RESEARCH

BIS | Department for Business Innovation & Skills

BIS RESEARCH PAPER NO. 48 Annex 1

Individualised Learner Record / Labour Force Survey earnings comparisons

JUNE 2011

A report by Frontier Economics and the Institute for Fiscal Studies (IFS)

Annex 1 to 'Reporting on employment and earnings using experimental matched data'

Department for Business, Innovation and Skills

1 Victoria Street

London, SW1H 0ET

www.BIS.gov.uk

BIS Research paper number 48 (Annex 1)

June 2011

Contents

Executive summary

BIS asked Frontier Economics and the Institute for Fiscal Studies to assess the feasibility of comparing earnings in the ILR-HMRC/DWP matched data with earnings derived from other sources: the Annual Survey of Hours and Earnings ("ASHE") and the Labour Force Survey ("LFS"). These comparisons help cast light on questions such as:

- *Do the earnings distributions for a given group calculated in ILR-HMRC/DWP look reasonable?* This will help us to assess the validity of the earnings distributions in the matched data.

- *How do earnings differentials in the ILR-HMRC/DWP (by qualification group) compare to earnings differentials found in other data sources?* This will provide a means of validating the impacts we have estimated in previous work.

- *Can other data sources be used to construct a control group?*

Our assessment of the data sources suggests that:

1. ASHE does not provide appropriate comparisons.

- Since the ASHE survey does not record the type or level of individuals' qualifications, groups in ASHE cannot reliably be compared to learners in the ILR.

- ASHE focuses only on employed individuals, whereas the learners in the matched data will be involved in a variety of economic activities.

- The earnings distribution for a typical comparator group looks very dissimilar to the matched data.

2. The LFS may produce some valid comparisons.

- We can define groups within both the LFS and matched data in terms of age, level and subject of qualifications and demographic characteristics.

- The LFS contains individuals in a mix of economic activities, like learners in the matched data.

- Earnings distributions look very similar between datasets for some comparator groups.

Defining the comparator groups correctly is a reasonably complex exercise. The challenge is to define the groups narrowly enough so that the characteristics of the groups in each dataset are similar. On the other hand, if the groups are defined too narrowly, then sample sizes may become too small, so that the results become inaccurate and the comparisons become invalid. Our attempt at finding comparator groups finds some cases where the earnings distributions look fairly similar between datasets.

However, we should also draw attention to the following:

- With the matched data we do a genuine before and after comparison. With the LFS data we are merely comparing individuals with different levels of qualification. We therefore risk conflating changes over time with differences between individuals.

- In the LFS we cannot easily pin down the time of learning. If the impact of learning changes over time for an individual, we would risk comparing the impact at the 'wrong' time.

- The LFS is representative of the UK population as a whole, not the specific population of FE learners. Although we can compare groups defined by gender, age, and other demographic factors, the composition of the two comparator groups are unlikely to be the same. In theory, one could reweight the LFS sample to make it more representative of FE learners, but this task might take considerable effort.

Introduction

In previous work[1], we used newly matched data from the Department for Business, Innovation and Skills (BIS) to attempt to cast light on the employment and earnings outcomes of individuals who undertake training. The unique dataset used for this analysis was created through the linkage of Individualised Learner Record (ILR) data held by BIS (which records information about learners in the Further Education sector) and the Work and Pensions Longitudinal Study (WPLS) held by the Department for Work and Pensions (DWP), which records individuals' benefits information, as well as their employment records from Her Majesty's Revenue and Customs (HMRC).

In our previous work, we relied on the WPLS data to give an indication of the earnings distributions of different groups of learners, and the growth in earnings they experienced over time (notably before and after receiving training). The data we used contained information on all individuals who have been in learning in the Further Education (FE), Train to Gain (TTG) and Apprenticeship funding streams in 2005/06 and 2006/07.

We have no reason to believe that WPLS data is anything other than representative of the sample it covers – those who have experience of FE. However, given the novelty of the matched ILR-HMRC/DWP dataset, it is important to establish whether the new dataset provides robust, reliable measures of earnings for this population. It is therefore instructive to compare the earnings distribution (and, if possible, the growth in earnings) among groups in the matched ILR-HMRC/DWP data with the earnings of similar groups in other national datasets.

However, creating such comparisons is far from straightforward. There are no other national datasets which explicitly set out to measure the earnings of individuals with experience of the FE system, so we must instead draw comparison groups from surveys created for other purposes. However, such surveys may contain limited information regarding individuals' education and qualifications (making it difficult to define similar comparator groups between surveys), or simply have very limited sample sizes for our population of interest.

In this report we assess the feasibility of making earnings comparisons with two of the UK's largest surveys of employment outcomes:

- *The Annual Survey of Hours and Earnings (ASHE).* This is a data source providing information about the levels, distribution and make-up of earnings and hours for employees in the UK, based on a one percent sample of employee jobs taken from HMRC Pay As You Earn (PAYE) tax records[2]. This sampling strategy means that ASHE does not cover the self-employed, nor employees who were not paid in the

[1] "How to make best use of the new matched ILR-HMRC/.DWP data for reporting on the employment and earnings outcomes of training", Frontier Economics and the Institute for Fiscal Studies, 2010.

[2] In 2007 and 2008, however, ASHE's sample size was reduced by 20% from its usual one percent of employers sampling frame.

reference period. The survey distinguishes between males and females, and between part-time and full-time workers, as well as disaggregating earnings by industry and occupation. However, the survey does not take account of learning or qualifications.

- *The Labour Force Survey (LFS).* This is a large household survey, conducted at quarterly intervals, designed to give information about the number of people with jobs, the characteristics of those jobs, the job-search activities of individuals in those households, and other characteristics of the UK's labour force. It contains detailed information regarding respondents' learning and qualifications, alongside a wealth of demographic data.

In **Section 2** we assess the feasibility of using each of these datasets for deriving comparator groups for learners in the ILR-HMRC/DWP. For the sake of brevity we have focussed on FE learners. However, we would expect our general findings on the validity of comparisons to be broadly similar for learners in the other funding streams. We find some scope for making earnings comparisons using the LFS, but little scope for such comparisons using ASHE.

In **Section 3** we present some earnings comparisons between the matched ILR-HMRC/DWP data and the LFS. For Level 2 and Below Level 2 learners our comparator groups are defined in terms of broad age bands. More detailed work could be done to weight these samples to make them more representative of FE learners, however. For Level 3 learners we have also sought to compare earnings by subject area. In some cases the sample sizes are sufficiently large to allow robust comparisons. Reassuringly, the earnings distributions look reasonably similar between datasets.

In **Section 4** we present the results of some preliminary earnings comparisons, using the comparator groups outlined in Section 3. We compare earnings disaggregating by gender, level of study and broadly defined age group.

In **Section 5** we present our conclusions.

Identifying appropriate comparator groups

In this section we explore ASHE and the LFS in order to identify the most appropriate groups to compare with the learners in the ILR-HMRC/DWP matched data

Annual Survey of Hours and Earnings

ASHE is a survey of employers, containing details on hours and earnings of individuals in employment. Summary data tables are published on the ONS website[3]. However, the microdata can only be accessed with a special user license from the Virtual Microdata Laboratory.

The ASHE data tables report earnings and hours by gender, whether an employee is working part-time or full-time, and the interaction of these two characteristics simultaneously. Various tables provide further levels of disaggregation, such as by occupation, region, industry and age. The most relevant tables for the purposes of comparison with the ILR-HMRC/DWP data would be:

- Table 20: Age by occupation (2-digit SOC)
- Table 21: Age by industry - SIC 2007

These disaggregate by age band (18-21, 22-29, 30-39, 40-49, 50-59, 60+) and by occupation or by industry. They do not, however, disaggregate by qualifications. One possibility, therefore, would be to match these industries/occupations to subject areas in the ILR and compare the earnings distributions. Of course, we would have to be confident that the qualifications and industry/occupation genuinely matched up (a far from trivial concern). Moreover, even if they did, it is not necessarily the case that an individual who has gained a qualification will then work in a corresponding industry/occupation. Hence, the link between qualifications in the ILR and occupations and industries in ASHE seems likely to be reasonably weak. Unless we were to have a very clear indication from other sources that the qualification mix was similar between groups, we would risk making false comparisons.

For each group, the ASHE tables present both mean earnings and the earnings at various percentiles in the distribution. It is worth noting that in the published tables, statistics are not provided if the coefficient of variation accompanying an observation is too high, or if the sample size is insufficient. Therefore, in some cases ASHE will not even provide data.

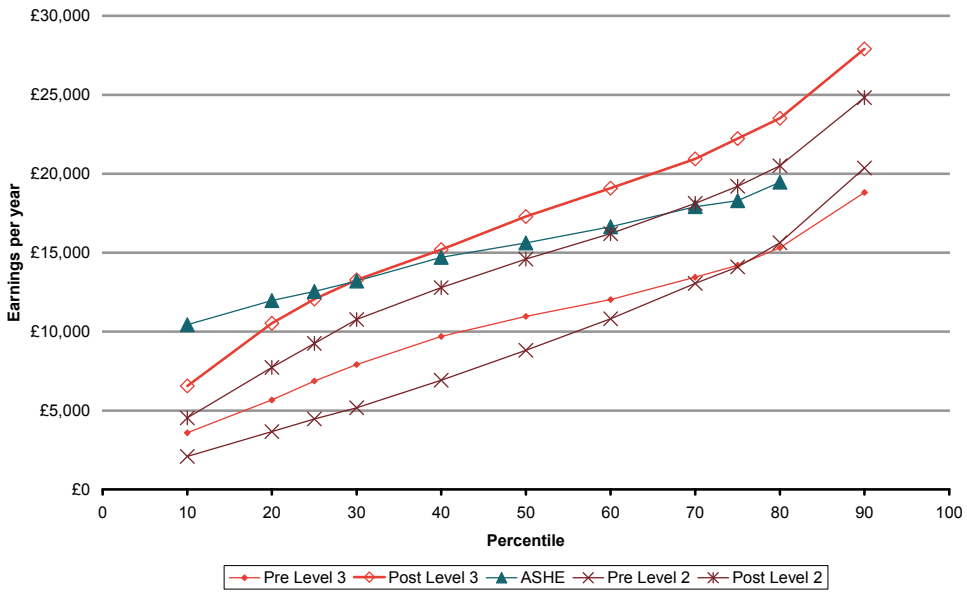
In the chart overleaf, for males aged 18-21 we plot the ASHE earnings distribution of the "skilled construction and building trades" occupation against earnings of those who are completing Level 2 and 3 "construction, planning and built environment" qualifications. In

[3] See e.g. www.statistics.gov.uk/statbase/product.asp?vlnk=13101

order to make the groups as comparable as possible, the individuals in FE we focus on are those who have been in employment for all of the relevant time period and earning positive amounts. However, this approach is made more complicated by the fact that employment spells recorded in the WPLS are far from precisely measured. In cases where the start or end date of a spell of employment is unknown, a default start date (April 6th) and/or end date (April 5th) is recorded for the corresponding tax year. Around 23% of employment spells in the P45 dataset have either uncertain start dates, uncertain end dates, or both. Even after substantial data cleaning (notably attempting to reconcile the HMRC data in the WPLS with the less noisy data in the National Benefits Database), our measure of individual employment in the WPLS remains noisy and imprecise.

Figure 1 below compares the earnings distributions from the different datasets. It is not clear which the relevant level for comparison with the ASHE group should be. In any case, the distributions are not very similar, with ASHE being rather flatter.

