CVER Discussion Paper Series - ISSN 2398-7553

CENTRE FOR VOCATIONAL EDUCATION RESEARCH

The Payoff to Vocational Qualifications: Reconciling Estimates from Survey and Administrative Data

Gavan Conlon, Sophie Hedges

Steven McIntosh, Damon Morris and Pietro Patrignani

Research Discussion Paper 009

November 2017

CENTRE for ECONOMIC PERFORMANCE

National Institute of Economic and Social Research

LE London Economics

The University Of Sheffield.

The Centre for Vocational Education Research (CVER) is an independent research centre funded by the UK Department for Education (DfE). CVER brings together four partners: the LSE Centre for Economic Performance; University of Sheffield; National Institute of Economic and Social Research and London Economics.

Any views expressed are those of the authors, and do not represent the views of DfE. For more details on the Centre, go to cver.lse.ac.uk

Published by:
Centre for Vocational Educational Research
London School of Economics & Political Science
Houghton Street
London WC2A 2AE

The Payoff to Vocational Qualifications: Reconciling Estimates from Survey and Administrative Data

Gavan Conlon[1], Sophie Hedges[1]

Steven McIntosh[2], Damon Morris[2] and Pietro Patrignani[1]

November 2017

Abstract: This paper takes as a starting point the wage/earnings differentials associated with vocational qualifications that have been estimated in the literature. Two distinct parts of that literature are identified, using survey and administrative data respectively. We describe the results of example papers of each type, and show how they have produced different estimates of the differentials, particularly for lower level qualifications. We then discuss the various differences in equation specification that have typically been estimated by users of the two types of data. We replicate existing results in our own analysis of the two data sources, and then show how the estimated differentials change when the specifications are changed, step-by-step, to more closely resemble each other. The results show that when estimated on the same specification, survey and administrative data produce estimated earnings differentials that follow the same pattern in terms of relative size across qualifications, and that for most qualifications are similar in absolute size. There does not seem to be one particular specification difference that was driving the differing results in the literature, with some specification changes being more important for some qualifications, but less so for others.

Keywords: Wage differentials; Vocational education; Administrative and survey data.

JEL codes: I26, J31

Acknowledgements: We would like to thank Sandra McNally and an anonymous referee for helpful comments on an earlier draft, as well as other CVER members attending an internal presentation of the paper. In addition, participants at a departmental seminar at the University of Sheffield in May 2017 and a Department for Education Evidence Seminar in June 2017 provided helpful comments and suggestions. The Labour Force Survey data used were kindly provided by the Data Archive at the University of Essex. We would especially like to thank the Department for Education (DfE) for their help with access to the administrative data used in this project.

Affiliations: Corresponding author: Steven McIntosh, s.mcintosh@sheffield.ac.uk

All authors are affiliated with the Centre for Vocational Education Research

1. London Economics, Somerset House, Strand, London WC2R 1LA
2. Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT

Executive Summary

Estimating the wage differentials associated with particular qualifications has increasingly become part of the general 'returns to education' literature. The focus in this paper is on vocational qualifications. Two particular strands of the literature looking at such qualifications can be identified. The first uses data on individuals obtained through representative sample surveys, such as the Labour Force Survey (LFS) in the UK (see for example, Dearden *et al.*, 2002, 2004; McIntosh, 2006; and Jenkins *et al.*, 2007), while the second uses data from administrative sources, such as the Individualised Learner Record (ILR) in the UK, often matched to other data sets (see for example, BIS, 2013, and Bibby *et al.*, 2014).

The starting point of this paper is to report the estimated wage/earnings differentials that have previously been obtained in these two strands of the literature. Differing results have often been observed between them, particularly for low level vocational qualifications, with the administrative data typically producing larger estimated differentials.

A possible reason for these differences is that the equations estimated to produce these results typically vary between data sources. In particular, there are differences in the sample used (full- or part-time, age group considered and when the qualification was acquired), the variables used (definition of the earnings variable used as the dependent variable e.g. hourly wages or daily earnings, how qualifications are observed, and the list of control variables included), the comparison group against whom the wage/earnings differentials are measured, and finally the specification used (whether qualifications entered into a single or separate equations).

Having shown that the two data sources have been used to estimate different specifications which then produce different results, this leads to the two research questions of this paper: (i) Do the two data sources produce more similar estimates of the differentials associated with vocational qualifications, when a common specification is used to obtain those estimates? (ii) Which of the differences in specifications can explain the variation in estimated results between data sources that has previously been observed in the literature?

In order to answer these questions, LFS and ILR data were obtained with which to estimate various wage/earnings equations. For the LFS, we use data for the period 2004-2016, while for the ILR, the sample used is the Key Stage 4 cohorts from 2001/02, 2002/03, 2003/04 (i.e. those aged 15 at the start of those academic year), with labour market outcomes observed at the age of 26 in matched HMRC tax revenue data.

We first estimate specifications that are typical of the usual equations estimated by researchers with the two types of data, and confirm that we can replicate the results usually found in the literature. We then alter the specifications, one change at a time, until they are as similar as possible across data sets.

The results show that changing the specification of the estimated equation does cause significant movement in the estimated differentials. In particular, changing the comparison group used with the ILR from non-achievers to those whose highest qualification is one level below increases the estimated earnings differential for most of the qualifications considered. While this pertains to the ILR and hence younger workers only, it suggests that at least for this age group, the choice of a non-achievers comparison group might not have generated as much (upward) bias as previously thought likely by critics of this strategy.

The exception is for Level 1 qualifications, for which there is a significant fall in estimated earnings differentials, when the comparison group is changed from being non-achievers for the same qualification. It would seem that the entry into Level 1 vocational qualifications involves very particular selection effects that are difficult to model.

Turning to the LFS estimates, there is more flexibility to vary the specification, and so more changes made in this case. Estimating different equations for each qualification, rather than a single equation with all qualifications entered, does not make any difference to the estimated differentials associated with any qualification. A more important change is to the list of control variables used in the estimated equations, which has varying effects for different qualifications, increasing differentials for some qualifications (NVQ2), reducing differentials for others (BTEC Levels 3 and 2, and intermediate apprenticeships), while having no effect for still others (advanced apprenticeships and NVQ3s). These differing effects reflect different relationships between each qualification and the various control variables. The next change is to restrict the sample to younger individuals, observed in the labour market aged below 30, to mirror the fact that the ILR results are for individuals observed at age 26. This again has varying effects, typically raising the estimated wage differentials associated with Level 2 qualifications, but reducing the differentials at Level 3. The final change to the LFS specification is to use weekly earnings rather than hourly wages as the dependent variable. This change is made to reflect that fact that the administrative data set does not contain a hourly wage measure, but has to rely upon a daily earnings measure derived as the ratio of annual earnings to annual days worked (as reported in annual tax records). This change to weekly earnings in the LFS tends to increase

estimated differentials, reflecting the fact that the more qualified tend to work longer hours, the effect of which is also picked up in a weekly earnings measure.

After making these various changes, the cumulative result is to produce specifications that are very similar using both types of data. This then allows us to answer the research questions set out above. The results show that when used to estimate similar specifications, survey and administrative data do produce very similar results for qualifications at all levels except Level 1 – at least for the younger workers, on whose data the equations are estimated. The relative rankings of qualifications in terms of their estimated earnings differentials are the same, with apprenticeships attracting the highest differentials, and NVQs earning higher differentials than BTECs at each level. To answer the second question, it is not the case that there is a single change in specification that was the main source of the variation in results between the 'typical' specifications for *all* qualifications. Some changes in specification are more important for some qualifications but less so for others. At Level 2, where much of the controversy over differing results in the literature has been found, the key specification changes which bring the LFS results up to those observed in the ILR are the use of weekly earnings rather than hourly wages, followed by restricting the sample to young workers.

Finally, it is important to stress that the common specification estimated in this report is not being claimed to be the optimal specification, but merely the one that can be estimated on both sets of data. Furthermore, for that reason, the results of the common specification are not necessarily the 'best' estimates of the earnings differentials for each qualification. What the paper does show is that we have confidence that either data set can be used for future research, as availability and the requirements of the researcher vary, and neither data set should be viewed as necessarily producing unrealistic estimates.

1. Introduction

A significant amount of research time and effort has been spent on estimating wage or earnings differentials between individuals with and without a particular level of education.[1] There are many estimates of such differentials available in the literature, however they do not always tell the same story, even when they are for the same country at the same point in time. This has particularly been the case for recent work studying vocational qualifications in the UK where a range of estimates has been found. In particular, different results seem to be obtained depending on whether survey or administrative data are used.

This paper will first replicate the results that have been obtained in the recent literature using the two types of data. It will then identify the differences in the specifications most frequently estimated in the literature when using the two data sources. Having done this, the aim is then to establish, as far as possible, a common specification across the two data sources. We will then be in a position to answer the following questions: (i) Do the two data sources produce more similar estimates of the wage differentials associated with vocational qualifications, when a common specification is used to obtain those estimates? (ii) Which of the differences in specifications can explain the variation in estimated results between data sources that has previously been observed in the literature?

The survey dataset most often[2] used in the UK literature has been the Labour Force Survey (LFS). This is a representative sample of UK households, undertaken quarterly, with a narrow longitudinal element in that respondents are interviewed for five consecutive quarters before dropping out of the sample. The survey is useful in that it provides detailed information on all qualifications held (not just highest qualification) including data on type, level and subject of qualification. In addition, information on wages and hours worked is available, and so an hourly wage variable can be derived. The survey also provides a wide array of individual and job, and a somewhat narrower range of firm-level, characteristics, to act as control variables.

Administrative data have become available through the matching of data on the population of Further Education learners in the Individualised Learner Record (ILR), forwards to labour market data in the HMRC tax records database, and backwards to school records in the National Pupil Database (NPD). Containing administrative data provided by learning providers, the ILR

[1] See Ashenfelter *et al.* (1999), Blundell *et al.* (1999) and Harmon and Oosterbeek (2000) for reviews of this literature.

[2] For example, Dearden *et al.* (2002, 2004), McIntosh (2006) and Jenkins *et al.* (2007)

supplies very detailed information on the actual learning programme being followed, as well as having the advantage of huge sample sizes allowing disaggregation of results at a detailed level. A key limitation of the merged administrative data set, for the present purpose, is that it does not contain information on hours of work and so it is not possible to construct an hourly wage variable. The information available is for annual earnings, which is affected by time spent working as well as the wage rate received. The approach usually adopted is to make use of information on days worked per year, to create a daily earnings variable as a proxy for the wage received, though this will clearly still be affected by variation in hours worked per day and so is a less than ideal solution.[3]

The next section discusses some of the results that have been produced by previous research using these two data sources. The following section sets out our methodology for reconciling these findings, with the results provided in Section 4. A final section concludes. We show that there are significant differences in specification between the two literatures, relating to the comparison group used, the age group considered, the definition of the dependent variable and the list of control variables included. When the specifications are equivalised, they do produce much more similar estimates of the earnings differentials associated with each qualification (though not hourly wage differentials, since such information is absent in the merged administrative data).

There is, however, no single specification difference that causes the variation in results between data sources for every qualification; some specification differences are important for some qualifications but less so for others.

[3] Throughout this paper, the term 'wages' will be reserved for the hourly wage rate received by individuals. Whenever a longer time period involving a variation in hours of work is considered, for example, daily, weekly or annual, then the term 'earnings' will be used.

2. Existing Results in the Literature

Much of the earlier work on wage differentials associated with vocational qualifications, and education in general, was undertaken with data obtained from sample surveys of the population, with the LFS being the most popular in the UK. Key papers in this area[4] include Dearden *et al.* (2004) and Jenkins *et al.* (2007).

Dearden *et al.* (2004) use LFS data from 1996 to 2002. Their focus is on Level 2 qualifications, held as the individuals' highest qualification, with no restrictions on current age. Results are presented for two different comparison groups, the first having no qualifications at all, and the second having either none or at best Level 1 qualifications. Their sample sizes in the second case (which is similar to that which we use here) is around 85,000 in each of the separate male and female equations estimated. The majority of these observations are from the comparison group; the numbers of treated observations with the various qualifications range from 13,700 males with an apprenticeship, through 2,800 males and 4,400 females with an NVQ2 qualification, to below 500 for each gender with BTEC and RSA qualifications at Level 2. The estimated equations control for age, age squared, region, ethnicity, employer type and firm size.

Jenkins *et al.* (2007) consider vocational qualifications at both Level 2 and 3, but other than that adopt a similar approach to Dearden *et al.* (2004). Their sample period is extended from 1997 to 2006, and they again include all individuals of working age. For Level 2 vocational qualifications, they choose as the comparison groups the same two as adopted by Dearden *et al*. (2004), as discussed above. For the Level 3 vocational qualifications, the comparison group used is all individuals with at most Level 2 qualifications (either academic or vocational). Control variables used in the estimated wage equations are age, ethnicity, region, year, part-time status, and whether a proxy responded for the individual. The authors report number of observations used as around 53,000 and 40,000 for males and females respectively for the Level 2 equations, and around 80,000 and 85,000 for the Level 3 equations. Few of these are treatment group observations, with all vocational qualifications bar apprenticeships and NVQs held by only around 1-2% of the sample.

