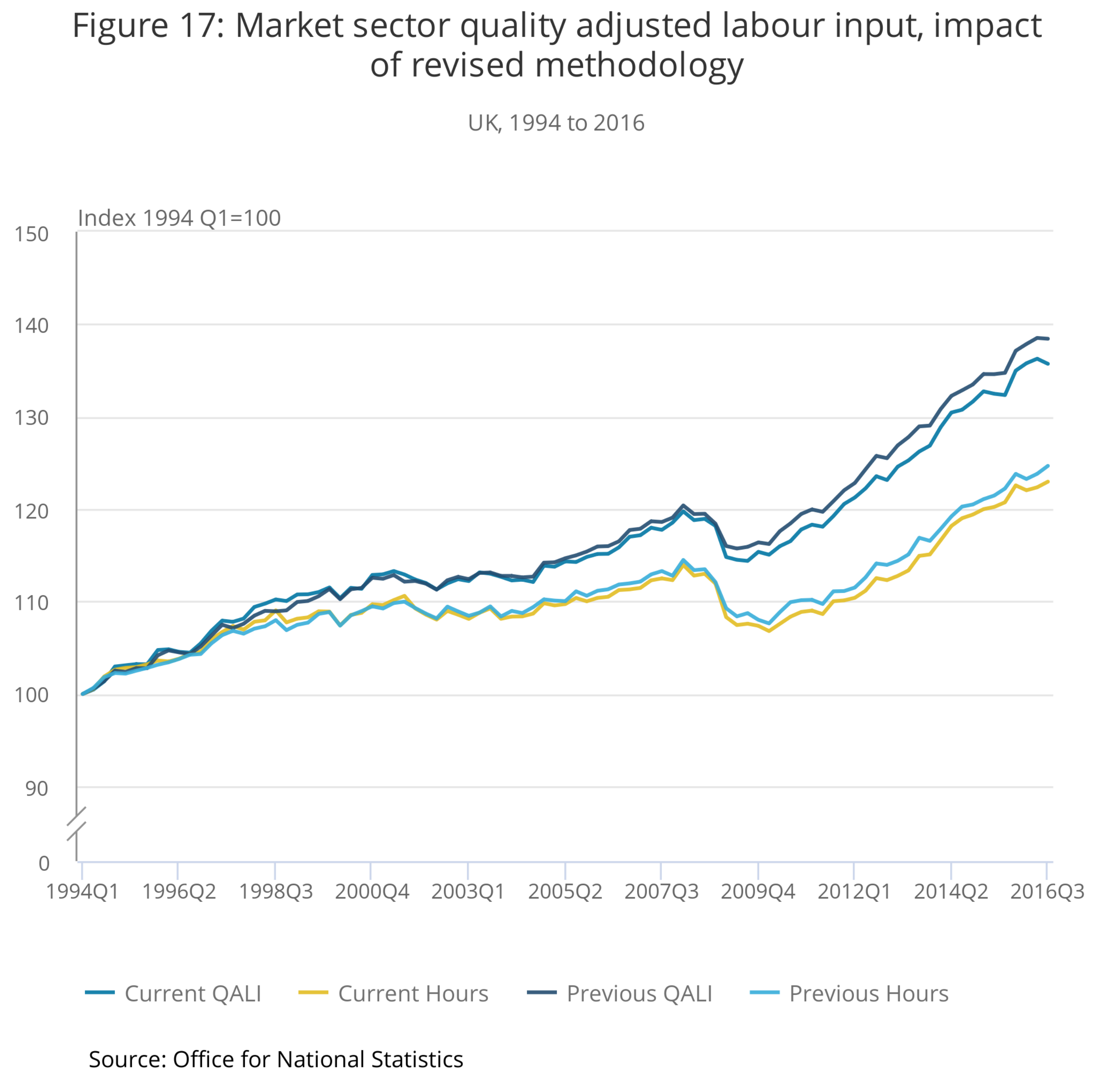
1. Top down estimates from Sector and Financial Accounts, sum of corporate and household sector COE; Bottom up estimates are sums of industry level market sector COE, sectorised as described in the text.

We prefer the bottom-up estimates because they are conceptually closer to the derivation of market sector hours worked. By contrast, the top-down series is derived from the sector and financial accounts, which are compiled at some distance from the compilation of industry level gross value added (GVA) and its income components.