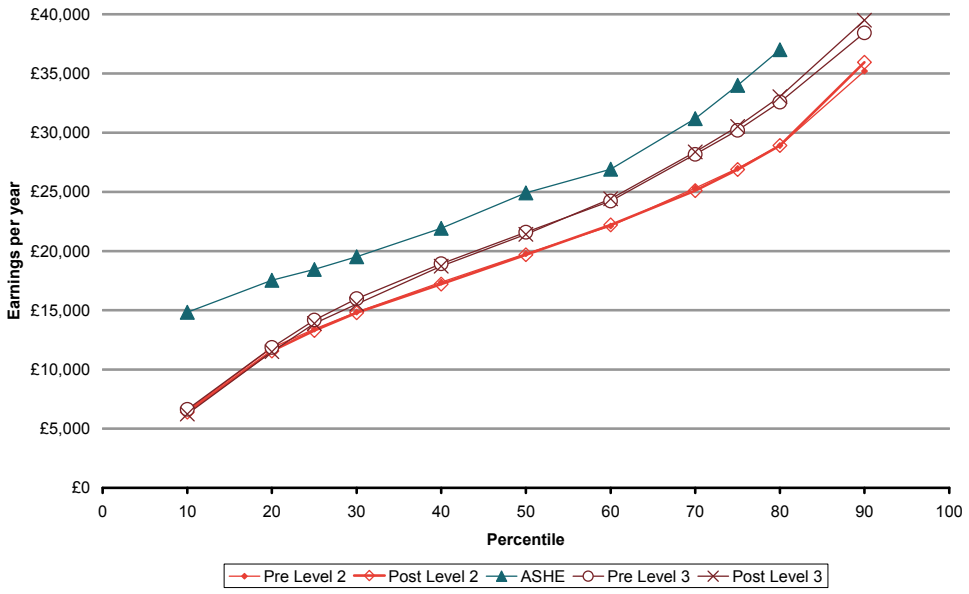
Figure 1: 18-21 males in construction and building, ASHE vs. ILR-HMRC/DWP



Source: Frontier analysis of ASHE 2010 table 20 and ILR-HMRC/DWP matched data. [ASHE data not reported for 90th percentile)

Another comparison is shown overleaf. This time the ASHE occupation is health and social welfare associate professionals (SOC 32) and we are comparing with those studying health, public services and social care. Again, the distributions do not look very similar.

Figure 2: Comparison of ASHE and ILR-HMRC/DWP - males aged 30-39 in health, public services and care



Source: Frontier analysis of ASHE 2010 table 20 and ILR-HMRC/DWP matched data. [ASHE data not reported for 90th percentile)

For other subject areas, it is very unclear what occupation to compare with in ASHE. For example, business, administration and law would be fairly difficult to pin down as a single occupation.

Overall, therefore, our view is that ASHE does not make an appropriate dataset for comparisons with the ILR-HMRC/DWP matched data. This is because it is unlikely that matching industry/occupation groups (in ASHE) with qualification types (in the ILR) will yield truly comparable groups. In addition ASHE only records individuals currently in employment, which is difficult to compare with learners in the ILR-HMRC/DWP whose annual earnings may include periods when they were not working.

The Labour Force Survey

The LFS is a quarterly survey of households in the UK, designed to provide information regarding the state of the labour market. Sampled household members are interviewed regardless of whether they are in employment, self-employment, unemployment, education, training, or economic inactivity. Earnings distributions derived from the LFS should therefore cover the same potential labour market outcomes as the ILR-HMRC/DWP matched data. One caveat, however, is that the P14 data does not necessarily include earnings below the lower earnings threshold for income tax. Neither dataset records the earnings of the self-employed, so in this respect they are consistent.

The LFS contains a wealth of demographic data, including qualifications, tenure in job, and the age at which respondents completed their full-time education. The questionnaire also records the industry and occupation in which respondents work. We can therefore

potentially compare individuals between datasets with the same type and level of qualification. The feasibility of such comparisons is discussed later in this section.

In our analysis below, we have focused on earnings comparisons for males. Our expectation was that the comparison would be less straightforward for females, due to their greater degree of part-time learning and heterogeneity in terms of economic activity. Unless the tendency of part-time working is the same in the LFS as in FE learners, this would make the groups less comparable.

The approach we take is to look at successively more finely disaggregated subgroups, to see how detailed we could make the earnings comparisons without running into sample sizes that are too small.

Wave structure

Each quarter's LFS sample is made up of five 'waves', each containing a roughly equal number of private households. Each wave is interviewed in five consecutive quarters, before dropping out of the sample. In recent years of the LFS there have been approximately 24,000 respondents in each wave, with a slight drop-off due to sample attrition. The number of respondents in each wave of the LFS is shown below for the April to July 2008 data.

Table 1: Wave structure of LFS

Wave	Respondents
1	25,581
2	26,050
3	24,362
4	23,472
5	22,584

Source: LFS - April to July 2008

Overleaf we examine the sample sizes which the LFS offers at various levels of disaggregation. The following analysis in this section uses all waves of the April to July 2008 LFS.

Level of qualification

The LFS variable we use to define comparator groups by qualification level is levqual – the individual's highest qualification. As the table below shows, in the April to July 2008 LFS data there were around 10,000 respondents at each of the main levels of interest (below level 2, level 2, level 3, no qualifications). Sample sizes for trade Apprenticeships are somewhat lower (around 4,000 individuals).

Table 2: Level of highest qualification (levqual)

Level	Male	Female	Total
Does Not Apply	20,729	21,273	42,002
No Answer	54	51	105
Nvq Level 4 And Above	10,213	11,432	21,645
Nvq Level 3	6,175	5,379	11,554
Trade Apprenticeships	3,262	637	3,899
Nvq Level 2	4,770	7,356	12,126
Below Nvq Level 2	4,502	6,083	10,585
Other Qualifications	3,866	3,254	7,120
No Qualifications	5,501	7,512	13,013

Source: LFS - April to July 2008

Age

Age is of course an important variable in defining the comparator groups. As the table below shows, for the qualification groups with larger sample sizes it may be possible to disaggregate the data by single year age groups (provided that age is not interacted with other characteristics). For trade Apprenticeships, however, the sample sizes are already too small to disaggregate the data by age in a single quarter of data.

Table 3: Sample sizes by age and level of qualification - males

Age	Level 3	Trade Apprenticeship	Level 2	Below Level 2	No Qualification
16	-	-	179	169	418
17	52	-	436	209	72
18	185	-	267	165	64
19	248	12	134	125	46
20	277	21	141	95	63
21	236	28	79	91	58
22	185	22	92	91	51
23	138	21	73	83	52
24	101	24	101	58	42

Age	Level 3	Trade Apprenticeship	Level 2	Below Level 2	No Qualification
25	111	28	92	65	42
26	97	30	72	74	40
27	102	25	76	71	49
28	129	26	73	68	35
29	96	23	89	69	46
30	113	19	78	74	39
Total	-	282	1982	1507	1117

Source: LFS - April to July 2008. Cells which are potentially disclosive have been marked with "-".

Vocational learning

Because the FE funding stream is largely vocational in nature, we should therefore seek to compare individuals in the ILR data with individuals in the LFS whose highest qualification is vocational in nature. The **subcod1** variable in the LFS records the subject type for vocational qualifications. It is only recorded for level 2 and above. For Level 3 learners, therefore, we could compare pre-training earnings with Level 2 qualified individuals in the LFS and post-training earnings with Level 3 qualified individuals in the LFS. There is less value in trying to do this comparison for Level 2 learners, because we would have no pre-learning comparison (i.e. Below Level 2 individuals in the LFS).

The subcod1 variable identifies 81 different subjects. We have attempted to map these subjects to study areas in the ILR. The full mapping is provided in the annex. Clearly this is a somewhat subjective exercise, and were the earnings comparison work to be taken further, we would recommend a detailed (albeit time-consuming) refinement and robustness testing of this mapping.

Table 4, below, shows the sample sizes for males aged 16 to 30, by vocational subject area in a single quarter of the LFS data. For some subject areas (e.g. Engineering and Manufacturing) sample sizes are encouragingly large, suggesting that a subject-level comparison may be possible.

Table 4: Sample sizes by subject area - males 16-30 with vocational qualification

	Level 3	Level 2
Health, Public Services and Care	48	42
Science and Mathematics	11	5
Agriculture, Horticulture and Animal Care	13	20
Engineering and Manufacturing Technologies	278	138
Construction, Planning and Built Environment	123	70
Information and Communication Technology	84	57
Retail and Commercial Enterprise	25	38
Leisure, Travel and Tourism	60	67
Arts, Media and Publishing	109	47
History, Philosophy and Theology	-	-
Social Sciences	-	-
Languages, Literature and Culture	-	7
Education and Training	5	-
Preparation for Life and Work	-	7
Business, Administration and Law	115	56

Source: LFS - April to July 2008. Cells which are potentially disclosive have been marked with "-".

Where the sample size is large enough, it may be feasible to disaggregate the subject area groups still further (e.g. into age bands). In the analysis in the next section we use two age bands for larger subject areas – 16 to 25 and 26-59.

Job tenure and time since completed learning

Ideally, in creating comparison groups for individuals who have recently finished their training (the individuals we observe in the ILR-HMRC/DWP data), we would like to focus on individuals in the LFS who have recently finished their learning. We are less interested in individuals who obtained their qualification many years ago and have built up tenure in their current job. It would therefore be desirable to find a filter to focus on recently completed learners. Three variables can potentially address this:

Length of time in current job (Emplen). We could remove individuals with tenure greater than 1 or 2 years. This would remove employed learners who completed their learning several years before. However, it would not remove learners who suffered a long spell of unemployment following their training, before ultimately finding a job. Nor would it remove individuals who completed their training many years ago, but have recently changed jobs.

This approach would therefore have an unpredictable biasing effect on the comparison groups.

Age completed continuous full time education (Edage). We could remove individuals who completed their continuous full time education more than a specified number of years ago. However, many learners taking vocational qualifications will not have been in education continuously, but will have had gaps between their spells in education/training. Continuously educated learners may be different from other vocational learners.

Completed qualification in last 12 months (qlsty601). We can use using the qlsty601 variable to remove individuals who completed their qualification 'too recently' to make a fair comparison group. However, as noted above, we do not have a straightforward means of removing those who completed a long time ago. This means we cannot define a comparable window of when learners completed in our two data sources. However, a different analysis facilitated by this variable would be to look at the earnings of recently completed learners. Although this will only give a very short term effect, the advantage is that we directly pin down the time of learning. We could then compare the earnings distribution of these individuals with those who recently completed in the ILR-HMRC/DWP dataset.

In summary, while each of these variables offers some scope for further filtering of comparison groups, they do not allow us to create perfectly comparable groups in the LFS and ILR-HMRC/DWP datasets. In the analysis which follows we have not filtered the data using these variables, for the reasons detailed above, but we suggest that they may warrant further investigation should earnings comparisons with the LFS be carried forward in future.

Suggested comparator groups

Based on the above analysis, Table 5 lists our suggested comparator groups for different qualification levels in the ILR. Comparison groups are most difficult to define for the lowest qualification levels, with below level 2 qualifications posing a particular problem. When comparing ¬pre-training earnings, the only possible comparison group for below level 2 groups in the ILR is individuals recorded in the LFS as having 'no learning.' This seems likely to be a poorly-matched comparison group, however, since it will contain many individuals who have no intention of taking training in the future. For the post-training earnings of individuals in this group, however, the comparison group is better defined (individuals in the LFS with below level 2 qualifications).

For level 2 and level 3 qualifications in the ILR data, comparison groups in the LFS are easier to define. However, due to restricted sample sizes, it is only in the case of some of the more 'popular' level 3 subject areas that we suggest disaggregating the comparison groups by subject

Table 5: Suggested degrees of disaggregation of comparator groups for learners at different levels

Learning in ILR	Before training (LFS)	After training (LFS)
Below level 2	No learning (levqual=7) No disaggregation by subject	Below level 2 (levqual=5) No disaggregation by subject

Learning in ILR	Before training (LFS)	After training (LFS)
Level 2	Below level 2 (levqual=5) No disaggregation by subject	Level 2 (levqual=3) Vocational only. No disaggregation by subject
Level 3	Level 2 (levqual=3) Vocational only. Disaggregated by subject (at least for larger study areas)	Level 3 (levqual=2) Vocational only. Disaggregated by subject (at least for larger study areas)

In any case, we consider that the sample sizes we would obtain using a single quarter of LFS would be not sufficiently large to allow for robust analysis. We therefore recommend pooling across more years to increase the sample size. The methodology for deriving datasets for comparison is described in the following chapter.

Constructing datasets for comparison

The Labour Force Survey

The sample sizes in the LFS considered in the previous section are not sufficient to generate robust comparison groups at any level of disaggregation. It is necessary to increase the sample size somehow. There are various ways to do this. The simplest methodology would be to use all waves appearing in each quarter of LFS and append several of quarters of LFS together. However, as we would be repeatedly sampling the same individuals at successive waves, the effective sample size would be overstated. We may also find idiosyncrasies in the data due to "lumps" from the same individuals being sampled repeatedly.