Both of these papers find similar results. The average wage differentials between individuals with and without Level 2 vocational qualifications are always very small, often zero and

[4] A summary of this literature can be found in McIntosh (2010).

sometimes even negative and statistically significant (particularly for NVQ2 qualifications). This last result received further attention, since it seems unlikely that one's wages would actually fall following the completion of an NVQ2. The argument therefore made is that there is a selection effect into low-level vocational qualifications, such that learners on such qualifications will have unobserved characteristics associated with lower wages on average. When the comparison group is adjusted to try to reflect the sort of individuals who would choose to do the NVQ2, by considering only those with no qualifications at all in the comparison group, then Dearden *et al.* show they can obtain a positive, albeit very small at around 2%, wage differential, while Jenkins *et al.* also report small positive differentials, including a statistically significant differential of 5% for women. At Level 3, Jenkins *et al*. (2007) report that wage differentials to Level 3 vocational qualifications are respectable and usually in double figures, if a little lower than those of their academic equivalents.

More recent work using the LFS, for example by McIntosh and Morris (2016) has replicated these earlier results on average, using the most up-to-date data, with a sample of LFS data covering the period 1997-2015. While a range of specifications are estimated, the preferred ones, and those most comparable to the earlier work discussed above, estimate the wage differentials associated with holding vocational qualifications at each level as highest qualifications, in each case relative to comparison groups of individuals whose highest qualification is specifically one level lower (with separate specifications for the comparison group to contain individuals with any, academic only and vocational only at the lower level). The number of observations varies across these different specifications, but is typically around 25,000 for the Level 3 equations, and 15,000 for the Level 2 equations. With the data set pooled across more years than the earlier work, cell sizes of treated individuals with vocational qualifications number at least several hundred in all cases, except for RSA qualifications for males.

McIntosh and Morris' findings are similar to those from earlier papers discussed above. In particular, large positive wage differentials are associated with most Level 3 vocational qualifications, with Level 2 differentials being considerably lower and in some cases (NVQ2) observed to be negative. The authors go on to show the variation in results around these averages, by subject area of the qualification, with the highest differentials being observed for engineering and construction qualifications. This latter result is mostly due to such qualifications leading to higher paying occupations. Looking *within* occupations, the

differentials between individuals with vocational qualifications and those qualified to one level below are more similar.

None of the research using the LFS discussed above has addressed the endogeneity issue, that the attainment of a qualification is not random, and that individuals select into particular qualifications. As mentioned above, in the case of low level vocational qualifications in particular, this could be negative selection, such that individuals of lower ability or other characteristics associated with lower wages are more likely to choose to undertake such qualifications. Thus, the observed wage differentials cannot be interpreted as causal. Other research has included childhood test scores as an indicator of ability, in an attempt to control for this selection, at least to some extent. For example, Dearden *et al.* (2004) use the British Cohort Study (BCS), which is a birth cohort data set that has followed a group of individuals born in a particular week in 1970 throughout their lives. Their wage equation specifications are the same as those they estimated with the LFS as described above, with the difference being the addition of controls for ability (childhood test scores) and family background. Numbers of observations are also considerably smaller, given the smaller overall size of the BCS, at around 1,300 in each of the male and female equations with no qualifications or at best Level 1 as the comparison group. In addition, since it is a birth cohort data set, the wages are observed for all observations at a specific age, in this case age 30, in the year 2000. They find similar results to those obtained in their LFS analysis, with observed wage differentials associated with NVQ2 qualifications still often being negative, unless a specific comparison group of individuals with no qualifications is used, in which case the differential is positive, but small and statistically insignificant.

Using the same data, De Coulon and Vignoles (2008) estimate first-differenced wage equations, looking at the *change* in wages following the attainment of a qualification, which will difference out any unobserved characteristics of the individuals that remain constant. The comparison is therefore with individuals who do not acquire the qualification. Their results, looking at the change in wages between the ages of 26 and 34 (1996 to 2004), are based on 202 individuals in the birth cohort who acquire an NVQ2 qualification in this period. In a first-differenced equation, only factors that vary over time are included as control variables, so the variables included differ from those used by the cross-sectional studies discussed above (although time-constant variables are implicitly controlled for). In particular, the control variables included are time spent unemployed and the number of children born in the study period. The authors find significantly higher wages for those who acquire an NVQ2

qualification in this period, in the order of around 22% on average, and slightly higher and statistically significant for females when split by gender. This is a large effect, though it should be noted that this is a single estimate for a single cohort, and other studies using the same methodology on other cohorts of individuals have not found similarly high results.[5]

Turning to the administrative data studies, these have grown in number in recent years, as matched administrative data sets have become more widely available. Examples of research that has studied vocational qualifications and *earnings* include BIS (2013) and Bibby *et al*. (2014), recalling again that such data sets do not have measures of hourly wages, and are restricted to using a derived daily earnings variable. These studies both use the non-achievers (i.e. those with the same learning aim as the qualification of interest) as the comparison group. The idea is that by comparing the treated individuals with the qualification of interest to this group, the selection effects into such qualifications are held constant. The downside of the methodology is that it introduces other potential biases, for example motivation or ability differences between the achievers and non-achievers in the treatment and control groups. Clearly, sample sizes are much larger when using administrative rather than survey data, since the former includes the population of learners. Hence, the two studies referenced here have sample sizes of 6.8 million and around 15 million learners respectively. BIS (2013) look at learners observed between 2003 and 2006 together with their earnings observed between 2004 and 2010, while Bibby *et al.* (2014) focus on learning spells of age 19+ learners between 2002 and 2011, and their observed earnings between 2005 and 2012. Control variables used by both studies include indicators of socio-demographic characteristics (such as age, gender, ethnicity and region) and course characteristics (delivery mode, number of hours etc).

The results in BIS (2013) show positive and significant earnings differentials associated with most vocational qualifications, with even those at Level 2 showing healthy differentials, for example 12% for City and Guilds and 6% for NVQs at this level, seven years after completion of the qualification. Bibby *et al.* (2014) follow a similar methodology, but include a wider range of control variables from the period before enrolment onto the learning programme, for example the number of days in the previous year spent on benefits or in employment. They also undertake analysis where the treatment and control groups are matched on such characteristics. The idea is that the prior labour market history will proxy the ability of the

[5] For example, Jenkins *et al.* (2002), using data from a 1958 birth cohort, the National Child Development Study (NCDS), do not find a statistically significant positive change in wages following the acquisition of any intermediate level qualification, between the ages of 33 and 42.

individuals, as valued by the labour market. The authors do not consider individual qualifications, as the other studies reviewed here do, but rather aggregate qualifications up into their respective levels. The results show large positive and statistically significant earnings differentials. For example, a full Level 2 qualification is associated with 11% higher earnings for achievers relative to non-achievers, averaged over the period 3-5 years after attainment.

More recent work using administrative data, for example Patrignani *et al*. (2017), has made use of the matched Longitudinal Education Outcomes (LEO) dataset, which merges NPD and ILR data sets together, and therefore removes the restriction that all individuals to be analysed must be in the ILR and so in some funded Further Education. Their control group is therefore based on individuals holding qualifications at the level immediately below in the Regulated Qualification Framework. They continue to find particularly strong and positive differentials for Level 4 vocational qualifications, Apprenticeships, and NVQs at Levels 2 and 3., though not for BTEC qualifications at Levels 2 and 3 for men, or for any vocational qualifications at Level 1 bar NVQs for men, in their fully controlled ('augmented plus') specification. A similar move away from the 'typical' non-achievers specification to one using the level below as a comparison will be adopted in this study.

In general, therefore, the message is that the studies based on administrative data produce higher estimated differentials than those typically produced when using survey data, particularly compared to the cross-sectional studies using the LFS.

3. Methodology

The discussion of the existing literature using the two data sources in the previous section makes clear that although there are minor differences between the specifications used within the 'survey' and 'administrative' groups[6], looking across groups there are significant differences between their specifications. We consider a 'generalised' specification that is typical of the specifications estimated with each type of data, and are broadly consistent within groups, and then list the differences between those specifications across groups. Table 1 summarises each of these differences, which are discussed in more detail in Appendix 1. As can be seen in Table 1, a number of differences exist. These differences relate to the sample

[6] The exception to this statement is the Patrignani *et al.* (2017) study, which by not using a non-achievers comparison group marks a significant change compared to previous work using administrative data.

used (full- or part-time, age group considered and when the qualification was acquired), the variables used (definition of the earnings variable used as the dependent variable, how qualifications are observed, and the list of control variables included), the comparison group against whom the wage differentials are measured, and finally the specification used (whether qualifications entered into a single or separate equations).

Once the differences have been identified, we can estimate wage/earnings differentials for both data sources so that they more closely resemble each other. We can then answer the research questions posed in the Introduction, concerning whether the estimated differentials are more similar when a common specification is estimated, and which differences in specification are causing the differences in results in the existing literature.

LFS and ILR data were therefore obtained, with which to estimate the various specifications of the wage/earnings equations. For the LFS, we used data for the period 2004-2016, with the starting point determined by the year a question was first introduced into the survey asking respondents when they acquired their highest qualification (since this variable is used in the typical ILR specification). Data from the quarterly surveys in each year were pooled together, with only the wave one response used in order to avoid multiple observations on the same individuals. For the ILR, the sample used was the Key Stage 4 cohorts from 2001/02, 2002/03, 2003/04 (i.e. those aged 15 at the start of those academic year), with labour market outcomes observed at the age of 26 in matched HMRC tax revenue data.

The starting point is to estimate a specification for each data source, that replicates the specifications typically found in the literature.[7] Changes to the specification are then made, along the dimensions listed in Table 1. These changes are made one at a time, with the changes added either cumulatively or singly, in two separate experiments (the latter approach to check whether the order of changes matters in the cumulative approach).

The equation to be estimated is of the form

$$Ln\,(Earnings_i) = \sum_{j-1}^{j=J} \gamma_j Q_{ji} + \boldsymbol{\beta X_i} + u_i$$

[7] Note that it is the specification that we are replicating, not the sample period that was used in earlier studies in the literature. The results are therefore not expected to match exactly those found in the literature. For example the LFS results reported here using the most recent data available differ slightly from those reported in our earlier paper (McIntosh and Morris, 2016), and suggest a fall in estimated wage differentials to some qualifications in recent periods.

Where Q_j represents the list of vocational qualifications of interest, with the γ_j the estimated coefficients of interest giving the wage/earnings differentials, $\mathbf{X_i}$ is a vector of control variables, and u_i is a disturbance term. The estimated equation takes the same basic form in every specification, with the variation coming in terms of the sample used, the definition of the variables, whether qualifications are entered all at once or in separate equations and the comparison group used.

The starting specification for the administrative data analysis, matching the specification typically found in the literature, is the 'non-achievers' comparison, with a separate equation estimated for each qualification. The first step is to change the comparison group to those individuals who hold a qualification at one level below. The only other change is to restrict the control variables to match those available in the LFS.

The LFS offers the flexibility to make more changes to the specification, so there are more changes made on that side. The starting point is again the typical specification in the literature, with all qualifications at a particular level held as highest qualifications included in the same equation, to be compared to the control group of those whose highest qualification is one level below, with a dependent variable of log hourly wages. The stepwise changes are then to: estimate a separate equation for each qualification; add in a variable that controls for time elapsed since a qualification was acquired (as typically found in administrative data specifications); restrict the other control variables to those available in the ILR (just ethnicity, once time elapsed since the highest qualification was achieved is also controlled for); restricting the age of individuals considered, in an attempt to mirror the ages of learners used in the ILR (by considering only those individuals who acquired their highest qualification by age 22, and who are aged below 30 at the time of their survey[8]); and change the dependent variable to weekly earnings rather than hourly wages.

As mentioned above, one of the issues to be addressed in creating a common specification is the choice of the comparison groups, against which to compare the treated group with the qualification of interest. The change to be made is the movement away from using a comparison group of non-achievers when using the administrative data, to one of using individuals qualified to one level below as a reference point. This does not define exactly who should be in the comparison group, however, and different possibilities exist. We therefore experiment

[8] Sample sizes do not allow a specific age to be considered, as was done in the ILR by considering individuals' earnings at the specific age of 26.

with two different comparison groups, to gauge the effect of such choices on the estimated results. The first contains all individuals whose highest qualification was one level below that of the treatment qualification, whatever the nature of that qualification, while the second contains all individuals whose highest qualification was a *vocational qualification* one level below that of the treatment qualification. The rationale for the second control group is that individuals acquiring a vocational qualification are more likely to have qualified for that via vocational qualifications from the prior level, so that the latter provides a more realistic estimate of the counterfactual of what would have happened to the treated group had they not reached a new higher level through further qualification attainment. Table 2 summarises the qualifications that are held by individuals in the comparison groups for each of the qualifications of interest.

4. Results

A large number of results have been generated in this project, given the number of specifications estimated across two data sets, all for each combination of two genders, two comparison groups and two means of making the changes to the specifications (sequentially or cumulatively). In this section, we present the findings and discuss the implications for only one set of results, also focussing only on a subset of qualifications. The full, comprehensive set of results can be found in the appendix.

The results chosen as the main illustration to be discussed in this section, are for Comparison Group 2 (those with vocational qualifications at the level below), with the changes to the specification made cumulatively. The reason for the latter is that we want to compare equivalent specifications estimated on survey and administrative data sets, and so need to have considered all of the changes cumulatively that got us to these equivalent specifications. The reason for focussing on Comparison Group 2 is that this is the most common route into vocational qualifications at the next level.[9] Thus, the estimated wage/earnings differentials between the treatment group and Comparison Group 2 represent the average difference in wages or earnings between an individual with a vocational qualification, and an individual in whose position the former would most likely have been had they not attained that qualification.

