



Department
for Business
Innovation & Skills

BIS RESEARCH PAPER NUMBER 105

Review of the Economic Benefits of
Training and Qualifications, as shown
by Research based on Cross-
Sectional and Administrative Data

FEBRUARY 2013

RESEARCH

A final report for the Department for Business, Innovation and Skills

Cambridge Econometrics

Covent Garden

Cambridge

CB1 2HT

UK

Tel 01223 533100 (+44 1223 533100)

Fax 01223 533101 (+44 1223 533101)

Email mmg@camecon.com

Web www.camecon.com

Warwick Institute for Employment Research

University of Warwick

Coventry

CV4 7AL

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Business, Innovation and Skills.

Department for Business, Innovation and Skills

1 Victoria Street

London SW1H 0ET

www.gov.uk/bis

Research paper number 105

February 2013

Contents

Contents	3
Executive Summary	5
The Research Papers	6
Estimating the Benefits of Training and Education	8
How does the research add to the existing evidence base?	9
Where is it better to use one set of results over another?	11
What are the implications of the report’s findings within the context of BIS’ NPV model?	13
Promising areas and methods for future research	15
1 Introduction - Aims and Objectives.....	19
2 Summary of the Papers Reviewed	21
2.1 Introduction.....	21
2.2 Summary of the Papers Reviewed	21
3 Theoretical Perspectives on the Returns to Training and Qualifications.....	25
3.1 Introduction.....	25
3.2 Estimating the Benefits of Training and Education.....	25
3.3 Limitations of Rates of Return Analysis.....	28
3.4 What are the Benefits of Training and Qualifications?	32
3.5 How Well do the RPs Represent the Education System?	35
4 Key Datasets and Statistical Methods – Strengths and Weaknesses.....	39
4.1 Introduction.....	39
4.2 Data Requirements	39
4.3 Data sources	40
4.4 Properties of the Data	42

4.5	Linking, Missing Data and Selection Bias	42
4.6	Suitability of the Indicators	46
4.7	Strengths and Weaknesses of the Data.....	47
4.8	Summary of the Methodological Approach	49
4.9	Specification of the Counterfactual.....	49
4.10	Suitability of the Dependent and Explanatory Variables	51
4.11	Significance of the Statistical Inference.....	54
4.12	Strengths and Weaknesses of the Econometric and Statistical Techniques	54
5	Key Findings.....	56
5.1	Introduction.....	56
5.2	Returns by Funding Stream, Level and Subject.....	59
	Sector.....	61
5.3	The Socio-Economic, Demographic and Educational Characteristics of Individuals	64
5.4	Evidence on the Long-Term Returns and Their Persistence	68
5.5	In Which Circumstances is it Preferable to Use Which Evidence?	73
6	Conclusions and Recommendations	74
6.1	Introduction.....	74
6.2	How does the research add to the existing evidence base?	74
6.3	Where is it better to use one set of results over another?	76
6.4	What are the implications of the report's findings within the context of BIS' NPV model? 78	
6.5	Promising areas and methods for future research	80
7	References.....	84
8	Glossary.....	86
Appendix A:	The Review Questionnaire.....	87

Executive Summary

Having set out in [The Plan for Growth](#) its ambition “To create a more educated workforce that is the most flexible in Europe”, the government has announced policy measures to further promote vocational education and training. To better incentivise training and assess employer demand, there is a need to understand the returns to various forms of vocational education and training.

There is a large and growing body of evidence on the relative returns to obtaining qualifications

There is now a large body of evidence, collated over the past decade, on the relative returns to obtaining qualifications in further¹ and higher education² based upon cross-section analysis (usually the Labour Force Survey) and longitudinal surveys. Whilst these studies employ robust econometric techniques, they are constrained by sample sizes and the problems posed by sample attrition.

Two recent studies³, however, have made use of a recently-developed experimental ‘matched’ database that combines administrative data from the Individualised Learner Record (ILR) with data from Her Majesty’s Revenue and Customs (HMRC) and data from the Department for Work and Pensions (DWP). This provides a rich source of information that allows some of the drawbacks of earlier analysis to be overcome.

This study was commissioned by BIS to update the evidence base on the economic returns to vocational qualifications

This study was commissioned by BIS to update the evidence base on the economic returns to vocational qualifications (VQs) in England by reviewing the two research papers

¹ For example see: Dearden, L., McIntosh, S., Myck, M. and Vignoles, A. (2002) ‘The Returns to Academic and Vocational Qualifications in Britain’, *Bulletin of Economic Research*, Vol. 54, p. 249- 274; . 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) ‘A Cost-Benefit Analysis of Apprenticeships and Other Vocational Qualifications’. DfES Research Report 834; McIntosh, S. (2010) “The Economic Value of Intermediate Vocational Education and Qualifications”, UK Commission for Employment and Skills, Wath-upon-Deerne.

² For example: Purcell, K., Elias, P., and Davies R., (2005) *The Class of ‘99: A study of the early labour market experience of recent graduates*, University of Warwick, October 2005; London Economics (2011) *The Returns to Higher Education Qualifications*, BIS Research Paper No.45; O’ Leary N., and Sloane, P., (2005), ‘The returns to a university education in great Britain’, *National institute Economic Review*, July 2005, vol. 193, no. 1, 75-89.

³ BIS Research Papers 47: London Economics (June 2011) ‘The Long Term Effect of Vocational Qualifications on Labour Market Outcomes’; and 48: Frontier Economics / IFS (June 2011) ‘Reporting on Employment and Earnings Using Experimental Matched Data’;

that use the matched administrative dataset, together with another recently-published research paper that applied well-established techniques to recent survey data.⁴

The different types of approach used in these recent analyses were reviewed so as to:

- provide an overall assessment of the approaches used;
- interpret the findings in the context of the extensive past literature in this area;⁵
- advise in which circumstances it is recommended to use estimates derived from the different analyses;
- comment on the implications of the findings within the context of the BIS 'Economic Impact of FE Model'; and so
- summarise what the recent research adds to the existing evidence base.

The review considered:

- the wage and employment benefits which learners gain from different qualifications;
- the value of qualifications held in the working age population; and the extent to which the returns might vary for different individuals and in different circumstances.

The Research Papers

RP53: Returns to Intermediate and Low Level Vocational Qualifications

In this paper, London Economics (September 2011) applied well-established techniques to recent survey data to derive updated estimates of returns to intermediate and low level qualifications. This paper used Labour Force Survey (LFS) data and estimated OLS cross-sectional regressions for each year to estimate marginal and average earnings returns. For the analysis of employment returns, a probit model was estimated for the probability that an individual is in employment versus unemployment or inactivity. In addition, a longitudinal dataset from the British Cohort Study (BCS) was used to estimate employment and earnings returns: separate cross-sectional regressions are estimated for each of the three waves as well as a fixed effects estimation using all three waves. The use of the BCS provided a counterfactual 'before attainment'.

⁴ BIS Research Paper 53: London Economics (September 2011) 'Returns to Intermediate and Low Level Vocational Qualifications'.

⁵ The past literature will be represented by the extensive review by McIntosh, S. (2009) 'The Economic Value of Intermediate Vocational Education and Qualifications', UKCES.

The authors' findings are consistent overall with previous studies which adopted similar approaches: there continue to be positive employment and earnings returns for the vast majority of qualifications gained in adulthood, but the route of acquisition of qualification has become less important in determining the earnings returns (previous estimates, and theory, suggested that the workplace route gave greater returns). The report notes the significant variation in returns by sector and occupation of employment.

RP48: Reporting on Employment and Earnings Using Experimental Matched Data

The development of a matched administrative dataset provided the basis for the innovative research published in RP48 and RP47:

Frontier Economics/IFS (June 2011) undertook analysis of matched administrative data comprising: learner information from the Individualised Learner Record (ILR); information on benefit recipients held by the Department for Work and Pensions (DWP); and information on employment status and earnings in work held by Her Majesty's Revenue and Customs (HMRC). The matched data were used to assess the labour market performance of learners both before and after learning, and disaggregated by qualification and subject area. The majority of the analysis is cross-tabular in nature involving a comparison of various descriptive statistics and sample proportions relating to earnings and employment outcomes for groups of learners in the 12 months pre- and post-qualification attainment. For the Further Education learning stream only, impact analysis is carried out using fixed effects estimation of earnings, benefit, and employment outcomes.

The authors conclude that, in general, there are improvements in the economic performance of learners following achievement. The innovative contribution made by this paper is the exploration of the matched dataset and the opportunities that it presents; for example, in the analysis, this paper presents findings by level of qualification and by *subject area*.

RP47: The Long-Term Effect of Vocational Qualifications on Labour Market Outcomes

In this paper, London Economics (June 2011) also utilised the matched administrative dataset (which while from the same original sources differed somewhat in the data used in the final analyses), and various econometric techniques and *alternative counterfactuals* were applied to estimate *long-term* earnings and employment effects.

A series of annual cross section OLS regressions are estimated. There are two main aspects to the research. A first stage analysis is conducted of the determinants of qualification completion. This is specified as a probit and ordered probit model and estimated by learning aim rather than by individual learner, over the entire sample of qualification completers and non-completers for the years 2002/3 to 2005/6. The significant explanatory variables from these regressions are used to inform the main analysis of earnings, employment and benefit dependency outcomes. This second stage of the analysis estimates a cross-section OLS regression for each year post-attainment of the qualification up to seven years afterwards.

The authors conclude that, no matter which counterfactual is used, there are strong positive and persistent returns to completing a qualification at a given level. Having said

that, the magnitude (and in some cases the sign) are sensitive to the choice of counterfactual.

Estimating the Benefits of Training and Education

Obtaining a vocational qualification, or any qualification for that matter, can be regarded as an investment by: (i) the person studying towards the qualification (in terms of the amount of effort and time required to obtain the qualification); and (ii) the parties which meet the costs of the individual working towards the qualification (typically some combination of the State, employers, and individual learners). Typically the returns to training are measured with respect to increases in wages and the probability of remaining in employment, though there are a range of other benefits typically listed under the category of the 'wider benefits of training' which relate to factors such as the individual's well-being, the intrinsic rewards of learning, reduction in benefit uptake/dependency etc

The overall aim of estimating the returns to vocational education and training therefore requires a comparison of people's transitions through the overall education system incorporating both the compulsory and post-compulsory elements. The determinants of the various transitions people make need to be determined. At each stage of the process – i.e. pre-school, compulsory schooling, further education and higher education – choices or selections made with respect to what education or learning institution is attended, subjects studied and to what level, and the type of qualifications, will be shaped, inter alia, by a range of socio-economic and demographic factors and by ability.

In a perfect world there is a need to control for all of these factors if the value-added obtained from any vocational qualifications is to be determined. In this way, it will be possible to determine, given the current structure of education and learning provision, the extent to which vocational qualifications of different types, in different subjects and at different levels give relatively high or low returns – compared with some comparator qualification(s) - to people with differing characteristics who have taken, for one reason or another, differing pathways through the compulsory and post-compulsory education and learning system.

Therefore, when seeking to measure the economic returns to vocational qualifications, the critical issue to be addressed is the extent to which, other things being equal, a given qualification (with respect to its subject and level) gives a better or worse return (typically measured with respect to employment and earnings) than a comparable qualification, or no qualification at all. Ideally, the analysis would also recognise that different qualifications are appropriate for different circumstances or learners (some are more relevant to younger college-based learners, while others are more relevant to older work-based learners) – so it is not necessarily the case that a better overall average return makes that qualification better for all circumstances. Because other things are seldom equal – and due to the absence of an experiment to test the critical issue identified above - researchers have had to be inventive in the use of survey and administrative data in order to control for a range of variables which may confound any analysis. Typically, the choice of course leading to a qualification which an individual may choose to study is determined by a range of factors other than ability. And a range of factors other than ability may determine whether a person successfully completes a course and their grade of achievement should they complete. In practice, researchers are trying to control for a range of socio-demographic,

educational, and ability factors, with imperfect data, in order to measure the returns which can be safely regarded as being derived from possession of a given qualification.

From the evidence presented in McIntosh (2009), and summarised later in this report⁶, it is apparent that the findings derived from either longitudinal or pooled cross-sectional analyses struggle to control for these type of effects. In short, like is generally not being compared with like. This simply reflects the complexity of the issue being addressed and the consequent analytical challenges. But, as the following report demonstrates, recent studies funded by BIS have been able to make substantial progress in tackling some of these issues.

How does the research add to the existing evidence base?

The findings of recent analysis of the LFS/BCS (RP53) are broadly consistent with previous studies

The existing evidence base on the returns to vocational qualifications draws principally on cross-section analysis of the Labour Force Survey (LFS) and longitudinal surveys such as the British Cohort Survey (BCS). RP53 applies these well-established techniques to recent survey data to derive updated estimates of returns to intermediate and low level qualifications, and, because the methods used are similar to those previously used, the results are the most readily comparable, of the three papers, with earlier work. The differences between results from RP53 and those of other studies (e.g. McIntosh 2007) are partially attributed to the use of hourly wages rather than weekly wages. The findings of RP53 are, however, broadly consistent with previous studies that applied similar analysis:

- There continue to be positive employment and earnings returns for the vast majority of qualifications gained in adulthood.
- The returns to vocational qualifications are higher at Level 3 than Level 2, and Apprenticeships have the highest overall wage and employment returns.
- Typically, the wage returns to vocational qualifications are found to be higher for men than women.
- Returns are higher for younger learners (aged under 25 years).
- There is significant variation in returns by sector and occupation of employment: for example, there are higher returns to qualifications found in lower-skilled occupations (e.g. men working in skilled trades and machine operative occupations, and for females working in personal services).

⁶ See Section 3.4.

- However, the route of acquisition of qualification has become much less important in determining the earnings returns (previous estimates, and theory, suggested that the workplace route gave greater returns).

The use of cross-sectional analysis of the LFS does not really allow for analysis in any detail beyond Level of attainment and qualification type (as recalled by the respondent) and even the relatively large sample sizes in the LFS soon become quite small when disaggregated by type of qualification and the year in which it was attained.

Because they rely on survey responses, the LFS and BCS are both subject to errors arising from the fact that the individual may not report correctly the qualifications that they are working towards or have attained.

The BCS longitudinal survey dataset suffers from the problems of small sample sizes and sample attrition over time.

The matched administrative dataset provides a rich source of information that allows some of the drawbacks of earlier analysis to be overcome

The other two papers reviewed (RP47 and RP48), made use of a recently-developed experimental 'matched' database that combines administrative data from the Individualised Learner Record (ILR) with data from Her Majesty's Revenue and Customs (HMRC) and data from the Department for Work and Pensions (DWP).

Where these linked databases add value is in the considerable amount of detail they provide about the type of FE in which the individual has engaged and the possibility via the HMRC and DWP databases of assessing the extent to which gaining a qualification at a given level via one route or another in a particular subject provides some relative long-term benefit to the individual. There do though remain a number of methodological and data issues to be overcome to improve the analyses:

- the data do not provide a complete record of the individual's path through education (for instance, there are no data relating to the individual's entry to and exit from higher education);
- the definition of an appropriate counterfactual proves difficult to define and construct;
- the research papers do not report fully on the potential biases introduced via the data matching and cleaning processes, and the consequent implications for the findings;
- conditions in the external labour market are not always controlled for.

It is encouraging that the results of the analysis using the new linked datasets do not, in general, contradict the existing evidence base, finding that higher levels of attainment are generally associated with higher returns in terms of wages and employment.

However, the magnitude of those returns can vary considerably between studies, with, for example, all RPs confirming earlier findings that there are positive wage returns associated with completion of Apprenticeship training, but the magnitude of the estimates differing across studies and even within studies depending on model specification or counterfactual used.

Table 0.1 summarises the relative strengths and weaknesses of the analysis using the two alternative datasets: cross-section and longitudinal survey data; versus the matched administrative dataset.

Table 0.1: Strengths and Weaknesses of the Alternative Datasets

	Survey data	Matched administrative data
Sample size	Small	Large
Sample attrition	Yes	No
Control variables for characteristics of individuals	Some Longitudinal surveys allow control for time-invariant individual characteristics	Few
Recall bias	Yes	No
'Categorisation' of qualifications	Less detail	Greater detail (eg subject)
Counterfactual	Outcomes of those holding various (and no) other qualifications Outcomes before attainment	Outcomes available only for those that enroll to undertake FE training. Alternative counterfactuals: 'non-completers'; simultaneous completers at adjacent levels of qualification; before attainment
Source(s): BIS Research Papers 47,48 and 53.		

The analysis using the matched administrative data is able to overcome some of the weaknesses of the survey-based analysis – the key strengths of the matched administrative data are: larger sample size; reduced recall bias; and the much richer level of detail by which to categorise the provision of vocational qualifications.

Where is it better to use one set of results over another?

The three papers reviewed in this report do not allow for a straightforward decision to be made on which is the best dataset to use to estimate the returns to qualifications, even for particular situations.

Merits of the variables in LFS versus administrative data

On the one hand, the LFS allows control for more of the individual-level characteristics of individuals such as other training/learning. On the other hand, the larger, linked, administrative datasets allow analysis at a much finer grain of detail, such as subject.

The relevance and interpretation of the counterfactual used

Recall that a key analytical challenge to estimating the economic returns to vocational qualifications is how to define the comparison group (counterfactual). This challenge is faced by all researchers, and neither of the existing datasets, the survey and the matched administrative data, are sufficient to provide the ideal counterfactual. The choice of counterfactual can be severely limited by data and so the researchers develop alternative ways of defining the counterfactual (see Table 0.1). Consequently, perhaps more important than deciding which dataset to use in which circumstance is the need to clearly define the counterfactual and to ensure that the comparison is: the most useful and relevant for a particular purpose (see Gambin et al, 2011).