To get around this we have used the first wave of each quarter and pooled together across LFS datasets. In doing this it is not obvious precisely what timescale to use. A timescale that is too long would risk including information that is out of date. On the other hand, it is imperative that the sample sizes be large enough. There are no hard and fast rules in striking this balance. We opted to use 4 years of LFS, running from October 2004 to September 2008. This would provide data that is contemporaneous with the ILR-HMRC/DWP matched data, which looks at learning from 2005 to 2007 and earnings a year either side of this.

This gives rise to the following sample sizes by level and age group:

Table 6: Sample sizes, males by age group and level (vocational for 2 and 3)

	16 to 25	26 to 60	Total
No qualifications	3,400	13,570	16,970
Below Level 2	4,086	10,174	14,260
Level 2	1,223	3,174	4,397
Level 3	1,494	8,596	10,090

Source: Labour Force Survey, 16 quarters from October-December 2004 to July-September 2008

These groups are reasonably large. But when we disaggregate by study areas (for Level 2 and Level 3), the sample sizes become too small for many groups. This is especially true for men in the 16 to 25 age group. In these cases there is no value in trying to do a comparison, since the data will not be robust. As a rule of thumb, we consider that if the group has a sample size less than 200 it is not worthwhile analysing it further. Therefore we only do earnings comparisons at the level of the subject area for a handful of study areas.

Table 7: Sample sizes by subject area, level and age group

Study area	16 to 25 males		26 to 60 males	
	Level 2	Level 3	Level 2	Level 3
Health, Public Services and Care	59	53	251	355
Science and Mathematics	6	14	17	37
Agriculture, Horticulture and Animal Care	28	22	137	179
Engineering and Manufact uring Technologies	258	381	788	3432
Construction, Plan ning and B uilt Environment	138	220	444	1753
Information and Commun ication Technology	150	141	157	224
Retail and Commercial Enterprise	65	32	202	122
Leisure, Travel and Tourism	126	102	384	446
Arts, Media and Publishing	112	193	126	378
Social Sciences	-	-	-	-
Languages, Literature and Culture	12	6	7	11
Education and Training	4	6	14	37
Preparation for Life and Work	12	15	42	50
Business, Administration and Law	74	79	216	467

Source: Labour Force Survey, 16 quarters from October-December 2004 to July-September 2008.
Cells which are potentially disclosive have been marked with "-".

The earnings measure we use is gross weekly pay ("grsswk"). We multiply these earnings by 52 in order to derive a measure of annual pay. Gross weekly pay only includes income from employment, and does not take into account the earnings of the self-employed. This matches up with the way in which the earnings measure is calculated for the ILR-HMRC/DWP data, in that the P14 only includes earnings in employment, not earnings of the self-employed.

A minor source of inconsistency between the earnings measures is that the P14 is not required for individuals whose earnings are below the lower earnings threshold. Some of these low-earning individuals might have P14s anyway, but not all of them. In any case, there will be some individuals earning low but positive amounts who show up in the ILR-HMRC/DWP matched data as having zero earnings.

We measure earnings at successive percentiles in the distribution. In doing this we use population weights in the LFS to weight respondents so that the groups are representative of the UK as a whole.[4] Of course it is unlikely to be the case that learners in FE would be exactly representative of the wider UK population (even if, for level 2 and 3, we focused only on those with vocational qualifications). The groups are therefore likely to be dissimilar to some extent.

[4] The weight we use is **pwt07**

Ideally one would construct an alternative set of weights so that the individuals in the LFS were representative of the population of FE learners. However, such reweighting is well beyond the scope of this exploratory work. The alternative to this is to manually disaggregate the groups further along demographic lines, but this would again make the sample sizes too small and also make the analysis less tractable.

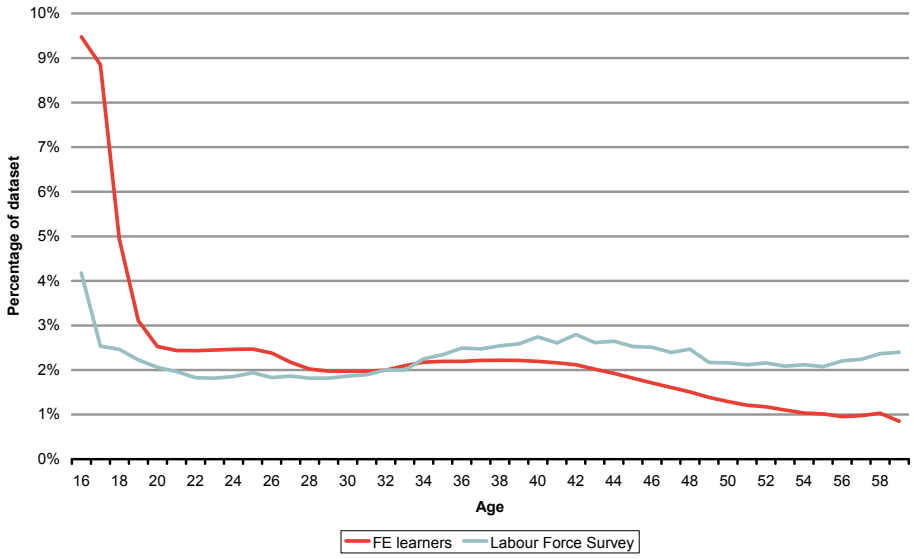
ILR-HMRC/DWP

The earnings measure used for FE learners is described in detail in our 2010 report. We used the data from the P45 and National Benefits Database to identify months when individuals were in work and combined this with annual earnings data from the P14. We then allocated the annual pay to the months that the individual was working, thus giving a measure of average monthly pay for the months in which they worked. We then summed these earnings for the 12 months before and the 12 months after learning.

In this work we have focused entirely on FE learners whose study was in 2005/06 and 2006/07. We could in principle do a similar analysis using learners from the other funding streams.

It can be seen that the datasets we are comparing are rather different in terms of their age compositions. These are compared in the chart below. There are many more 16 to 20 year olds in the ILR-HMRC/DWP data than in the LFS. There are less people aged over 40. We have not looked at other demographic characteristics that could potentially affect earnings. In a more detailed analysis, a significant task would be to ensure that the groups being compared had matching characteristics.

Figure 3: Composition of datasets by age



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Earnings comparison results

In this section we present the results of some preliminary earnings comparisons, using the comparator groups outlined in Table 5. For ease of exposition, we present results in a graphical format, but full tables of the results underlying these charts are provided in the Annex.

We present the following comparisons:

- Earnings distribution by level of achievement (men);
- Earnings distribution by level of achievement (women);
- Earnings distribution by subject area for Level 3 learners (men).

As discussed in the previous section, we focus on comparisons between male sample groups (since these are the most straightforward comparisons), and where the learning is vocational (for level 2 and level 3 qualifications).

Earnings by level of achievement and wide age band (men)

At the highest level we can compare the pre-learning and post-learning earnings of individuals in the ILR with the earnings of LFS respondents with similar qualification levels. Due to the importance of age as a determinant of earnings, we have disaggregated the groups into two age bands – the 16 to 25 age group and the over 25 age group.

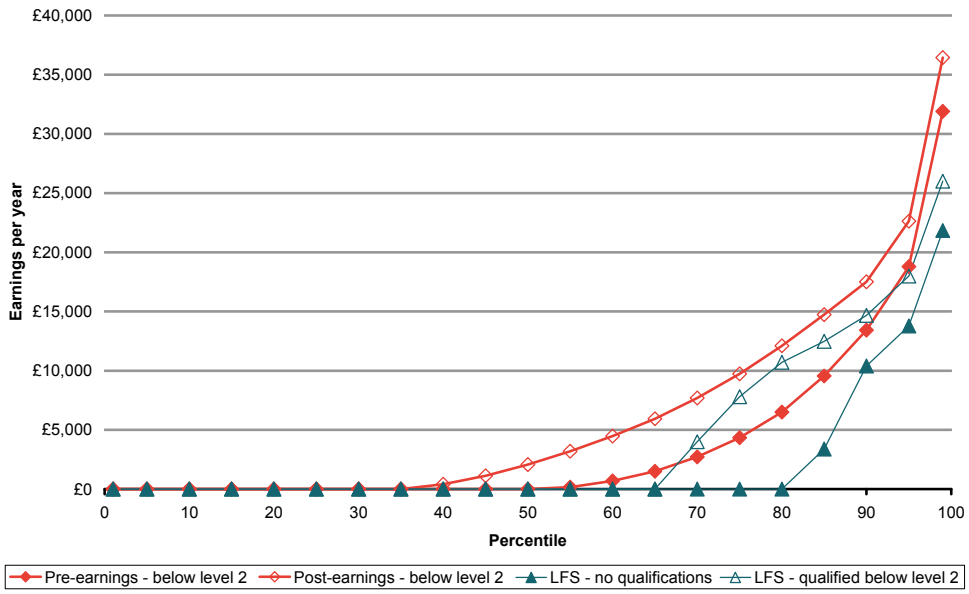
16 to 25 age group

We focus first on the 16 to 25 age group in Figure 2, overleaf. The red lines show the pre-learning and post learning earnings distributions of these learners. The blue lines show the LFS comparator groups. In both data sources we see clear evidence of the earnings distribution shifting upwards post-learning (though of course this should not be interpreted as a pure causal effect). Moreover, the earnings differentials[5] are of a similar magnitude in both data sources. In these respects, therefore, the data sources are similar.

However, the datasets differ in terms of the overall shape of the earnings distributions. The ILR-HMRC/DWP earnings rise very gradually from zero, possibly reflecting the fact that our earnings measure is based on monthly averages, and individuals may spend periods of the year out of work. By contrast, in the LFS earnings are reported for a given week and we then multiply these up to derive an annual total. An individual will either be working in that period or not, so there is much less of a gradual increase from zero earnings to the earnings of an individual working full time throughout the year.

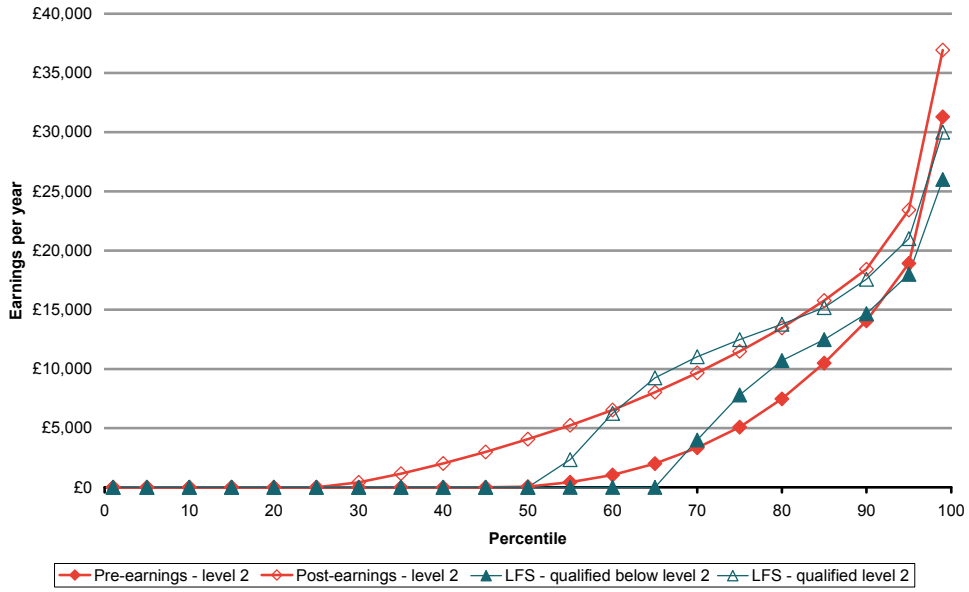
[5] For the ILR-HMRC/DWP data these are changes over time. However, for the LFS these represent differences between groups of individuals with different qualifications, so are not changes in the literal sense of the word. Nevertheless, the interpretation is that other things being equal we would see these differences in earnings arise in the event that an individual gained a higher qualification.