[9] Table A3 shows, for example, that very few holders of any vocational qualifications, except the very highest, have achieved many academic qualifications at Level 2 (GCSEs).

This choice of comparison group therefore tries to take the best aspects of the non-achievers and general comparison group approaches, in an attempt to minimise selection issues. In particular, the choice of the comparison group with vocational qualifications at a lower level attempts to focus the analysis on *the sort of individuals who select into vocational education* in both treated and comparison groups, as does the non-achievers approach, but without introducing the new selectivity issue between treated and comparison groups of the former being successful in their learning aim and the latter not, that the non-achievers approach introduces. As such, the results based on this specification are our preferred specification.[10]

In addition, in the full discussion in the text, we focus on apprenticeships, BTEC and NVQ qualifications at Levels 2 and 3, as these are the qualifications taken most frequently by current young learners (see Table A1), and have received the most attention in the literature. The results are displayed in Tables 3 and 4, for males and females respectively.[11] The structure of the tables is the same in each case. The first row in each table reports the coefficient observed for each qualification, relative to the comparison group, when estimated using the 'typical' specification seen in the literature when using LFS data. The subsequent rows then report the observed coefficients when a single change is made to the specification of the LFS equation, with the changes made cumulatively (i.e. the previous changes are not reversed).[12] The final two rows of each table report coefficients estimated using the administrative data set. The results in the lowest row are obtained from the estimation of the 'typical' specification observed in the literature when using administrative data. The row above this changes the comparison group from non-achievers to the relevant comparison group (1 or 2) who are qualified to one level

[10] Of course, other selections issues remain even with the use of the preferred Comparison Group 2 here. For example, qualification attainment can be related to selection into employment and to fertility decisions. However, to the extent that this paper is a comparison between the results using two data sources, there is no reason to suppose that such selection issues should be more prevalent in one dataset than in the other, and so should not influence the *difference* in their results.

[11] The full set of results when changes are made cumulatively, for both comparison groups and for both genders, and for all qualifications, are shown in the appendix, Tables A4-A7.

[12] Tables A8-A11 in the appendix show the results when the changes are not made cumulatively, but rather each is made in turn as a single change relative to the baseline 'typical' specification. This is done as a robustness check, to ensure that the order the changes are enacted is not important for driving any results. We therefore check whether any change to the specification has a large effect on the results when enacted in isolation, in Tables A8-A11, but does not have any significant effect on the results when it is enacted as part of a longer chain of specification changes. Comparing the results in Tables A4-A7 to those in Tables A8-A11 show that there are virtually no such cases, suggesting that the order the specification changes were made when enacted cumulatively is not important for driving their relative effects. The exception to this is that the effect of changing the dependent variable is, in most cases, much smaller when this change is enacted on its own rather than at the end of chain of changes (compare row 6 results to row 1 results in Tables A8-A11 and they are often not too dissimilar). This implies that the difference in results between using daily wages and weekly earnings in the LFS is greater when applied to the age-restricted sample, than when applied to the full sample.

below. Hence the lowest row of LFS results and the highest row of ILR results (i.e. rows 6 and 7 in Tables 3 and 4) report the results from equivalent specifications using LFS and ILR data respectively. These two rows are highlighted in each table as the rows to be compared across data sets.

4.1 Comparison Group 2 – those with a vocational qualification one level below as a highest qualification

Replicating existing specifications in the literature

Consider first males (Table 3) and the specifications that replicate those typically found in the literature using the two approaches (row 1 for the LFS and row 8 for the ILR). In terms of the general patterns, the earnings differentials associated with apprenticeships are the largest of all qualifications when using the administrative data. In the LFS, the wage differential is higher for other Level 3 qualifications than for the Advanced Apprenticeship. However, the Intermediate Apprenticeship is still the best performing qualification at Level 2 for males.

Adjusting the ILR specification

Comparing the results in row 7 to those in row 8 in Table 3 reveals that when Comparison Group 2 is substituted for the non-achievers comparison group, then the estimated earnings differentials increase for qualifications. It therefore seems clear that the high earnings differentials observed in administrative data results are not being driven by the use of the non-achievers comparison group, when this pertains to younger workers. Higher differentials remain, and indeed are larger in most cases, when a comparison is made with those whose highest qualification is one level lower.

Adjusting the LFS specification

Rows 2-6 of Table 3 introduce, cumulatively, the various changes to the LFS specification. Estimating the wage differentials using a separate equation for each qualification makes virtually no difference to the results. Similarly, in most cases changing the list of control variables makes little difference. The main exceptions are the BTEC qualifications, at both Level 2 and Level 3. The wage differentials for these qualifications fall when the control variable for time elapsed since qualification acquisition is added, and fall again when the usual

LFS control variables (age, whether the individual works full-time, regional dummy variables, and an indicator of whether the individual works in the public sector) are dropped from the equation. This would suggest that the holders of BTEC qualifications are more likely to have obtained their qualifications a longer time ago than individuals in the comparison group acquired their qualifications. This longer period of holding the qualification is then part of the reason for the higher differentials for BTEC qualifications observed above (assuming a positively-sloped wage profile over time), so that the estimated wage differential falls when time elapsed since acquisition is held constant in the comparison between groups. On the other hand, removing the LFS control variables leading to a lower differential suggests that BTEC holders have other characteristics associated with lower wages, so that when these characteristics are not held constant, then the observed differential again falls.

Restricting the sample to those aged under 30, and who obtained their highest qualification by age 22, has a variable effect on the estimated wage differentials, which increase for some qualifications and fall for others, with an increase being the more common outcome. At Level 2, all differentials are observed to be larger when the sample is restricted to this age group. This either suggests that such wage gains are short-lived and fade out as workers get older or that differentials are larger amongst current younger workers because of other changes in education and the labour market. The two apprenticeship levels also see higher wage differentials amongst the young sample, showing the early benefit to wages from involvement in such employment-based training. The only qualifications to experience a significant fall in their wage differential when switching to the young sample are BTEC Level 3. This is consistent with a steeper age-earnings profile for these qualifications, with higher wages being received when prime-aged rather than when young.

The final change in the LFS specification is to use weekly earnings rather than hourly wages as the dependent variable. For all Level 2 and 3 qualifications, this change results in larger differentials.[13] The changes are very large in some case, with those qualifications most associated with longer hours per week seeing big increases in their observed earnings differential. Looking at the distribution of weekly hours for Level 3 qualifications for males (Figure 1), it is difficult to discern differences between them. Looking at just the mean number of hours per week by qualification and gender (Figure 2), the shortest male hours amongst holders of Level 3 qualifications are for those with a BTEC Level 3, and this is the group for

[13] Though note the full results in Table A6 for all qualifications show that Level 1 and Level 4 qualifications experience a fall in estimated differentials when making this change.

whom moving from daily wage differentials to weekly earnings differentials makes no difference, with no evidence of an increase. At Level 2, the longest hours are worked by those with an Intermediate Apprenticeship and the shortest by those with a BTEC Level 2 qualification, and these are the qualifications that see the largest and smallest change in estimated differential, respectively, when moving to weekly earnings. The changes are therefore all consistent with observed hours of work. The change in point estimates are greater for lower level qualifications than higher level qualifications. It is thus potentially important for understanding why the source of data matters for those with low level qualifications in particular.

Comparing ILR and LFS results

Having adjusted the LFS specification so that it matches the ILR specification as closely as possible, how do the estimated differentials compare across data sets using such equivalent specifications? Comparing rows 6 and 7 in Table 3 (highlighted) shows that, on the whole, the results tell a consistent story in terms of ranking of qualifications by size of earnings differential. At both Level 2 and Level 3, both data sets agree that the earnings differential associated with NVQs is larger than that associated with BTEC qualifications. For example at Level 3, on the common specification, both data sets agree that the NVQ earnings differential is more than twice the size of the BTEC differential. This is despite the fact that the 'typical' LFS specification has a BTEC differential that is 50% larger than the NVQ differential, at Level 3. Looking at the various rows of specification changes for the LFS, the NVQ3 differential barely changes across the initial four changes, with the big increase due to the switch from hourly wages to weekly earnings, which in turn is due to the above average weekly hours of male NVQ3 holders, as noted above. For BTEC Level 3 qualifications, on the other hand, the change in the dependent variable makes little difference, and it is the cumulative effect of the other four changes that each reduce the estimated wage differential, in particular the two step changes in the control variables used, and also, to a slightly smaller extent, the change to the young age-restricted sample. Overall, this example is a very useful one to make an important point. Though both the NVQ3 and the BTEC Level 3 earnings differentials in the LFS end up similar to their ILR equivalents once the specification has been equalised as far as possible, the importance of the various changes to the specification differs across the two qualifications. It is therefore not the case that there is a common cause of variation in results between data sets.

The cause of the difference in results between data sets differs according to the qualification being considered.

Similarly, comparing the NVQ and BTEC differentials at Level 2 across data sets when estimated on a common specification, for the former qualification, the two data sets produce almost identical results (differing by just 0.009). For the BTEC Level 2 qualification, however, the ILR estimate (7% earnings differential) is substantially smaller than the LFS estimate (19%).[14] Thus, although the changes to the LFS specification have successfully reversed a dominant BTEC advantage in the 'typical' LFS specification (earnings differentials of 12% versus 5% for BTEC Level 2 and NVQ2 respectively) into a larger differential for the NVQ2, the difference between the earnings differentials for the two qualifications is much smaller in the LFS than in the ILR, when estimated on the common specification. However, on closer inspection, the penultimate LFS row, reporting the results from the age-restricted specification *but still using hourly wages*, does find a very similar estimated differential for BTEC level 2 qualifications as that found in the ILR (around 7% in each case).

This last finding raises an interesting point. The ILR dependent variable is the derived *daily* earnings for each individual. In attempting to replicate this variable in the LFS we have used *weekly* earnings, it not being possible to derive a daily measure in the LFS. However, in some cases, hourly wages may actually be a closer proxy than weekly earnings, for the daily earnings observed in the ILR. As argued above, part of the reason for the higher differentials observed when using earnings measures is due to longer time spent working by the more qualified. Such labour supply effects are not picked up by hourly wage measures. If the variation across workers in hours of work supplied is in terms of hours per day (some working 4 hours per day, others working 8 hours per day, for example) then the daily earnings measure is influenced by such variation in hours, and its differential will resemble that measured using weekly earnings. However, if the variation in hours of work supplied is in terms of number of days worked per week, with no variation across workers in hours worked per day (part-time workers working 3 full days per week, for example) then the weekly earnings measure would pick up this variation, but the daily earnings measure would not. In such a case, the ILR differential estimated using daily earnings would more closely resemble the LFS differential measured on hourly wages, and it is the latter that should be the comparison.

[14] All figures reported in the text are calculated as $e^{\beta} - 1$, where β is the estimated coefficient, as reported in the tables.

Hence, we should expect the LFS hourly wage differential to be close to the ILR daily earnings differential, for those qualifications where any variation in hours worked across individuals is found more in terms of days worked per week rather than hours worked per day. The LFS contains some questions that may shed a little light on this. These questions ask whether the individual works at weekends, and whether they have a zero hours contract, both of which would be indicative of greater variation in days per week. The proportion of individuals in these two work situations is shown in Figures 3 and 4 respectively, by highest qualification, and separately by gender. Consider the male BTEC Level 2 qualification discussed above. Given the ILR daily earnings differential for this qualification is closer to the hourly wage differential than to the weekly earnings differential in the LFS, the prediction above suggests there should be more variation in days worked per week for this qualification. This is indeed what we observe using the two indicators in Figures 3 and 4. Amongst men, above Level 1, BTEC Level 2 holders have the highest rate of working at weekends, and, though the numbers are small, by far the highest proportion working under a zero hours contract. It therefore seems that in such cases, with variable days worked per week, then the hourly wage measure is the more appropriate match for the daily earnings measure from the ILR in the common specification.

Turning to the results for women using Comparison Group 2, these are presented in Table 4. The relative size of the differentials for NVQ and BTEC qualifications are in agreement across the two data sets even in the 'typical' specifications (with BTECs having higher differentials in both, unlike for men). There are fewer differences in the 'typical' results between data sets to be explained, and therefore less work for the specification changes to do. Most of the initial specification changes do not have a large effect on the estimated differentials, NVQ2 being an exception where the changes to the control variables increase the wage differential from essentially zero, to positive and significant in the range 3-6%.

There are larger changes to the estimated wage differentials for women when the sample is restricted to the under 30s (with highest qualification attained by age 22) in row 5. For the Level 2 qualifications, the observed differentials generally seem to be larger amongst the young sample compared to the sample covering the whole working age. For the Level 3 qualifications the reverse is true, and the younger sample, at least at the point of their lives that they are observed, receive lower differentials than older individuals.

The final change (row 6) is the alteration of the dependent variable to weekly earnings, and this does have a large effect on the estimated differentials, in general increasing them in size by varying degrees, sometimes substantially. This is due to the fact that women with higher level qualifications tend to work more hours per week than women at lower levels, as can be seen in Figure 2. It is also clear in Figure 5, which shows the full distribution of hours by qualification, and reveals the higher spikes in the distributions around 40 hours for women with Level 4, Advanced Apprenticeship and NVQ3 qualifications, and the lower likelihood of working 20 hours or fewer for women with these qualifications.