Impact on QALI

Figure 17: Market sector quality adjusted labour input, impact of revised methodology

UK, 1994 to 2016



Source: Office for National Statistics

Notes:

1. Previous is market sector QALI at 05/04/17. Current shows impact on QALI index and index of hours worked of revised estimates of market sector hours worked and labour remuneration as described in the text.

2. Q1 refers to Quarter 1 (Jan to Mar)
   Q2 refers to Quarter 2 (Apr to June)
   Q3 refers to Quarter 3 (July to Sept)
   Q4 refers to Quarter 4 (Oct to Dec)

Figure 17 summarises the impact of revised market sector estimates of hours worked and labour remuneration on market sector QALI, where the baseline estimates are as in our multi-factor productivity release dated 5 April 2017. Note that these impacts do not include the impacts of the minor methodological changes described in the Appendix 2. The main impact is on hours worked, reflecting the use of ASHE data on NPISH employment, which pushes up non-market estimates of hours worked and pushes down our market sector estimates.

There is some variation in the impact on hours worked across different QALI categories and some limited impact on labour composition in a few QALI categories including industries OPQ, RSTU, females and workers with A-levels and equivalent qualification. There are, of course, no impacts on industries that are entirely market sector, because in these industries, hours worked and total labour income are both unchanged.

7 . Appendix 2: Changes to treatment of LFS respondents who do not report their level of education

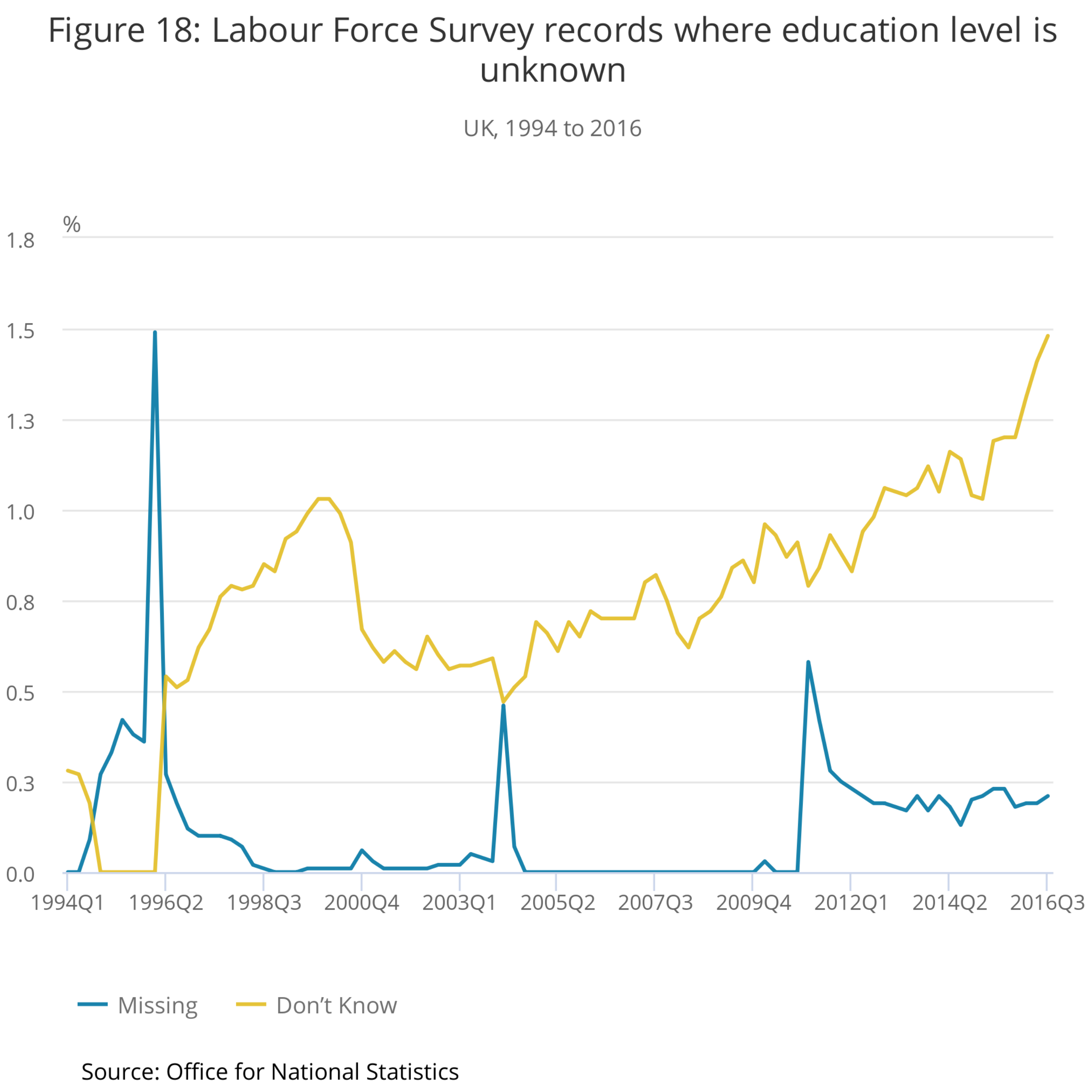
Quality Adjusted Labour Index (QALI) currently drops Labour Force Survey (LFS) records with a missing response for highest educational qualification and reassigns “don’t know” responses proportionately among different education groups according to the proportions within each QALI category of those who do report their level of education. So for instance if in a calendar quarter 40% of the hours worked by men aged 16 to 29 in industry F (construction) were those with no qualifications and 60% of the hours were worked by those with GCSEs (highest qualification 2(HQ2)), then 40% of the hours worked by those with a “don’t know” response for education would be reallocated to HQ1 of that category and the other 60% to HQ2. The same process would be used to redistribute pay; by reallocating the pay of those with missing education records according to the pay proportions for each QALI category.