RP53 defines the counterfactual in a similar way to much of the earlier work that utilises survey data. LFS data are used to estimate marginal returns by comparing the returns to individuals with a qualification at a particular level with returns to similar individuals with the next lowest (or no) qualification. Individuals can be identified as similar (by a limited subset of characteristics) using the control variables that are available in the survey data. This counterfactual is a well-established approach which is considered both meaningful and intuitively appealing. However, it fails to take account of differences in ability and motivation of individuals studying for different levels of qualification.

The analyses using the linked administrative datasets (in RP47 and RP48) also have particular weaknesses in their definition of the counterfactual. These datasets only include those enrolled on FE courses and so do not provide the basis for comparison with those who have not undertaken training. The researchers develop and test alternative approaches to defining the counterfactual.

RP48 uses outcomes before attainment as the counterfactual. RP47 also investigates the use of 'before attainment' plus two other approaches to constructing a counterfactual. The authors of RP47 conclude that, no matter which counterfactual is used, there are strong positive and persistent returns to completing a qualification at a given level. There are weaknesses to the counterfactuals used, and we note that the magnitude (and in some cases the sign) of the estimates are sensitive to the choice of comparison:

1. Before and after qualification acquisition – fails to take account of factors that change over time (e.g. age effect).
2. Simultaneous completers at adjacent levels of qualification – fails to take account of differences in ability and motivation of individuals studying for different levels of qualification.
3. Non-completers – fails to fully take account of: factors that affect the likelihood of course completion which may also affect labour market outcomes (e.g. the learner may have failed to complete because of ill-health); all types of non-completion (e.g. transfer to other course).

Enhancements to the administrative datasets that would provide a basis to improve the counterfactuals used in RP47 and RP48 are discussed under 'Promising areas and methods for future research' below.

Persistence of returns

The main analysis in RP53 in effect estimates the average returns over a fairly long period of time, but says nothing about how those returns may change post-achievement. Some of the results based on analysis using the BCS do suggest that some indication of returns over time could be established, but the robustness of the results would need to be investigated further.

RP47 finds that for many vocational qualifications the impact on earnings and employment is positive and increases over time (up to 7 years post-completion), and that achieving the qualifications reduces benefit dependency. Apprenticeships have some of the largest long-term impacts on learners (as was estimated also by RP53), but there is some erosion of the earnings premium over time. It should be noted that this analysis of the persistence of returns is experimental and there are methodological issues that need to be overcome and warrant further research.

Further development and analysis of the linked datasets will prove beneficial

The use of the LFS data is important to ensure that the analysis found in earlier studies can be replicated and results compared over time, but, despite a number of shortcomings, our general belief is that further development and analysis of the linked datasets is the route that will prove most beneficial in producing practical findings because of the substantial advantages on most counts set out in Table 0.1.

In any case, having both survey and administrative data provides opportunities to explore numerous facets and features of the returns provided by FE qualifications. The results of the three papers are promising in that they do not contradict each other in terms of the overall picture.

What are the implications of the report's findings within the context of BIS' NPV model?

The NPV model⁷ developed for BIS by CE and IER includes estimates of wage premia and employment premia that were drawn from the analysis by McIntosh⁸ and developed through consultation between BIS and IER.

RP53 provides updated estimates of marginal returns

⁷ Cambridge Econometrics and IER (2011) Measuring the Economic Impact of Further Education, BIS Research Paper Number 38.

⁸ McIntosh, S. (2009). The Economic Value of Intermediate Vocational Education and Qualifications, UKCES, Evidence Report 2011, December 2009.

RP53 effectively updates much of that earlier analysis, and so the estimates of wage and employment premia in RP53 could be used to update those in the NPV model, at least for vocational qualifications.

Although the analyses of the administrative data suffer a some drawbacks...

Our overall feeling at this stage, however, is that the analysis in both RP47 and RP48 is too experimental for us to recommend making use of these estimates in the NPV model. It would be preferable first to overcome some of the shortcomings to make further improvements to the analyses.

...they offer potential to update and further develop the NPV model in the future

When the NPV model was originally developed, much thought was given to the trade-off between complexity and manageability. There is a need for the model to be: sufficiently complex to capture the pertinent relationships and impacts within the FE system (with the associated risk that the model will not be at all user friendly or transparent); and sufficiently manageable, and easy to understand and operate (with the associated risk that it is over-simplified to the extent that it is not suitable to undertake the kind of analysis required).

The decision was made to build into the NPV model a fairly high degree of disaggregation and detail to make the model more flexible in terms of the analysis that can be undertaken. The high level of disaggregation, with 15 provision types, 7 prior qualifications, gender, and two age bands identified separately in many of the calculations, is reflected in the complexity of the model. This complexity is at the bounds of what can be effectively and transparently handled within the software Excel. The alternative would be to migrate the model to a more flexible programmable language (but this would limit the accessibility of the model to analysts at BIS). The complexity of the model is also reflected in the robustness of some of the existing assumptions, which in many cases were made based on data at higher levels of aggregation. It was the intention that the assumptions could be updated and improved over time as better evidence became available.

The NPV model incorporates assumptions on the marginal benefit (i.e. relative to having a lower level qualification) of the highest qualification held, and so as long as any analysis reveals robust estimates on that basis then they could be used to update the assumptions in the model. What the matched administrative data provide is the opportunity to make estimates in more detail than previously available (from LFS) and so once robust estimates have been made, then they can be used to better populate the assumptions and parameters used in the NPV model.

Where the linked administrative datasets have the potential to add value to the NPV is to provide estimates of returns by subject of study, alongside information about Level, funding stream, and qualification type (e.g. vocational versus academic). The current NPV model does not distinguish qualifications by subject area, and adding this further dimension would substantially increase the complexity of the model. The current findings (in RP47 and RP48) by sector are not particularly robust, but, given these findings, and the theory, it may be desirable to consider ways in which to add this information to the model if more robust estimates become available in the future. One way in which to handle the additional dimension might be to use sub-models for each sector which are replicas, in

structure, of the core model, but populated with assumptions by sector. This would overcome the challenge of introducing an additional dimension in Excel, but:

- not all of the problems of complexity would be eliminated – e.g. assumptions of returns would be required by sector in combination with all the other existing dimensions (15 provision types, 7 prior qualifications, gender, and two age bands).
- the sub-model would undertake analysis for a single sector in isolation – it would likely be desirable to compare findings across sectors and so a number of sub-models would need to be set up and maintained.

Another area in which analysis of the new datasets might offer improved evidence is in the pattern of returns over time. The NPV model assumes that the returns to training and qualifications are persistent and constant over the working lifetime of each learner. This assumption was made because there was little evidence available about the persistence of returns, and to simplify the framework of the model. RP47 suggests that returns are not constant, over the first seven years post-qualification at least. The returns are found to increase over time, with negative returns in the first few years after achievement in some cases (the strong negative impact of Level 3 qualifications in the first four years following attainment reflects the inclusion of GCE A' Levels in the analysis, where completers are more likely to then progress to further or higher education and so remain as relatively low earners for a number of years). Further analyses to support these tentative conclusions would provide better evidence to further develop the NPV model to incorporate variability of returns over time. Adding the ability to vary over time the assumptions for returns would, however, add an additional layer of complexity to the model.

Promising areas and methods for future research

There are limits to what survey data can tell us about returns to qualifications

Over the past 20 years there has been no shortage of analyses looking at the returns to qualifications of one kind or another. The most recent studies - based on cross-sectional and longitudinal data - were appositely summarised by McIntosh (2009). In some respects, the limits to what survey data, either cross-sectional or longitudinal, can tell about the returns to qualifications have been reached. It is possible to provide relatively robust findings relating to academic versus vocational, level, and type of qualification, whilst controlling for a variety of socio-demographic and educational factors which have some bearing on the level or rate of return.

Whilst there is merit in repeating analysis using cross-sectional and longitudinal survey data in updating findings over time, especially so given that the vocational education and training system is subject to periodic overhaul, there is also a need to progress beyond what these types of data can provide.

The matched datasets provide the most promising opportunity to further our knowledge of returns to qualifications

The use of matched and linked data as reported in RP47 and RP48 provides a considerable opportunity to progress the analysis even if, to date, there are a variety of

conceptual, econometric and statistical issues to be resolved. In looking at how to further develop the approach adopted in these research reports there are a number of issues which future research might address.

In an ideal world, if the aim is to gauge how the pathway through the FE and skills system affects labour market outcomes relating to wages, spells in and out of employment, benefit dependency, etc. then there is a wish list of data that is ideally required, and a list of methodological issues that would ideally be addressed to achieve fuller insights and more robust results. These issues include the following:

- providing a comprehensive database which allows an individual to be tracked through post-compulsory education. This will necessitate supplementing the existing databases with data which captures an individual's higher education achievements;
- there is also a need to consider how data can be obtained about the outcomes of people who are not captured by the ILR such as those who do not enter post-compulsory education. At the moment the analysis consists solely of those who have entered the further education and skills system;
- the lack of measures (to be included as control variables) to take account of ability – such as the reading, mathematics, verbal and non-verbal reasoning test scores which some longitudinal surveys collect. But in many respects the key question is what value-added do qualifications obtained in the FE and skills system produce given an individual's educational attainment on entry to it;
- greater attention might be given to ensuring that there is a comprehensive record of the individual's educational attainment on entry to the further education system. The National Pupil Database may be useful in this regard if it were linked to the ILR (which has been achieved in some analyses). But alongside educational qualifications on entry, it would be useful to know something about an individual's motivation or aspirations for choosing a particular course within the system because these may well be associated with eventual success within the FE system (i.e. passing a course with commendation or merit), the predilection to continue with learning, and eventual labour market outcomes;
- an individual's passage through the FE and skills system is not included in the analyses, yet the individual's experiences whilst in the system may be a determinant of their labour market outcomes. A person who was generally satisfied with their educational experience in FE, who received fairly broad based vocational education, with plenty of opportunity to train and practice skills learnt, may well, other things being equal, fare relatively well in the labour market when compared with someone who had a less positive experience;
- the selection of an appropriate counterfactual is of critical importance since the question about any return is always that of "compared with what?". A flexible approach can be adopted here (as has been in the papers reviewed) such that the counterfactual is selected on the basis of what particular question is being addressed (Gambin et al, 2011) , but as noted above, there is a need to have data for those who are not captured in the ILR for one reason or another;

- consideration also needs to be given to constructing a consistent time series of data relating to the measurement of employment spells and the measurement of earnings over time based on the use of HMRC data. Related to this, if estimates are to be compared between studies and over time there is a need to control for conditions in the labour market and economy more generally, since the returns to some qualifications will to a greater or lesser extent simply reflect growth in the economy.

If it were possible to make the various improvements to the database used to provide the analysis in RP47 and RP48 then the analysis will begin to provide detailed information about the subject of study, alongside information about Level, funding stream, and qualification type. By taking an approach which looks at the value-added by different types of qualification for people who otherwise share similar characteristics then much progress will be made towards answering questions such as what qualifications give the highest return for people with a given set of characteristics. Both RP47 and RP48 provide many valuable insights but, for the reasons outlined in the main body of this report, we recommend that caution be exercised in interpreting the results of these initial studies.

In looking more to the future, any data wish list can be divided between possessing information about:

- an individual's experience before entry to the FE and skills system such that there is a comprehensive record of their achievements on entry;
- their experiences in passing through the FE and skills system;
- the outcomes achieved at the point of exiting the FE and skills system;
- any further education and learning post-FE; and
- economic status over a period of, say, seven years post-FE.

Table 0.2 contains the wish list of data. What is apparent is that much of the data is already collected but in separate surveys so there needs to be a means sought of being able to link various data sets. In this way a much more comprehensive set of data will be available with which to gauge the relative success of following different pathways within the FE and Skills system. It will also potentially explain why there is variation in outcomes within a particular pathway.

The data set out in Table 0.2 is what would be required for an approach which allows the passage through the FE system – and beyond where it leads to higher education or other learning – to be modelled at each stage and then to determine the overall outcome with respect to employment. As noted above, the linking of relevant databases is close to making this realisable. The data needed are held in various administrative databases and surveys so that the major issue is not so much about collecting more data but being able to match and link data sources. A unique identifier is required to do this ideally – though the probabilistic matching techniques used to link ILR and HMRC data appears to obtain a high degree of match – which would require being able to include the unique learner code in surveys of learners. From a technical perspective, if the ILR is being used as a

sampling frame for surveys of learners, this is not an insurmountable barrier. Respondents, however, will need to give their permission for survey responses to be linked to other databases. And where permission is not given, or where probabilistic matching is used, there is a need to consider what types of bias may be being introduced into any analysis. But the fact remains, as evidenced by RP47 and RP48, that the ability to link and match is improving all the time such that the comprehensive analysis envisaged above is close to being achieved. The move to a unique identifier to link databases which record individuals' pathways through the education, further education, and higher education systems would make this much more achievable.

Table 0.2: Ideal Data Requirements

Stage in the education system	Type of data required	Availability of data
Entry to FE and Skills system	<ul style="list-style-type: none"> • Type of programme studying / working towards • Subject(s) / framework of study • Prior educational attainment • Socio-demographic characteristics • Aspirations / motivations • Date of entry 	<ul style="list-style-type: none"> • NPD • ILR • Surveys of Prior Educational Attainment • Data on aspirations and motivations from ad hoc surveys
Passage through FE and Skills system	<ul style="list-style-type: none"> • Indicators of training received • Satisfaction with learning • Satisfaction with work environment (if applicable) 	<ul style="list-style-type: none"> • Principally available in survey data such as Apprenticeship Pay Survey and National Learning Satisfaction Surveys. Other ad hoc surveys; e.g. of apprentices
Situation on exit from FE & Skills system	<ul style="list-style-type: none"> • Whether completed • Qualification / education award obtained (and level) • Date of award • Employment status on exit 	<ul style="list-style-type: none"> • ILR
Experience of HE	<ul style="list-style-type: none"> • Type of course pursued • Qualification obtained • Date of award 	<ul style="list-style-type: none"> • UCAS / HESA databases
Initial labour market position over 7 years post-FE	<ul style="list-style-type: none"> • Employment status Year 1 - 7 (years 1 - 7 respectively) • Wage (years 1 - 7 respectively) • Benefit status (years 1 - 7 respectively) 	<ul style="list-style-type: none"> • HMRC / DWP databases

1 Introduction - Aims and Objectives

This is the final report for the study commissioned in March 2012 by the Department for Business, Innovation and Skills (BIS) to update the evidence base on the economic returns to vocational qualifications. The research team was led by Cambridge Econometrics (CE), in collaboration with the University of Warwick Institute for Employment Research (IER).

Having set out in The Plan for Growth its ambition “To create a more educated workforce that is the most flexible in Europe”, the government has announced policy measures to further promote vocational education and training. To better incentivise training and assess employer demand, there is a need to understand the returns to various forms of vocational education and training.

There is now a large body of evidence, collated over the past decade, on the relative returns to obtaining qualifications in further⁹ and higher education¹⁰ based upon cross-section analysis (usually the Labour Force Survey) and longitudinal surveys. Whilst these studies employ robust econometric techniques, they are constrained by sample sizes and the problems posed by sample attrition.

Two recent studies¹¹, however, have made use of a recently-developed experimental ‘matched’ database that combines administrative data from the Individualised Learner Record (ILR) with data from Her Majesty’s Revenue and Customs (HMRC) and data from the Department for Work and Pensions (DWP). This provides a rich source of information that allows some of the drawbacks of earlier analysis to be overcome.

⁹ For example see: Dearden, L., McIntosh, S., Myck, M. and Vignoles, A. (2002) ‘The Returns to Academic and Vocational Qualifications in Britain’, *Bulletin of Economic Research*, Vol. 54, p. 249- 274; . 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) ‘A Cost-Benefit Analysis of Apprenticeships and Other Vocational Qualifications’. DfES Research Report 834; McIntosh, S. (2010) “The Economic Value of Intermediate Vocational Education and Qualifications”, UK Commission for Employment and Skills, Wath-upon-Deerne.

¹⁰ For example: Purcell, K., Elias, P., and Davies R., (2005) *The Class of ‘99: A study of the early labour market experience of recent graduates*, University of Warwick, October 2005; London Economics (2011) *The Returns to Higher Education Qualifications*, BIS Research Paper No.45; O’ Leary N., and Sloane, P., (2005), ‘The returns to a university education in great Britain’, *National institute Economic Review*, July 2005, vol. 193, no. 1, 75-89.

¹¹ BIS Research Papers 47: London Economics (June 2011) ‘The Long Term Effect of Vocational Qualifications on Labour Market Outcomes’; and 48: Frontier Economics / IFS (June 2011) ‘Reporting on Employment and Earnings Using Experimental Matched Data’;

This study was commissioned by BIS to update the evidence base on the economic returns to vocational qualifications (VQs) in England by reviewing the two research papers that use the matched administrative dataset, together with another recently-published research paper that applied well-established techniques to recent survey data.¹²

The different types of approach used in these recent analyses were reviewed so as to:

- provide an overall assessment of the approaches used;
- interpret the findings in the context of the extensive past literature in this area;¹³
- advise in which circumstances it is recommended to use estimates derived from the different analyses;
- comment on the implications of the findings within the context of the BIS 'Economic Impact of FE Model'; and so
- summarise what the recent research adds to the existing evidence base.

The review considered:

- the wage and employment benefits which learners gain from different qualifications;
- the value of qualifications held in the working age population; and
- the extent to which the returns might vary for different individuals and in different circumstances.

This report presents the findings of the study. To set the scene for the subsequent detailed review of the papers, Chapter 2 provides a brief summary of each of the papers. Chapter 3 discusses the theoretical perspectives and analytical challenges to the empirical estimation of the economic returns to vocational qualifications. Chapter 4 discusses the key datasets and statistical methods used in the papers. Chapter 5 presents the key findings of the research papers and Chapter 6 draws together the conclusions and recommendations of this study, including recommendations for future analysis.

¹² BIS Research Paper 53: London Economics (September 2011) 'Returns to Intermediate and Low Level Vocational Qualifications'.

¹³ The past literature will be represented by the extensive review by McIntosh, S. (2009) 'The Economic Value of Intermediate Vocational Education and Qualifications', UKCES.