As far as the comparison of results across data sets is concerned, as with the men, for some qualifications, the LFS differential closest to that estimated on daily earnings in the ILR is obtained using hourly wages (with BTEC and NVQ Level 2 qualifications providing good examples) whilst for other qualifications, it is the weekly earnings measure in the LFS that produces the closest match to the ILR estimate (for example, Advanced Apprenticeship and NVQ3, which both produce very similar estimates when the common specification is used). As discussed with the male results above, we would expect the hourly wage measure in the LFS to produce the closest results, for those qualifications where individuals typically work full days, and any variation in hours per week is more in terms of days worked per week. Looking at Figures 3 and 4, however, the story is not as clear as it was for the men above. Whilst it is true that BTEC holders, particularly at Level 2, are more likely to have a zero hours contract (as was found for males) it is not the case that they are more likely to work at weekends as might have been expected according to the above hypothesis, with the holders of apprenticeships and NVQs at both Level 2 and 3 being more likely to work at weekends. It therefore seems as though patterns of work for women are more complex than can be captured by the somewhat crude indicators used here, making it difficult to explain why hourly wages match the ILR results best for some qualifications, while weekly earnings match best for others. What we can say, at least, is that we can find a specification which produces similar results to those seen in the ILR for most qualifications.

Level 1 Results

Level 1 qualifications were not included in Tables 3 and 4 with the other example qualifications earlier, but their results can be found in the Appendix, in Tables A6 and A7 for Comparison Group 2 and cumulative changes to specifications. The results for Level 1 qualifications, in

terms of comparing specifications, do not follow the general pattern of results discussed above, which is why they were not included in the main tables.

Unlike the other qualifications discussed above, when the ILR specification is changed from one with non-achievers as the comparison group to one with individuals whose highest qualification is one level below, there is a significant fall in the estimated earnings differential, though for NVQ1 and Other Level 1 qualifications, it does remain positive and statistically significant. At least part of the large positive earnings differentials that have been estimated with administrative data in the literature is therefore due to the use of the non-achievers comparison group used by such papers. The difference between administrative and survey data results at Level 1 is not purely, or even largely, down to the use of the non-achievers comparison group by the former, however. All of the changes to the LFS specification to produce a consistent specification merely succeed in producing large negative earnings differentials associated with Level 1 qualifications, that are statistically significant in the case of NVQs for both genders. The difference between the specifications that remains is that with the ILR administrative data, the comparison group contains individuals with qualifications below Level 1 (i.e. Entry Level qualifications), while the analysis with the LFS data uses individuals with no qualifications as the comparison group.

These results suggest that the negative selection effects into vocational qualifications at Level 1 are particularly strong, and unless individuals with such qualifications are compared to others who have chosen or selected into low level vocational learning, then the estimated earnings differentials are likely to be biased downwards, due to the unobserved characteristics of such individuals.

4.2 Comparison Group 1 – all with highest qualification one level below

As argued earlier, Comparison Group 2 is considered to be the more appropriate comparison for individuals with intermediate vocational qualifications, since vocational qualifications at the lower level are more likely to be what individuals would hold in the absence of further vocational attainment. For completeness, though, full results for all qualifications when using comparison group 1 are reported in Tables A4 and A5 in the appendix, for males and females respectively. This allows us to determine whether the conclusions reached in the previous section are specific to using Comparison Group 2, or whether they can be generalised to using other comparison groups.

Looking at Tables A4 and A5, although the estimated coefficients are different to earlier, (since the comparison group has changed), the general pattern of results, and the implications that can be drawn from them, are very similar. In particular, the results with Comparison Group 1 suggest that, for males (Table A4): the largest differentials are associated with apprenticeships; the 'typical' LFS results suggest that within levels, BTEC qualifications earn higher differentials than NVQs, but that this is reversed once the specification is adjusted to match that used with the administrative ILR data; and that Level 1 qualifications earn positive earnings differentials compared to non-achievers in the administrative data set, but remain large and negative when using the survey data.

When comparing the equivalent specification results across data sets (the highlighted results in rows 6 and 7), the results overall are less similar than those produced using Comparison Group 2. While the common specification produces very similar estimates in both data sets for both apprenticeships and for NVQ3s, for other qualifications the results are further apart. In particular, changing the control variables has much more of an effect on the results using Comparison Group 1. This shows that there are larger differences in the observed control characteristics between treatment and comparison groups when Comparison Group 1 is used, which would suggest that there are also likely to be larger differences in unobserved characteristics between the treatment group and Comparison Group 1.

For women (Table A5) the results using Comparison Group 1 are even more erratic. Only two qualifications (Level 4 and Other Level 3) attract statistically significant coefficients in the LFS version of the common specification, and in general there is much less similarity between the ILR and LFS common specification results than was observed for Comparison Group 2. This reinforces our view that individuals with academic qualifications at a lower level are not the most appropriate people against whom to compare those with vocational qualifications at the higher level, apparently particularly for women. It seems that the 'sort of people' who take vocational qualifications, particularly lower level ones, are typically not the 'sort of people' who hold academic qualifications at a lower level.

Another possible reason for such larger differences between data sources when Comparison Group 1 is used is the sources of information about lower level academic qualifications held (self-reported in the LFS relative to the National Pupil Database with the administrative data).

5. Conclusions

The analysis in the preceding section has presented a large number of estimates of wage/earnings differentials associated with the full range of vocational qualifications, estimating a range of specifications, and using the two principal data sources that have been employed in the literature. One, the ILR, is an administrative data source, matching information on the population of learners on vocational courses in Further Education, to tax records that provide information on their earnings once in work after the period of learning. The other, the LFS, is a survey-based data set of a random sample of individuals of working age. Previous research in the literature using these two data sources has produced differing estimates of the wage/earnings differentials associated with some vocational qualifications, particularly those at lower levels. The aim of this paper was to see whether such variation in results can be explained in terms of differences in specification of the equations typically estimated with the two types of data. If such specification issues are relevant in explaining the difference in results in the literature, then the second aim of the paper was to identify which differences in specification in particular were driving the variation in results.

The first conclusion to draw from the results is that the specification used for the wage/earnings equation is important, and when a common specification (or as close to a common specification as the two data sources allow) is estimated, the ILR and LFS produce similar estimates of the differentials associated with most qualifications – at least for the younger workers considered here. It can therefore be stated that there is nothing inherent in one data source or the other that produces biased results. The nature of the differences in results is not the datasets themselves, but rather in how they are used and interpreted. This is a re-assuring result. It means that there is nothing fundamentally wrong with using either data set, and both can continue to be used to estimate wage/earnings differentials associated with qualifications, allowing researchers to take advantage of the benefits of each. Thus, administrative data can continue to be used, and indeed developed further, as new possibilities for matching administrative data sets become available, while survey data will also continue to play a role, with for example their greater number of variables with which to examine interacting effects on the differentials.

The exception to this is Level 1 qualifications, for which the two data sources continue to tell a different story, with the administrative ILR data presenting positive (when non-achievers form the comparison group) or neutral (when Entry Level individuals form the comparison

group) estimates. The survey data results, on the other hand, continue to show large negative differentials. These differences are likely to be due to the different comparison groups used, which in turn are important on account of the particular selection issues into Level 1 vocational education. The administrative data results show that when compared to others who have selected into similarly low level vocational qualifications, but who did not achieve, then positive earnings differentials are observed (though these could also be influenced by differences in, for example, motivation between achievers and non-achievers). Unlike for the intermediate levels, the use of the non-achievers comparison group does therefore produce higher differentials for Level 1 qualifications. Even when using the 'level below' comparison group, however, there is still no equivalence of ILR and LFS estimates at Level 1, given the ILR's superior information on qualifications below Level 1. For this particular case, Level 1, the LFS data, using a 'no qualifications' comparison group, are unlikely to produce an accurate estimate of the gains to be made by the individuals who choose to do such qualifications. Those who have selected into Level 1 Further Education will remain a special group, unlike other FE learners, and modelling their earnings benefit will therefore continue to require special care.

Answering the second research question proved to be more difficult. The results have shown that the specification changes have different effects for different qualifications. For some qualifications, for example BTECs, the control variables included are particularly important, and the changes to the controls, in an attempt to match the ILR specification, tend to reduce the estimated differentials. For other qualifications, particularly those at Level 2, the move to restrict the age of the sample to young people aged under 30 has a large effect on the estimated differentials, typically to increase them. Finally, the change to using weekly earnings rather than hourly wages can have dramatically different effects on different qualifications' differentials, depending on their relationships with hours of work.

It is already clear, therefore, that there is not one single story to be told in terms of the impact of the various specification changes on the estimated wage/earnings differentials. Changes to the specification have different effects, not just in terms of size but also in terms of sign, for different qualifications and genders. It is not the case, therefore, that a particular specification change always affects an estimated differential upwards or downwards. The story is more nuanced than that, reflecting the differing characteristics of individuals with different qualifications (and of different genders).

A third research question that could have been asked, but deliberately was not as it was not the focus of this report but will be considered in future work, is what is the 'correct' specification for the wage/earnings equation. It is important to stress that we are not necessarily saying that the common specification estimated here is the 'correct' specification, or that the similar result produced for most qualifications when the common specification is estimated is the 'correct' estimate of the wage/earnings differential associated with those qualifications. Rather, the common specification was estimated as the 'lowest common denominator', that is, the most appropriate specification that could be estimated given the information available in the two data sets. There is no particular reason why such a specification should turn out to be the 'right' one.

Having established that either data source can be used, and that they will produce similar results when estimated across a common specification, the challenge now in forthcoming research is to identify the 'best' specification that can be estimated, given the data available to us. We simply make some initial observations here. First, in terms of data sources, administrative data sets represent an exciting new area for research, and will provide fruitful avenues for new research in this area. Survey data sets such as the LFS will also remain important, however, with the wider range of information they contain ensuring they will remain of interest. The inclusion of hours of work information, for example, means hourly wage measures can be constructed, allowing analysis of differentials that are not affected by differences in time spent in work. New additions to the survey, for example the inclusion of questions on parental occupation, will allow new analysis of social mobility. The administrative data sets, on the other hand, provide information on prior attainment at a young age, and also follow individuals over time. Such features allow researchers to control, for differences in ability between treatment and comparison groups, via the early attainment measures and potentially by differencing out unobserved characteristics that remain constant over time, using difference-in-difference and matching techniques. The challenge will be to develop specifications for the estimated equations, that exploit such possibilities, and others that become available as the potential to match data sources grows. Finally, in terms of comparison group, our preference is for Comparison Group 2 used here, that is, individuals whose highest qualification is a vocational qualification one level lower. Using this group limits the selectivity issues related to individuals who choose to do vocational qualifications, since the comparison group are on a similar route, but does not introduce additional selectivity issues between success and failure

in attainment, inherent in the non-achievers approach. The suitability of control groups will be further addressed directly in forthcoming research.

References

Ashenfelter, O., Harmon, C. and Oosterbeek, H. (1999). 'A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias.' *Labour Economics,* 6, 453-470.

Bibby D., Buscha, F., Cerqua, A., Thompson, D., and Urwin D. (2014). *Estimation of the labour market returns to qualifications gained in English Further Education*. Department for Business, Innovation and Skills Research Report 195, December 2014.

Blundell, R., Dearden, L., Meghir, C. and Sianesi, B. (1999). Human capital investment: the returns from education and training to the individual, the firm and the economy. *Fiscal Studies*, *20*(1), 1-23.

Dearden, L., McGranahan, L. and Sianesi, B. (2004). 'An In-Depth Analysis of the Returns to National Vocational Qualifications Obtained at Level 2'. Centre for the Economics of Education Discussion Paper 46.

Dearden, L., McIntosh, S., Myck, M. and Vignoles, A. (2002) 'The returns to academic and vocational qualifications in Britain,' *Bulletin of Economic Research*, 54, 249-274.

De Coulon, A. and Vignoles, A. (2008) 'An Analysis of the Benefit of NVQ2 Qualifications Acquired at Age 26-34'. Centre for the Economics of Education Discussion Paper 106.

Department for Business, Innovation and Skills (2013). *A Disaggregated Analysis of the Long Run Impact of Vocational Qualifications.* BIS Research Report 106. London Economics.

Harmon, C. and Oosterbeek, H. (2000). 'The Returns to Education: A Review of Evidence, Issues, and Deficiencies in the Literature.' Centre for the Economics of Education Discussion Paper 5.

Jenkins, A., Vignoles, A., Wolf, A. and Galindo-Rueda, F. (2002). 'The Determinants and Effects of Lifelong Learning'. Centre for the Economics of Education Discussion Paper 19.

Jenkins, A., Greenwood, C. and Vignoles, A. (2007). 'The Returns to Qualifications in England: Updating the Evidence Base on Level 2 and Level 3 Vocational Qualifications'. Centre for the Economics of Education Discussion Paper 89.

McIntosh, S. (2006) 'Further analysis of the returns to academic and vocational qualifications,' *Oxford Bulletin of Economics and Statistics*, 68, 225-251.

McIntosh, S. (2010). 'The Economic Value of Intermediate Vocational Education and Qualifications,' UKCES Evidence Report 11.

McIntosh, S. and Morris, D. (2016). 'Labour Market Returns to Vocational Qualifications in the Labour Force Survey'. Centre for Vocational Education Research Discussion Paper 002.