There are three changes that we are looking to make with the way that we treat data with missing educational records. The first is to treat missing education records in the same way that “don’t know” responses for education are treated. The second change is to use previous and subsequent LFS responses where possible, to determine education level where there is no data on education. The third and final proposed change is to alter the way in which missing education responses are reallocated, by moving them to HQ2; as opposed to proportionately reallocating them according to the hours worked and pay in each QALI category. The reasoning behind each of the proposed changes will be set out and then the effects on QALI will be analysed.

The first proposed change is that responses with missing education are not dropped but treated in the same manner as “don’t know” responses. Figure 18 shows the percentage of hours worked by LFS respondents with “don’t know” responses and those that do not report their education level (missing). It shows that from the fourth quarter (October to December) of 1994 to the first quarter (January to March) of 1996 it appears that records with a “don’t know” response were recorded as missing, as there are no “don’t know” responses in this period. For long periods there are no missing records for highest educational qualification. Given that there does not appear to be a consistent collection of “don’t know” responses and missing responses over time, there is little justification for treating them differently in QALI.

Figure 18: Labour Force Survey records where education level is unknown

UK, 1994 to 2016



Source: Office for National Statistics

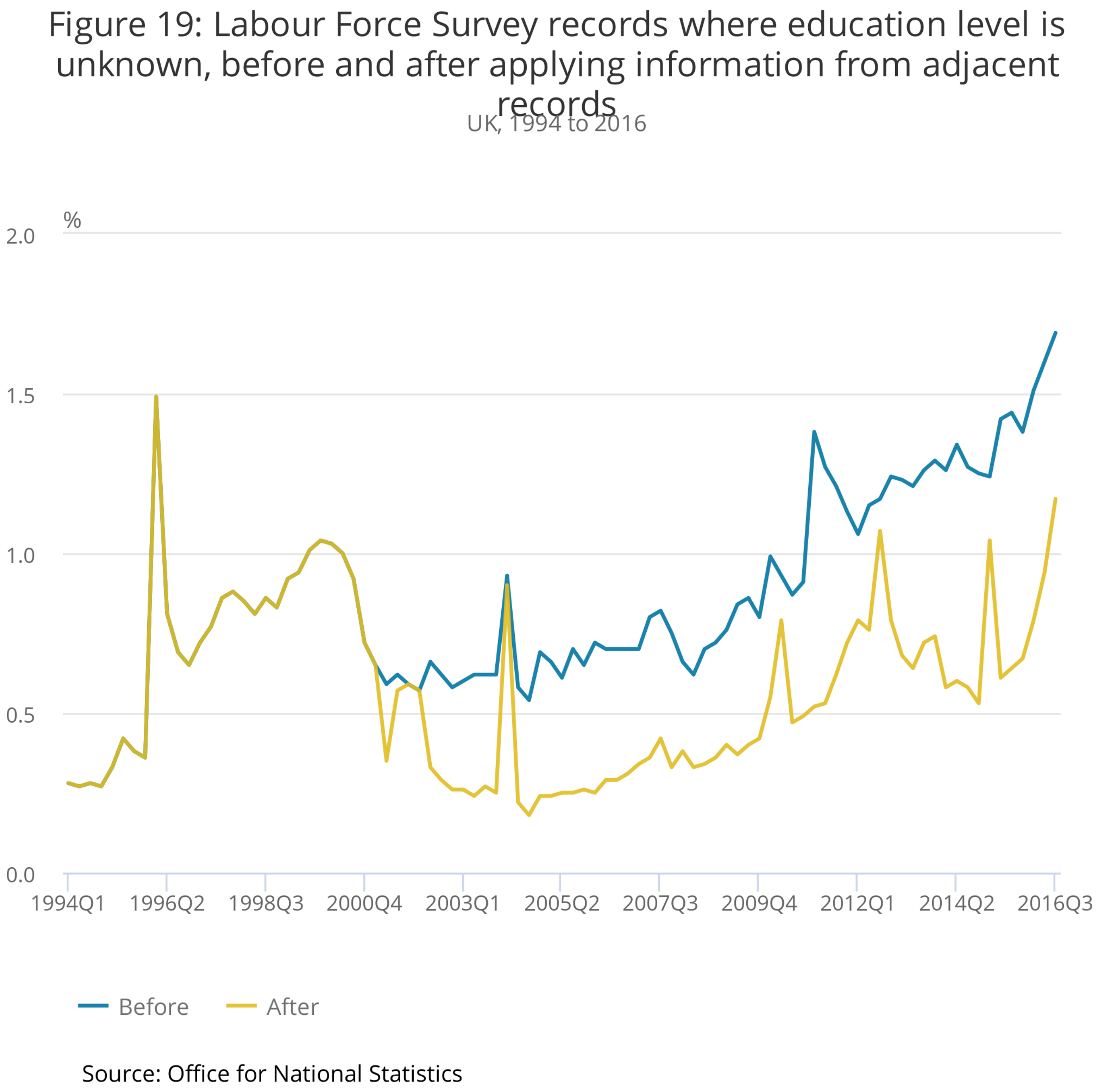
Notes:

1. Q1 refers to Quarter 1 (Jan to Mar)
   Q2 refers to Quarter 2 (Apr to June)
   Q3 refers to Quarter 3 (July to Sept)
   Q4 refers to Quarter 4 (Oct to Dec)