2 Summary of the Papers Reviewed

2.1 Introduction

The objective of this project is to update the evidence base on the economic returns to vocational qualifications in England by reviewing three specific BIS research papers (RP) recently published. To set the scene for the subsequent detailed review of the papers, this chapter provides a brief summary of the papers reviewed.

2.2 Summary of the Papers Reviewed

Table 2.1 below summarises for each of the three BIS research papers: the economic benefits modelled; the datasets used; the different ways in which qualifications are categorised; the estimation methods used; the associated comparison groups (counterfactuals); and the key findings.

RP53: Returns to Intermediate and Low Level Vocational Qualifications

In this paper, London Economics (September 2011) applied well-established techniques to recent survey data to derive updated estimates of returns to intermediate and low level qualifications. This paper used Labour Force Survey (LFS) data and estimated OLS cross-sectional regressions for each year to estimate marginal and average earnings returns. For the analysis of employment returns, a probit model was estimated for the probability that an individual is in employment versus unemployment or inactivity. In addition, a longitudinal dataset from the British Cohort Study (BCS) was used to estimate employment and earnings returns: separate cross-sectional regressions are estimated for each of the three waves as well as a fixed effects estimation using all three waves. The use of the BCS provided a counterfactual 'before attainment'.

The authors' findings are consistent overall with previous studies which adopted similar approaches: there continue to be positive employment and earnings returns for the vast majority of qualifications gained in adulthood, but the route of acquisition of qualification has become less important in determining the earnings returns (previous estimates, and theory, suggested that the workplace route gave greater returns). The report notes the significant variation in returns by sector and occupation of employment.

RP48: Reporting on Employment and Earnings Using Experimental Matched Data

The development of a matched administrative dataset provided the basis for the innovative research published in RP48 and RP47:

Frontier Economics/IFS (June 2011) undertook analysis of matched administrative data comprising: learner information from the Individualised Learner Record (ILR); information on benefit recipients held by the Department for Work and Pensions (DWP); and information on employment status and earnings in work held by Her Majesty's Revenue

and Customs (HMRC). The matched data were used to assess the labour market performance of learners both before and after learning, and disaggregated by qualification and subject area. The majority of the analysis is cross-tabular in nature involving a comparison of various descriptive statistics and sample proportions relating to earnings and employment outcomes for groups of learners in the 12 months pre- and post-qualification attainment. For the Further Education learning stream only, impact analysis is carried out using fixed effects estimation of earnings, benefit, and employment outcomes.

The authors conclude that, in general, there are improvements in the economic performance of learners following achievement. The innovative contribution made by this paper is the exploration of the matched dataset and the opportunities that it presents; for example, in the analysis, this paper presents findings by level of qualification and by *subject area*.

RP47: The Long-Term Effect of Vocational Qualifications on Labour Market Outcomes

In this paper, London Economics (June 2011) also utilised the matched administrative dataset (which while from the same original sources differed somewhat in the data used in the final analyses), and various econometric techniques and *alternative counterfactuals* were applied to estimate long-term earnings and employment effects.

A series of annual cross section OLS regressions are estimated. There are two main aspects to the research. A first stage analysis is conducted of the determinants of qualification completion. This is specified as a probit and ordered probit model and estimated by learning aim rather than by individual learner, over the entire sample of qualification completers and non-completers for the years 2002/3 to 2005/6. The significant explanatory variables from these regressions are used to inform the main analysis of earnings, employment and benefit dependency outcomes. This second stage of the analysis estimates a cross-section OLS regression for each year post-attainment of the qualification up to seven years afterwards.

The authors conclude that, no matter which counterfactual is used, there are strong positive and persistent returns to completing a qualification at a given level. Having said that, the magnitude (and in some cases the sign) are sensitive to the choice of counterfactual.

Table 2.1: Summary of the Papers Reviewed

SUMMARY OF THE PAPERS REVIEWED			
	RP53: London Economics (September 2011) <i>Returns to Intermediate and Low Level Vocational Qualifications</i>	RP47: London Economics (June 2011) <i>The Long Term Effect of Vocational Qualifications on Labour Market Outcomes</i>	RP48: Frontier Economics / IFS (June 2011) <i>Reporting on Employment and Earnings Using Experimental Matched Data</i>
Economic benefits modelled	Marginal and average earnings Marginal employment NPV to individual NPV to Exchequer	Earnings Employment outcomes (over a seven year period post-completion) Long-term benefit dependency	Earnings Employment Benefit dependency
Nature of datasets	Survey data (cross-sectional and longitudinal)	Linked administrative data	Linked administrative data
Datasets used	LFS; BCS	ILR; HMRC P14; HMRC P45; DWP National Benefit Database	ILR; HMRC P14; HMRC P45; DWP National Benefit Database
'Categorisation' of qualifications	Intermediate and low level (levels 2 and 3) vocational qualifications by learning stream Employment sector Employment occupation	Levels 1-4 by learning stream	Funding streams: Further Education; Apprenticeships (in work-based learning); Train to Gain Subject (study areas)
Estimation methodology	1. OLS 2. Quantile regression (for NVQ level 2 only) 3. First difference estimator	OLS	First difference estimator

SUMMARY OF THE PAPERS REVIEWED			
	RP53: London Economics (September 2011) <i>Returns to Intermediate and Low Level Vocational Qualifications</i>	RP47: London Economics (June 2011) <i>The Long Term Effect of Vocational Qualifications on Labour Market Outcomes</i>	RP48: Frontier Economics / IFS (June 2011) <i>Reporting on Employment and Earnings Using Experimental Matched Data</i>
Counterfactuals (comparison groups)	<ol style="list-style-type: none"> 1. Outcomes of those holding various (and no) other qualifications 2. Outcomes before attainment (for BCS only) 	<ol style="list-style-type: none"> 1. Outcomes of ‘Non-Completers’ 2. Outcomes of simultaneous completers at adjacent levels of qualification 3. Outcomes before attainment 	Outcomes before attainment (controlling for fixed effects)
Key findings	<p>Concurs overall with some previous studies: there continue to be positive employment and earnings returns for the vast majority of qualifications gained in adulthood.</p> <p>The route of acquisition of qualification has become much less important in determining the earnings returns.</p> <p>There is significant variation in returns by sector and occupation. Earnings returns have a negative sign for higher level occupations.</p>	No matter which counterfactual is used, there are strong positive and persistent returns to completing a qualification at a given level. (although the magnitude, and in some cases the sign, are sensitive to the choice of comparison group).	In general, there are improvements in economic performance of learners following achievement; these improvements vary by level and by subject area.

3 Theoretical Perspectives on the Returns to Training and Qualifications

3.1 Introduction

This chapter sets out the theoretical perspectives and analytical challenges to the empirical estimation of the economic returns to vocational qualifications. The critical issue to be addressed is the extent to which a given qualification (with respect to its subject and level) gives a better or worse return (typically measured with respect to employment and earnings) than a comparable qualification. Section 3.2 highlights the main conceptual and methodological issues regarding the estimation of the returns to vocational education and training. Section 3.3 outlines a number of analytical challenges and the limitations which often prevent studies from producing unbiased estimates. Section 3.4 reviews recent evidence regarding the returns to training and qualifications, including both pecuniary and non-pecuniary benefits, and highlights a number of caveats of the empirical estimation. Section 3.5 discusses how well the three BIS reports represent the theoretical perspectives and address the analytical challenges outlined in this chapter.

3.2 Estimating the Benefits of Training and Education

Obtaining a vocational qualification, or any qualification for that matter, can be regarded as an investment by: (i) the person studying towards the qualification (in terms of the amount of effort and time required to obtain the qualification); and (ii) the parties which meet the costs of the individual working towards the qualification (typically some combination of the State, employers, and individual learners). Typically the returns to training are measured with respect to increases in wages and the probability of remaining in employment, though there are a range of other benefits typically listed under the category of the 'wider benefits of training' which relate to factors such as the individual's well-being, the intrinsic rewards of learning, reduction in benefit uptake/dependency *etc.*

The overall aim of estimating the returns to vocational education and training therefore requires a comparison of people's transitions through the overall education system incorporating both the compulsory and post-compulsory elements. The determinants of the various transitions people make – as outlined in Figure 3.1 – need to be determined. Figure 3.1 presents a simplified view of the stages through which individuals (might) progress from entry into the formal school system, through further education, and then potentially into higher education, before entering the labour market. For the sake of simplicity, lifelong learning which leads to the award of a vocational qualification has been excluded. Figure 3.1 seeks to demonstrate that at each stage of the process – *i.e.* pre-school, compulsory schooling, further education and higher education – choices or selections made with respect to what education or learning institution is attended, subjects studied and to what level, and the type of qualifications, will be shaped, *inter alia*, by a range of socio-economic and demographic factors and by ability.

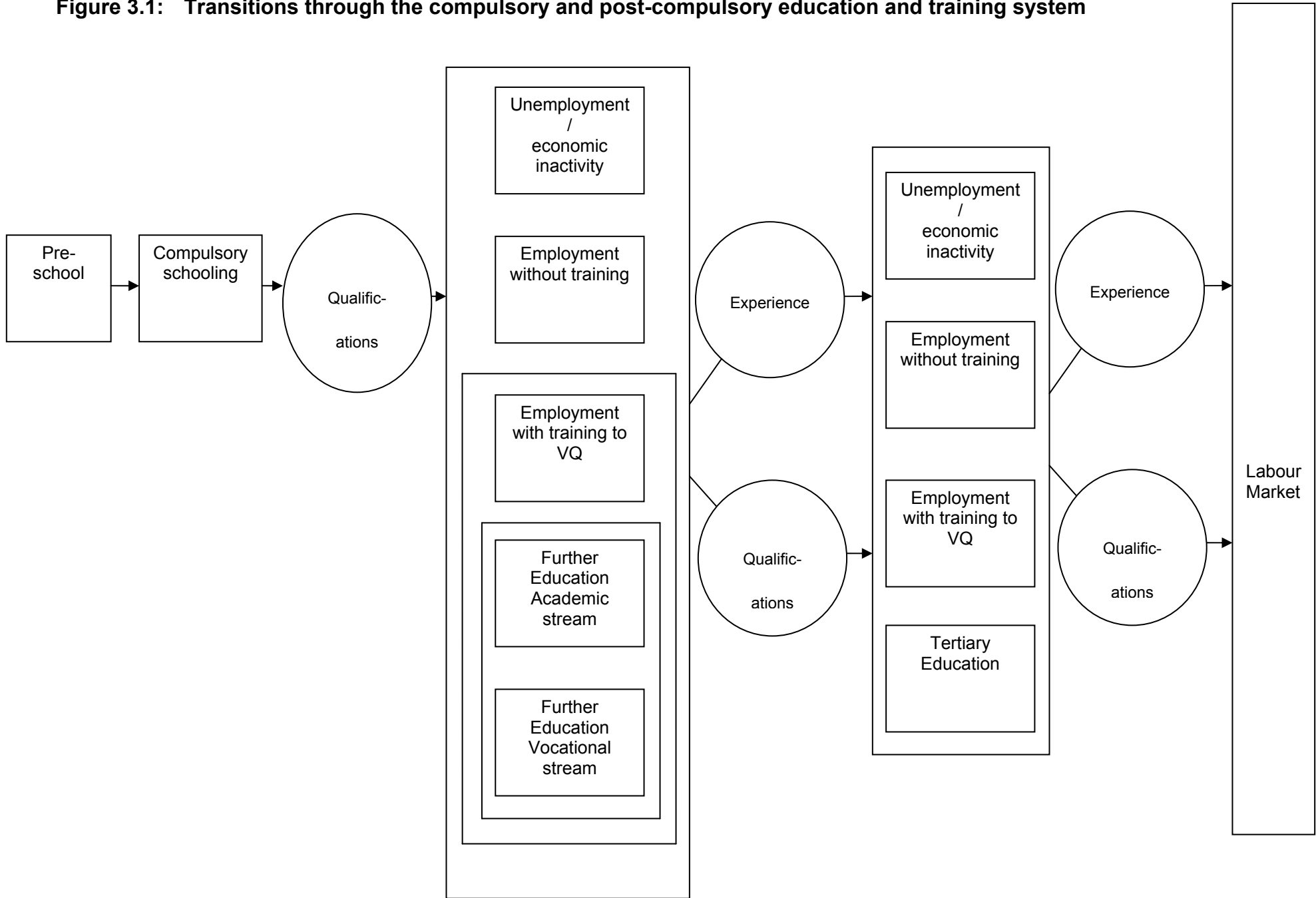
In a perfect world there is a need to control for all of these factors if the value-added obtained from any vocational qualifications is to be determined. In this way, it will be possible to determine, given the current structure of education and learning provision, the extent to which vocational qualifications of different types, in different subjects and at different levels give relatively high or low returns – compared with some comparator qualification(s) - to people with differing characteristics who have taken, for one reason or another, differing pathways through the compulsory and post-compulsory education and learning system.

Therefore, when seeking to measure the economic returns to vocational qualifications, the critical issue to be addressed is the extent to which, other things being equal, a given qualification (with respect to its subject and level) gives a better or worse return (typically measured with respect to employment and earnings) than a comparable qualification, or no qualification at all. Ideally, the analysis would also recognise that different qualifications are appropriate for different circumstances or learners (some are more relevant to younger college-based learners, while others are more relevant to older work-based learners) – so it is not necessarily the case that a better overall average return makes that qualification better for all circumstances. Because other things are seldom equal – and due to the absence of an experiment to test the critical issue identified above - researchers have had to be inventive in the use of survey and administrative data in order to control for a range of variables which may confound any analysis. Typically, the choice of course leading to a qualification which an individual may choose to study is determined by a range of factors other than ability. And a range of factors other than ability may determine whether a person successfully completes a course and their grade of achievement should they complete. In practice, researchers are trying to control for a range of socio-demographic, educational, and ability factors, with imperfect data, in order to measure the returns which can be safely regarded as being derived from possession of a given qualification.

From the evidence presented in McIntosh (2009), and summarised later in this Chapter¹⁴, it is apparent that the findings derived from either longitudinal or pooled cross-sectional analyses struggle to control for these type of effects. In short, like is generally not being compared with like. This simply reflects the complexity of the issue being addressed and the consequent analytical challenges (as discussed in the next section). But, as Section 3.5 demonstrates, recent studies funded by BIS have been able to make substantial progress in tackling some of these issues.

¹⁴ See Section 3.4.

Figure 3.1: Transitions through the compulsory and post-compulsory education and training system



3.3 Limitations of Rates of Return Analysis

Selection and comparison problems

The principal issue of interest is the extent to which, other things being equal, achieving a given qualification is better than achieving some alternative qualification. This raises the important issue of whom and what is being compared. This is a problem which besets all research in the area of vocational education and training.¹⁵ A number of approaches are evident in the literature which involves comparing the returns to people:

- i. acquiring various qualifications at the same level whilst trying to control for a variety of socio-demographic characteristics;
- ii. acquiring a given qualification at a given level to those acquiring a qualification at a lower level; and
- iii. completing a qualification to those who drop-out.

Each approach – especially (i) and (ii) - have much to recommend them but they are also flawed in various ways. The principal problem is that like is not being compared with like. There are a number of problems which analysts recognise but struggle to control:

- there is no guarantee that qualifications which are nominally classified to a given level in the NQF are actually the same with respect to level of difficulty or ability required to achieve a pass. Five GCSEs at grades A to C might not be the same as an NVQ at Level 2;
- even at the same level of qualification there may be some unobserved characteristics – usually *ability* as measured by reading, mathematics, verbal reasoning and non-verbal reasoning tests – which leads a person to study qualification x rather than y;
- in contemporaneous studies there is no guarantee that one is observing the highest qualification attained which potentially distorts the returns over the initial post-completion phase of the qualification under observation;
- the comparison of qualifications at different levels (to measure the marginal gain from acquiring an additional qualification) struggles to control for why one group has attained a qualification at the lower level (for example, do they lack the ability to complete a qualification at the higher level);

¹⁵ Grubb, W.N. and Ryan, P. (1999) *The Role of Vocational Education and Training: Plain talk on the field of dreams*, ILO/Kogan Page, London.

- where a person fails to complete a qualification, it is known that people drop-out for a variety of reasons not all of which are related to the qualification being studied,¹⁶ e.g. ill-health, such that one is unsure of what is being compared (for example, one may be comparing the returns of those who suffer ill-health with those who do not).

Whilst in an ideal world one can identify all the variables one might like to include in a statistical model, it is unrealistic to expect that such data should be available. That said, as will be made clear below, the use of linked administrative databases provides the potential to overcome many of these issues. In general, there is a need to recognise the limitations of any analysis, though it is often possible to find measures which proxy – however imperfectly – missing variables (such as, for instance, using prior educational attainment as a proxy measure of ability in models which sufficiently control for the socio-economic determinants of who studies what in different types of school in the compulsory education system).

In many respects, the comparison in (i), where there the highest level of qualification attained is the dependent variable, has most to recommend it, since it is at least nominally comparing people with the same level of qualification.

Biases in estimating returns

In reviewing the evidence on the returns to vocational qualifications, McIntosh (2009) draws attention to a series of problems which relate to estimation of any returns. The essential problem is that of identifying people who are otherwise identical, however measured, where one has studied towards a vocational qualification and another has studied towards some comparator non-vocational qualification or not studied towards a qualification at all. If such a comparison were available then the value-added by the qualification would be obvious, though even in the ideal case, as can be seen above, there is some debate over what should be the comparator: studying towards a comparable non-vocational qualification or no qualification at all?

Because, in practice, it is not possible to provide the ideal comparison and the best has to be made of the data readily available, McIntosh points to two types of potential bias introduced into the rates of return analysis:

- i. an upward bias to the returns associated with those who study towards qualifications. This results from the more able being more likely to study towards qualifications of a higher level and, consequently, being rewarded with better paying jobs. There are other factors too which come into play such as the pay awards which might be associated with obtaining a qualification;

¹⁶ Hogarth T., Gambin L., Hasluck C., de Hoyos, M., and Owen D. (2009) *Maximising Apprenticeship Completion Rates*. Learning and Skills Council, Coventry.