Figure 4: Comparison for 16-25 age males - below level 2



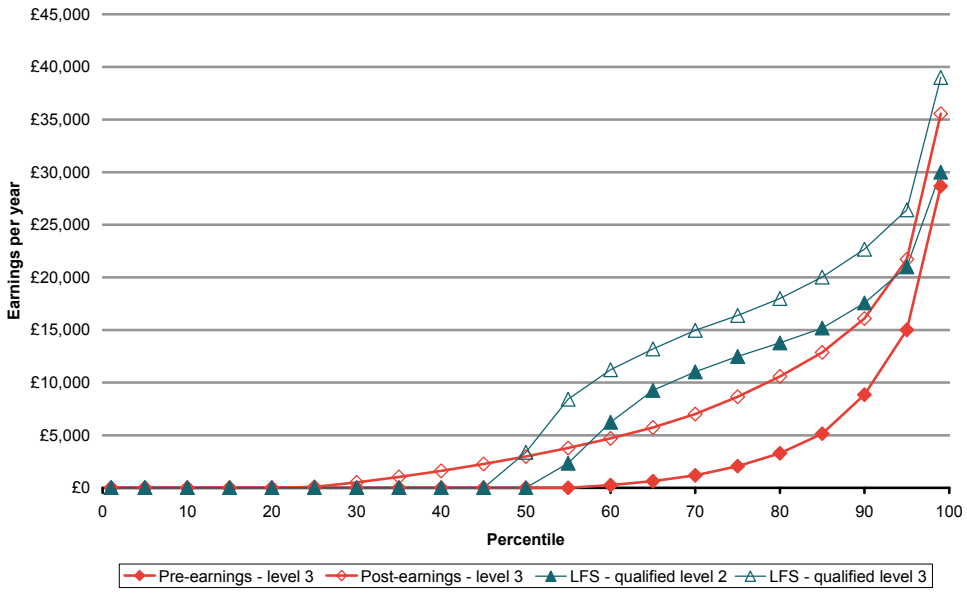
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 5: Comparison for 16-25 age males - level 2



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 6: Comparison for 16-25 age males - level 3



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

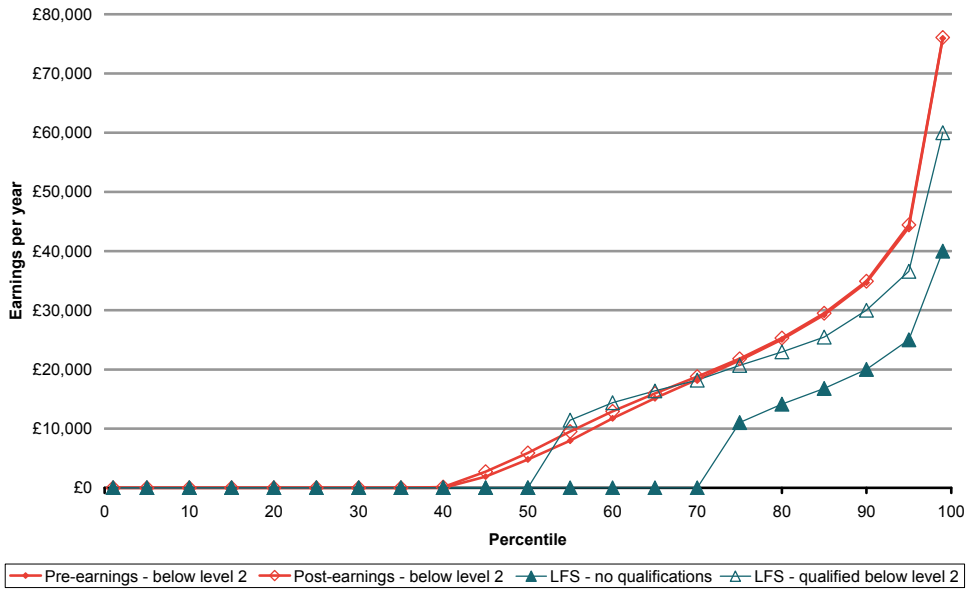
Over 25 age group

We now turn to the earnings comparison for males aged over 25.

Here we see very little change in earnings after learning for individuals in the ILR-HMRC/DWP. By contrast, in the LFS data we see some differences between individuals in terms of their level of qualification. Between no qualification and below level 2 qualification we see a large uplift in earnings, presumably driven by higher employment rates. Comparing Level 2 and Level 3, the LFS shows a positive differential at the upper part of the distribution, but would seem to indicate lower employment rate of Level 3 workers.

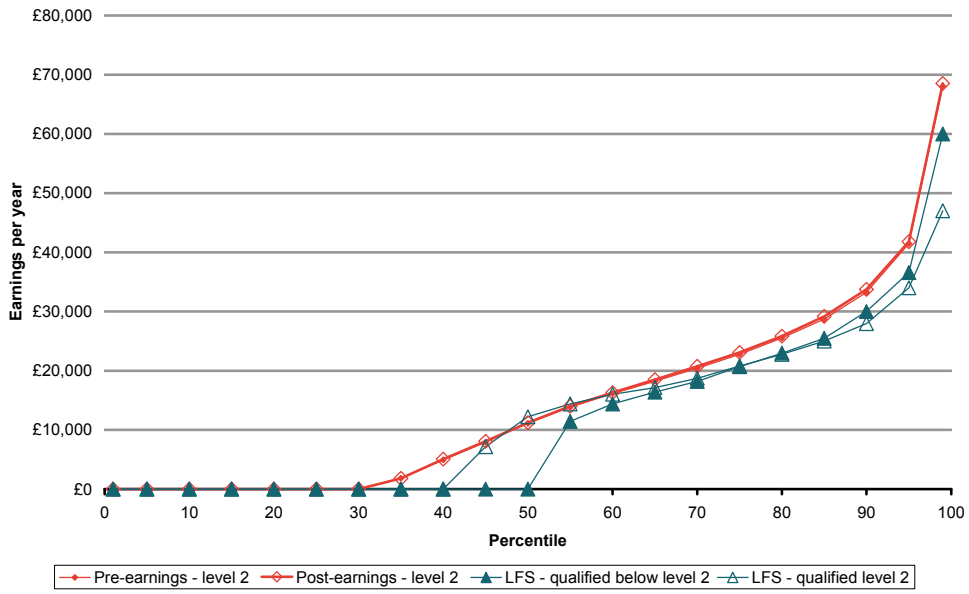
The ILR-HMRC/DWP earnings distributions are very close to the relevant post-earnings comparator groups in the LFS. At any percentile the difference in earnings between datasets is rarely ever more than several thousand pounds per annum.

Figure 7: Comparison for >25 males - below level 2



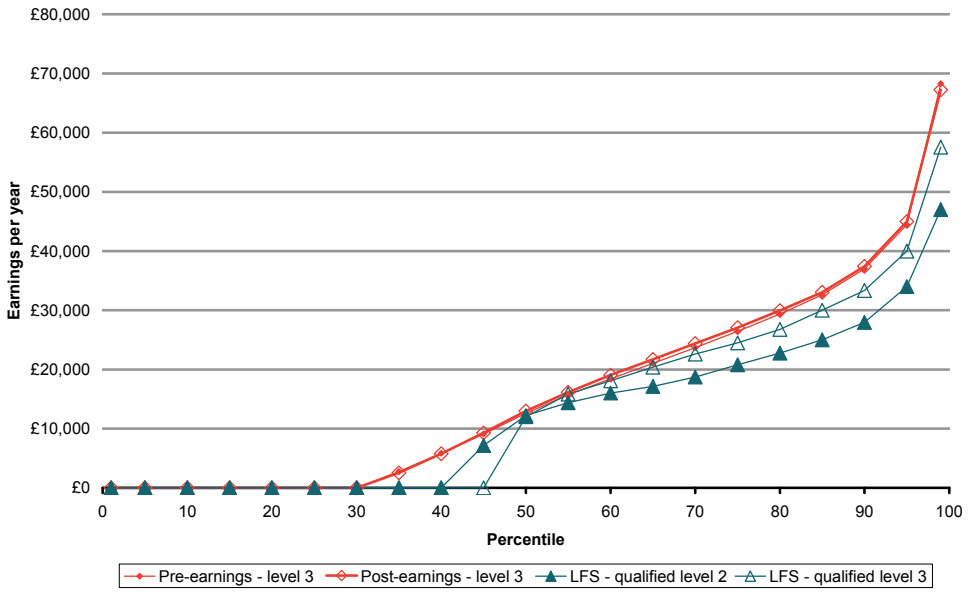
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 8: Comparison for >25 age males - Level 2



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 9: Comparison for >25 age males - level 3



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

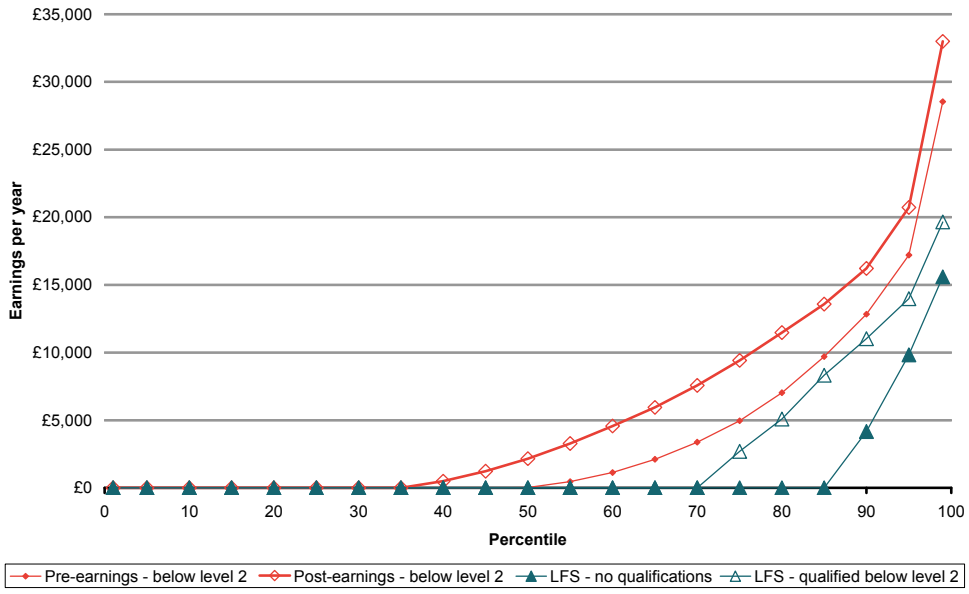
Earnings distribution by level of achievement (women)

In this section we analyse the earnings distribution by level of achievement for women. We had initially been sceptical as to the value of this analysis, as part-time working is more of a complicating factor in relation to women than it is in relation to men. As a result we had expected the comparisons between datasets to be less robust, as the extent of part-time working could give greater scope for FE learners not to be representative of the wider UK population. However, the distributions presented in this section suggest such comparisons are less problematic than we had feared.

16 to 25 age group

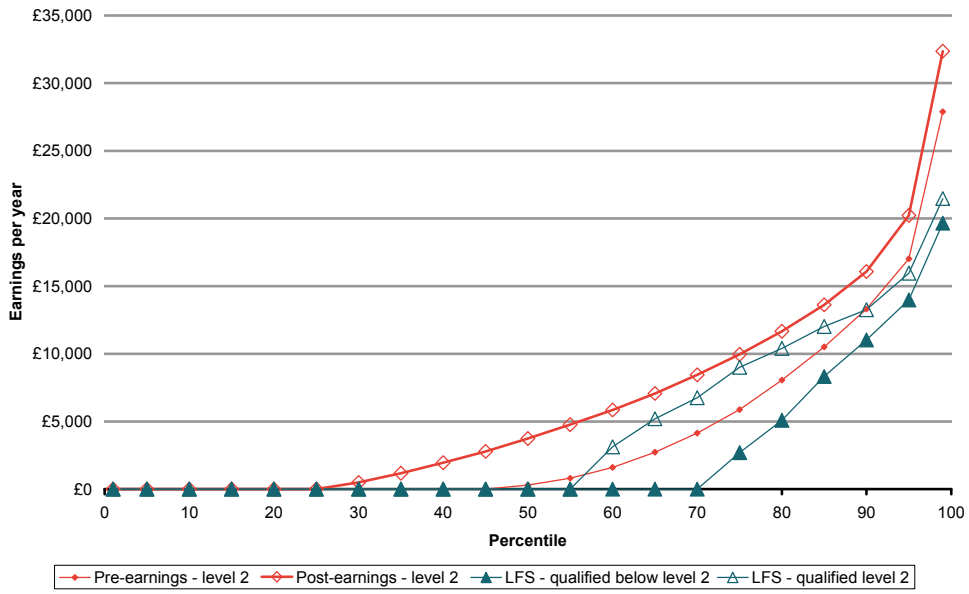
In both data sources we see clear evidence of the earnings distribution shifting upwards post-learning (though of course this should not be interpreted as a pure causal effect). There is also higher employment. The earnings differentials are larger in the LFS. As with men in this age group, for below level 2 learners, earnings are higher for the ILR-HMRC/DWP group than in the LFS. At level 3 this has reversed and the learners are earning less than the LFS comparator group. This effect is stronger in women than in men. It is not obvious exactly what is causing it, but presumably it hinges on differences between FE learners and the wider UK population.