In order to reduce the number of observations with missing or “don’t know” educational data, the LFS person identifier was used to identify if there were educational qualifications records for the same individual in their previous or subsequent responses. Given that the highest educational qualification achieved is fairly stable over time, it is likely that this is a fairly accurate predictor of education. Figure 19 illustrates the number of LFS records without data on education before and after previous and subsequent records are taken into account. This new method results in a significant reduction in the percentage of hours worked for records where education is unknown after 2001 (there is no person identifier before 2001).

Figure 19: Labour Force Survey records where education level is unknown, before and after applying information from adjacent records

UK, 1994 to 2016



Source: Office for National Statistics

Notes:

1. Q1 refers to Quarter 1 (Jan to Mar)
   Q2 refers to Quarter 2 (Apr to June)
   Q3 refers to Quarter 3 (July to Sept)
   Q4 refers to Quarter 4 (Oct to Dec)

The third proposed change is to add those with missing or “don’t know” responses to those with GCSEs, rather than reallocate them to different educational groups according to each QALI category. A multiple linear regression of hourly remuneration controlling for age, sex, industry and year reveals that there is no significant difference between the hourly remuneration of those with missing or “don’t know” educational data and those with GCSEs.

The combined effect of including missing entries for education, using adjacent responses for educational data and then reassigning those with missing education data to HQ2, results in a small fall of 0.17% in the QALI index. All these changes to the QALI index are as a result of labour quality, as hours worked are constrained using labour productivity figures. There are also a number of small changes in the QALI index by industry, education, age and sex.

The only industries with an increase in the QALI index are R (arts, entertainment and recreation) by 0.42% and K (financial and insurance activities), which increases by 0.18%. The industries with the largest falls are A (agriculture, forestry and fishing) by 0.37% and industry H (transport and storage) also by 0.37%. The changes by industry are also entirely as a result of labour quality changes as hours worked are constrained by industry.