- ii. a downward bias to the returns associated with those who study towards lower level qualifications or do not study any qualification at all because they are potentially less able.

These are essentially selection effects relating to ability. One may be simply measuring the returns to ability rather than the value-added resulting from acquiring a qualification.

Selection effects in practice relate not just to ability but also to a range of socio-economic factors which affect an individual's access to schooling and training. Some individuals, for instance, as a consequence of their socio-economic position are less likely to acquire qualifications or perform well in schools because, regardless of their ability, they attend less successful schools, are less likely to have parental support for their learning, experience less peer group pressure to perform well educationally, etc. (Rutter, *et al.*, 1979; Smith and Tomlinson, 1984). In summary, those more academically successful in the school system are more likely to adopt the academic pathway through further education rather than the vocational one.

The upshot of the problems identified by McIntosh (2009), and others (e.g. Garrett *et al.*, 2010), suggest that great care needs to be taken in making comparisons between groups (Ryan and Grubb, 1991). Perhaps more importantly it points to the need for a longitudinal perspective to understanding the returns to qualifications since there is a need to control, as far as is feasible, for the types of selection effects described above.

Measuring persistence

Persistence potentially has three meanings in the context of understanding the returns to qualifications:

- that the returns persist for a period of time after the award of a qualification;
- the relative returns between qualifications persist in differing economic contexts;
- the returns persist regardless of the socio-demographic characteristics of the individual.

The first issue relates to the period of time over which the returns are to be measured. If the aim is to measure the rate of return then there needs to be some consideration over what period the returns are to be measured. Over recent years there has been considerable interest in a lifecycle approach to understanding how various factors affect outcomes experienced by individuals at different stages of their lives.¹⁷ Such an approach in relation to the acquisition of a particular qualification poses a number of problems. On the one hand, there are a number of jobs where a specific qualification is required to

¹⁷ For example, Anxo D., Fagan C., Cebriuan I. and Moreno G. (2006) '*Patterns of Labour Market Integration in Europe: a Life Course Perspective on Time Policies*'. The Socio-Economic Review, Vol. 5, No 2, Oxford University Press, 2007.

access a given job; the most obvious example being a medical degree in order to practise medicine. In such cases, acquisition of a given qualification is likely to be a determinant of returns obtained over the lifecycle. On the other hand, there are very many more jobs where the qualification is used to screen applicants but subsequent progression will be determined by other factors. In such cases, the impact of the qualification or training undertaken in the FE sector will diminish quite rapidly. From a comparative aspect, therefore, perhaps the most important period is that immediately post-qualification where individuals use qualifications acquired to obtain employment and progress their careers. Any choice of a specific period is likely to be arbitrary but in the recent analysis undertaken in BIS' RP47, the results seem to show that it is possible to discern an effect from a specific qualification at around seven years post-qualification.

In relation to the second point there has been relatively little consideration given to how the economic climate affects the returns to qualifications. There is no reason why, for example, a downturn in the economy should affect the returns to different qualifications at the same level similarly: during a period of weak consumer demand there may be much less demand for those with qualifications in retail and much more for, say, qualifications in construction where infrastructure projects are being used to boost the economy. In McIntosh's studies, the returns to Apprenticeships appear to have increased over the middle years of the last decade. Whether this is due to the intrinsic merits of Apprenticeships or simply a consequence of the economic growth over the middle part of the last decade is a moot point.¹⁸ The macroeconomic environment is a factor seldom considered in the microeconomic studies which assess the returns to qualifications.

Finally, there is a need to consider whether the returns to a qualification are the same across all groups. Who studies what, and where, has a strong socio-economic aspect to it. Studies in higher education, for example, reveal that the principal determinant of the returns to a degree is subject but the institution attended can also have an influence.¹⁹ And access to study a given subject at a given university is determined at least in part by a range of socio-economic factors.²⁰ There is, therefore, at least prima facie evidence that the returns to a given qualification will vary by socio-economic and demographic characteristics.

Returns to what qualification and when

In the studies reviewed by McIntosh (2009) most provide results disaggregated by level and qualification type (e.g. City & Guilds, BTEC, Apprenticeship, etc.) but little attention is

¹⁸ McIntosh, S. (2009). *The Economic Value of Intermediate Vocational Education and Qualifications*, UKCES, Evidence Report 2011, December 2009.

¹⁹ Metcalf, H. (1997). *Class & Higher Education: the Participation of Young People from Lower Social Classes*, CIHE, London.

²⁰ Werfhorst, Herman Van der and Sullivan, Alice and Cheung, Sin Yi (2003) *Social Class, Ability and Choice of Subject in Secondary and Tertiary Education in Britain*. *British Educational Research Journal*, 29 (1). pp. 41-62.

given to subject of study. This is a consequence of either the subject of study not being available in various data sets, or there being an insufficient number of observations to provide statistically robust results. The data which are available suggest that the returns do vary by subject of study. This is observed in higher education, for instance, where the returns to subjects such as mathematics, IT, medicine are relatively high compared with arts subjects (Purcell et al., 2005).

It is also apparent that where individuals are asked to provide information about their qualifications, including subject and type of qualification, this is often based on recall. This potentially introduces a number of biases into the analysis:

- firstly, a qualification in subject x studied thirty years ago, is likely to be different in its content, delivery and examination to that which is delivered in the same subject today;
- second, where information is based on recall there is likely to be some error in that recall which may not even itself out over a large number of responses. For example, someone working in an environment which is now dominated by graduates may be reluctant to admit that their highest level of qualification was, for instance, an NVQ. This may result in more people being able to recall that they have higher qualifications rather than lower ones;
- third, some people may have simply forgotten which qualifications they have.

These types of error or bias have implications for the analysis of the returns to qualifications. One possible bias is that people are more likely to report that they have a qualification which they regard as conferring prestige upon them which may result in an upward bias to the returns to qualifications. And given problems with recall, and changes in the qualification system at various points since the 1944 Education Act – which founded the modern day education system in England - any analysis of the returns over time is likely to be compromised too.

Through the use of administrative data it is possible to overcome many of the biases and estimation problems identified above. The two papers reviewed in Section 3.5 provide an indication of the potential this type of analysis provides, but before reviewing their findings, a brief summary, drawing on McIntosh's fine overview of the current state of play,²¹ is provided so that it is possible to see the extent to which a consistent story emerges.

3.4 What are the Benefits of Training and Qualifications?

There is an enormous literature that has attempted to evaluate and estimate the economic impact of education and training on individuals. This section reviews the most recent literature and discusses the evidence on returns to education regarding levels of training

²¹ McIntosh, S. (2009). *The Economic Value of Intermediate Vocational Education and Qualifications*, UKCES, Evidence Report 2011, December 2009.

and qualifications, returns of “mature” students, returns of life-long training as well as wider returns, which are not specific to pay.

Returns to training and qualifications

The evidence from most studies examining the returns to training and qualifications suggest that there are higher returns for high levels of NVQs, however the returns to low levels (i.e. NVQ 1 and NVQ 2) remain relatively low and in some instances have been estimated to be negative.

McIntosh (2009) reviewed recent evidence on the economic value of intermediate vocational qualifications. In general his review reveals that:

- the returns to vocational qualifications are higher at Level 3 than Level 2;
- returns vary according to how qualifications were acquired, with those acquired whilst at work generally having higher returns;
- returns are highest for those aged under 25 years;
- returns vary by gender though the results depend upon Level, occupation, and sector;
- Apprenticeships at either Level 2 or Level 3 have relatively high returns compared to other routes; and
- returns vary according to the occupation or industry in which an individual is employed.

More specifically, McIntosh draws the following conclusions:

- Where individuals have obtained an NVQ2 qualification as their new highest qualification they earn around 4% more than individuals with no qualifications.
- Other Level 2 vocational qualifications have larger returns ranging from around 7% for City and Guilds Craft qualifications up to 20% for RSA Level 2, when compared to individuals who have no qualifications.
- At Level 3, all vocational qualifications are associated with positive and statistically significant returns when held as highest qualifications compared with a Level 2 qualification: ranging from 11% for an NVQ Level 3 up to 25% for an ONC/OND.
- An individual with an NVQ Level 2 held as a highest qualification and obtained through workplace learning earns approximately 10% more than an individual with no or lower level qualifications.
- Apprenticeships in their modern guise show positive and statistically significant returns. Relative to a comparison group with Level 2 qualifications (either academic or vocational), those who complete a Level 3 Apprenticeship earn a wage premium

of 22% for men and 14% for women. Relative to a comparison group with Level 1 or Level 2 qualifications, those who complete a Level 2 Apprenticeship earn a wage premium of 20% for men and 4% (statistically insignificant) for women.

- The evidence shows that returns to qualifications are sensitive to the occupation in which the individual eventually finds employment. Compared to individuals with no or only lower level qualifications working in the same occupation, individuals with an NVQ Level 2 as their highest qualification obtain statistically significant higher returns in some occupations: namely skilled occupations (8% return); personal services (5%); sales (3%); machine operatives (4%); and elementary occupations (4%).
- Similarly, certain industries see positive and statistically significant returns to NVQ Level 2 as the highest qualifications: e.g. construction (9% return); public administration, education and health (3%); and other services (7%).
- Where attempts are made to control for ability through use of longitudinal data sets which contain childhood test scores, the “available evidence is not strong” but tends towards lower ability individuals obtaining higher returns to intermediate vocational qualifications than those of higher ability.
- Additionally, longitudinal data have been used to estimate by how much individuals’ wages change following the acquisition of an intermediate vocational qualification. The evidence is mixed, with one birth cohort, the BCS showing large wage changes of around 22% for those who acquired an NVQ Level 2 between 1996 and 2004 (age 26 and 34 years respectively) and 36% for other Level 2 qualifications (academic and vocational combined), but researchers using the National Child Development Study (NCDS) found no change in wages between 1991 and 2000 (age 33 and 42 years respectively) for individuals who acquired intermediate vocational qualifications between this period.
- In general, all returns are greater for those aged under 25 years.

The literature appears to point to lower returns to vocational education and training than academic education at the same level of the NQF, perhaps with the exception of the highest level of vocational qualifications.

A wide body of literature has examined non-pecuniary aspects of returns to education. Starting with labour market outcomes, from economic theory it can be inferred that in most developed countries there exists an inverse relationship between education and unemployment. The fact that individuals with higher education experience relatively low levels of unemployment is well established, by evidence, in many developed countries (OECD, 2000). In addition, a number of empirical studies (Nickel, 1979; Mincer, 1991; Woolbers, 2000) suggest that as the level of education increases, the probability of unemployment decreases.

Returns to lifelong learning

Lifelong learning is a generic term referring to a variety of forms of human capital accumulation occurring after the end of the first period of continuous education. Lifelong

learning most commonly refers to vocationally-oriented and general adult education and training (Green et al., 2000; Aspin et al., 2001; Wolf et al., 2006). Jenkins, et al. (2002) used the NCDS to identify the effect of lifelong learning on wages and employment, controlling for a wide range of other factors. The main effects of lifelong learning are as follows (ibid. p. i):

- individuals out of the labour market in 1991 had a higher probability of being in work in 2000 if, in the intervening period, they had undertaken lifelong learning;
- an episode of lifelong learning increased the probability of the individual undertaking more learning;
- failing a course between the ages of 33 and 42 years reduced the individual's likelihood of being a current learner;
- individuals with no qualifications in 1991 who undertook lifelong learning between 1991 and 2000, earned higher wages in 2000 than those who did not engage in lifelong learning over this period;
- the authors found only limited evidence of positive wage effects from lifelong learning for other groups.

Wider returns

A stream of literature focuses on the effects on health of extra years of schooling, suggesting that there is a casual positive relationship between education and health, as measured by mortality rates, morbidity rates, work-days lost and self-assessed health status. For example Silles (2009) finds for the UK a strong relationship running from more schooling to better health. A study for the Netherlands (Hartog and Oosterbeek 1998), however, suggests that the highest level of education does not lead to the highest wealth, health or happiness. On the other hand, Oreopoulos and Salvares (2011) list a number of non-pecuniary benefits of education, including more opportunities for self-accomplishment, inter- social interaction, action, independence, and occupational prestige. They also highlight that education reduces the chance of ending up on welfare or unemployed, it improves success in the labour market and the marriage market, it leads to better decision-making leading to better health, happier marriages, and more successful children. Schooling, according to the same authors, also encourages patience and long-term thinking. In addition, it decreases teen fertility, criminal activity, and other risky behaviour. It also promotes trust and civic participation and it teaches students how to enjoy a good book and manage money.

3.5 How Well do the RPs Represent the Education System?

Two of the three papers reviewed break new ground with respect to understanding the relationship between attaining vocational qualifications and labour market outcomes (RP47 and RP48) whilst a third (RP53) provides a point of reference insofar as it employs a more conventional approach to deriving the benefits of acquiring vocational qualifications. Recall that in estimating returns to qualifications, the critical issue to be addressed is the extent to which a given qualification (with respect to its subject and level) gives a better or

worse return (measured with respect to employment and earnings) than a comparable qualification. A brief summary is given below of how each research paper represents the education system and which comparison groups (counterfactuals²²) are used in the analyses.

RP47: The Long-Term Effect of Vocational Qualifications on Labour Market Outcomes

The study provides an analysis of earnings, employment, and benefit dependence by using matched ILR, HMRC and DWP databases to consider the long-term impacts of obtaining vocational qualifications (of different types) by comparing:

- before and after training (the authors recognise that this does not tell you the impact of the learning itself just the return associated with the person who achieved a given learning aim);
- simultaneous attainment of people acquiring qualifications at adjacent Levels, say, Level 2 and Level 3, respectively (the authors regard this as flawed because there may be differences in ability between the groups studying towards the respective qualifications);
- completers versus non-completers (though it is important to note that people drop out of courses for a variety of reasons and the differing reasons may have a differential impact on outcomes – see Hogarth, et al., 2010).

RP48: Reporting on Employment and Earnings Using Experimental Matched Data

The study uses the matched ILR, HMRC and DWP databases to provide an analysis of the impact of obtaining a vocational qualification by comparing the before and after situation with respect to earnings, employment, and benefit dependence, for people who obtained vocational qualifications through different funding routes disaggregated by course type, subject and level. There is some economic modelling which controls for individual characteristics and prior educational attainment. The authors recognise that the before and after comparison does not tell you the impact of the learning itself just the return associated with the person who achieved a given learning aim. Using fixed effects they try to control for unobserved individual level time-invariant heterogeneity (e.g. ability, gender, etc.) but the data are limited in the availability of individual-level characteristics that may be controlled for.

RP53: Returns to Intermediate and Low Level Vocational Qualifications

The study uses pooled cross-sectional LFS data for 1996 – 2009 to reveal the impact on earnings and employment of obtaining a number of different types of vocational qualifications. Because the analysis uses LFS data it is able to control for various

²² For further discussion, see [Section 4.8](#) Specification of the Counterfactual.

personal characteristics and the age at which qualifications were acquired. The analysis also uses data from the British Cohort Survey (1970) to provide a longitudinal perspective which allows the inclusion of a number of socio-demographic characteristics, including household characteristics, and test scores relating to ability.

With RP47 and RP48 a detailed picture begins to emerge of the returns to vocational qualifications of different types. Information is available on Level, funding stream, type of course, and subject to provide a before and after comparison containing a degree of detail which accurately captures the heterogeneity in vocational qualification provision. The data has a longitudinal aspect to it insofar as it is possible to track an individual through further education and then into the labour market (up to seven years after completing their vocational qualifications).

The authors readily acknowledge some of the weaknesses of the analysis using the linked administrative datasets:

- it is not clear what biases have been introduced via the data matching exercise²³ (or rather few details are provided in the research papers);
- there are no data relating to the higher education experiences of individuals (so it is not clear whether the analysis captures the highest qualification possessed by the individual);
- there are relatively limited data on educational experiences before entering further education;
- there are limited data on the socio-demographic and economic situation of the individuals other than via their location and the use of the Index of Multiple Deprivation (IMD);
- ability is not controlled for in any of the analyses; and
- there is no analysis of the returns relative to academic qualifications at the same level.

What RP47 and RP48 provide is a longitudinal view which captures the heterogeneous pathways through further education. In this respect, that part of the data jigsaw required to provide information about the further education stage in Figure 3.1 above is provided. RP53 based on LFS and BCS data produces an analysis which in many respects is more conventional in that it uses data sources and techniques used in some of the reports reviewed by McIntosh (2009). But it is at once apparent that one loses much of the detail contained in RP47 and RP48 either because the data are not available in the LFS (e.g. subject of study) or because sample sizes in BCS are too small.

²³ Potential biases are discussed in more detail in Section **Error! Reference source not found. Error! Reference source not found.**, below.

Despite the caveats, concerns, and groans reviewers are wont to make when considering the scientific evidence placed before them, the analyses reviewed in Section 3.4 which draw mainly on survey data - mainly cross-sectional data which has been occasionally pooled over time – provide a remarkably consistent set of results with those (in RP47 and RP48) that use the linked administrative databases. Together they provide a much clearer and detailed picture than hitherto about the returns to qualifications. The following Chapters consider in more detail the strengths and weaknesses of using survey versus administrative data in measuring the returns to qualifications.