Patrignani, P., Conlon, G. and Hedges, S. (2017). 'The earnings differentials associated with vocational education and training using the Longitudinal Education Outcomes data'. Centre for Vocational Education Research Discussion Paper 002.

Figures

Figure 1: Distribution of Total Weekly Hours by Highest Qualification (Male)

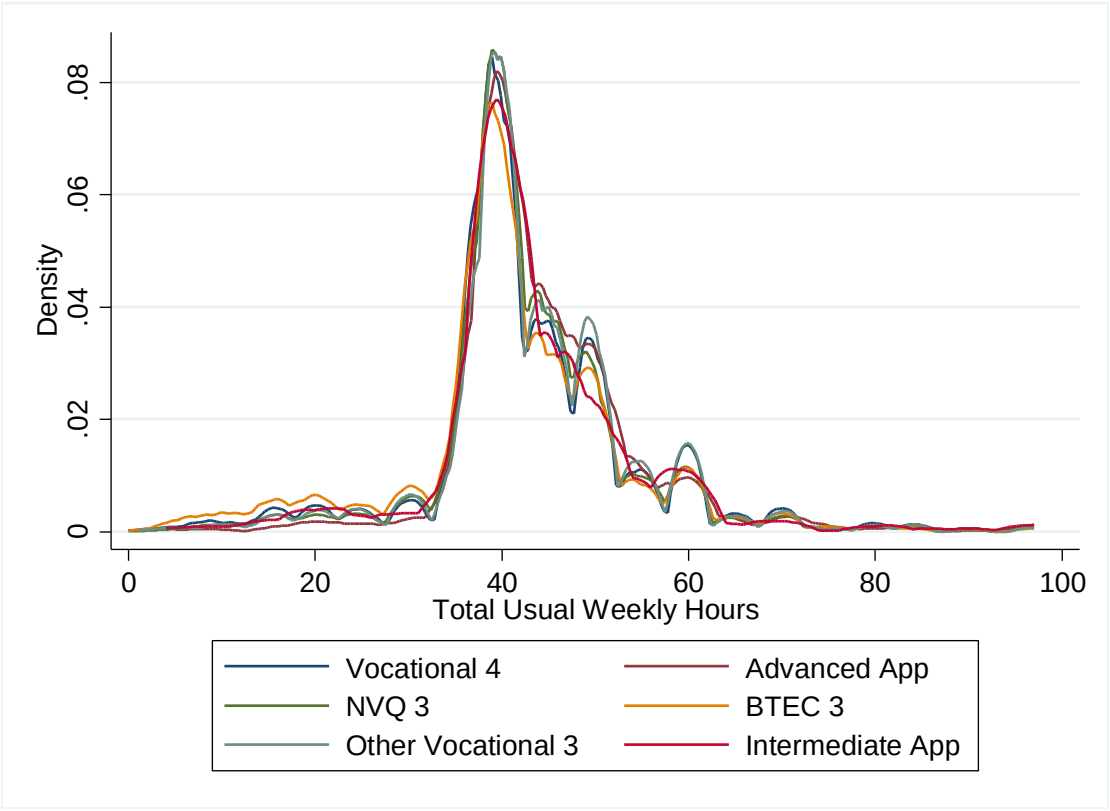


Figure 2: Mean Total Weekly Hours by Highest Qualification and Gender

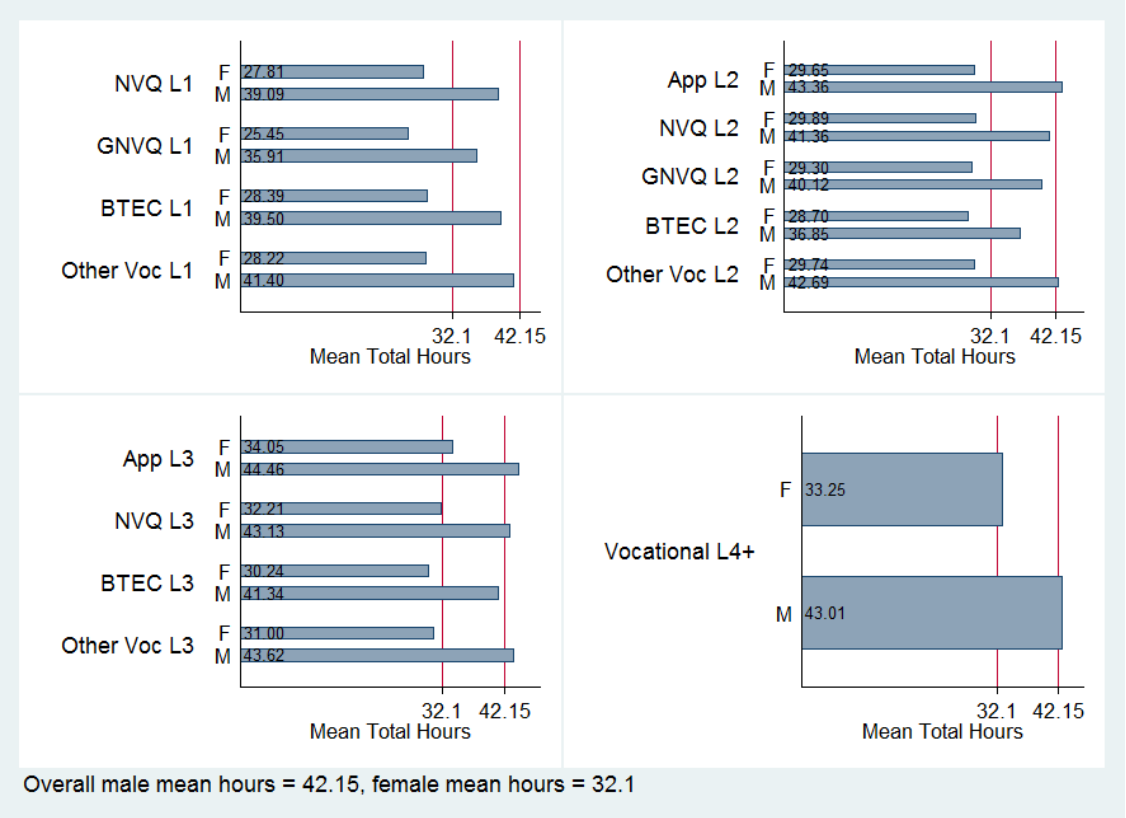


Figure 3: Proportion usually working weekends by Highest Qualification and Gender

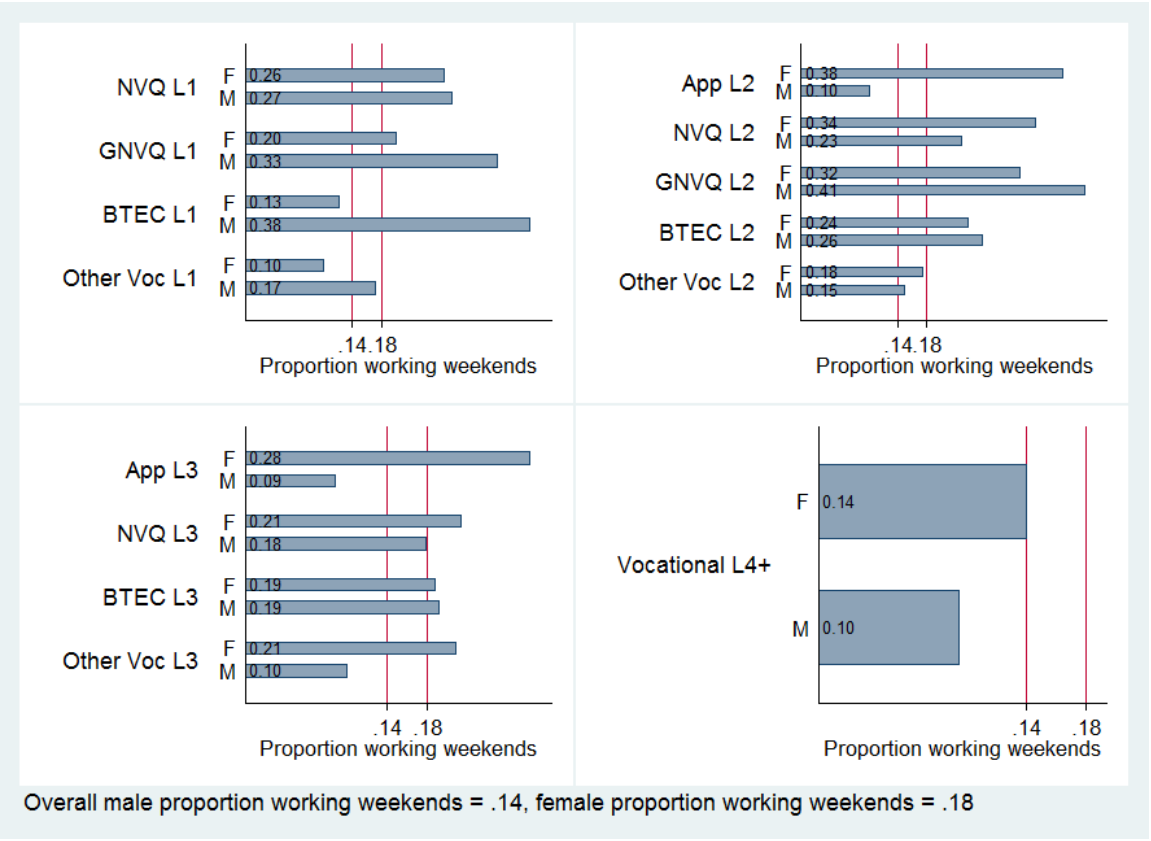


Figure 4: Proportion on Zero-hours Contracts by Highest Qualification and Gender

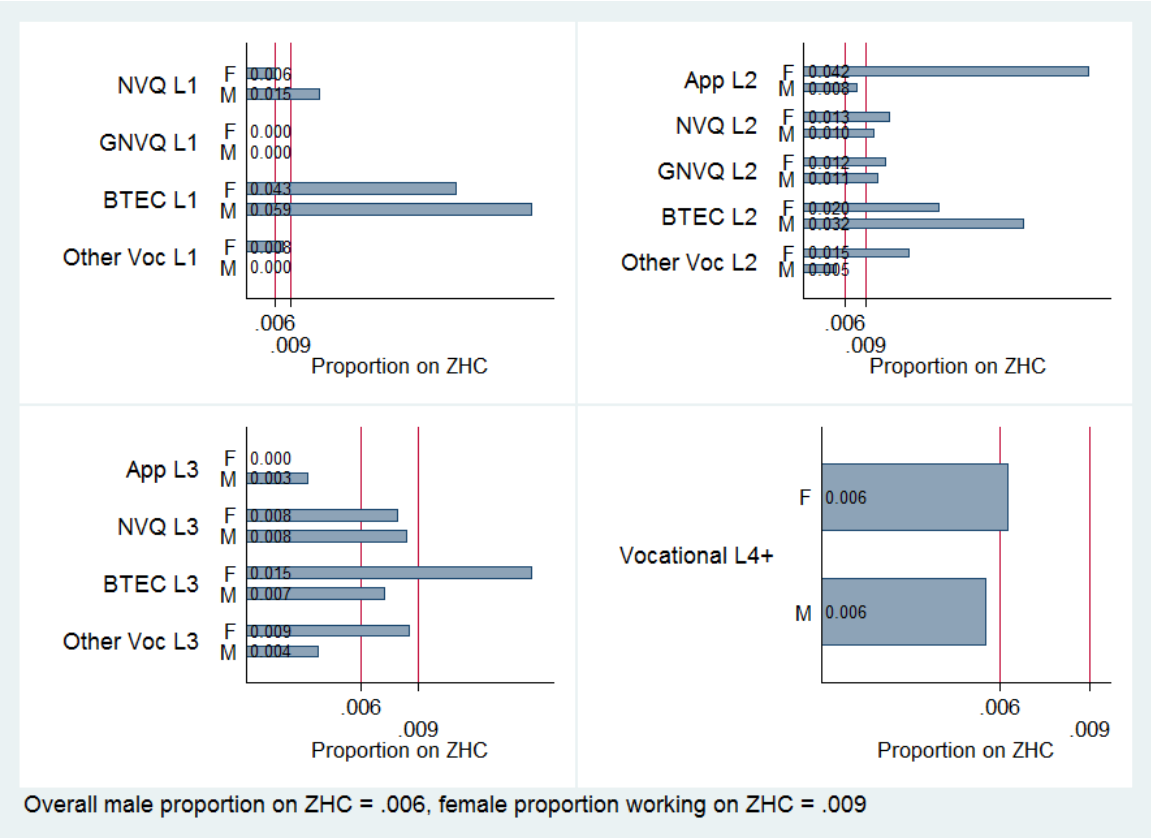
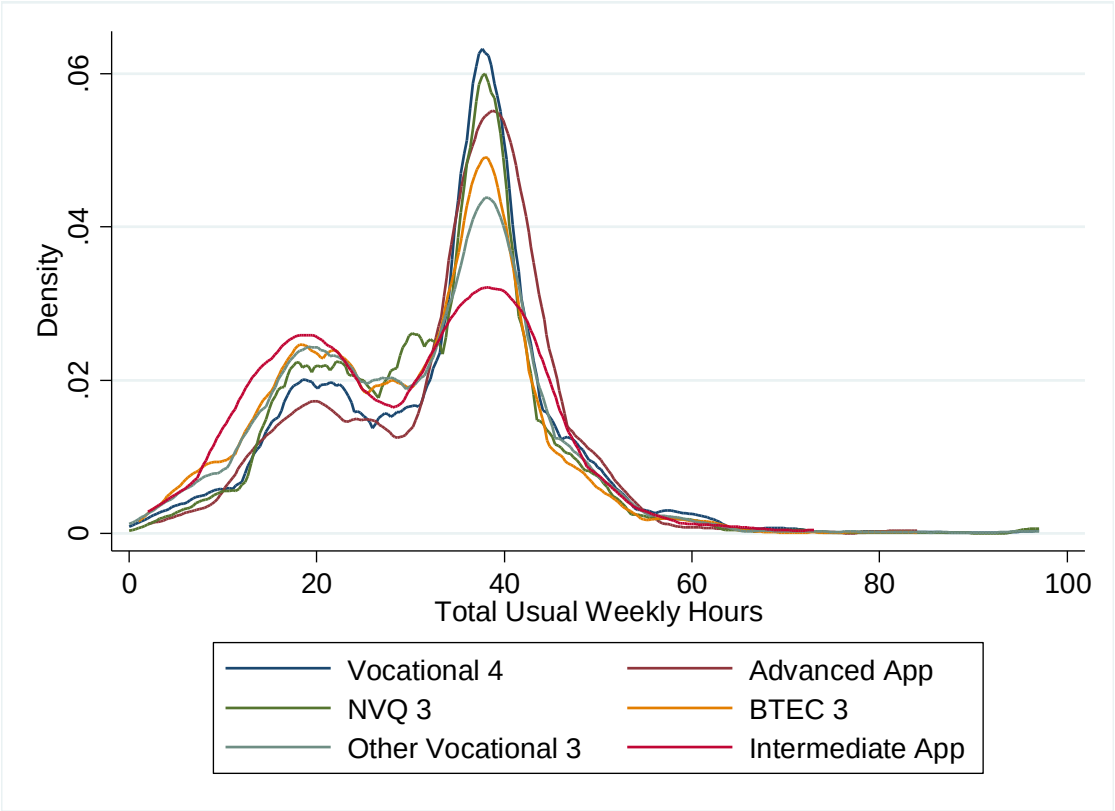