Figure 10: Comparison for 16-25 age females – below level 2



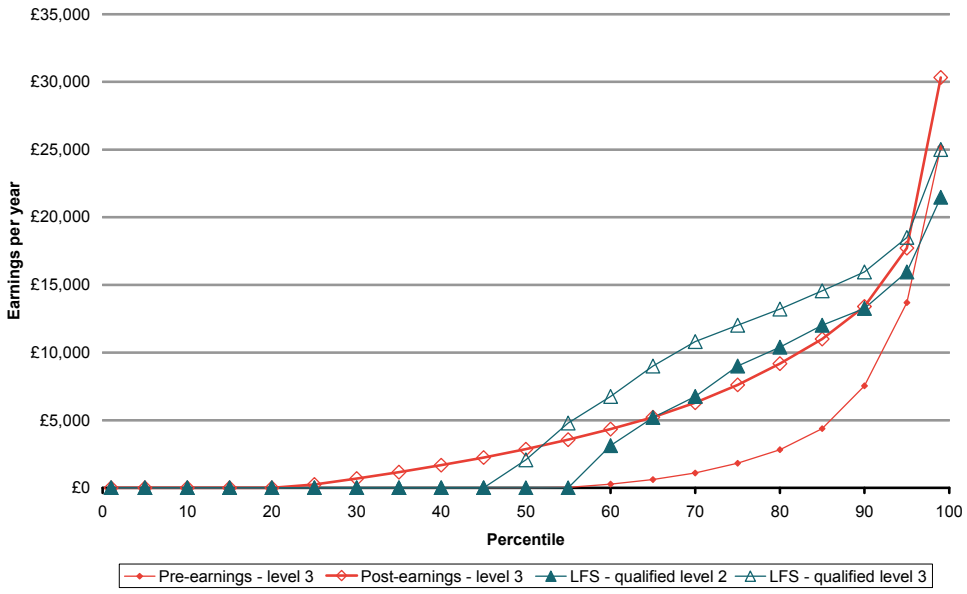
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 11: Comparison for 16-25 age females – level 2



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 12: Comparison for 16-25 age females – level 3

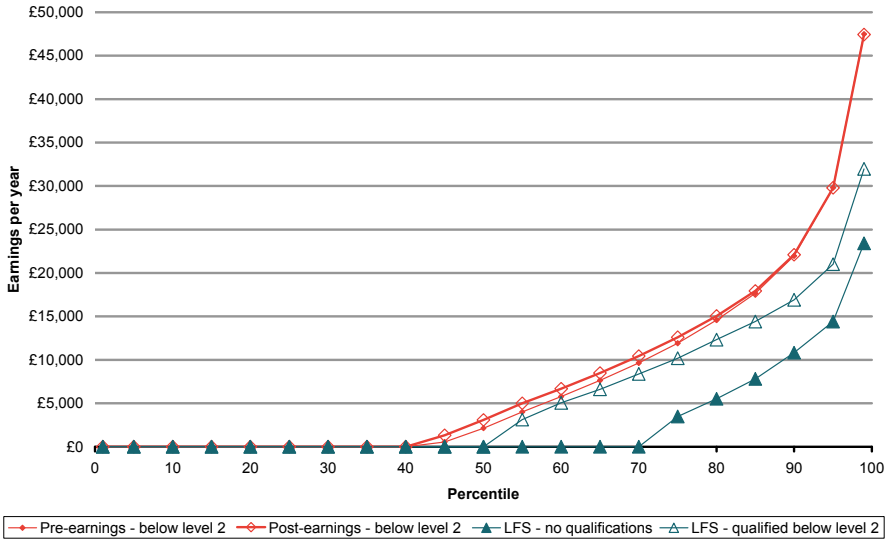


Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Over 25 age group

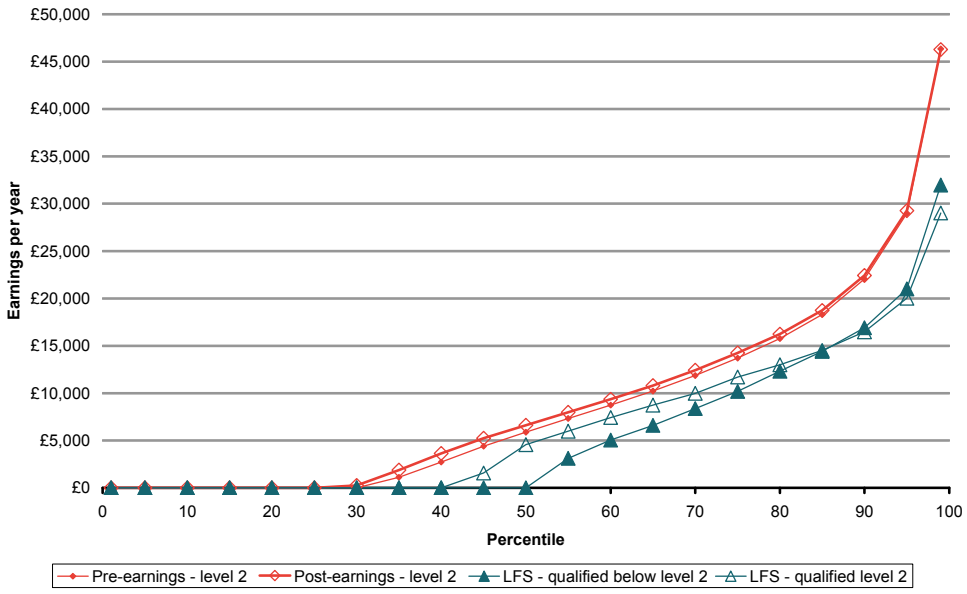
Amongst this age group earnings are uniformly higher in the ILR-HMRC/DWP than in the LFS. The gap is larger for women than it is for men. This could suggest that among women FE learning is a stronger signal of economic activity than it is with men. That is, the very fact these women are learning may imply their greater extent of participation in the labour market.

Figure 13: Comparison for >25 females – below level 2



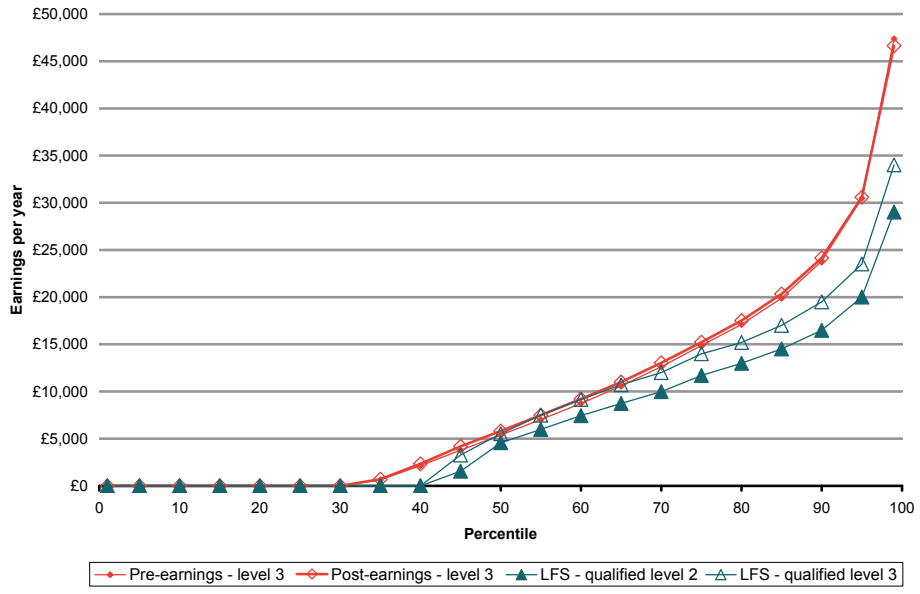
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 14: Comparison for >25 females – level 2



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 15: Comparison for 16-25 age females – level 3



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Study area comparisons

Using our mapping of the subno1 variable in the LFS to study areas in the ILR, in some cases there appears to be a sufficiently large sample size within the LFS to make comparisons at the level of the study area. The subno1 variable in the LFS is only

provided for respondents qualified at level 2 (i.e. the comparator group for level 3 pre-earnings) or higher. Therefore the only comparison we can make in terms of study area is for level 3.

The quality of this comparison does, of course, depend on the validity of the mapping we have used. This mapping is provided in the **Annex**. As before, we use two separate age groups (16-25 and Over 25), so that results do not suffer materially from comparator groups having substantially different age compositions (naturally, this reduces the available sample sizes). We only provide a comparison if the LFS sample size is in excess of 150 individuals.

Comparisons for the 16-25 age group

For this age group, sample sizes mean that comparison between ILR-HMRC/DWP and LFS is only appropriate for three study areas. These are:

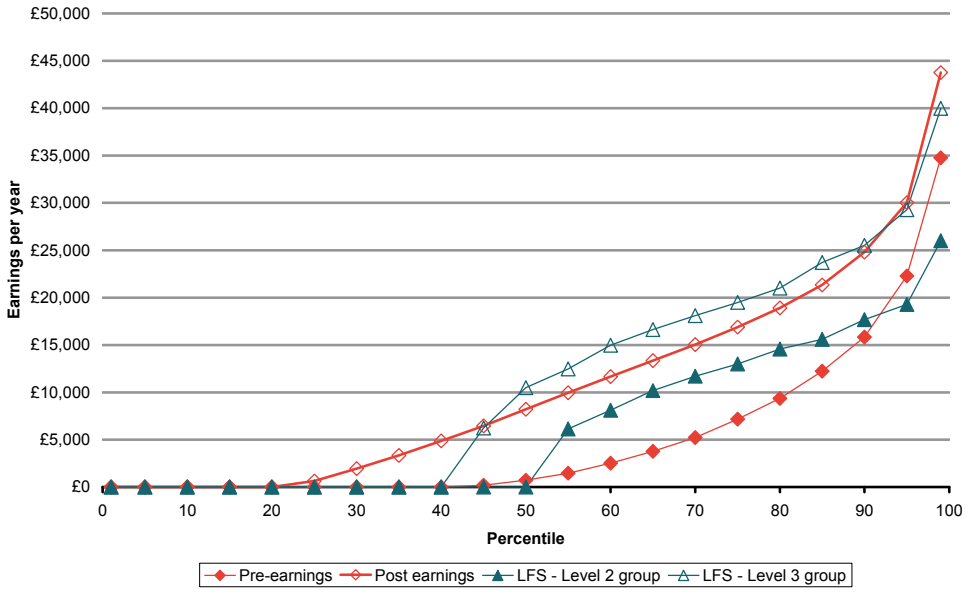
- Engineering and Manufacturing Technologies;
- Construction, Planning and Built Environment;
- Information and Communication Technology.

We examine these in turn.

In both data sources and all study areas we see clear evidence of the earnings distribution shifting upwards post-learning. The earnings differentials (pre and post-learning) are always larger in the ILR-HMRC/DWP group than in the LFS.