4 Key Datasets and Statistical Methods – Strengths and Weaknesses

4.1 Introduction

In this chapter we provide an overview of the data and statistical methods used in the reviewed papers, and their associated strengths and weaknesses. Section 4.2 first presents the requirements of the ideal dataset to support empirical estimation of the economic returns to vocational qualifications. We then summarise the data sources used in the reviewed papers (Section 4.3), and the properties of the data (Section 4.4). Section 4.5 discusses issues of potential bias due to linking and missing data and Section 4.6 comments on the suitability of the indicators used to measure economic returns. Section 4.7 summarises the relative strengths and weaknesses of the alternative datasets used. This is followed by an assessment of the statistical methods applied and the extent to which they address the analytical challenges discussed in Chapter 3. Section 4.8 summarises the methodological approaches and Section 4.9 the counterfactuals used. In this context, Section 4.10 comments on the variables included in the analyses and how well they represent the learning system. Section 4.11 comments on the suitability of the sample sizes to support statistical inference and Section 4.12 concludes with the strengths and weaknesses of the econometric and statistical techniques used.

4.2 Data Requirements

The empirical estimation of the economic returns to vocational qualifications seeks to measure the extent to which a given qualification gives a better or worse return than a comparable qualification. The essential problem is that of identifying people who are otherwise identical, however measured, where one has studied towards a vocational qualification and another has studied towards some comparator non-vocational qualification or not studied towards a qualification at all. Following on from the main conceptual and methodological issues presented in Chapter 3, the requirements of an ideal dataset to support such analysis can be summarised as follows.

Economic returns

The ideal dataset would measure for each person the outcomes of interest, i.e. economic returns, typically measured with respect to employment, earnings and benefit dependency. With regards to employment, in addition to measuring whether or not an individual is in employment, it would be preferable to assess the ‘quality’ of that employment (e.g. in terms of whether it is long-term, stable, full-time or part-time). With regards to earnings, analyses that use annual and monthly earnings indicators need also to take account of the number of hours worked.

The persistence of returns to qualifications has an important bearing on the long-term benefits of qualifications and their overall value to the individual learner, firms, the

economy and the exchequer. It is therefore preferable to have a consistent time series of data relating to the measurement of employment spells and earnings over time.

Qualifications

The dataset would include people who do and do not enter post-compulsory education. In order to make assessments of 'what qualifications give the highest return for people with a given set of characteristics' the dataset would identify for each person detailed information about the subject of study, alongside information about level, funding stream, and qualification type (e.g. vocational versus academic).

Controls for confounding effects

The dataset must include sufficient indicators to control for a range of variables which may confound the analyses, e.g. the characteristics of the individual, such as age gender, ethnic and socio-economic group.

To counter selection bias, the dataset would include indicators to take account of ability and other socio-economic factors (e.g. parental support, attendance at successful school) which affect an individual's access to schooling and training.

To best account for the impact of schooling on post-compulsory education, and the impact of learning choices on the scale and persistence of lifetime returns, the ideal would be a comprehensive database which allows an individual to be tracked through post-compulsory education, including a record of their achievements on entry.

There is also a need to include indicators to best control for conditions in the labour market and economy more generally, since the returns to some qualifications will to a greater or lesser extent reflect growth in the economy.

Sample size

The analysis cuts across many dimensions of interest, such as the rich variety of learning pathways (e.g. qualification subject, level, funding stream) and the characteristics of individuals. Sufficient sample size will be required to support the analysis of returns classified across these many dimensions.

4.3 Data sources

Here we describe which data sources the reviewed papers used (as summarised in Table 4.1) to best match the requirements.

RP47 and RP48 use matched administrative data from the ILR, HMRC and DWP records, but for different periods

The analyses undertaken in research papers RP47 and RP48 use data drawn from the Individualised Learner Record (ILR). The ILR provides information by learning aim on the number of guided learning hours, funding source and delivery mode as well as a number of conditioning variables such as gender, disability and ethnic origin. There is a difference in the time periods used for each report. RP47 uses records spanning the 2002/3 to 2005/6 academic years whereas RP48 uses records covering the 2005/6 and 2006/7

period. Both RP47 and RP48 also use data from the HMRC P14 and P45 administrative records and the DWP National Benefits Database.

For the HMRC and DWP administrative records, RP47 draws data for the period 1999/00 to 2009/10 whereas RP48 uses records spanning the 2003/04 to 2008/09 financial years. In both cases, information on annual earnings and employment spells are taken from the HMRC P14 and P45 records, respectively. Information on benefits spells from the DWP National Benefits Database is used by both research papers. RP48 additionally uses the Index of Multiple Deprivation (IMD) at postcode level as a conditioning variable in some parts of the analysis.

RP53 uses survey data from the LFS and BCS

In contrast to the use of administrative datasets, RP53 uses data from two separate surveys to carry out its analysis. These are the 1996, 2000 and 2004 waves of the 1970 British Cohort Study (BCS) and pooled cross-sections from the Labour Force Survey (LFS) for the years 1996 to 2009. The LFS provides information on qualification attainment, hourly earnings and employment status as well as information pertaining to the socioeconomic and demographic characteristics of individuals and households. The BCS also provides information on qualification attainment, employment status, and hourly earnings as well as detailed information about the respondents' childhoods, including family background and ability scores.

Table 0.1: Comparison of the Datasets

	RP47	RP48	RP53
Datasets used	ILR HMRC P14 HMRC P45 DWP National Benefit Database	ILR HMRC P14 HMRC P45 DWP National Benefit Database	LFS 1970 BCS
Nature of Datasets	Administrative	Administrative	Survey (LFS) Cohort study/ survey (BCS)
Matched	Yes	Yes	No
Aggregate Size	6,868,830	5,651,517	518,802 (LFS) 6,475 (BCS)
Time Span	2002/3–2005/6 (ILR) 1999/00–2009/10 (HMRC/DWP)	2005/6–2006/7 (ILR) 2003/04–2008/09 (HMRC/DWP)	1996–2009 (LFS) 1996, 2000, 2004 (BCS)
Source(s): BIS Research Papers number 47,48 and 53.			

4.4 Properties of the Data

The aggregate datasets used for RP47 and RP48 contain 5-7 million observations, but final sub-samples used for estimation depend on the unit of analysis

The ILR used in RP47 and RP48 is organised by learning aim but has been reshaped to make the individual learners the unit of observation for the purposes of these analyses. The HMRC and DWP administrative datasets are also arranged on an individual basis and in many cases include repeated observations of individuals over time. The size of the final dataset used in the analysis for RP47 and RP48 is ultimately determined by the number of matches between learners in the ILR and the administrative datasets. After linking and cleaning, the aggregate dataset used in the analysis for RP47 contains 6,868,830 observations. The final dataset used in the analysis in RP48 contains around 5,651,517 observations.

In both research papers the sizes of the sub-samples used for the individual regressions depend on the unit of analysis. At detailed levels of analysis the cell sizes are relatively small: for example, in RP48 the cell size for the filtered analysis for males with Level 2 NVQ in science and mathematics courses is only 73 observations.

The survey datasets used in RP53 provide much smaller estimation samples

The underlying structure of the LFS, as used in RP53, is a five quarter revolving panel at the individual level, but it is pooled cross-sections of the quarterly version of the LFS survey that are used in this analysis. Pooled over all years the size of the aggregate sample is 518,802 observations.

The BCS is a longitudinal dataset at the individual level, and three waves are used for the analysis undertaken in RP53. The size of the sample is considerably smaller than both of the administrative datasets and the government administered social surveys. The size of the aggregate sample used in the analysis is 6,475 and this is considerably reduced when disaggregated analysis is carried out.

4.5 Linking, Missing Data and Selection Bias

RP47 and RP48 do not report the details of the matching procedures used

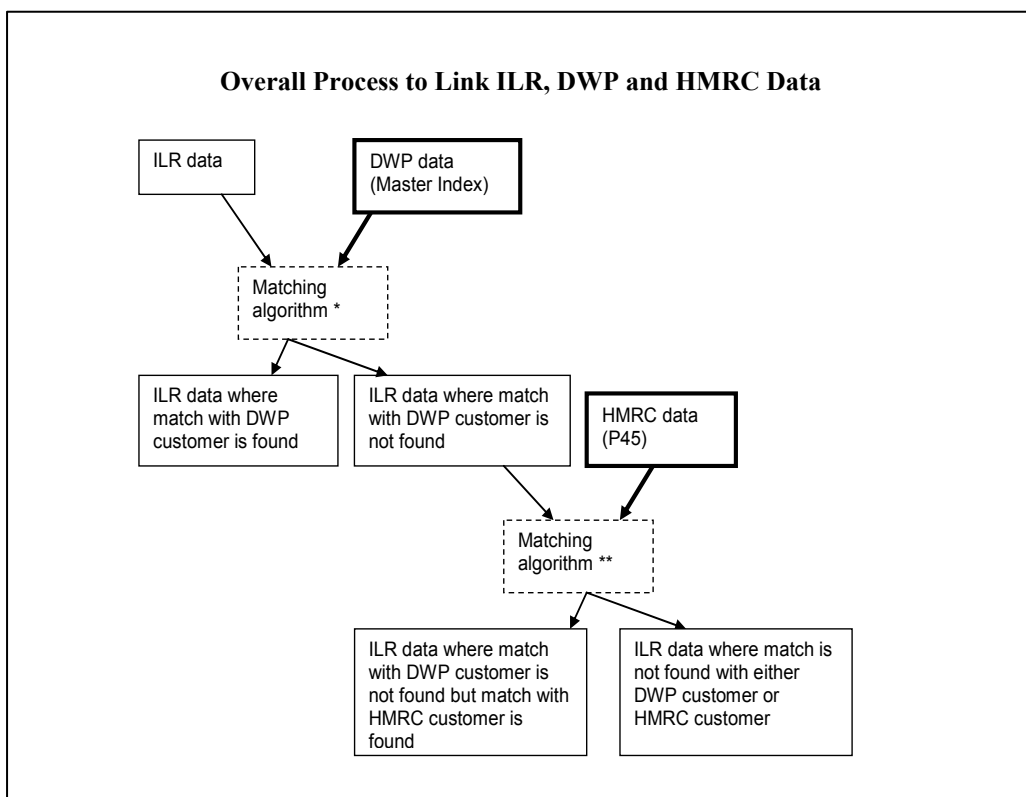
The linked administrative datasets used in RP47 and RP48 were constructed by matching ILR records with administrative tax (HMRC) and benefits (DWP) data. The research papers are somewhat vague over the precise details of how this linking process, and cleaning of the data, was carried out.

The initial matching exercise was implemented and overseen by DWP and BIS, initially in the 'Data Matching Project' during 2009-10. BIS provided the matched dataset to the researchers. We have consulted with BIS analysts to learn more about the matching process and any assessments that were undertaken to evaluate the quality of the

matching achieved. The overall matching process follows the steps as set out in Figure 4.1²⁴.

The ILR and DWP data were first matched, by using National Insurance Number (nino) and other personal details (e.g. initial of forename, date of birth) as the fields for linking. The data were linked using a “fuzzy” data matching algorithm as illustrated in Figure 4.2 below. As the algorithm runs on, the quality of the matches diminishes (from “Green and Amber” to “RedAm”). Having developed and tested the algorithm for matching ILR data to the DWP data, the same one was then used to match the remainder of the ILR data (i.e. those where a match to the DWP data was not found) against the HMRC data.

Figure 0.1: Overall Process to Link ILR, DWP and HMRC Data



Potential biases introduced by the matching procedures

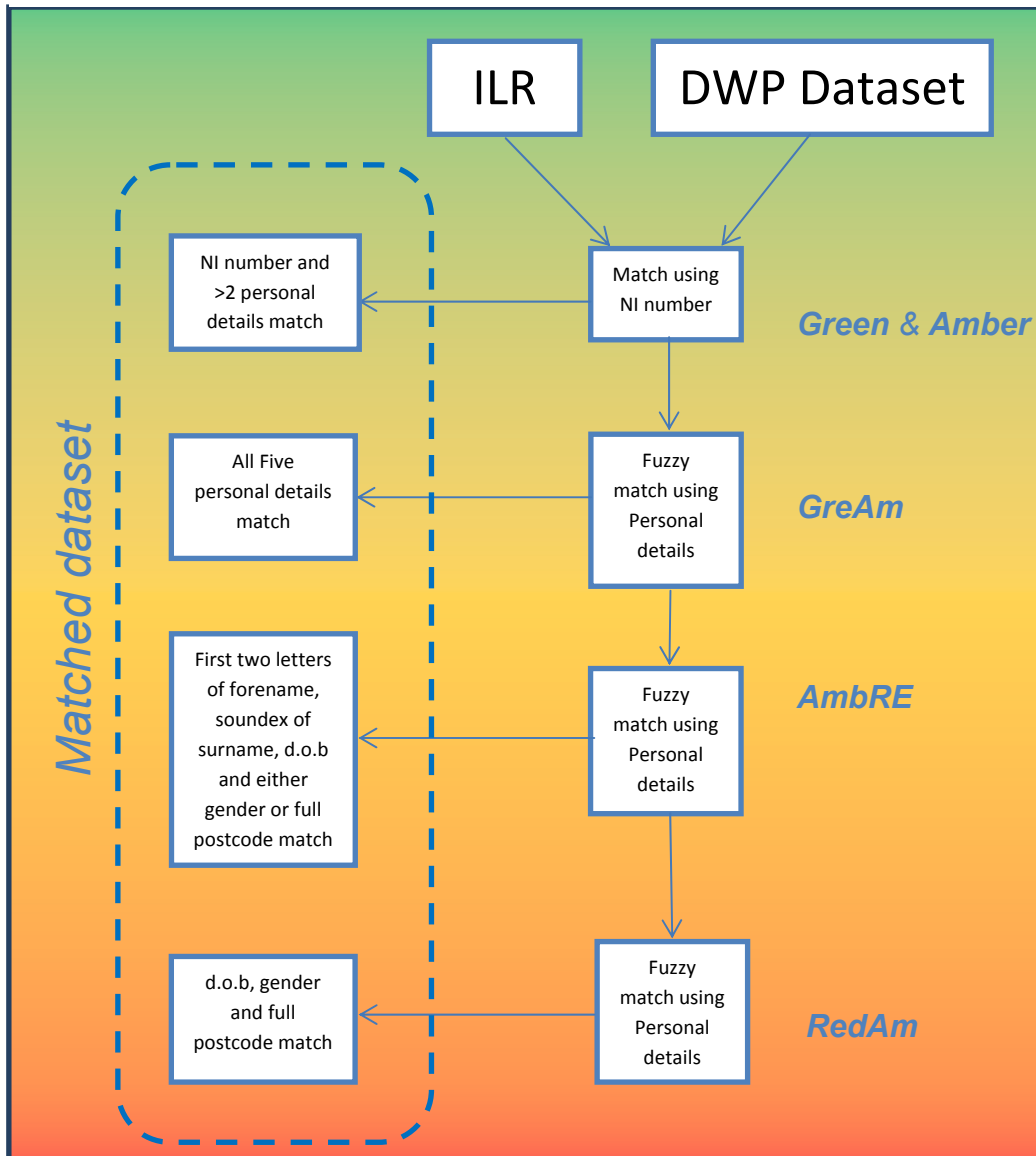
If those individuals that were not successfully matched by the process share certain characteristics then this might introduce bias to the estimates derived from the dataset. For example, bias would be introduced if some characteristic of the non-matched individuals, such as unstable housing tenure, was correlated with the conditional

²⁴ Source: BIS' *Matching Final Technical Report.doc*.

expectation of the dependent variable in the regression analysis (wages, benefit dependency, or employment).

It has been difficult to assess in detail the direction and extent of any biases this may introduce without full details of the matching procedure and information on the characteristics of those individuals who were not matched. However, consultation with and documentation provided by BIS has given some insight into this issue.

Figure 0.2: Example of Fuzzy Data Matching Algorithm



Socio-economic characteristics

The match rate by age was generally similar for most cohorts, except a poor match for those under 16 years, and a higher match rate for the oldest learners (a greater proportion of which are matched to DWP data because they draw pension-related benefits). These age groups account for a relatively small proportion of learners and so this was not considered cause for concern. With regards to other socio-economic characteristics:

match rates for men and women were similar; and by ethnic group, match rates were generally similar except poor rates for 'White Other' and 'Chinese'.

Characteristics of learning

There was some variation in match rates by funding stream. Work-based learning and Train to Gain yielded the highest match rates (around 90%) with adult and community learning (ACL) and European Social Fund (ESF) much lower (at around 60%). By training provider, match rates were generally good, with a small number of institutions with match rates that BIS considered to be poor. By level of qualification, there were some notable differences in match rates: the higher-level qualifications tend to have higher match rates because these learners are more likely to have an employment record in the HMRC data.

BIS concludes²⁵ that *'The matching has shown to be very successful and there are no significant problems with bias. There are small groups where there may be relatively poor coverage but this is unlikely to provide any significant restrictions in using the data.'* It is worth noting that the authors of the research papers used different methods to clean the dataset which would introduce differences to the final data actually used in the analyses. This is evident in the final number of observations that each paper reports: 6,868,830 observations for RP47; and 5,651,517 for RP48. The papers provide little comment on the implications for the potential bias introduced by data matching and cleaning.

There is a large incidence of non-matches to the DWP database

RP47 does explicitly mention a large incidence of matches to the HMRC datasets but non-matches to the DWP National Benefits Database (NBD). This is attributed to a significant proportion of the individuals in the ILR sample not having ever claimed benefits. As part of the data cleaning and merging process, the researchers assumed that those individuals with no match to the NBD did not claim benefits during the period considered. This potentially understates the number of individuals on benefits. It would be informative to assess the characteristics of the individuals not matched to NBD to better understand the potential for bias in the regression results.

RP53 does not address the possibility of attrition and wave non-response bias in the BCS

RP53 uses the BCS to carry out some panel-data analysis, but the possibility of systematic attrition and wave non-response bias is not mentioned in the text of the report. Attrition bias is a type of sample selection problem which is present when the reasons for sample drop-out are correlated with the error term of the regression, i.e. correlated with the dependent variable after conditioning on observables. Wave non-response in the BCS as a proportion of the longitudinal target sample is 30.3% for the 1996 wave and 11.8% for the 2000 wave²⁶. Furthermore 1.8% and 0.4% of the observed sample, for 1996 and 2000

²⁵ In *Initial Assessment of Match Coverage.doc*.