Figure 5: Distribution of Total Weekly Hours by Highest Qualification (Female)



Tables

Table 1: Specifications Used When Estimating Wage Differentials Using Administrative and Survey Data

	Administrative	Survey
1. Sample	All workers (irrespective of hours worked).	Usually full-time only.
2. Age group considered	All ages, but since analysis considers recent learners, they will mostly be young.	All ages.
3. Time since achievement	Up to 9 years	Could be long time ago, depending upon age of respondent. Observe when highest qualification was obtained.
4.Dependent variable	Annual wages – can deduce daily wages.	Hourly wages for those paid by the hour, or weekly /annual wages –can deduce hourly wages.
5. Vocational qualifications observed	All vocational qualifications taken in available period. Focus on highest qualification	All qualifications.
6. Control variables	Limited individual characteristics	Extensive individual characteristics.
	Prior ability (with NPD)	Nothing on prior ability.
	Nothing on job/firm characteristics.	Detailed job characteristics. Limited firm characteristics.
7. Control/ comparison group	Non-achievers studying for the same qualification.	All without the qualification ('all qualifications' specification)
		Those with no qualifications (highest qualification specification).
8. Specification	Estimate separate equations for different qualifications.	Estimate single equation with multiple qualifications.

Table 2: Comparison Group Used for Each Qualification of Interest

Type of qualification (treatment)	Comparison group 1	Comparison group 2
Level 4 vocational (incl. Higher Appren)	Level 3 (excluding 1 single A lev)	FL 3 vocational
Advanced Apprenticeship	Level 2 (all, including GCSEs)	Intermed. Apprenticeship
NVQ level 3	Level 2 (all, including GCSEs)	Level 2 vocational
BTEC level 3	Level 2 (all, including GCSEs)	Level 2 vocational
Other full level 3 vocational	Level 2 (all, including GCSEs)	Level 2 vocational
Intermediate Apprenticeship	Level 1 (all)	Level 1 vocational
NVQ level 2	Level 1 (all)	Level 1 vocational
BTEC level 2	Level 1 (all)	Level 1 vocational
Other full Level 2 vocational	Level 1 (all)	Level 1 vocational
NVQ level 1	Below Level 1/ no quals	Entry level/ no quals
BTEC level 1	Below Level 1/ no quals	Entry level/ no quals
Other level 1 vocational	Below Level 1/ no quals	Entry level/ no quals

Table 3: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Cumulative, Comparison Group 2, Males

	Advanced apprenticeship	NVQ3	BTEC3	Intermediate apprenticeship	NVQ2	BTEC2
1. LFS – typical spec	0.140***	0.157***	0.221***	0.193***	0.045***	0.114***
	(0.021)	(0.007)	(0.008)	(0.022)	(0.014)	(0.025)
2. LFS – separate equations	0.131***	0.155***	0.221***	0.209***	0.043***	0.132***
	(0.022)	(0.007)	(0.008)	(0.024)	(0.015)	(0.028)
3. LFS – add time elapse	0.132***	0.164***	0.176***	0.185***	0.068***	0.120***
	(0.022)	(0.007)	(0.009)	(0.024)	(0.015)	(0.028)
4. LFS – match ILR controls	0.127***	0.174***	0.111***	0.135***	0.074***	0.002
	(0.024)	(0.008)	(0.010)	(0.023)	(0.016)	(0.028)
5. LFS – restrict age	0.142***	0.163***	0.073***	0.217**	0.091**	0.073
	(0.034)	(0.016)	(0.019)	(0.050)	(0.043)	(0.049)
6. LFS – weekly earnings	0.195***	0.241***	0.080***	0.385***	0.194**	0.172*
	(0.042)	(0.022)	(0.025)	(0.091)	(0.085)	(0.099)
7. ILR – control group lower level	0.260***	0.212***	0.091***	0.290***	0.203***	0.064***
	(0.005)	(0.010)	(0.005)	(0.006)	(0.007)	(0.010)
8. ILR – typical (non-achievers)	0.225***	0.071***	0.087***	0.162***	0.128***	0.030*
	(0.007)	(0.016)	(0.007)	(0.006)	(0.010)	(0.018)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. The dependent variable is the natural logarithm of wages/earnings.
See Tables A1 and A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year. The 'typical' ILR specification controls for ethnicity and time elapsed since highest qualification was obtained.

Table 4: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Cumulative, Comparison Group 2, Females

	Advanced apprenticeship	NVQ3	BTEC3	Intermediate apprenticeship	NVQ2	BTEC2
1. LFS – typical spec	0.116***	0.093***	0.190***	-0.050*	-0.018*	0.070***
	(0.032)	(0.005)	(0.008)	(0.026)	(0.009)	(0.022)
2. LFS – separate equations	0.131***	0.093***	0.191***	0.041	-0.019**	0.135***
	(0.033)	(0.005)	(0.008)	(0.032)	(0.009)	(0.023)
3. LFS – add time elapse	0.132***	0.104***	0.136***	0.035	0.060***	0.121***
	(0.034)	(0.005)	(0.008)	(0.032)	(0.011)	(0.023)
4. LFS – match ILR controls	0.145***	0.118***	0.113***	-0.105***	0.034***	0.053**
	(0.035)	(0.005)	(0.009)	(0.028)	(0.011)	(0.022)
5. LFS – restrict age	0.068	0.070***	0.107***	0.078	0.100**	0.147**
	(0.046)	(0.014)	(0.018)	(0.051)	(0.042)	(0.060)
6. LFS – weekly earnings	0.158	0.147***	0.156***	0.184*	0.218***	0.344***
	(0.099)	(0.025)	(0.032)	(0.100)	(0.075)	(0.105)
7. ILR – control group lower level	0.173***	0.161***	0.258***	0.187***	0.141***	0.182***
	(0.008)	(0.007)	(0.007)	(0.009)	(0.010)	(0.015)
8. ILR – typical (non-achievers)	0.150***	0.091***	0.125***	0.138***	0.152***	0.076***
	(0.010)	(0.013)	(0.010)	(0.009)	(0.012)	(0.028)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. The dependent variable is the natural logarithm of wages/earnings.
See Tables A1 and A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year. The 'typical' ILR specification controls for ethnicity and time elapsed since highest qualification was obtained.

Appendix 1: Differences in 'Typical' Specifications Estimated in the Literature Using Administrative and Survey Data

Table 1 summarised the main differences in specifications typically estimated in the literature by researchers using survey and administrative data respectively. In this appendix, more detail is provided on these differences.

The first difference is in terms of the sample of observations used. The labour market information in the administrative data set is obtained from tax records, for which hours of work are not reported. Analyses using such data therefore have no option but to consider all workers. Researchers using survey data, on the other hand, have access to information on hours worked per week, and so can determine whether each individual is a full-time or part-time worker. Their wage equations are then most often estimated on full-time workers only, since there is less variation in hours of work amongst full-time workers than part-time and so it is not necessary to include the likely-endogenous variable 'hours of work' as an explanatory variable, which would bias the results.

The sample used to derive the results is also defined by the age group of individuals involved. Individuals of any age can potentially be found in either data set, and the LFS does indeed provide a representative cross-section of all those of working age, who are therefore used in the analysis. However, given that the available administrative datasets contain information on cohorts of learners from 2003 onwards, then younger age groups are likely to dominate.

For the same reason that linked administrative data sets are only available for relatively recent cohorts of learners, then any observed qualifications by definition will have been acquired relatively recently too. The LFS, on the other hand, asks respondents, of all ages, to report all qualifications that they hold, and so some of those reported qualifications will have been acquired some considerable time ago.

Turning now to the variables used in the estimated equations, the absence of hours information in the administrative data set, means that an hourly wage rate is not observed and cannot be derived. Earnings information is in the form of annual earnings received during the tax year. It is possible to derive a daily earnings measure, since the start and end date of each employment spell, and hence the total days spent in employment during the tax year, are observed. In the LFS, respondents report their earnings according to the usual period of payment (hourly, daily, weekly, monthly, and annually). For those who do not report an hourly wage rate, additional information on hours of work allow an hourly wage to be calculated.

The key explanatory variables are the qualification variables, and in this case the information available is similar across data sources. In both cases, information on all qualifications held is, at least in principle, available to the researcher. Learner identifier codes in the ILR allow the same individual to be tracked across different cohorts, so that all their learning aims, within the given sample period, will be observed, and the one that leads to the highest qualification can be identified. In addition, if the ILR is matched to the NPD, then academic qualifications in school are also observed. The LFS asks respondents to report all the qualifications that they hold, including all levels of each particular type of qualification, so that similarly a highest qualification can be identified. Thus, either data set could be used to estimate an equation

including variables for every qualification held (producing so-called 'average wage differentials') or an equation that includes a variable for just the highest qualification of each individual (producing so-called 'marginal wage differentials').

Although the information available on qualifications is similar across data sources, the way such information has been used has differed, for apparently no particular reason other than habit and precedent. In particular, wage equations using the administrative data have typically estimated a separate equation for each qualification of interest, whereas those estimated with the survey data have traditionally included an indicator for all qualifications, or at least all those at the same level, in the same equation.

The remaining variables in the estimated wage equations comprise the control variables. The ILR contains very limited information on individuals' characteristics, with which to derive control variables to hold those characteristics constant. Specifically, the list is restricted to gender and age. Additional information can be obtained if the ILR is matched to the NPD, such as ethnicity, and markers to indicate those individuals with Special Educational Needs (SEN) and who are entitled to free school meals (FSM). Additionally, the NPD could provide prior attainment information, such as GCSE results and Key Stage test scores.[15] The LFS contains more information on individual characteristics and those of their jobs, plus limited data on the characteristics of their firms, such as size.

Finally, but potentially importantly, there is the issue of the control or comparison group used, against which to compare the wages of the vocational qualification holders. The early work using the administrative data sets faced the issue that all observed individuals were learners, since the ILR was the basis of the dataset. Therefore there were no non-learners to form a comparison group. The group usually chosen as the control group were therefore non-achievers, that is, individuals with the same learning aim as the treatment group of interest, but who failed to complete their course or achieve the qualification. In the LFS, different control groups have been used. In specifications that consider individuals' highest qualification, in an equation with all qualifications included, the comparison group is the excluded qualification category, which is typically those with no qualifications. Other analyses have restricted the sample to holders of qualifications at two consecutive levels (say Level 3 and Level 2). The lower level then forms the excluded category, so that the wage differentials earned by those with the included Level 3 qualifications are estimated relative to those whose highest qualification is at Level 2.

[15] GCSEs are national examinations taken at the age of 16 at the end of compulsory full-time schooling. Key stage tests are again national tests taken by all pupils at age 7 (Key Stage 1), age 11 (Key Stage 2) and age 14 (Key Stage 3).