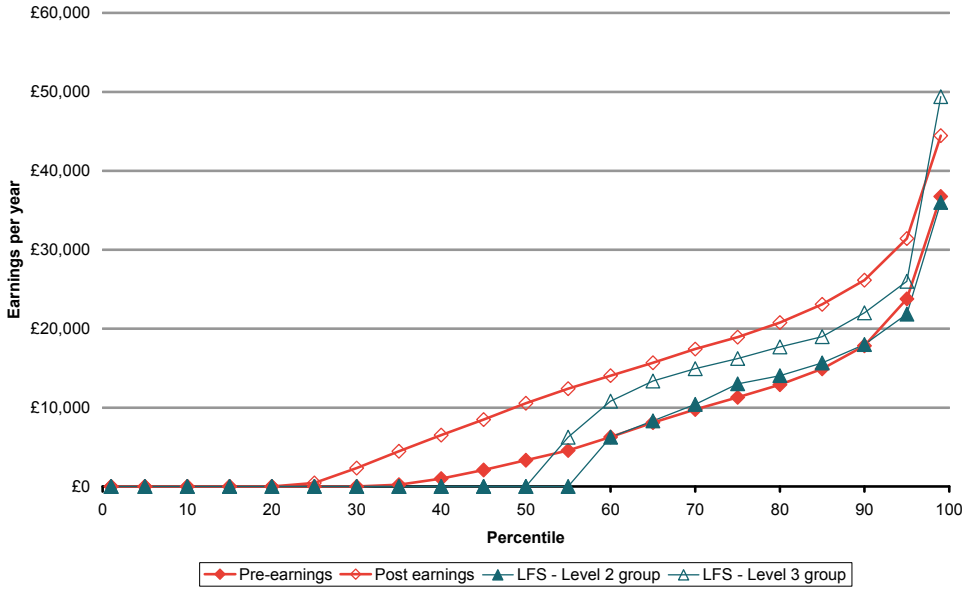
In Engineering and Manufacturing Technologies and Information and Communication Technology earnings are higher for the LFS group both pre and post-learning. However, in Construction, Planning and Built Environment earnings are very similar pre-learning but are greater for the ILR-HMRC/DWP group than in the LFS post-learning.

Figure 16: Engineering and Manufacturing Technologies - males16 to 25



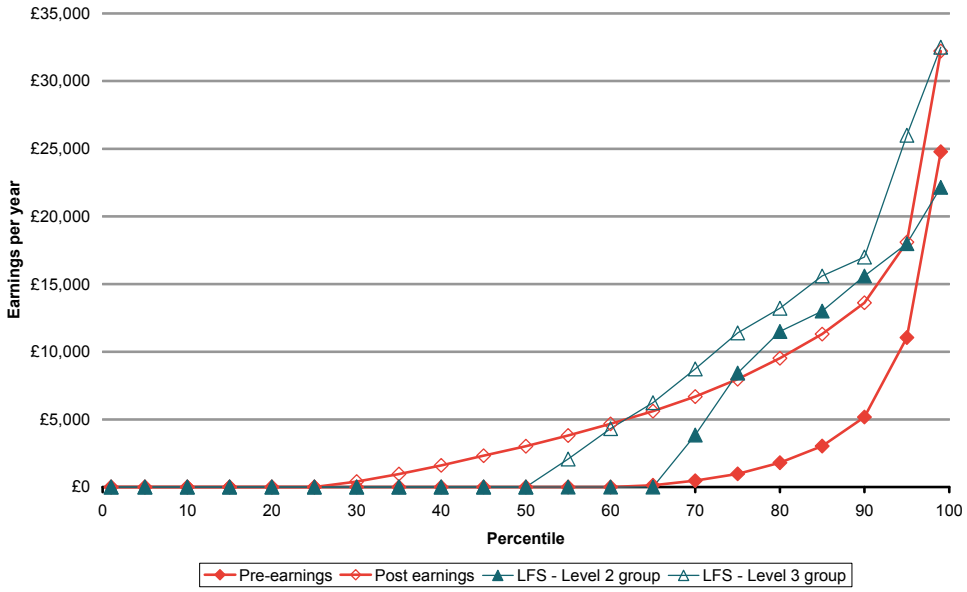
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 17: Construction, Planning and Built Environment - males16 to 25



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 18: Information and Communication Technology - males16 to 25



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Comparisons for the over 25 age group

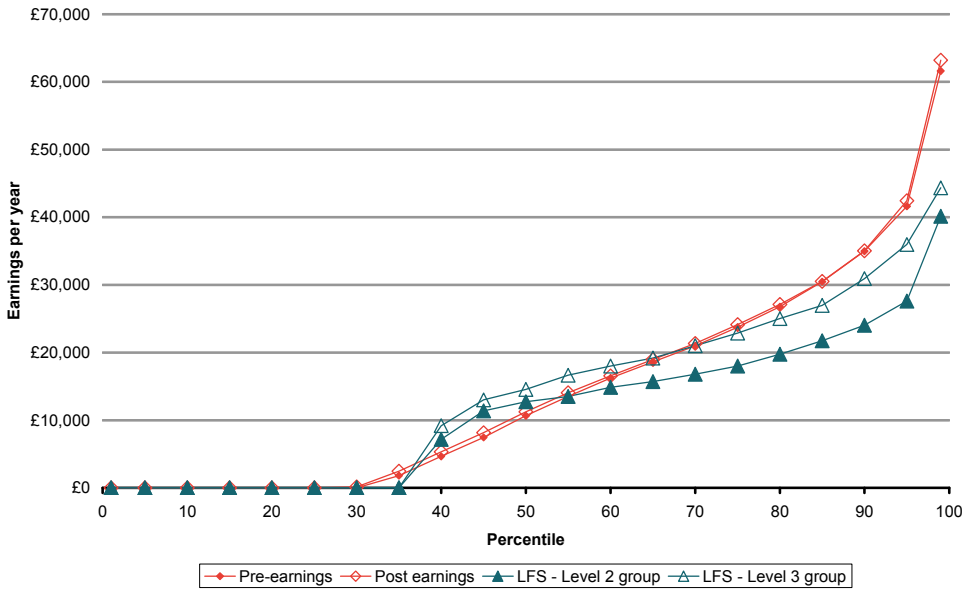
By pooling all ages together, we naturally increase sample sizes – though this comes at the cost of comparison groups which are likely to be far less similar to those found in the ILR data. For this age group sample sizes are larger, so it is feasible to compare earnings for more study areas.

- Health, Public Services and Care;
- Engineering and Manufacturing Technologies;
- Construction, Planning and Built Environment;
- Information and Communication Technology;
- Leisure, Travel and Tourism;
- Business, Administration and Law.

In both data sources and most study areas we see evidence of the earnings distribution shifting upwards post-learning. However, relative to the 16-25 age group, the shifts in the earnings distributions are very modest. Indeed in Leisure, Travel and Tourism, the earnings distribution of the ILR-HMRC/DWP group appears to be shifting downwards.

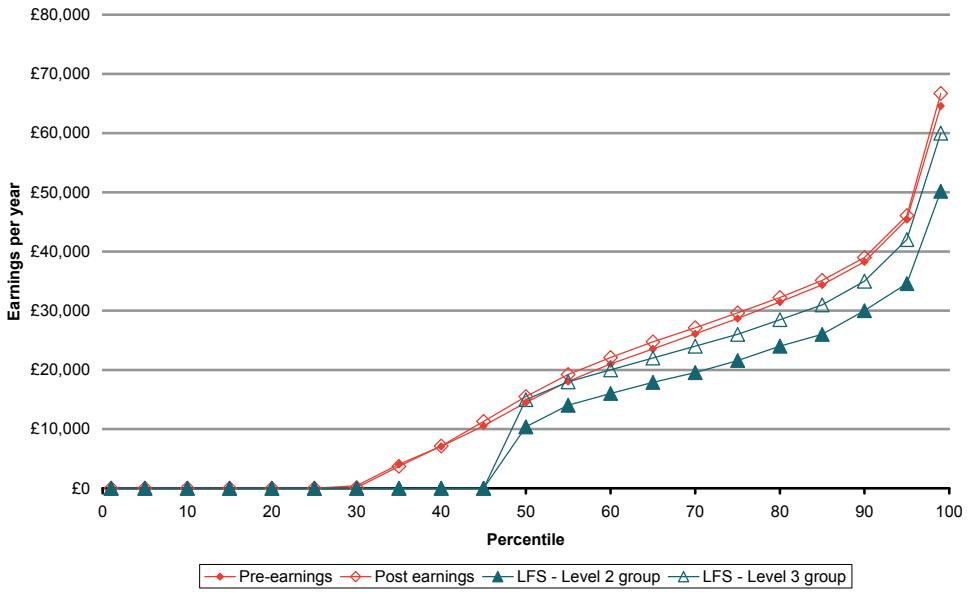
Interestingly, the earnings differentials (pre and post-learning) are always larger in the LFS than the ILR-HMRC/DWP group and earnings are higher in the ILR-HMRC/DWP group than the LFS group. This pattern is most evident in Construction, Planning and Built Environment.

Figure 19: Health, Public Services and Care – males aged over 25



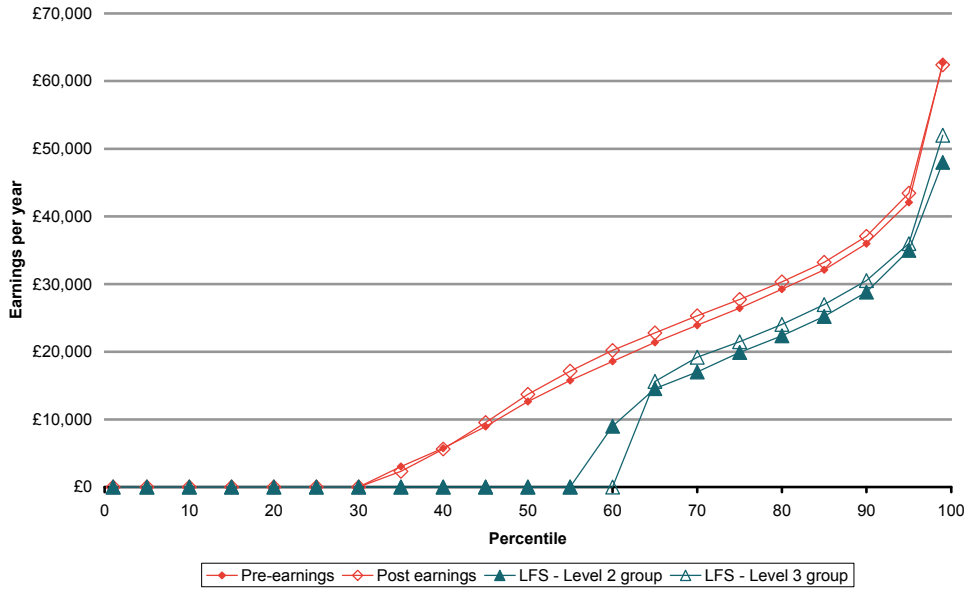
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 20: Engineering and Manufacturing Technologies – males aged over 25



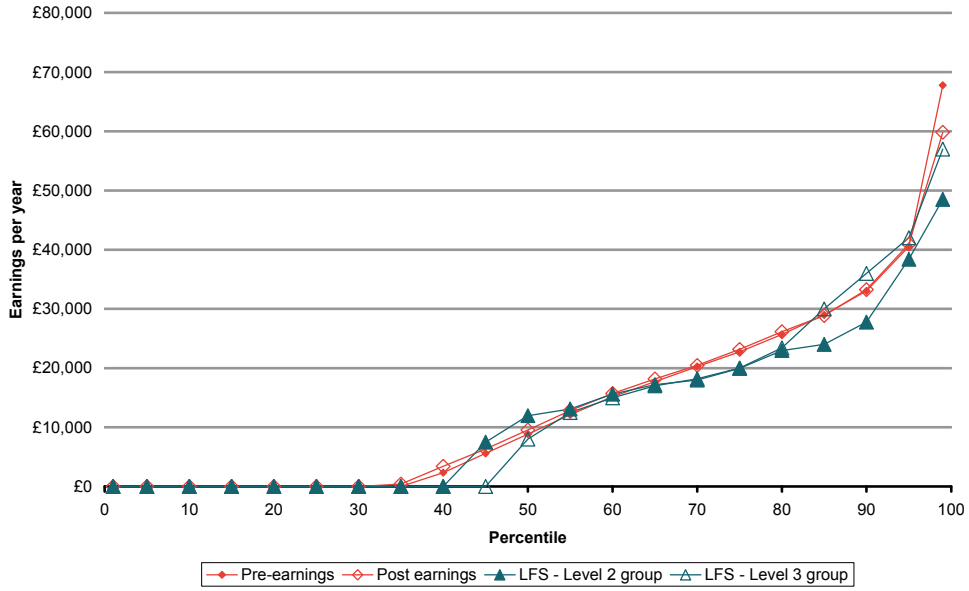
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 21: Construction, Planning and Built Environment – males aged over 25