²⁶ Plewis, I., Calderwood, L., Hawkes, D. and Nathan, G. (2004), "Changes in the NCDS and BCS70 Populations and Samples Over Time, *National Child Development Study and 1970 British Cohort Study Technical Report*, Centre for Longitudinal Studies, Institute of Education, University of London.

respectively, did not answer any questions in the employment and financial situation domain which contains those variables related to employment and earnings outcomes.

There are known biases in the LFS measurement of adult educational attainment

Another notable source of potential bias in the LFS is in the measurement of adult educational attainment as discussed in detail by RM Data Solutions (2010). In addition to attrition, recall bias might introduce upward bias if respondents have a tendency to overstate their qualifications. Note also that statistical methods are applied to best impute and apportion responses for: the one-fifth of responses that are made by proxy on behalf of others; and the 10% of responses that report 'Other' as the highest qualification level. The LFS is thought to under report the number of Apprenticeships. If the under-representation is in lower level and less traditional sectors, because individuals were less aware that they had engaged in Apprenticeships, then this might introduce upward bias to estimates of returns to Apprenticeships.

4.6 Suitability of the Indicators

RP47 and RP48 use proportion of the year in employment as the indicator of employment outcomes...

To take account of employment outcomes, both RP47 and RP48 model the proportion of the year spent in employment, calculated from employment spell data from the HMRC P45 records. While this has been found by reviewers to be a generally acceptable indicator, it is noted that it fails to capture broader dimensions of employment outcomes such as the quality of employment (in terms of whether it is long-term, stable, full-time or part-time) and more generally the number of hours worked. Furthermore, it is not possible to determine whether the spells of non-employment were involuntary or whether the individual was out of the labour force.

...whereas RP53 uses a binary employment /non-employment indicator

RP53 uses an alternative measure of employment outcome, estimating the point-in-time probability that an individual is in employment. Again, a problem with this approach is that it groups the economically inactive with the unemployed when they are in fact two fundamentally different groups. Furthermore, an estimate of the probability of being in employment at a given point in time measures an even narrower dimension of employment outcome than the proportion of the year in employment.

Measures of benefit dependency

To measure the impact of qualification attainment on benefit dependency both RP47 and RP48 use the proportion of the year spent in receipt of Job Seekers Allowance (JSA) and Incapacity Benefit. This is generally accepted by the reviewers as an adequate indicator, but, with respect to the inclusion of incapacity benefit, the extent to which qualification

attainment can influence an individual's physical health is an open question. RP53 does not attempt to measure any benefit dependency impacts.

RP47 and RP48 use annual earnings and average monthly earnings, respectively, as outcome indicators

RP47 uses annual earnings as an indicator of earnings outcomes. A problem with this approach is that it is not possible to distinguish between whether an individual is in full or part time employment. A rise in annual earnings as a result of a switch from part-time to full-time employment could therefore be falsely attributed as a rise in wages as a result of qualification attainment.

RP48 uses an average monthly earnings indicator constructed from annual earnings and the number of months in employment, but it is still not possible to distinguish whether a change in average monthly earnings is due to a change in working hours or a change in hourly wages. The authors of RP48 fully acknowledge this weakness of the indicator in the report. The report also contains a filtered version of the analysis, which includes men and women (with analyses presented separately by gender) working only 11 months and with average monthly earnings of between £4,800 and £80,000, in order to try and remove individuals from the analysis with unrealistic average monthly pay indicated. RP53 is able to use hourly earnings, which is preferable to the annual and average monthly earnings indicators.

The survey-based analysis of RP53 benefits from the inclusion of a wide variety of conditioning variables

The equation specifications for the regression analysis undertaken in RP48 and RP53 contain annual dummies to account for the trend in earnings growth and changes over time. The analysis performed using the LFS in RP53 also includes quarterly dummies to account for seasonal effects on earnings. The regression specifications in RP53, in being estimated using survey data, benefit from the inclusion of a wide variety of conditioning variables to account for socioeconomic and demographic characteristics.

4.7 Strengths and Weaknesses of the Data

The main weakness of the administrative datasets is the lack of a robust counterfactual

The principal weakness of the data used in RP47 and RP48 is the lack of availability of a comparison group of individuals that have not enrolled in FE. This prohibits the use of difference-in-difference or matching techniques to develop a counterfactual. Both reports are also careful to point out the existence of gaps in the database and missing data. Some individuals in the ILR could be matched to neither employment nor benefit records. Where matches could be made there was often only partial or conflicting information, for example on the start and end date of employment spells. In terms of the matched ILR records some important information is not available such as progression to higher education. A final weakness of administrative data in general which affects this analysis is the lack of variables available to control for heterogeneity between individuals.

The variables in the administrative databases (especially ILR) are also subject to some changes over time which can cause problems when carrying out longer analyses, with the categories of some variables being changed and other fields being introduced or dropped.

Similarly, with the introduction of Universal Credits, it may be the case that it will be impossible or increasingly difficult to distinguish between the uptake of benefits for reasons that might have some bearing on labour market or learning outcomes.

The main strength of the matched administrative datasets is their size and the information they provide on subject area and acquisition

The key strength of these administrative datasets is their size, containing millions of matched observations. This feature is crucially important if robust analysis at highly disaggregated levels is to be carried out in order to identify the relative returns to the rich variety of learning pathways. Furthermore the match with the ILR provides information on the subject area of the qualification and details on acquisition route, funding type and guided learning hours that is not contained in other datasets. Another key strength is the feature of administrative data that it is not subject to recall bias, which is a well-documented problem with respect to survey data pertaining to earnings and employment histories.

The main weaknesses of the survey data are the small sample sizes and limited information about qualification attainment

A weakness of the LFS and BCS datasets pertains to the limited information they provide about qualification attainment; in the RP47 and RP48 analysis this information is obtained from the ILR. The general problems of recall bias relating to survey estimates of earnings also apply to LFS and BCS. A second major drawback of these survey sources is the much smaller sample size, which limits the level to which the analysis can be disaggregated. This is a particular problem with the BCS analysis. The LFS has a weakness pertaining to a discontinuity in variables over time, including educational attainment variables, as well as the change in administration and collection of the data as a result of the move from seasonal to calendar quarters in 2005/06. This also limits the comparability with previous estimates.

The main strength of the LFS and BCS is the wide range of conditioning variables they contain

The major benefit of the LFS and BCS is the wide range of conditioning variables available to control for the socioeconomic and demographic characteristics of individuals. The LFS also contains information on the occupation and industrial grouping of the individual, thereby allowing analysis by these specific dimensions (although it is not clear how these variables are defined for individuals not currently in employment). The BCS has the specific strength of containing a measure of childhood ability, which can be used to attempt to control for ability bias in the estimates of earnings and employment impacts. Furthermore, it has a panel structure which enables the application of techniques to control for time-invariant individual characteristics which might affect earnings and employment outcomes.

4.8 Summary of the Methodological Approach

In this section, the different methodologies applied are summarised and compared (see Table 4.2).

RP47 estimates a cross-section regression for each year post-attainment of the qualification

There are two main aspects to the research contained in RP47. A first stage analysis is conducted of the determinants of qualification completion. This is specified as a probit and ordered probit model and estimated at learning aim rather than individual level, over the entire sample of qualification completers and non-completers for the years 2002/3 to 2005/6. The significant explanatory variables from these regressions are used to inform the main analysis of earnings, employment and benefit dependency outcomes. This second stage of the analysis estimates a cross-section OLS regression for each year post-attainment of the qualification up to seven years.

Most of the analysis in RP48 is descriptive

The majority of the analysis contained in RP48 is cross-tabular in nature involving a comparison of various descriptive statistics and sample proportions relating to earnings and employment outcomes for groups of learners in the 12 months pre and post-qualification attainment. For the Further Education learning stream only, impact analysis is carried out using fixed effects estimation of earnings, benefit and employment outcomes.

RP53 estimates repeated cross-sections regression using the LFS and a first difference model using the BCS

For the earnings analysis conducted using the LFS in RP53 a series of OLS cross-sectional regressions are estimated for each year. This estimation is conducted on the sample of individuals in full-time employment only. For the analysis of employment returns, a probit model is estimated for the probability that an individual is in employment (the ILO definition) versus unemployment or inactivity. For the earnings analysis conducted using the BCS, separate cross-sectional regressions are estimated for each of the three waves as well as a fixed effects estimation.

The methodological approach adopted in all three research papers does not involve the use of any matching techniques or difference-in-difference estimation. The use of these approaches is limited by the datasets used in RP47 and RP48 as there is limited information available on individual characteristics and there is only information on those who have enrolled for a qualification in the FE system.

4.9 Specification of the Counterfactual

RP47 uses the outcomes of ‘non-completers’ as its preferred counterfactual

As set out in Chapter 3, to measure the returns to training a key methodological challenge is to determine an appropriate comparison group, or counterfactual. The different datasets offer alternative ways of defining the counterfactual (see Table 4.2).

The analysis contained in RP47 considers three approaches to constructing a counterfactual. The first of these is a comparison of outcomes before and after qualification acquisition. The report acknowledges that the returns estimated using this approach will include the time trend and age effect of earnings and not just the impact of the qualification itself. The second approach to the counterfactual is to use the outcomes of individuals simultaneously attaining qualifications at different levels. The report acknowledges the weaknesses of this approach in terms of failing to account for differences in ability and motivation between individuals studying for different levels of qualification. The interpretation of estimated returns is also more complicated using this counterfactual.

The report identifies its preferred counterfactual as individuals enrolled on a course for the same qualification that have failed to complete it. The challenge of this approach is trying to account for factors that affect the likelihood of course completion which may also affect labour market outcomes (e.g. the learner may have failed to complete because of ill-health). In an attempt to do this, the significant explanatory variables from the first stage analysis of the determinants of qualification completion are included as conditioning variables in the “completers versus non-completers” estimations. It is noted that the range of explanatory variables included in the first stage analysis are limited and so would not capture all the factors that might influence non-completion. There are also potentially unobserved factors that determine the completion of qualification attainment, such as motivation and ability that may also impact on labour market outcomes (Hogarth et al, 2010). Another weakness of this approach is that the first stage analysis does not consider all types of course non-completion, such as transfer to another course. Furthermore no account is taken of the possibility that non-completers have undertaken other activities to improve their human capital and labour market prospects or where non-completion has occurred as a result of individuals dropping out to take up viable employment opportunities (Winterbotham et al, 2000).

Table 0.2: Comparison of Methodologies

	RP47	RP48	RP53
Counterfactual	Outcomes of “Non-Completers” Outcomes of simultaneous completers at adjacent levels of qualification Outcomes before attainment	Outcomes before attainment (controlling for fixed effects)	Outcomes of those holding various (and no) other qualifications Outcomes before attainment (for BCS only)
Earnings equation specification	Levels	Differences	Levels Differences
Earnings Equation Estimation	OLS	First difference estimator	OLS Quantile regression (for NVQ level 2 only)

	RP47	RP48	RP53
Methodology			First difference estimator
Source(s)	BIS Research Papers 47, 48 and 53.		

RP48 uses pre-acquisition outcomes as a counterfactual

RP48 rejects the use of non-completers and simultaneous attainers as a counterfactual due to the reasons outlined above. In the absence of a “suitable” counterfactual, fixed effects models of the difference in outcomes before and after qualification attainment are estimated to control for individual-specific, time-invariant unobserved heterogeneity. Annual dummies are included in a variant of this specification to account for time trend effects.

RP53 uses the outcomes of individuals with different (and no) qualifications as counterfactuals

The approach to the counterfactual adopted in RP53 typically involves a comparison of the earnings of individuals with a given qualification to those holding a qualification at an adjacent level down. These have been identified by the reviewers as both meaningful and intuitively appealing. Fixed effects analysis of earnings is also carried out using three waves of the BCS.

4.10 Suitability of the Dependent and Explanatory Variables

Table 4.3 lists and compares the variables included in the research papers reviewed.

The specification in RP47 controls for acquisition route and source of funding

For the earnings analysis conducted in RP47 the dependent variable is the log of annual earnings. A “basic” and “extended” equation is specified, the former containing only individual and socioeconomic control variables and the latter including also the benefit and employment outcomes. The problems with using (the log of) annual earnings as a dependent variable without including hours worked as an explanatory variable have been previously discussed in Section 4.6 Suitability of the Indicators. Furthermore it has been noted that prior educational attainment does not appear to have been accounted for in the specification and that the highest educational attainment is not properly captured because information on whether a learner progressed to higher education is not included. The control for progression to higher education and the number of hours worked are also missing in the equation specification used in RP48 which uses a derived measure of (the log of) average monthly earnings as a dependent variable.

The specifications in RP53 contain a more comprehensive range of conditioning variables

The analysis conducted in RP53 uses survey data which offers a wider range of explanatory variables. The inclusion of childhood ability test scores in the BCS analysis

goes some way towards addressing the problem of ability bias. The use of (the log of) hourly earnings as a dependent variable in the earnings outcome analysis means that it is not necessary to control for the number of hours worked. A weakness of the specification of the dependent variable in the employment outcome analysis is that the conceptually different labour market states of unemployment and inactivity are grouped together.

Table 0.3: Comparison of Variables

	RP47	RP48	RP53
Earnings Analysis Dependent Variable	Log of annual earnings	Log of average monthly earnings	Log of hourly earnings
Employment Analysis Dependent Variable	Proportion of year in employment	Proportion of year in employment	Binary variable indicating: Employment Non-employment
Benefit Dependency Analysis Dependent Variable	Proportion of year in receipt of benefits	Proportion of year in receipt of benefits	n/a
Explanatory Variables in Earnings Analysis	<p>Basic Specification:</p> <p>Age</p> <p>Age squared</p> <p>Gender</p> <p>Disability</p> <p>Ethnic origin</p> <p>Funding type</p> <p>Acquisition route</p> <p>Index of Multiple Deprivation</p> <p>Region of residence</p> <p>Year</p> <p>No. of guided learning hours</p> <p>Extended Specification:</p> <p>All variables in basic specification</p> <p>Proportion of year in receipt of benefits</p> <p>Proportion of year in employment</p>	<p>Basic Specification:</p> <p>Time-invariant individual-specific characteristics are differenced out</p> <p>Alternative Specification:</p> <p>Annual dummies</p> <p>Time-invariant individual-specific characteristics are differenced out</p>	<p>Basic Specification:</p> <p>Age[†]</p> <p>Age squared[†]</p> <p>Gender[†]</p> <p>Ethnic origin</p> <p>Region of usual residence[†]</p> <p>Qualification[†]</p> <p>Marital status</p> <p>No. of dependent children</p> <p>Temporary/permanent contract[†]</p> <p>Public/private sector</p> <p>Full-time/part-time[‡]</p> <p>Labour market experience[‡]</p> <p>Disability or long-term illness[‡]</p> <p>Interaction terms[†]</p> <p>Seasonal dummies[†]</p> <p>Annual dummies[†]</p> <p>Extended Specification (BCS average return estimation only):</p> <p>Variables in basic specification</p> <p>Parental social class</p> <p>Parental education</p> <p>Parental interest in child's education</p> <p>Childhood ability score</p> <p>Free school meals aged 10</p> <p>Financial hardship aged 16</p>
Note(s):	<p>† Denotes variables included in the LFS specification only.</p> <p>‡ Denotes variables included in the BCS specification only.</p>		
Source(s):	BIS Research Papers 47,48 and 53.		

4.11 Significance of the Statistical Inference

The size of the estimation samples used in RP47 and RP48 results in statistically significant estimates in most cases

The large size of the datasets used in RP47 and RP48 leads to statistically significant results in most cases, although the estimations at some disaggregated levels of analysis fail to produce statistically significant relationships. It should be borne in mind that there is a limit to the inference that can be drawn from the results. The nature of the data, in particular the lack of information on those without qualifications, inhibits causal interpretation of the results. The LFS datasets used in RP53 provide much smaller estimation samples and as a result there is a much higher incidence of effects that are not significant at conventional levels. This is particularly the case for the estimates of the earnings effects by occupation and sector. The BCS dataset provides an aggregate sample of only 6,475 observations and most coefficients estimated using this dataset are not statistically significant.

4.12 Strengths and Weaknesses of the Econometric and Statistical Techniques

The main weakness of the analysis carried out in RP47 and RP48 is the specification of the counterfactual

The econometric techniques employed in RP47 and RP48 are limited by the nature of the datasets used, in particular that the observations are only over those having enrolled on a qualification in the FE system. Furthermore the administrative nature of the datasets provides a less than comprehensive set of conditioning variables, as discussed in Section 4.10 Suitability of the Dependent and Explanatory Variables. RP47 uses “non-completers” as its preferred counterfactual and attempts are made to control for the determinants of course completion. As discussed in Section 4.9 Specification of the Counterfactual, biases are likely to remain.

The majority of the analysis carried out in RP48 is descriptive in nature and therefore cannot be used as a robust estimate of the earnings, employment or benefit outcomes of qualification attainment. Some impact analysis is undertaken for further education attainment only. This uses a fixed effects approach to difference out individual-specific, time-invariant unobserved heterogeneity. A constant is also included in the equation specification to take account of the time trend in earnings. Clustered standard errors are used to address correlation in error terms for a given individual over time. Nevertheless, the identification of causal effects is not possible from the econometric approach.

The lack of information on hours worked combined with the use of a derived measure of average monthly earnings in RP48 results in some unrealistic values of the dependent variable (see Section 4.6 Suitability of the Indicators for a discussion of this). To deal with this, the authors of the report also perform the analysis on a filtered version of the sample containing observations over a restricted range of the dependent variable. Quantile regression estimates of the earnings returns may have proven useful in this situation given that they are more robust to outlying observations than estimates of the conditional mean.

With respect to the analysis of benefit dependency outcomes, the text of RP47 contains a comment on the dependent variable containing a large proportion of zeros, in part because of the assumption made by the researchers that individuals with no match to the DWP's NBD did not claim benefits during the period considered (in RP48 reference is not made to the large proportion of zeros even though it uses data from the same sources and the same indicator). The econometric methodology applied does not take account of this and the estimates are likely to be biased as a result.