There is an increase in the QALI index for HQ2s of 2.64% and by 0.84% for HQ1s, with hours worked increasing by 2.45% and 0.29% respectively. By contrast all other education groups experience falls in the QALI index by 0.83 to 1.11% and falls in hours worked from 0.82 to 1.00%. The increase in the HQ1 and HQ2 indices are mainly as a result of the increase in hours worked by those with missing education data, with the changes in methodology resulting in a larger proportion of hours worked in the lowest two education groups. The increase in the hours worked of lower-educated workers at the expense of higher-educated groups, leads to the small reduction in the QALI index.

The QALI index for each age group falls, with the largest fall of 0.32% for 16 to 29 year olds and smaller falls of 0.13% for 30 to 49 year olds and 0.21% for those over 50. There is a slightly greater fall in the female QALI index of 0.26% compared with a fall of 0.13% for the male QALI index.

8 . Links to related publications

5 July 2017: UK productivity introduction: Jan to Mar 2017 draws together the headlines of the productivity releases into a single release, providing additional analysis of our productivity statistics.

5 July 2017: Labour productivity: Jan to Mar 2017 contains the latest estimates of labour productivity for the whole economy and a range of industries, together with estimates of unit labour costs.

5 July 2017: Introducing industry-by-region labour metrics and productivity presents new, experimental industry-by-region productivity metrics. This includes measures of hours worked, jobs, and accompanying productivity measures for the SIC letter industries in the NUTS1 regions.

5 July 2017: Who are the “laggards”? Understanding firms in the bottom 10% of the labour productivity distribution in Great Britain examines the characteristics of businesses in the bottom 10% of the labour productivity distribution in terms of their size, age, industry and location, between 2003 and 2015.

5 July 2017: Developing improved estimates of Quality Adjusted Labour Inputs using the Annual Survey of Hours and Earnings: A progress report describes work to improve the precision of income weights used in quality adjustment and to develop finer industry granularity of quality adjusted labour input for multi-factor productivity.

5 July 2017: Developing new measures of infrastructure investment: July 2017 is the first in a series of papers on infrastructure statistics, focusing on definitional and data challenges in measuring infrastructure investment.

5 July 2017: Quarterly public service productivity (experimental statistics): Jan to Mar 2017 presents experimental estimates for quarterly UK total public service productivity, inputs and output to provide a short-term, timely indicator of the future path of annual public service productivity estimates.

5 April 2017: International comparisons of UK productivity (ICP), final estimates: 2015 presents an international comparison of labour productivity across the G7 nations, in terms of growth in GDP per hour and GDP per worker.

5 April 2017: Multi-factor productivity estimates: Experimental estimates to 2015 decomposes output growth into the contributions that can be accounted for by labour and capital inputs. The contribution of labour is further decomposed into quantity (hours worked) and quality dimensions.

5 April 2017: Labour productivity measures from the Annual Business Survey, 2006 to 2015 presents an analysis of detailed productivity trends and distributions among businesses in the UK from 2006 to 2015, using firm-level data from the Annual Business Survey (ABS).

5 April 2017: Introducing quarterly regional labour input metrics provides a first look at the new experimental quarterly regional labour input metrics. Hours and jobs for the NUTS1 regions.

5 April 2017: Exploring labour productivity in rural and urban areas in Great Britain investigates differences in rural and urban labour productivity in Great Britain using firm-level microdata analysis of the business economy.

6 January 2017: Regional and sub-regional productivity in the UK: Jan 2017 provides statistics for several measures of labour productivity. Statistics are provided for the NUTS1, NUTS2 and NUTS3 subregions of the UK, and for selected UK city regions.

6 January 2017: Regional firm-level productivity analysis for the non-financial business economy: Jan 2017 provides experimental analysis on the sources of regional differences in labour productivity in the non-financial business economy in Great Britain.

6 January 2017: Volume index of UK capital services (experimental): estimates to 2015 provide estimates of the contribution of the capital stock to production in the economy, split by asset and industry.

6 January 2017: Public service productivity estimates: total public service, UK: 2014 presents updated measures of output, inputs and productivity for public services in the UK between 1997 and 2013, in addition to new estimates for 2014. Includes service area breakdown, as well as impact of quality adjustment and latest revisions.

6 January 2017: Public service productivity estimates: healthcare, 2014 presents updated estimates of output, inputs and productivity for public service healthcare in the UK between 1995 and 2013, and new estimates for 2014.

6 October 2016: Quality adjusted labour input: UK estimates to 2015 includes estimates of changes in the number of hours supplied in the UK economy adjusted for changes in the quality of the labour supply.

9 . Authors

Mark Franklin and Kris Johannsson.