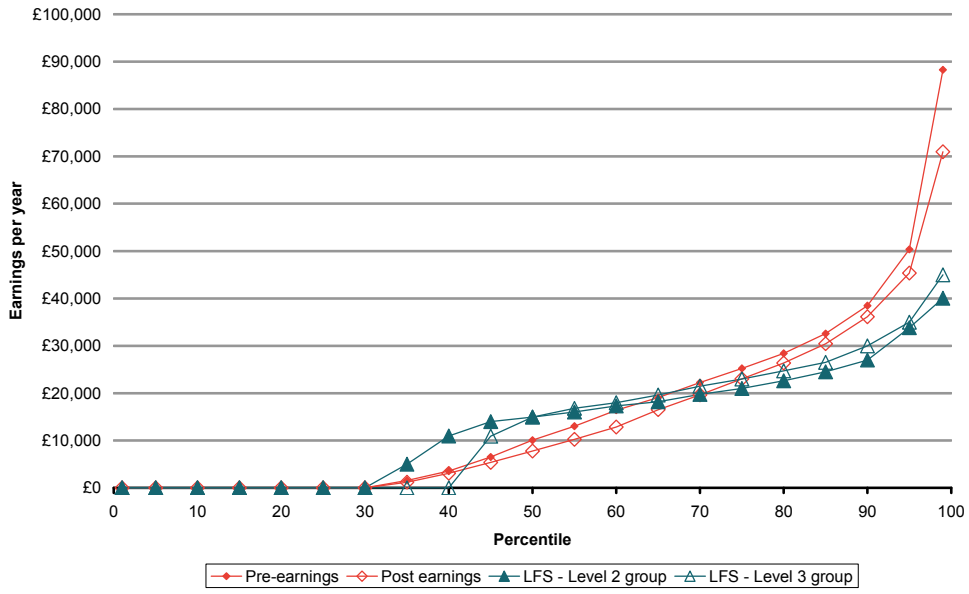
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 22: Information and Communication Technology – males aged over 25



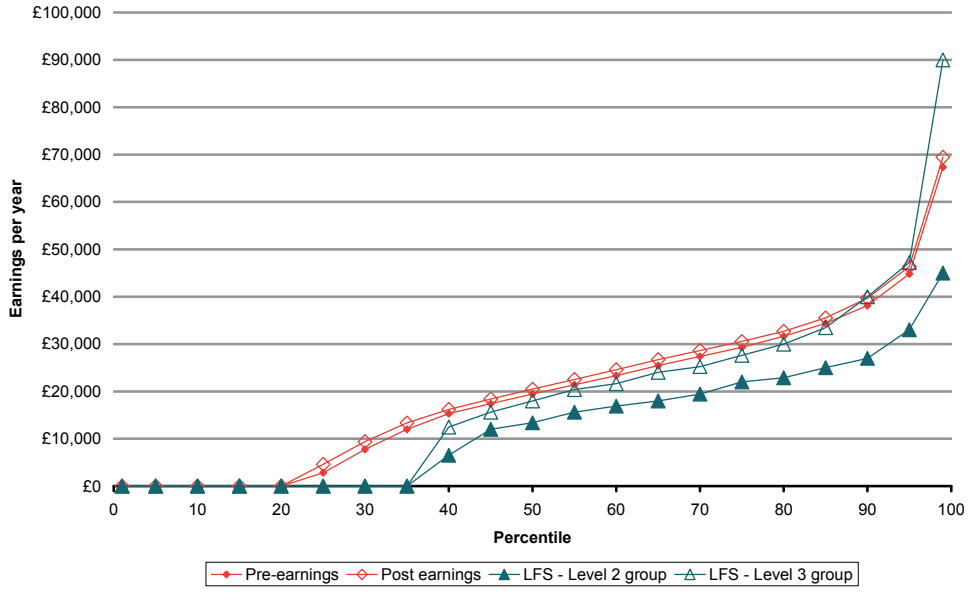
Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 23: Leisure, Travel and Tourism – males aged over 25



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Figure 24: Business, Administration and Law – males aged over 25



Source: Frontier analysis of ILR-HMRC/DWP matched data and LFS

Conclusions

This is an exploratory piece of work. Therefore, it would not be sensible to derive from it conclusions that are overly substantive. Nonetheless, this work has shown that the following considerations will be relevant in finding comparator groups from external data sources.

- At a conceptual level, the LFS would seem more fruitful than ASHE as a source of comparator groups. Most importantly, subjects and levels of qualification can be compared fairly easily, whereas in ASHE this information is not accessible. The other advantage of the LFS is that it covers the spectrum of economic activities. Moreover, at an empirical level, the earnings comparisons look more similar with the LFS than they do for ASHE.

- By pooling data from several years of LFS it is possible to generate sufficiently large sample sizes to analyse groups at a fairly disaggregated level. However, there will be differences between the characteristics of the groups in the two samples. Unless these are addressed, they will bias any comparison. This would ideally be done by reweighting the LFS.

- The LFS does not pin down the time of learning. To the extent that earnings depend on the time of learning this will reduce the validity of the comparisons.

- Initial comparisons between the ILR-HMRC/DWP matched data and the LFS find some groups for which the earnings distributions are similar. With further work making the groups in each dataset match each other more closely, the degree of similarity between earnings distributions will increase.

Annex: Mapping LFS subjects to ILR study areas

The LFS variable subcod1 identifies the subject of the respondent's main qualification, and is recorded for qualifications at NQF level 2 and above. The table below shows our initial mapping of these to study areas in the ILR.

Table 8: Mapping of LFS subject areas to ILR study areas

subod1	LFS subject	ILR study area
1	Basic programmes	N/A
14	Teacher training and educ science	Education and Training
14.2	Educ science	Education and Training
14.6	Training for vocational teachers	Education and Training
21	Arts	Arts, Media and Publishing
21	Arts	Arts, Media and Publishing
21.1	Fine arts	Arts, Media and Publishing
21.2	Music and performing arts	Arts, Media and Publishing
21.3	Audio visual and media production	Arts, Media and Publishing
21.4	Design	Arts, Media and Publishing
21.5	Craft skills	Arts, Media and Publishing
22	Humanities	Social Sciences
22.2	Foreign languages	Languages, Literature and Culture
22.3	Mother tongue	Languages, Literature and Culture
31	Social and behavioural science	Social Sciences
32.1	Journalism and reporting	Arts, Media and Publishing
32.2	Library information archive)	Arts, Media and Publishing
34	Business and admin	Business, Administration and Law
34	Business and admin	Business, Administration and Law
34.1	Wholesale and retail sales	Retail and Commercial Enterprise
52.3	Electronics and automation	Engineering and Manufacturing Technologies

subod1	LFS subject	ILR study area
52.4	Chemical process	Engineering and Manufacturing Technologies
52.5	Motor vehicles ships and aircraft	Engineering and Manufacturing Technologies
54	Manufacturing and production	Engineering and Manufacturing Technologies
54.1	Food processing	Engineering and Manufacturing Technologies
54.2	Textiles clothes footwear leather	Engineering and Manufacturing Technologies
54.3	Materials eg wood pa per glass plastic	Engineering and Manufacturing Technologies
54.4	Mining and extraction	Engineering and Manufacturing Technologies
58	Architecture and building	Construction, Planning and Built Environment
58.1	Architecture and town planning	Construction, Planning and Built Environment
58.2	Building and civil engineering	Construction, Planning and Built Environment
62	Agriculture forestry and fishery	Agriculture, Horticulture and Animal Care
62	Agriculture forestry and fishery	Agriculture, Horticulture and Animal Care
62.1	Crop and livestock production	Agriculture, Horticulture and Animal Care
62.2	Horticulture	Agriculture, Horticulture and Animal Care
62.3	Forestry	Agriculture, Horticulture and Animal Care
64	Veterinary	Agriculture, Horticulture and Animal Care
72	Health me dicine n ursing dentistry etc	Health, Public Services and Care
72.1	Medicine	Health, Public Services and Care
72.3	Nursing and caring	Health, Public Services and Care
72.4	Dental studies	Health, Public Services and Care
72.5	Medical di agnostic an d treat ment tech	Health, Public Services and Care
72.7	Pharmacy	Health, Public Services and Care
76.1	Child care and youth services	Health, Public Services and Care
76.2	Social work and counselling	Health, Public Services and Care
8	Literacy and numeracy	Preparation for Life and Work
81	Personal services	Preparation for Life and Work

subod1	LFS subject	ILR study area
81	Personal services	Preparation for Life and Work
81.1	Hotel restaurant and catering	Leisure, Travel and Tourism
81.2	Travel tourism and leisure	Leisure, Travel and Tourism
81.3	Sports	Arts, Media and Publishing
81.4	Domestic services	Health, Public Services and Care
81.5	Hair and beauty	Health, Public Services and Care
84	Transport services	Leisure, Travel and Tourism
85.2	Natural environments and wildlife	Construction, Planning and Built Environment
85.3	Community sanitation services	Construction, Planning and Built Environment
86	Security services	Health, Public Services and Care
86.1	Protection of persons and property	Health, Public Services and Care
86.2	Occupational health and safety	Health, Public Services and Care
86.3	Military and defence	Health, Public Services and Care
9	Personal skills	Health, Public Services and Care

Source: Frontier qualitative assessment

Annex: Data Tables

ILR-HMRC/DWP by level

Table 9: Distribution of annual earnings by achievement level, men 16-25 in FE

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
<20	0	0	0	0	0	0
25	0	0	0	0	0	99
30	0	0	0	427	0	528
35	0	0	0	1,152	0	1,035
40	0	417	0	2,012	0	1,614
45	0	1,137	0	2,995	0	2,262
50	0	2,064	52	4,070	0	2,975
55	168	3,205	447	5,240	0	3,784
60	688	4,479	1,053	6,535	258	4,706
65	1,501	5,946	1,995	8,033	633	5,745
70	2,716	7,697	3,333	9,672	1,198	7,018
75	4,341	9,740	5,091	11,485	2,070	8,653
80	6,507	12,113	7,472	13,467	3,283	10,595
85	9,562	14,735	10,492	15,768	5,144	12,870
90	13,425	17,518	14,071	18,419	8,863	16,093
95	18,794	22,630	18,913	23,431	15,019	21,720
99	31,910	36,437	31,305	36,939	28,664	35,559

Source: ILR-HMRC/DWP matched data

Table 10: Distribution of annual earnings by achievement level, men 26 and over in FE

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
<20	0	0	0	0	0	0

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
25	0	0	0	0	0	0
30	0	0	0	0	0	0
35	0	0	1,854	1,834	2,762	2,551
40	0	138	4,948	5,068	5,842	5,757
45	1,907	2,723	7,912	8,088	9,113	9,292
50	4,793	5,889	11,125	11,212	12,583	12,982
55	7,965	9,522	13,856	14,022	15,725	16,170
60	11,753	12,923	16,129	16,293	18,420	19,049
65	15,142	15,978	18,215	18,492	21,047	21,715
70	18,320	18,817	20,395	20,729	23,689	24,410
75	21,526	21,822	22,784	23,072	26,434	27,080
80	25,062	25,320	25,534	25,847	29,370	29,935
85	29,242	29,498	28,710	29,190	32,620	33,046
90	34,659	34,902	33,221	33,762	36,986	37,412
95	43,982	44,425	41,463	41,815	44,467	45,010
99	75,943	76,092	68,087	68,531	68,293	67,224

Source: ILR-HMRC/DWP matched data

Table 11: Distribution of annual earnings by achievement level, women 16-25 in FE

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
<20	0	0	0	0	0	0
25	0	0	0	0	0	251
30	0	0	0	492	0	687
35	0	0	0	1,170	0	1,157
40	0	495	0	1,951	0	1,675
45	0	1,229	0	2,793	0	2,238
50	30	2,154	300	3,740	0	2,864