Appendix Tables

Table A1: Number of Observations in Treatment and Comparison Groups, by Highest Qualification, in Administrative Dataset (ILR)

	Males			Females		
	Treatment	Comp 1	Comp 2	Treatment	Comp 1	Comp 2
Vocational L4+	6,108	90,090	29,548	3,934	95,247	38,054
Advanced Appren.	31,134	74,261	30,233	16,357	51,806	24,400
NVQ 3	4,241	74,261	51,865	13,882	51,806	32,420
BTEC 3	21,545	74,261	51,865	16,335	51,806	32,420
Other Vocational L3	15,223	74,261	51,865	18,237	51,806	32,420
Intermediate Appren.	30,233	87,879	23,776	24,400	45,093	10,692
NVQ 2	17,903	87,879	23,776	13,922	45,093	10,692
BTEC 2	5,043	87,879	23,776	3,008	45,093	10,692
Other Vocational L2	28,919	87,879	23,776	15,490	45,093	34,401
NVQ 1	2,681	21,547	5,719	1,799	8,151	1,803
BTEC L1	433	21,547	5,719	301	8,151	1,803
Other Vocational L1	20,662	21,547	5,719	8,592	8,151	1,803

Table A2: Number of Observations in Treatment and Comparison Groups, by Highest Qualification, in Survey Dataset (LFS)

	Males			Females		
	Treatment	Comp 1	Comp 2	Treatment	Comp 1	Comp 2
Vocational L4+	8,274	21,421	13,255	5,918	22,698	13,521
Advanced Appren.	802	17,928	487	227	27,738	162
NVQ 3	4,379	17,928	5,068	2,940	27,738	7,974
BTEC 3	3,325	17,928	5,068	2,940	27,738	7,974
Other Vocational L3	5,551	17,928	5,068	1,750	27,738	7,974
Intermediate Appren.	487	13,277	715	162	17,970	1,684
NVQ 2	3,481	13,277	715	6,628	17,970	1,684
BTEC 2	260	13,277	715	298	17,970	1,684
Other Vocational L2	1,073	13,277	715	758	17,970	1,684
NVQ 1	353	9,010	8,489	513	7,117	6,479
BTEC L1	29	9,010	8,489	34	7,117	6,479
Other Vocational L1	320	9,010	8,489	1,120	7,117	6,479

Table A3: Prior Achievement at Key Stages 2 and 4, by Qualification (ILR)

Qualification	KS2 Maths test score	KS2 English test score	Number of GCSEs A*-C
Postgraduate qualification	68.3	73.8	9.7
First degree	64.4	69.3	8.7
Level 4 vocational	58.7	63.8	7.0
Advanced Apprenticeship	53.3	57.5	4.6
NVQ level 3	49.3	52.5	3.6
BTEC level 3	53.0	56.6	4.7
Other Level 3 vocational	54.7	51.8	5.3
GCE A/AS/A2 level	62.6	66.2	7.9
Intermediate Apprenticeship	48.0	50.0	2.3
NVQ level 2	44.4	45.4	1.6
GNVQ level 2	47.2	44.5	1.6
BTEC level 2	44.0	46.6	1.2
Other Level 2 vocational	45.1	46.8	1.7
5 GCSEs A-C	59.8	63.4	7.1
NVQ level 1	37.5	35.7	0.4
GNVQ level 1	38.7	35.3	0.3
BTEC level 1	32.6	29.7	0.3
Other level 1 vocational	39.9	39.6	0.6
Level 1 academic	45.8	45.6	1.0
Entry/Other Leve	30.7	28.6	0.0
KS4 record but n	36.9	36.7	0.1
Total	53.2	55.6	4.7

Source: Individualised Learner Record. Key Stage 2 test scores recorded in national tests at age 11. The number of GCSEs passed at grade C or above is the number of subject examinations passed at the end of lower secondary schooling at age 16.

Table A4: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Cumulative, Comparison Group 1, Males

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.139*** (0.005)	0.166*** (0.013)	0.048*** (0.006)	0.108*** (0.007)	0.062*** (0.006)	0.091*** (0.018)	-0.064*** (0.007)	0.029 (0.022)	-0.000 (0.012)	-0.117*** (0.018)	0.076 (0.060)	-0.049** (0.020)
2.LFS: separate equations	0.139*** (0.005)	0.175*** (0.013)	0.049*** (0.006)	0.107*** (0.007)	0.063*** (0.006)	0.093*** (0.018)	-0.064*** (0.007)	0.037* (0.022)	-0.004 (0.012)	-0.118*** (0.018)	0.076 (0.060)	-0.049** (0.020)
3. LFS: add time elapse	0.150*** (0.005)	0.187*** (0.013)	0.117*** (0.006)	0.104*** (0.007)	0.054*** (0.006)	0.113*** (0.018)	0.010 (0.009)	0.056** (0.022)	0.021* (0.012)	-0.117*** (0.018)	0.076 (0.060)	-0.050*** (0.021)
4.LFS: match ILR controls	0.203*** (0.006)	0.200*** (0.014)	0.225*** (0.007)	0.117*** (0.008)	0.081*** (0.007)	0.127*** (0.019)	0.124*** (0.008)	0.027 (0.024)	0.084*** (0.013)	-0.159*** (0.020)	0.047 (0.081)	-0.047** (0.022)
5.LFS: restrict age	0.120*** (0.015)	0.235*** (0.019)	0.123*** (0.014)	0.034** (0.016)	0.161*** (0.018)	0.183*** (0.032)	0.047*** (0.017)	0.029 (0.030)	0.175*** (0.032)	-0.156*** (0.050)	-0.182 (0.210)	0.017 (0.087)
6.LFS weekly earnings	0.127*** (0.020)	0.360*** (0.024)	0.202*** (0.017)	0.044** (0.021)	0.230*** (0.021)	0.269*** (0.039)	0.066*** (0.024)	0.036 (0.055)	0.219*** (0.046)	-0.284*** (0.113)	-0.346 (0.473)	-0.082 (0.140)
7.ILR: control group lower level	0.348*** (0.0098	0.400*** (0.005)	0.197*** (0.010)	0.072*** (0.005)	0.126*** (0.010)	0.220*** (0.005)	0.137*** (0.006)	-0.002 (0.009)	0.095*** (0.007)	-0.018 (0.026)	-0.136** (0.055)	-0.040*** (0.012)
8.ILR: typical (non-achievers)	0.131*** (0.013)	0.225*** (0.007)	0.071*** (0.016)	0.087*** (0.007)	0.073*** (0.015)	0.162*** (0.006)	0.128*** (0.010)	0.030* (0.018)	0.114*** (0.011)	0.088*** (0.020)	-0.005 (0.055)	0.069*** (0.011)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. * significant at 10%. The dependent variable is the natural logarithm of wages/earnings. See Tables A1 and A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year. The 'typical' ILR specification controls for ethnicity and time elapsed since highest qualification was obtained.

Table A5: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Cumulative, Comparison Group 1, Females

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.127*** (0.006)	-0.001 (0.021)	-0.016** (0.004)	0.087*** (0.007)	0.037*** (0.009)	-0.075*** (0.024)	-0.066*** (0.005)	0.028 (0.020)	-0.003 (0.012)	-0.121*** (0.014)	0.026 (0.069)	0.047*** (0.011)
2.LFS: separate equations	0.127*** (0.006)	0.003 (0.021)	-0.015*** (0.004)	0.089*** (0.007)	0.036*** (0.009)	-0.060*** (0.025)	-0.066*** (0.005)	0.035* (0.020)	-0.008 (0.012)	-0.121*** (0.014)	0.023 (0.070)	0.049*** (0.011)
3. LFS: add time elapse	0.132*** (0.006)	0.023 (0.021)	0.062*** (0.004)	0.086*** (0.007)	0.031*** (0.009)	-0.046* (0.025)	0.005 (0.007)	0.050** (0.020)	0.010 (0.012)	-0.103*** (0.014)	0.004 (0.070)	0.006 (0.012)
4.LFS: match ILR controls	0.165*** (0.006)	0.005 (0.023)	0.124*** (0.005)	0.105*** (0.008)	0.049*** (0.010)	-0.092*** (0.026)	0.047*** (0.006)	0.058*** (0.021)	0.062*** (0.013)	-0.143*** (0.015)	-0.042 (0.073)	-0.023** (0.012)
5.LFS: restrict age	0.083*** (0.015)	0.004 (0.032)	0.007 (0.011)	0.044*** (0.016)	0.111*** (0.019)	-0.022 (0.035)	0.005 (0.015)	0.063 (0.045)	-0.015 (0.067)	-0.190*** (0.043)	-0.006 (0.318)	-0.072 (0.065)
6.LFS weekly earnings	0.108*** (0.025)	0.010 (0.066)	0.008 (0.020)	0.015 (0.028)	0.193*** (0.033)	-0.049 (0.073)	-0.007 (0.029)	0.119 (0.080)	0.105 (0.112)	-0.563*** (0.083)	-0.043 (0.321)	-0.261* (0.153)
7.ILR: control group lower level	0.288*** (0.010)	0.226*** (0.007)	0.115*** (0.007)	0.213*** (0.007)	0.132*** (0.008)	0.104*** (0.007)	0.057*** (0.008)	0.084*** (0.014)	0.037*** (0.012)	-0.026 (0.035)	-0.074 (0.071)	-0.022 (0.021)
8.ILR: typical (non-achievers)	0.162*** (0.017)	0.150*** (0.010)	0.091*** (0.013)	0.125*** (0.010)	0.093*** (0.014)	0.138*** (0.009)	0.152*** (0.012)	0.076*** (0.028)	0.097*** (0.021)	0.133*** (0.031)	0.132 (0.086)	0.096*** (0.021)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. * significant at 10%. The dependent variable is the natural logarithm of wages/earnings. See Tables A1 and A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year. The 'typical' ILR specification controls for ethnicity and time elapsed since highest qualification was obtained.

Table A6: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Cumulative, Comparison Group 2, Males

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.176***	0.140***	0.157***	0.221***	0.182***	0.193***	0.045***	0.114***	0.125***	-0.119***	0.072	-0.045**
	(0.006)	(0.021)	(0.007)	(0.008)	(0.007)	(0.022)	(0.014)	(0.025)	(0.017)	(0.018)	(0.061)	(0.020)
2.LFS: separate equations	0.176***	0.131***	0.155***	0.221***	0.194***	0.209***	0.043***	0.132***	0.131***	-0.119***	0.071	-0.046**
	(0.006)	(0.022)	(0.007)	(0.008)	(0.007)	(0.024)	(0.015)	(0.028)	(0.018)	(0.018)	(0.061)	(0.020)
3. LFS: add time elapse	0.179***	0.132***	0.164***	0.176***	0.164***	0.185***	0.068***	0.120***	0.115***	-0.119***	0.071	-0.046**
	(0.006)	(0.022)	(0.007)	(0.009)	(0.009)	(0.024)	(0.015)	(0.028)	(0.018)	(0.018)	(0.061)	(0.021)
4.LFS: match ILR controls	0.202***	0.127***	0.174***	0.111***	0.189***	0.135***	0.074***	0.002	0.137***	-0.168***	0.040	-0.049**
	(0.006)	(0.024)	(0.008)	(0.010)	(0.009)	(0.023)	(0.016)	(0.028)	(0.019)	(0.020)	(0.082)	(0.022)
5.LFS: restrict age	0.140***	0.142***	0.163***	0.073***	0.203***	0.217**	0.091**	0.073	0.210***	-0.169***	-0.193	0.004
	(0.016)	(0.034)	(0.016)	(0.019)	(0.020)	(0.050)	(0.043)	(0.049)	(0.050)	(0.051)	(0.209)	(0.087)
6.LFS weekly earnings	0.102***	0.195***	0.241***	0.080***	0.270***	0.385***	0.194**	0.172*	0.330***	-0.306***	-0.367	-0.102
	(0.021)	(0.042)	(0.022)	(0.025)	(0.025)	(0.091)	(0.085)	(0.099)	(0.093)	(0.113)	(0.473)	(0.140)
7.ILR: control group lower level	0.371***	0.260***	0.212***	0.091***	0.152***	0.290***	0.203***	0.064***	0.162***	0.046***	-0.085	0.021***
	(0.009)	(0.005)	(0.010)	(0.005)	(0.010)	(0.006)	(0.007)	(0.010)	(0.008)	(0.017)	(0.056)	(0.013)
8.ILR: typical (non-achievers)	0.131***	0.225***	0.071***	0.087***	0.073***	0.162***	0.128***	0.030*	0.114***	0.088***	-0.005	0.069***
	(0.013)	(0.007)	(0.016)	(0.007)	(0.015)	(0.006)	(0.010)	(0.018)	(0.011)	(0.020)	(0.055)	(0.011)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. * significant at 10%. The dependent variable is the natural logarithm of wages/earnings. See Tables A1 and A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year. The 'typical' ILR specification controls for ethnicity and time elapsed since highest qualification was obtained.