The primary weakness of the earnings analysis carried out in RP53 is the failure to model the process of selection into the estimation sample

A key weakness of the econometric approach adopted in the analysis of earnings returns in RP53 is the failure to account for labour market participation. The estimation sample is over those in full-time employment only and the process of selection into the sample is not modelled, for example using a Heckman-type selection model. The strength of the econometric techniques applied in RP53 is the inclusion of a wide range of conditioning variables, in particular the use of childhood ability scores in the BCS analysis to attempt to control for ability bias. The longitudinal structure of the BCS was also used to carry out fixed effect analysis of earnings returns. However the panel nature of the dataset was not fully exploited in a way that could give rise to causal inference. The issue of attrition and wave non-response bias in the BCS was also not considered, or even mentioned in the text of the report.

In terms of the employment outcome analysis undertaken in RP53, a binary outcome of either employment or non-employment is modelled using a probit regression. A conceptually preferable approach would be to model the states of unemployment and inactivity separately using, for example a multinomial logit or sequential logit model.

A final comment relating to the analysis undertaken in all three research papers is that the full equation specifications for each type of analysis have not been explicitly set out and this is a limitation to a review of the appropriateness of the econometric methodology.

5 Key Findings

5.1 Introduction

This chapter presents the key findings of the reviewed research papers and, where possible, draws comparisons across the papers and with the existing evidence base for the economic returns to vocational qualifications.²⁷

Issues to consider when comparing the evidence

In comparing the results obtained by the authors of RP47, RP48 and RP53, it is not possible to directly compare the magnitude of their estimates due to methodological differences as well as differences in the datasets used (particularly comparing RP53 to the other two RPs).

There are differences in the comparison groups (counterfactuals) used

One major difference which impacts on the estimates is the comparison or counterfactual used in the analysis by each study. The choice of counterfactual can be severely limited by data and so the researchers develop alternative ways of defining the counterfactual. Whilst earlier work in this area has focussed mainly on the returns to qualifications by comparing individuals with a qualification at a particular level to a similar individual with the next lowest qualification, the papers under consideration here use a variety of comparisons including:

- individual learners' outcomes before and after completion of their qualification ('before and after');
- individuals who attained different levels of qualification at the same time ('simultaneous attainment') ; and
- individuals who completed a particular qualification compared to those who started but did not complete the same qualification ('completers v non-completers').

RP53 defines the counterfactual in a similar way to much of the earlier work that utilises survey data. LFS data are used to estimate *marginal* returns by comparing the returns to individuals with a qualification at a particular level with returns to similar individuals with the next lowest (or no) qualification. Individuals can be identified as similar (by a limited subset of characteristics) using the control variables that are available in the survey data. This counterfactual is a well-established approach which is considered both meaningful and intuitively appealing. However, it fails to take account of differences in ability and motivation of individuals studying for different levels of qualification.

²⁷ Section 3.4 *What are the Benefits of Training and Qualifications?* summarises the existing evidence base.

RP47 investigates the use of all three of the above counterfactuals, but although the authors conclude that “no matter which counterfactual is used, there are strong positive and persistent returns to completing a qualification at a given level”, there are weaknesses to the counterfactuals used. The magnitude (and in some cases the sign) of the estimates are sensitive to the choice of comparison group:

1. Before and after qualification acquisition – fails to take account of factors that change over time (e.g. age effect).
2. Simultaneous completers at adjacent levels of qualification – fails to take account of differences in ability and motivation of individuals studying for different levels of qualification.
3. Non-completers – fails to fully take account of: factors that affect the likelihood of course completion which may also affect labour market outcomes (e.g. the learner may have failed to complete because of ill-health); all types of non-completion (e.g. transfer to other course).

Thus, the use of different counterfactuals can make it difficult to compare results between different studies. Indeed, different counterfactuals can be used to answer different questions, in which case it is not relevant to make direct comparisons.

There are differences in the outcome variables (returns) used

As discussed in Section 4.6, in addition to differences in the comparison group, the outcome variables also differ between some of the analyses with hourly, weekly and monthly pay figures being used in some instances and the impact of qualifications on employment considered in terms of the proportion of the year an individual was in employment and whether or not an individual is in employment at a particular point in time. RP53 is the most readily compared to previous work in this area because it uses LFS data and employs similar methods in estimating the returns. The authors also refer to the existing estimates of returns set out by Jenkins and De Coulon (2007) and McIntosh (2007). The differences in approach and measurement should be kept in mind throughout when considering particular estimates of the returns to qualifications and especially when trying to compare results across groups.

The use of hourly earnings is preferable for identifying the potential impact of training on an individual’s productivity. For example, when comparing the results for the impact of training on earnings, we might expect the results of analysis using weekly or monthly earnings as the outcome variable to be higher than the results using hourly earnings as the outcome variable. This is because it is not possible to distinguish between whether an individual is in part-time or full-time employment, and so an increase in weekly or monthly earnings which is due to a switch from part-time to full-time employment could be falsely attributed to training.

Potential biases

As discussed in Section 4.5, it is possible that the data matching and cleaning process used to construct the datasets used for the analysis in RP47 and RP48 could have

introduced a number of biases. Although BIS has stated that *'The matching has shown to be very successful and there are no significant problems with bias'*²⁸, it is possible that the data cleaning conducted after the matching could have introduced biases to the data.

RP47 mentions a significant number of non-matches of ILR to DWP data, where HMRC data were matched to ILR, which are all assumed to be due to those individuals never having claimed benefits. This could potentially underestimate the number of individuals on benefits, as there could be other reasons for the non-matches.

The BCS data used in RP53 is susceptible to attrition and wave response bias, and LFS data used for that paper is also known to be subject to recall bias, including under-reporting of Apprenticeship qualifications.

These biases could all impact, to a greater or lesser extent, on the accuracy of the results in each RP. For example, current under-reporting of Apprenticeships in the LFS (which is apparent because the number of respondents that report holding an apprenticeship has not increased in recent years despite the rapid expansion of the programme) could be due to individuals not recognising the qualification as an apprenticeship. If there is correlation between those that do not recognise their Apprenticeship qualification and receiving a poorer standard of training, then this could potentially bias the estimated LFS-returns upwards, because those with lower returns are excluded from the analysis.

Estimates of the returns must take account of the heterogeneity across FE

Within further education (FE) there is a range of programmes and qualifications which vary not only in content, delivery and funding but also in terms of the outcomes obtained by those who undertake such qualifications. There is also wide variation in the characteristics of learners themselves and the sectors in which training takes place. As was highlighted in Chapter 3, the heterogeneity across FE, within programmes and amongst individuals, must be taken into account when considering and estimating the returns to FE. This is vital in order to establish the programmes or types of training and education which lead to the highest returns and for which types of individuals the returns are greatest.

The extent to which economic returns might vary for different individuals and in different circumstances is of key interest, and so this chapter presents the findings as follows. Section 5.2 discusses the findings for different types of qualification (broken down by funding stream, level, sector and subject). Note that the issue of 'first' versus 'non-first' qualifications of a particular stream or Level is not fully addressed in the papers reviewed; however, the descriptive analysis in RP48 (Annex 2) does provide some assessment of the returns to 'first' qualifications. In Section 5.3 the results for different types of individual (distinguished by socio-economic group, age, gender etc) are presented. Section 5.4 comments on the evidence on long-term returns and their persistence. Section 5.5 considers in which circumstances it is preferable to use the evidence yielded by which approach.

²⁸ In *Initial Assessment of Match Coverage.doc*.

5.2 Returns by Funding Stream, Level and Subject

In a previous review of the data required to assess the value of Apprenticeships, Gambin *et al.*, (2011) drew attention to the need for analysis to be conducted by Framework or subject, and Level. Since the content of training varies so much according to the subject of study and the Level at which it is studied, the returns are likely to vary accordingly. Cross-sectional surveys such as the Labour Force Survey (LFS) or the various longitudinal cohort studies which have been periodically conducted over the past 30 years either do not collect data on subject studied or due to sample sizes are unable to provide robust results. Administrative data sets, such as the ILR, in contrast, provide a wealth of information about the subject studied, which has been reported in-depth in RP47 and RP48.

Previous research has highlighted the variation in the returns associated with different funding streams, levels and subject areas. Apprenticeships, for instance, have been associated with notably higher returns in terms of wages and employment (McIntosh, 2009, 2007). Within the Apprenticeship programme itself, the returns again vary between different sectors, levels and frameworks, with higher returns found for more traditional apprenticeship sectors (e.g. engineering, construction) and for Level 3 compared to Level 2.²⁹

The returns to qualifications by different funding streams (e.g. Apprenticeship, FE, Skills for Life) and Levels have been explored to varying extents in the three BIS RPs considered in this report. Also considered, albeit in a relatively limited fashion, is the sector in which learning has taken place. The advantage of administrative data comes to the fore in this regard as the ILR provides the greatest level of detail and accuracy on learning aims within the FE system, including sector subject area, level and funding stream (along with other features). For example, the administrative data provide the opportunity to investigate the economic returns to basic skills.

Funding stream and level

On the whole, the three RPs find that higher levels of attainment are associated with higher returns in terms of wages and employment. With respect to reductions in benefit dependence the results are more mixed with lower level qualifications, including Level 1, resulting in relatively greater reduction in benefit usage (RP47).

RP53 finds significant returns in terms of wage gains for most qualifications with the gain being higher for Level 3 qualifications compared to those at Level 2. Individuals with a Level 3 BTEC had an average 20% higher wage than individuals with a similar qualification at Level 2. The result was 16% for RSA Level 3, and 10% for NVQ Level 3. For Level 2 BTEC, the gain was 12%, 16% for RSA and 1% for NVQ Level 2 compared with similar qualifications at below Level 2. Table 5.1 summarises the findings by qualification and level from RP53. These results are very much in line with the results from

²⁹ We noted above the potential bias that might be introduced due to under-reporting of Apprenticeships in the LFS; if those that do not recognise their Apprenticeship qualification are more likely to train in less traditional sectors, this

previous work including Jenkins et al (2007). The analysis found no significant difference in the returns according to whether qualifications were obtained in the workplace or in school/college. The highest employment returns are found for NVQ Level 3 (15% increase in probability of being employed) and NVQ Level 2 (11.8% increase compared to less than Level 2). Apprenticeships have the highest overall wage and employment returns at Level 3 (22% and 14%, respectively) whilst the returns to Level 2 apprenticeships also give rise to significant positive wage and employment returns (12% and 10%, respectively).

Table 0.1: Wage and employment returns by qualification type and level using LFS, RP53

Qualification	Wages (%)		Employment (percentage points)	
	Level 3 v Level 2	Level 2 v less than Level 2	Level 3 v Level 2	Level 2 v less than Level 2
BTEC	20	12	7.6	8.7
RSA	16	14	5.8	9.1
NVQ	10	1	15.2	12.5
C&G	15	7	13.6	11.8
Apprenticeship	22	12	14.0	10.0

Source(s): based on estimates provided in RP 53.

Some of the differences in RP53 compared to the results of other studies (e.g. McIntosh 2007) are partially attributed to the use of hourly wages rather than weekly wages as the latter includes then the effects of qualifications on hours worked as well as increase in hourly pay.

RP47, which compares mainly completers to non-completers, also finds positive and significant wage returns to various types and levels of qualification. The analysis considers the returns for a number of years after completion of a particular qualification. They find that for most qualifications, wages are higher for those who completed compared to those who did not. At Level 3 (all qualifications), the wage returns are negative for the first five years after completion but are positive in the sixth and seventh years after completion of the qualification (in the years following attainment completers of GCE A' levels are more likely to then progress to further or higher education and so remain as relatively low earners for a number of years). Similarly at Level 2 (all qualifications) the results are negative for the first two years after completion and are positive from the fourth to seventh years post-completion. The wage returns to Level 4 and Level 1 are positive for all years considered (1 to 7 post-completion) and increase from 3% in the first year to 11% in the seventh year for Level 4 qualifications. At Level 3, NVQ completers were found to have 14% higher wages in the seventh year post-completion compared to individuals who started but did not complete the same qualification.

Obtaining any qualification was found to increase the proportion of the year an individual was in employment. Larger 'impacts' were found for NVQ qualifications than for other types of qualifications. In general, a Level 2 qualification was found to increase the portion of the year a person was in employment by 2.2% in the first year post-completion. For

Level 3 qualifications, the employment return was negative in the first 2 years after completion but by the seventh year, the proportion of the year in which a person was employed was 6.1% higher for completers versus non-completers. Considering a similar comparison group as that explored by other studies and by RP53 (simultaneous attainment), RP47 also finds significant, yet smaller, results in the effect of qualifications on the proportion of the year in employment.

In terms of benefit dependence, Level 3 qualifications were associated with a 1 to 2% decrease in benefits when comparing completers to non-completers. Completing any qualification was found to be associated with lower benefit dependency.

Apprenticeships receive attention in these RPs as it is a substantial funding stream within the FE system and previous work has shown different returns to this programme compared to others (e.g. Btec, NVQs). The RPs confirm earlier findings that there are positive wage returns associated with completion of Apprenticeship training but the magnitude of the estimates differs across studies and even within studies by model specification.

On the whole then, the estimates of labour market returns, especially wage returns, using the linked administrative data follow a remarkably similar pattern to the estimates that use LFS data, as summarised in BIS' own analysis in Table 5.2 below. The exception is the case of Level 2 NVQs, for which the LFS analyses estimate negligible wage effects whereas the linked administrative analyses show significant labour returns.

Table 0.2: Wage and employment returns by qualification type – comparison of LFS and matched data estimates

Level	Qualification type	Wage returns (%)		Employment returns (%)	
		LFS	Matched data ¹	LFS	Matched data ¹
Level 3	Advanced Apprenticeship	22	17	14	5
	BTEC	20	13	8	9
	City and Guilds	15	5	14	6
	NVQ/SVQ	10	11	15	7
Level 2	Intermediate Apprenticeship	12	12	10	5
	BTEC	12	12	9	9
	City and Guilds	7	6	12	5
	NVQ/SVQ	1	10	13	7

Source(s): BIS, from RP47.

Note(s): 1 The 'average' figures for the matched data are a simple arithmetic mean of the annual returns over years 4 to 7 after attainment of the qualification. Omitting the first three years from the average allows some time for the impact to stabilise.

Sector

A distinction needs to be made between sector and subject. In some analyses, wage returns are estimated for the sector in which the individual is employed at a given point in time. This provides no information about the subject in which an individual obtained a

given level of qualification, simply that they are employed in that sector. This section considers the results in relation to sector whilst the next section looks at the subject of study.

RP53 – based on pooled cross-sectional analysis of the LFS and analysis of the British Cohort Survey 1970 - finds that the returns to NVQs vary significantly by sector with more substantial returns found for men in such sectors as manufacturing and construction whilst greater returns are found for women in service sectors. In the analysis presented in RP53, those with C&G qualifications at Level 3 are found to have the highest marginal returns (compared to Level 2 qualifications) in the Manufacturing sector (15%), though there are marginal returns of more than 10% in Energy and Water, Construction, Distribution and Transport and Communication.

The average returns to training (which incorporates information on all qualifications held, not just the highest) reported in RP53 vary across sectors. A Level 3 C&G is reported to give rise to larger average returns in manufacturing (14%), energy (12%) and transport (12%) than in other sectors. BTEC L3 shows noticeably high returns in agriculture (26%), energy (19%) and manufacturing (17%). For the NVQ, positive returns are observed for only L3; this is highest in construction (14%). Trade apprenticeships give rise to significant positive average returns only in manufacturing (7%), construction (8%) and transport (7%).

The marginal returns to qualifications, which consider only the highest qualification held, are estimated to be higher than the average returns in most cases. At Level 3, the RSA gives the highest marginal return in transport (21%), the returns to C&G Level 3 are highest in manufacturing (15%) and energy (14%). The returns to BTEC are highest in manufacturing (18%), construction (18%) and transport (17%). The highest return to NVQ Level 3 is in construction (16%).

The marginal returns to Level 2 qualifications (compared to Level 1) vary considerably across the sectors but notably high returns are found for RSA in energy (59%) and manufacturing (30%). Similarly, the marginal returns to BTEC Level 2 are relatively high in construction (34%) and the C&G Level 2 gives rise to a 15% marginal return in energy.

Subject

RP48 considers the wage and employment returns by qualification level and subject as well but using only descriptive analysis of the data – their empirical analysis focuses on gender and level. In interpreting the differences by sector (or any other characteristic or factor) based on descriptive analysis only, it must be kept in mind that a number of additional factors have not been controlled for and that there is no basis on which to infer causal relationships – the authors note that *“observed difference in outcomes by a certain characteristic cannot be taken as evidence that the characteristic has an effect on those outcomes.”* The descriptive analysis does however allow for patterns in the data to emerge.

Mean (filtered³⁰) earnings are found to increase by 9% for those who complete FE qualifications in construction, planning and built environment subjects and by 7% for those completing engineering and manufacturing qualifications. Mean (filtered) earnings were found to decrease for those completing studies in leisure, travel and tourism; arts, media and publishing; and history, philosophy and theology (by 1% in each).

Based on the descriptive analysis presented in RP48, the employment rate before and after learning improves most (on average) for those who have undertaken the Preparation for Life and Work course (increase of 10pp in the share of learners in employment). Individuals obtaining FE qualifications in education and training exhibit a 7pp increase in the employment rate. The benefit rate at six months before and after training is found to decrease by between 1 and 4pp for learners aged 19 years and older - there is little variation by sector in the change in benefit rate.

Focusing mainly on earnings, Table 5.3 summarises the change in mean (filtered) earnings by sector and level of FE learning as set out in the descriptive statistics in RP48. Across many of the qualifications and levels considered, the gain in earnings after learning is highest in construction, planning and the built environment. Health, public services and care is associated with some of the lowest gains for many qualifications and levels as is the case for leisure, travel and tourism. The largest earnings increases are found for Apprenticeships in construction, planning and the built environment and in particular at Level 3.