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
55	459	3,276	811	4,766	40	3,560
60	1,132	4,568	1,600	5,855	280	4,340
65	2,110	5,945	2,727	7,069	612	5,225
70	3,381	7,572	4,138	8,440	1,092	6,294
75	4,963	9,423	5,875	9,967	1,818	7,609
80	7,030	11,468	8,057	11,667	2,823	9,178
85	9,700	13,584	10,514	13,626	4,382	11,007
90	12,840	16,212	13,313	16,078	7,538	13,408
95	17,206	20,731	17,029	20,231	13,693	17,722
99	28,551	33,001	27,888	32,372	25,190	30,328

Source: ILR-HMRC/DWP matched data

Table 12: Distribution of annual earnings by achievement level, women >25 in FE

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
<20	0	0	0	0	0	0
25	0	0	0	0	0	0
30	0	0	0	257	0	0
35	0	0	1,117	1,871	636	705
40	0	0	2,724	3,632	2,124	2,344
45	548	1,312	4,398	5,231	3,776	4,181
50	2,135	3,087	5,880	6,619	5,403	5,819
55	4,006	4,996	7,331	7,984	7,032	7,492
60	5,777	6,671	8,735	9,350	8,743	9,202
65	7,632	8,483	10,224	10,805	10,593	11,004
70	9,643	10,434	11,855	12,432	12,637	13,040
75	11,914	12,604	13,703	14,250	14,878	15,234
80	14,590	15,041	15,768	16,233	17,115	17,523

Percentile	Below Level2		Level 2		Level 3	
	Before	After	Before	After	Before	After
85	17,645	17,921	18,296	18,737	19,893	20,322
90	21,961	22,099	22,064	22,423	23,814	24,146
95	29,825	29,806	28,986	29,264	30,479	30,589
99	47,492	47,428	46,354	46,269	47,380	46,649

Source: ILR-HMRC/DWP matched data

Labour Force Survey by level

Table 13: Earnings distribution by level of highest achievement (vocational), men 16-25

Percentile	No qualifications	Below Level 2	Level 2	Level 3
0 to 35	0	0	0	0
40	0	0	0	0
45	0	0	0	0
50	0	0	0	3,380
55	0	0	2,340	8,424
60	0	0	6,240	11,232
65	0	0	9,256	13,208
70	0	4,004	11,024	14,976
75	0	7,800	12,480	16,380
80	0	10,712	13,780	17,992
85	3,380	12,480	15,184	20,020
90	10,400	14,664	17,576	22,672
95	13,780	17,992	21,008	26,416
99	21,840	26,000	30,004	39,000
Sample size	3,400	4,086	1,223	1,494

Source: LFS

Table 14: Earnings distribution by level of highest achievement (vocational), men over 25

Percentile	No qualifications	Below Level 2	Level 2	Level 3
0 to 35	0	0	0	0
40	0	0	0	0
45	0	0	7,176	0
50	0	0	12,220	12,012
55	0	11,440	14,404	15,860
60	0	14,404	16,016	18,096
65	0	16,380	17,160	20,384
70	0	18,200	18,720	22,620
75	11,024	20,696	20,800	24,492
80	14,144	22,984	22,776	26,780
85	16,796	25,480	25,012	30,004
90	20,020	30,004	27,976	33,332
95	25,012	36,608	34,008	39,988
99	39,988	60,008	47,008	57,564
Sample size	13,570	10,174	3,174	8,596

Source: LFS

Table 15: Earnings distribution by level of highest achievement (vocational), women 16-25

Percentile	No qualifications	Below Level 2	Level 2	Level 3
0 to 35	0	0	0	0
40	0	0	0	0
45	0	0	0	0
50	0	0	0	2080
55	0	0	0	4784
60	0	0	3120	6760
65	0	0	5200	8996
70	0	0	6760	10816
75	0	2704	8996	12012
80	0	5096	10400	13208
85	0	8320	12012	14560
90	4160	11024	13260	15964
95	9828	13988	15964	18512
99	15600	19656	21476	25012
Sample size	2903	3327	1674	1553

Source: LFS

Table 16: Earnings distribution by level of highest achievement (vocational), women over 25

Percentile	No qualifications	Below Le vel 2	Level 2	Level 3
0 to 35	0	0	0	0
40	0	0	0	0
45	0	0	1560	3276
50	0	0	4576	5512
55	0	3120	5980	7488
60	0	5044	7436	9152
65	0	6604	8736	10712
70	0	8372	9984	12012
75	3484	10192	11700	13988
80	5512	12324	13000	15184
85	7800	14404	14508	17004
90	10816	16900	16484	19500
95	14404	21008	20020	23504
99	23400	31980	29016	34008
Sample size	15538	14279	4967	6005

Source: LFS

Table 17: Earnings comparison, Level 3 men 16-25 studying Engineering and Manufacturing Technologies

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	0	632	0	0
30	0	1,948	0	0
35	0	3,349	0	0
40	0	4,884	0	0
45	183	6,461	0	6,240

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
50	721	8,214	0	10,504
55	1,474	9,969	6,136	12,480
60	2,521	11,665	8,112	14,976
65	3,776	13,362	10,192	16,640
70	5,222	15,044	11,700	18,096
75	7,170	16,883	13,000	19,500
80	9,359	18,903	14,560	21,008
85	12,231	21,330	15,600	23,712
90	15,844	24,800	17,680	25,532
95	22,287	30,024	19,292	29,276
99	34,767	43,754	26,000	39,988

Source: ILR-HMRC/DWP matched data and LFS

Table 18: Earnings comparison, Level 3 men 16-25 studying Construction, Planning and Built Environment

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
20	0	0	0	0
25	0	464	0	0
30	0	2,327	0	0
35	226	4,462	0	0
40	988	6,509	0	0
45	2,094	8,476	0	0
50	3,326	10,540	0	0
55	4,577	12,386	0	6,240
60	6,269	14,042	6,240	10,816
65	8,095	15,683	8,320	13,364
70	9,772	17,423	10,400	14,924
75	11,284	18,916	13,000	16,224

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
80	12,900	20,774	14,040	17,680
85	14,916	23,073	15,652	18,980
90	17,826	26,150	17,992	21,996
95	23,764	31,422	21,840	26,000
99	36,760	44,445	35,984	49,400

Source: ILR-HMRC/DWP matched data and LFS

Table 19: Earnings comparison, Level 3 men 16-25 studying Information and Communication Technology

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0-20	0	0	0	0
25	0	0	0	0
30	0	406	0	0
35	0	965	0	0
40	-	1,606	0	0
45	0	2,318	0	0
50	0	3,020	0	0
55	0	3,810	0	2,080
60	0	4,668	0	4,316
65	128	5,605	0	6,240
70	476	6,689	3,848	8,736
75	978	7,975	8,424	11,388
80	1,797	9,515	11,492	13,208
85	3.029	11,303	13,000	15,600
90	5,180	13,620	15,600	17.004
95	11.047	18,092	17,992	26,000

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
99	24,787	32.219	22.152	32.500

Source: ILR-HMRC/DWP matched data and LFS

Table 20: Earnings comparison, Level 3 men 26-59 studying Health, Public Services and Social Care

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	0	0	0	0
30	0	145	0	0
35	1,806	2,453	0	0
40	4,646	5,344	7,176	9,204
45	7,459	8,147	11,388	13,000
50	10,659	11,240	12,740	14,560
55	13,523	14,070	13,520	16,640
60	16,235	16,555	14,872	17,992
65	18,621	18,975	15,704	19,188
70	20,984	21,358	16,796	21,008
75	23,769	24,118	17,992	22,880
80	26,786	27,115	19,760	25,012
85	30,483	30,513	21,736	26,988
90	34,993	35,047	24,024	30,940
95	41,617	42,444	27,612	35,984
99	61,631	63,217	40,144	44,356

Source: ILR-HMRC/DWP matched data and LFS

Table 21: Earnings comparison, Level 3 men 26-59 studying Engineering and Manufacturing Technology

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	0	0	0	0
30	450	119	0	0
35	4,061	3,720	0	0
40	7,096	7,146	0	0
45	10,551	11,319	0	0
50	14,502	15,493	10,400	14,976
55	18,075	19,217	14,040	17,992
60	21,001	22,097	16,016	20,020
65	23,556	24,734	17,888	21,996
70	26,079	27,147	19,552	24,024
75	28,675	29,634	21,580	26,000
80	31,472	32,241	24,024	28,496
85	34,363	35,122	26,000	30,992
90	38,238	39,000	30,004	34,996
95	45,405	46,073	34,580	42,016
99	64,595	66,710	50,180	60,008

Source: ILR-HMRC/DWP matched data and LFS

Table 22: Earnings comparison, Level 3 men 26-59 studying Construction, Planning and Built Environment

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	0	0	0	0
30	0	0	0	0
35	3,029	2,316	0	0
40	5,754	5,625	0	0
45	8,959	9,540	0	0
50	12,655	13,706	0	0
55	15,746	17,128	0	0
60	18,580	20,192	8,996	0
65	21,380	22,779	14,560	15,600
70	23,908	25,303	17,004	19,188
75	26,436	27,726	19,916	21,476
80	29,243	30,326	22,360	24,024
85	32,128	33,217	25,220	26,988
90	35,980	37,058	28,808	30,524
95	42,094	43,444	34,996	35,984
99	62,702	62,397	47,996	52,000

Source: ILR-HMRC/DWP matched data and LFS

Table 23: Earnings comparison, Level 3 men 26-59 studying Information and Communication Technology

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	0	0	0	0
30	0	0	0	0
35	0	403	0	0

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
40	2,337	3,416	0	0
45	5,592	6,352	7,488	0
50	8,843	9,544	11,960	8,008
55	12,112	12,835	13,104	12,480
60	15,161	15,714	15,600	14,976
65	17,699	18,160	17,212	17,004
70	20,264	20,449	17,992	18,200
75	22,760	23,166	19,968	20,020
80	25,713	26,155	22,984	23,400
85	28,978	28,880	24,024	30,004
90	33,032	33,261	27,768	35,984
95	40,549	40,965	38,376	42,016
99	67,788	59,849	48,516	56,992

Source: ILR-HMRC/DWP matched data and LFS

Table 24: Earnings comparison, Level 3 men 26-59 studying Leisure, Travel and Tourism

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	0	0	0	0
30	0	0	0	0
35	1,556	1,195	4,992	0
40	3,559	3,054	10,972	0
45	6,516	5,391	13,988	10,920
50	10,063	7,794	14,976	14,976
55	13,043	10,243	16,016	16,796
60	16,401	12,857	17,316	17,992
65	19,050	16,500	18,200	19,604

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
70	22,218	19,658	19,760	21,476
75	25,225	23,028	21,008	22,984
80	28,391	26,377	22,620	24,700
85	32,587	30,455	24,492	26,520
90	38,463	36,112	26,988	30,004
95	50,346	45,375	33,800	34,996
99	88,273	70,965	40,040	44,980

Source: ILR-HMRC/DWP matched data and LFS

Table 25: Earnings comparison, Level 3 men 26-59 studying Business, Administration and Law

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
0 -20	0	0	0	0
25	2,831	4,612	0	0
30	7,773	9,416	0	0
35	11,999	13,338	0	0
40	15,321	16,217	6,500	12,480
45	17,395	18,405	12,012	15,600
50	19,454	20,491	13,364	17,992
55	21,415	22,501	15,600	20,384
60	23,345	24,570	16,900	21,632
65	25,439	26,699	17,992	24,024
70	27,364	28,659	19,448	25,220
75	29,291	30,556	21,996	27,612
80	31,634	32,673	22,880	30,004
85	34,353	35,555	25,012	33,488
90	38,136	39,665	26,988	39,988
95	44,828	46,370	33,020	47,216

Percentile	Pre-learning	Post-learning	Level 2 LFS	Level 3 LFS
99	67,360	69,452	44,980	90,012

Source: ILR-HMRC/DWP matched data and LFS

This publication is also available on our website at www.bis.gov.uk

Any enquiries regarding this publication should be sent to:

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

If you require this publication in an alternative format, email enquiries@bis.gsi.gov.uk, or call 020 7215 5000.

URN 11/1038