Table A7: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Cumulative, Comparison Group 2, Females

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.181***	0.116***	0.093***	0.190***	0.148***	-0.050*	-0.018*	0.070***	0.060***	-0.118***	0.032	0.049***
	(0.006)	(0.032)	(0.005)	(0.008)	(0.009)	(0.026)	(0.009)	(0.022)	(0.014)	(0.015)	(0.070)	(0.011)
2.LFS: separate equations	0.181***	0.131***	0.093***	0.191***	0.149***	0.041	-0.019**	0.135***	0.079***	-0.118***	0.030	0.051***
	(0.006)	(0.033)	(0.005)	(0.008)	(0.009)	(0.032)	(0.009)	(0.023)	(0.015)	(0.015)	(0.070)	(0.011)
3. LFS: add time elapse	0.169***	0.132***	0.104***	0.136***	0.105***	0.035	0.060***	0.121***	0.067***	-0.103***	0.003	0.002
	(0.006)	(0.034)	(0.005)	(0.008)	(0.010)	(0.032)	(0.011)	(0.023)	(0.015)	(0.015)	(0.070)	(0.012)
4.LFS: match ILR controls	0.182***	0.145***	0.118***	0.113***	0.108***	-0.105***	0.034***	0.053**	0.067***	-0.149***	-0.050	-0.033**
	(0.006)	(0.035)	(0.005)	(0.009)	(0.011)	(0.028)	(0.011)	(0.022)	(0.015)	(0.015)	(0.073)	(0.013)
5.LFS: restrict age	0.142***	0.068	0.070***	0.107***	0.178***	0.078	0.100**	0.147**	0.088	-0.191***	-0.009	-0.072
	(0.015)	(0.046)	(0.014)	(0.018)	(0.021)	(0.051)	(0.042)	(0.060)	(0.084)	(0.044)	(0.319)	(0.066)
6.LFS weekly earnings	0.193***	0.158	0.147***	0.156***	0.343***	0.184*	0.218***	0.344***	0.350**	-0.581***	-0.069	-0.275*
	(0.026)	(0.099)	(0.025)	(0.032)	(0.036)	(0.100)	(0.075)	(0.105)	(0.137)	(0.084)	(0.320)	(0.155)
7.ILR: control group lower level	0.361***	0.173***	0.161***	0.258***	0.182***	0.187***	0.141***	0.182***	0.119***	-0.003	-0.065	0.002
	(0.011)	(0.008)	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.015)	(0.013)	(0.038)	(0.073)	(0.025)
8.ILR: typical (non-achievers)	0.162***	0.150***	0.091***	0.125***	0.093***	0.138***	0.152***	0.076***	0.097***	0.133***	0.132	0.096***
	(0.017)	(0.010)	(0.013)	(0.010)	(0.014)	(0.009)	(0.012)	(0.028)	(0.021)	(0.031)	(0.086)	(0.021)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. * significant at 10%. The dependent variable is the natural logarithm of wages/earnings. See Tables A1 and A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year. The 'typical' ILR specification controls for ethnicity and time elapsed since highest qualification was obtained.

Table A8: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Singular, Comparison Group 1, Males

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.139*** (0.005)	0.166*** (0.013)	0.048*** (0.006)	0.108*** (0.007)	0.062*** (0.006)	0.091*** (0.018)	-0.064*** (0.007)	0.029 (0.022)	-0.000 (0.012)	-0.117*** (0.018)	0.076 (0.060)	-0.049** (0.020)
2.LFS: separate equations	0.139*** (0.005)	0.175*** (0.013)	0.049*** (0.006)	0.107*** (0.007)	0.063*** (0.006)	0.093*** (0.018)	-0.064*** (0.007)	0.037* (0.022)	-0.004 (0.012)	-0.118*** (0.018)	0.076 (0.060)	-0.049** (0.020)
3. LFS: add time elapse	0.150*** (0.005)	0.179*** (0.013)	0.107*** (0.006)	0.105*** (0.007)	0.049*** (0.006)	0.110*** (0.018)	0.005 (0.008)	0.053** (0.022)	0.023* (0.012)	-0.116*** (0.018)	0.076 (0.060)	-0.050** (0.020)
4.LFS: match ILR controls	0.203*** (0.006)	0.186*** (0.013)	0.207*** (0.007)	0.122*** (0.008)	0.080*** (0.007)	0.123*** (0.019)	0.114*** (0.008)	0.015 (0.024)	0.084*** (0.013)	-0.159*** (0.020)	0.048 (0.081)	-0.048** (0.022)
5.LFS: restrict age	0.096*** (0.015)	0.190*** (0.018)	0.066*** (0.013)	0.005 (0.015)	0.079*** (0.017)	0.140*** (0.030)	-0.004 (0.015)	0.027 (0.031)	0.104*** (0.031)	-0.118*** (0.049)	0.107 (0.068)	0.035 (0.075)
6.LFS weekly earnings	0.125*** (0.005)	0.249*** (0.014)	0.079*** (0.007)	0.120*** (0.008)	0.077*** (0.006)	0.122*** (0.019)	-0.049*** (0.007)	0.026 (0.029)	0.014 (0.013)	-0.166*** (0.023)	0.115* (0.061)	-0.062*** (0.022)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. * significant at 10%. The dependent variable is the natural logarithm of wages/earnings. See Table A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year.

Table A9: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Singular, Comparison Group 1, Females

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.127*** (0.006)	-0.001 (0.021)	-0.016*** (0.004)	0.087*** (0.007)	0.037*** (0.009)	-0.075*** (0.024)	-0.066*** (0.005)	0.028 (0.020)	-0.003 (0.012)	-0.121*** (0.014)	0.026 (0.069)	0.047*** (0.011)
2.LFS: separate equations	0.127*** (0.006)	0.003 (0.021)	-0.015*** (0.004)	0.089*** (0.007)	0.036*** (0.009)	-0.060*** (0.025)	-0.066*** (0.005)	0.035* (0.020)	-0.008 (0.012)	-0.121*** (0.014)	0.023 (0.070)	0.049*** (0.011)
3. LFS: add time elapse	0.132*** (0.006)	0.026 (0.021)	0.059*** (0.004)	0.088*** (0.007)	0.031*** (0.009)	-0.043* (0.024)	0.004 (0.006)	0.056*** (0.020)	0.021* (0.012)	-0.103*** (0.014)	0.007 (0.070)	0.004 (0.012)
4.LFS: match ILR controls	0.165*** (0.006)	0.001 (0.023)	0.122*** (0.005)	0.104*** (0.008)	0.049*** (0.010)	-0.094*** (0.026)	0.045*** (0.006)	0.056** (0.021)	0.062*** (0.013)	-0.145*** (0.015)	-0.041 (0.073)	-0.024** (0.012)
5.LFS: restrict age	0.060*** (0.015)	-0.001 (0.030)	-0.021** (0.010)	0.017 (0.015)	0.052*** (0.019)	-0.019 (0.032)	-0.024* (0.013)	0.020 (0.042)	-0.022 (0.062)	-0.079* (0.048)	0.167 (0.399)	-0.057 (0.052)
6.LFS weekly earnings	0.124*** (0.008)	0.112*** (0.027)	0.030*** (0.006)	0.118*** (0.010)	0.064*** (0.012)	0.019 (0.035)	-0.019*** (0.007)	0.045* (0.027)	-0.007 (0.018)	-0.112*** (0.022)	0.140 (0.090)	0.066*** (0.016)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. * significant at 10%. The dependent variable is the natural logarithm of wages/earnings. See Table A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.
The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year.

Table A10: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Singular, Comparison Group 2, Males

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.176*** (0.006)	0.140*** (0.021)	0.157*** (0.007)	0.221*** (0.008)	0.182*** (0.007)	0.193*** (0.022)	0.045*** (0.014)	0.114*** (0.025)	0.125*** (0.017)	-0.119*** (0.018)	0.072 (0.061)	-0.045** (0.020)
2.LFS: separate equations	0.176*** (0.006)	0.131*** (0.022)	0.155*** (0.007)	0.221*** (0.008)	0.194*** (0.007)	0.209*** (0.024)	0.043*** (0.015)	0.132*** (0.028)	0.131*** (0.018)	-0.119*** (0.018)	0.071 (0.061)	-0.046** (0.020)
3. LFS: add time elapse	0.179*** (0.006)	0.148*** (0.021)	0.163*** (0.007)	0.188*** (0.009)	0.138*** (0.008)	0.175*** (0.022)	0.064*** (0.015)	0.112*** (0.025)	0.102*** (0.017)	-0.118*** (0.018)	0.072 (0.061)	-0.046** (0.021)
4.LFS: match ILR controls	0.202*** (0.006)	0.112*** (0.023)	0.174*** (0.008)	0.143*** (0.010)	0.136*** (0.009)	0.140*** (0.023)	0.076*** (0.016)	-0.005 (0.028)	0.116*** (0.019)	-0.167*** (0.020)	0.041 (0.082)	-0.050** (0.022)
5.LFS: restrict age	0.139*** (0.016)	0.140*** (0.034)	0.134*** (0.016)	0.074*** (0.018)	0.149*** (0.019)	0.238*** (0.049)	0.095** (0.040)	0.129*** (0.049)	0.200*** (0.048)	-0.128*** (0.049)	0.072 (0.081)	0.019 (0.075)
6.LFS weekly earnings	0.153*** (0.006)	0.168*** (0.023)	0.165*** (0.008)	0.199*** (0.009)	0.181*** (0.008)	0.213*** (0.024)	0.053*** (0.017)	0.088*** (0.033)	0.133*** (0.020)	-0.175*** (0.023)	0.100 (0.063)	-0.062*** (0.022)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. The dependent variable is the natural logarithm of wages/earnings. See Table A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.

The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year.

Table A11: Estimated Wage Differentials for Vocational Qualifications: Specification Changes Singular, Comparison Group 2, Females

	Level 4	Advan appren	NVQ3	BTEC3	Other L3	Inter appren	NVQ2	BTEC2	Other L2	NVQ1	BTEC1	Other L1
1.LFS: typical spec	0.181*** (0.006)	0.116*** (0.032)	0.093*** (0.005)	0.190*** (0.008)	0.148*** (0.009)	-0.050* (0.026)	-0.018* (0.009)	0.070*** (0.022)	0.060*** (0.014)	-0.118*** (0.015)	0.032 (0.070)	0.047*** (0.011)
2.LFS: separate equations	0.181*** (0.006)	0.131*** (0.033)	0.093*** (0.005)	0.191*** (0.008)	0.149*** (0.009)	0.041 (0.032)	-0.019** (0.009)	0.135*** (0.023)	0.079*** (0.015)	-0.118*** (0.015)	0.030 (0.070)	0.049*** (0.011)
3. LFS: add time elapse	0.169*** (0.006)	0.130*** (0.032)	0.102*** (0.005)	0.142*** (0.008)	0.087*** (0.010)	-0.003 (0.026)	0.054*** (0.010)	0.097*** (0.021)	0.060*** (0.014)	-0.104*** (0.015)	0.007 (0.070)	0.004 (0.012)
4.LFS: match ILR controls	0.182*** (0.006)	0.142*** (0.034)	0.119*** (0.005)	0.124*** (0.008)	0.078*** (0.010)	-0.103*** (0.027)	0.033*** (0.011)	0.052** (0.022)	0.068*** (0.015)	-0.151*** (0.015)	-0.049 (0.073)	-0.024** (0.012)
5.LFS: restrict age	0.116*** (0.016)	0.069 (0.044)	0.057*** (0.013)	0.096*** (0.018)	0.128*** (0.021)	0.102* (0.052)	0.096** (0.043)	0.149** (0.058)	0.111 (0.073)	-0.061 (0.050)	0.157 (0.395)	-0.057 (0.052)
6.LFS weekly earnings	0.177*** (0.008)	0.128*** (0.043)	0.118*** (0.007)	0.189*** (0.011)	0.154*** (0.013)	0.024 (0.037)	0.021 (0.014)	0.078*** (0.030)	0.053** (0.021)	-0.113*** (0.022)	0.139 (0.090)	0.066*** (0.016)

Standard errors in parentheses. *** significant at 1%, ** significant at 5%. The dependent variable is the natural logarithm of wages/earnings. See Table A2 for information on the numbers of observations in the treatment and comparison groups for each qualification.

The 'typical' LFS specification controls for age, age squared, ethnicity, full time, public sector, region, year.

CVER PUBLICATIONS

Research Papers

Peer Effects and Social Influence in Post-16 Educational Choice, *CVER Research Paper 008*, November 2017
Sophie Hedges and Stefan Speckesser,

The earnings differentials associated with vocational education and training using the Longitudinal Education Outcomes data, *CVER Research Paper 008,* October 2017
Gavan Conlon, Sophie Hedges and Pietro Patrignani

The longer the better? The impact of the 2012 apprenticeship reform in England on achievement and other outcomes, *CVER Research Paper 006,* September 2017
Vahé Nafilyan and Stefan Speckesser

Post 16 remedial policies: a literature review, *CVER Research Paper 005*, April 2017
Clémentine Van Effenterre

Young people in low level vocational education: characteristics, trajectories and labour market outcomes
CVER Research Paper 004, March 2017
Sophie Hedges, Vahé Nafilyan, Stefan Speckesser and Augustin de Coulon

Vocational vs. General Education and Employment over the Life-Cycle: New Evidence from PIAAC
CVER Research Paper 003, November 2016
Franziska Hampf and Ludger Woessmann

Labour Market Returns to Vocational Qualifications in the Labour Force Survey, *CVER Research Paper 002,* October 2016
Steven McIntosh and Damon Morris

Post-Compulsory Education in England: Choices and Implications, *CVER Research Paper 001,* July 2016
Claudia Hupkau, Sandra McNally, Jenifer Ruiz-Valenzuela and Guglielmo Ventura

Briefing Notes

An analysis of the duration and achievement of apprenticeships in England, *CVER Briefing Note 004,* September 2017
Matthew Bursnall, Vahé Nafilyan, Stefan Speckesser

The incidence of publicly funded training in England, *CVER Briefing Note 003*, March 2017
Gavan Conlon, Sophie Hedges, Daniel Herr and Pietro Patrignani

The Decision to Undertake an Apprenticeship: A Case Study, *CVER Briefing Note 002,* March 2017
Steven McIntosh

Further Education in England: Learners and Institutions, *CVER Briefing Note 001,* February 2017
Claudia Hupkau and Guglielmo Ventura