Table 0.3: Changes in mean (filtered) earnings pre- and post-learning by study area / sector and level, RP48

Qualification Level	Average	Sector with greatest increase in earnings	Sector with lowest increase in earnings
FE BL2	2%	Social Sciences (5%)	History, philosophy and theology (-1%)
Skills for Life	5%	n/a	n/a
FE FL2	4%	Construction, planning and the built environment (8%)	Leisure, travel and tourism (-4%)
FE L3	5%	Construction, planning and the built environment (15%)	Leisure, travel and tourism (-9%)
FE FL3	8%	Construction, planning and the built environment (28%); Engineering and manufacturing technologies (27%)	Social Sciences (-3%)
Apprenticeship FL2	30%	Construction, planning and the built environment (53%)	Health, public services and care (21%)
Apprenticeship FL3	43%	Construction, planning and the built environment (80%)	Health, public services and care (28%); Information and communication technology (28%)

³⁰ The authors of RP48 filter earnings in order to eliminate the effect of those individuals whose earnings were zero before learning. The authors limit their analysis to those who earn at least £4,800 per annum and calculate separately the average earnings of men and women who were employed for at least 11 months either side of the learning window under consideration.

Qualification Level	Average	Sector with greatest increase in earnings	Sector with lowest increase in earnings
Train to Gain FL2	-2%	Health, public services and care (2%)	Agriculture, horticulture and animal care (-3%); Leisure, travel and tourism (-3%); Preparation for life and work (-3%); Business, Administration and Law (-3%)
Train to Gain FL3	-4%	Engineering and manufacturing technologies (0%); Construction, planning and the built environment (0%); Information and communication technology (0%)	Health, public services and care (-7%)
Source(s): Adapted from RP 48.			

5.3 The Socio-Economic, Demographic and Educational Characteristics of Individuals

The returns to qualifications by the characteristics of individuals have also been considered, though in a limited fashion, in the three RPs. Most of the analyses presented in the RPs consider the returns to men and women separately. So, all the papers provide comment on gender but little attention has been paid to other characteristics, most significantly prior educational attainment has largely been overlooked in the analyses. The LFS analyses (e.g. RP53) usually compare highest qualification relative to someone at the level below, and so implicitly assume that the learners and comparison group have equivalent prior attainment. Note that BIS research currently underway, making use of the detail available in the matched administrative dataset, is looking into the impact of controlling for prior attainment.

The data contained in the linked ILR-HMRC/DWP dataset provides only a limited number of control variables related to the individual. The data allows for age, sex, region (and associated measures of deprivation or other economic conditions) to be included. The LFS data on the other hand, presents more individual-level data such as occupation, other training undertaken (in reference) period and other qualifications held. In RP53, further analysis is carried out using data from the BCS70 which in addition to the control variables available in the LFS, also allows for control of such factors as parental social class, parental education level, parental interest in child's education, ability as child, and measures of financial hardship and free school meals. Differences in the returns to qualifications by these various individual level factors are not the focus of the analysis presented in RP53.

Age

RP53 considers returns by age group with a greater return in earnings associated with those who obtain vocational qualifications when under 25 years of age. Table 5.4 below indicates the returns to various Level 2 and Level 3 qualifications by age group. Only for BTEC Level 2 are the returns for those over the age of 25 years higher than for younger groups.

RP48 (Annex 2) provides more detailed analysis based on the administrative data sets. The findings are broadly consistent with those in RP53 in that the returns to learning are highest for the 16-19 year old age group (i.e. the age at which qualifications are obtained) and lowest for the 25 year plus age group. The findings are also disaggregated by level and subject of study such that a general pattern emerges of the returns being higher, in general, across subjects, bearing in mind that some subjects are associated with higher returns in general.

So if one looks at the employment rate of men, with a first Level 3, in construction, the raw difference in monthly earnings (filtered) is £602 for 16-19 year olds, £559 for 20-24 year olds, and £286 for those aged over 25 years. The comparable findings for a first Level 3 in retail are £477 for 16-19 year olds, £309 for 20-24 year olds, and £50 for those aged over 25 years.

Table 0.4: Wage returns to Level 3 qualifications by age group (%), RP53

Age	Level 3			Level 2		
	C&G	BTEC	NVQ	C&G	BTEC	NVQ
15-16	11	12	1	3	-12	1
17-18	11	15	9	5	7	4
19-20	14	15	11	12	2	4
21-25	14	13	11	11	10	7
26-30	6	12	6	9	17	0
>30	4	7	-1	6	0	-6

Source(s): RP53.
Note(s): shaded cells denote lack of statistical significance.

Gender

Typically, the wage returns to vocational qualifications are found to be higher for men than women. In terms of employment returns, RP47 reported a higher impact at Level 3 for men but a greater impact of Level 2 qualifications for women.

RP53 provides the most comprehensive and meaningful analysis by individual characteristics whilst RP48 provides, for the most part, descriptive analysis. RP47 uses a comparison of completers to non-completers thus the results are not directly comparable to previous studies and such an approach is not intuitively appealing as such a comparison will indeed be prone to biases as individuals who do not complete a course are likely to differ significantly from those who do complete.

As noted in the previous section, the sectors in which returns are relatively higher differ between men and women with more substantial returns for men in the traditional industries of manufacturing and construction and strong returns found for women in service-oriented sectors. RP53 also reports that returns to C&G Level 3 are typically higher for men than women in all sectors. For men in possession of NVQs, there are significant returns associated with employment in the Construction sector (16%). Relatively speaking, women appear to achieve strong benefits from NVQ Level 3 qualifications in the Manufacturing sector, and the Distribution, hotels and restaurants sector. As with the results more

generally, women with RSA Level 3 qualifications achieve significant earnings premiums across most sectors of the economy.

RP53 also reports significant variation in wage returns by occupation, with higher returns to qualification found in lower skilled occupations e.g. men working in skilled trades and machine operative occupations, and for females working in personal services.

RP47 presents results from additional econometric analysis for three different comparators in Annex 3. Within this, the analysis is performed for all learners as well as for men and women separately. Table 5.5 sets out the estimated earnings effects for men and women by type and level of qualification and for each of the three comparison definitions. The estimated returns are much higher for both sexes under the 'before and after' comparison relative to the other two approaches.

Table 0.5: Earnings returns by qualification type and level for men and women under different comparison groups (RP47)

Qualification level	Before and after (at t+4)		Completers versus non-completers (t+7)		Simultaneous attainment (t+7)	
	men	women	men	women	men	women
Level 4 (all)	55%	40%	11%	11%	13%	22%
Level 3 (all)	34%	35%	8%	7%	6%	10%
Level 2 (all)	23%	27%	4%	7%	-2%	4%
Level 1 (all)	19%	21%	5%	1%	0%	0%
NVQ Level 3	35%	49%	12%	17%	8%	14%
BTEC Level 3	88%	79%	10%	8%	7%	13%
C&G Level 3	74%	23%	21%	-3%	17%	20%
Academic Level 3	22%	19%	3%	15%	-1%	15%
Other Level 3	25%	27%	8%	0%	4%	7%
NVQ Level 2	-21%	30%	10%	16%	0%	2%
BTEC Level 2	-19%	32%	0%	10%	-3%	1%
C&G Level 2	-17%	39%	10%	-1%	15%	9%
Skills for Life Level 2	-10%	31%	4%	8%	2%	5%
Academic Level 2	-5%	42%	3%	2%	0%	1%
Other Level 2	-14%	24%	2%	5%	-4%	4%
Skills for Life Level 1	-12%	27%	6%	-5%	--	--
Skills for Life Entry Level	-13%	33%	-4%	7%	--	--
Other Level 1	-14%	19%	5%	2%	--	--

Source(s): Adapted from Annex 3, BIS RP47.
Note(s): shaded cells indicate estimates not statistically significant at 10% level or less
Bold typeface indicates higher returns for particular sex for same qualification.

The 'simultaneous attainment' estimates tend to be the lowest amongst the three comparatives, particularly for men. Comparing 'before and after' earnings shows returns greater for women than men for a number of qualifications and levels, mainly those at Level 2 and lower. At Level 3, the returns are higher for men than women for all

qualifications except NVQ and other qualifications. This result is similar for the ‘completers v non-completers’ analysis with the returns for men being higher than women’s for Level 3 BTEC and C&G. Under the simultaneous attainment analysis however, all Level 3 qualifications presented in the table are associated with higher returns for women than men. The results shown in Table 5.4 highlight the sensitivity of estimates of the returns to qualifications to the counterfactual group that is used in any analysis.

Ethnic Group

Only RP48 presents any results by ethnicity. These results are based only on descriptive statistics of the dataset rather than econometric analysis thus they should be interpreted with caution. Across all sectors earnings increased by 4% for FE achievers from ethnic minorities compared to 3% for white learners and 3% overall. Differences between ethnic minorities and white learners vary by the type and level of qualification. For FE BL2 and FL2, the change in earnings is greater for learners from ethnic minority groups than for whites. For FE FL3, earnings increased by 7% for ethnic minorities but by 9% for whites. Similarly, the proportion of the year spent in employment increased by 24% for whites and 22% for ethnic minority learners. The reduction in benefit dependence after learning was higher for those from ethnic minority groups compared to whites (-3% v -1%). Level 2 Apprenticeships are found to be associated with a greater increase in earnings for whites than for ethnic minority groups (26% v 25%) and similarly for employment (30% v 26%). Again, benefit dependence decreases by more for learners from ethnic minorities than for whites who complete apprenticeships at Level 2 and Level 3. For those who completed Train to Gain, at FL2 wages were found to decrease by 2% for whites compared with a 1% increase for learners from ethnic minorities. The proportion of the year a person was in employment increased by 12% on average for ethnic minorities, compared with just 3% for whites.

Socio-economic group

The socio-economic group of learners is measured in two ways across the three studies:

- i. using the Index of Multiple Deprivation (IMD); and
- ii. the socio-economic group of parents (SEG).

To a large extent the SEG of interest is that of the learner’s parents given that a large percentage of learners are young people with limited or no work experience on which to determine their SEG. Where administrative data are used, the IMD is used as a proxy measure of parental / learner SEG based on the reasonable assumption that people living in an area with a given IMD score will share the socio-economic characteristics of people in that area. Using survey data – principally the British Cohort Surveys in RP53 - it is possible to include direct measures of parental SEG in the data analysis along with other indicators relating to parental education and economic hardship, and the individual learner’s ability at age 10 in mathematics and reading.

In RP47, the IMD - which combines a number of indicators relating to economic, social and housing into a single deprivation score for each small area where a high IMD score indicates relative deprivation – was used as a proxy measure of the SEG of individuals’ backgrounds. The results reveal:

- the IMD score is negatively related to attaining a qualification at all Levels (1 – 4) – $\beta = -0.001$ for Levels 1 and 2 and $\beta = -0.002$ for Levels 3 and 4 with all coefficients significant at the $p < 0.01$ level – suggesting that attainment is dependent upon living in a relatively less deprived area;³¹
- whilst variables were included in the modelling of employment and earnings outcomes which captured the socio-economic characteristics of individuals these are not reported in the detailed findings.

The analysis in RP53, in summary, suggests that there are positive returns which increase with the level of qualification which persist over time (for those born in 1970, that will have acquired their qualifications c.1986-1990, there are still positive employment returns, especially at Level 3 and above, associated with those qualifications in 2004). Annex 2 of RP53 presents the findings relating to SEG without commentary, but from Tables 88 to 95 it is possible to discern, other things being equal, that:

- in 2000, where parental socio-economic group was I (higher managerial, professional, etc.) there was a significant and positive impact on gross earnings ($\beta = 0.147$), and where the parental SEG was III non-manual (supervisory, clerical, junior managerial) there was also a positive and significant effect ($\beta = 0.154$). The other SEGs had no significant impact;
- by 2004, the only significant effects were those of parent SEG II (Intermediate managerial, administrative, professional) where there was a negative significant effect ($\beta = -0.659$) and SEG III non-manual (supervisory, clerical, junior managerial) there was also a positive and significant effect ($\beta = 0.0784$).

What is apparent is that the positive impact of SEG on returns (to gross earnings) that was evident in 2000 was no longer clear in 2004. That said, it is difficult to disentangle the effects in the modelling given that financial hardship at age 16 is significantly and negatively associated with gross earnings in 2004 ($\beta = -0.0842$).

How might one interpret the findings from the two studies? One interpretation is that there is: (i) a selection effect about who studies what courses; (ii) once a selection has been made those from higher SEG groups are more likely to attain the qualifications for which they enrolled; and (iii) upon entry to the labour market, there are higher returns for those from SEG groups I and III (non-manual). The data in RPs 47 and 53 provide evidence in relation to (ii) and (iii) but not in relation to (i) which results in the data needing to be carefully interpreted with respect to any overarching conclusions about the impact of SEG on participation in, and obtaining returns from, vocational qualifications.

5.4 Evidence on the Long-Term Returns and Their Persistence

Persistence of returns is key to estimating the long-term benefits

³¹ Table 5 in RP47, p. 38.

The persistence of returns to qualifications has an important bearing on the long-term benefits of such qualifications and their overall value to the individual learner, firms, the economy and the exchequer. If the returns diminish from a particular value over time (from completion) then clearly the cumulated long-term returns will be less than if the returns remained constant at that value over time.

The model developed by CE and IER for BIS to measure the Net Present Value (NPV) of Further Education assumes that the impacts on wages and employment persist (i.e. are constant) over the individual's working lifetime. If the more-recent evidence suggests that returns diminish over time then it might be sensible to make adjustments to the model.

Although it appears plausible that the returns to education could diminish over time, for example as the knowledge gained at a particular point in time becomes out of date, there are also plausible reasons why the returns could actually increase, at least in the short/medium term. For example, if the qualification gained is a stepping stone to another qualification that has a positive impact on wages/employment then the return on the first qualification may be zero, or even negative, until the second qualification is achieved. If the second qualification cannot be taken without the first, then the first qualification is an important one (by allowing progression) even if it appears to have zero or negative returns.

RP48 does not investigate long-term returns

The analysis in RP48 only considers the returns to training between three months and one year following completion. It is therefore not possible to estimate the long-term benefits of training from these results.

RP53 does not explicitly examine persistence, but there does appear to be some evidence in the results presented

There is no explicit analysis in RP53 of whether returns to qualifications persist over time following completion. However, the use of the British Cohort Survey means that such analysis might be possible. For example, from Tables 90, 91 and 92 in RP53 there is some evidence of persistence of effects. Table 5.6 shows the coefficients attached to selected qualifications derived from the BCS analysis in RP53. It shows that for some qualifications the sign on the coefficient changes over time and in other cases the size of the effect increases over time. It is apparent that there is a strong positive coefficient attached to Trade Apprenticeships in all three years, the size of which does not change.

Table 0.6: Coefficients attached to selected qualifications (RP53)

	1996	2000	2004
NVQ3	0.0031	0.0263	0.0580
NVQ2	0.0082	-0.0624	-0.0293
NVQ1	-0.1010**	-0.0267	-0.0445
Trade Apprenticeship	0.0971**	0.0904**	0.0951**

Source: RP53 Tables 90, 91 and 92
 Note: Average net wage returns men and women with controls / **P<0.05

Having said that, there does not appear to be a variable in the model specification that captures employment growth over the period, and so these estimates could be biased in a number of ways.

In some sense, the main analysis in RP53 estimates the average returns over what could be a long period of time, as the time that has passed since each individual achieved the qualification could be many years. Notwithstanding the apparent implications of the BCS estimation results above, the main analysis does not give any indication as to how the returns may change each year, say, after the qualification has been achieved.

When calculating NPV, RP53 assumes that returns persist over time

RP53 does make some estimates of the Net Present Value and Internal Rate of Return associated with vocational qualifications, for both individuals and the Exchequer. In this analysis, it appears that returns are assumed to be constant and persist over the working lifetime of the individual (as in the BIS NPV model).

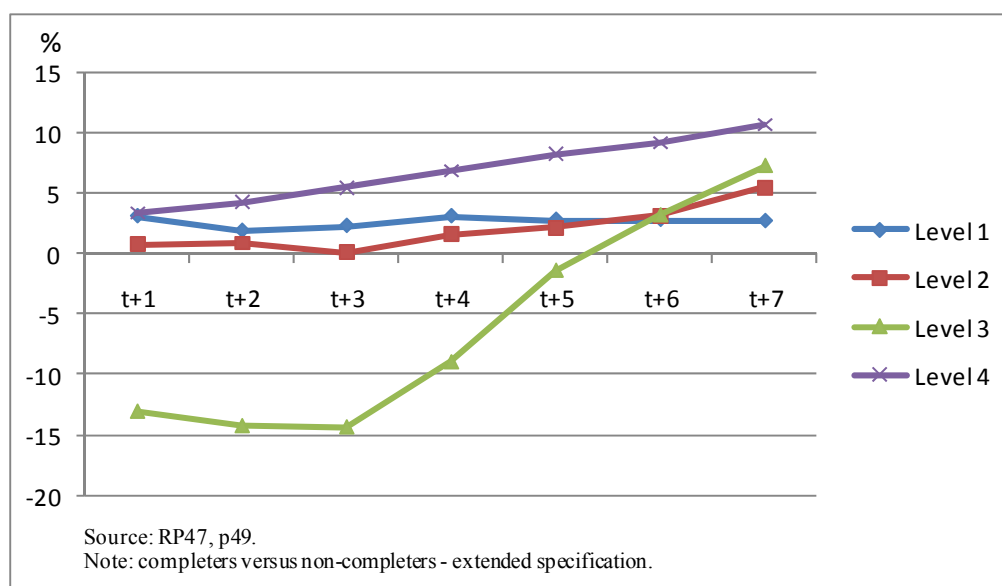
The focus of RP47 is on estimating long-term returns

The main focus of RP47 is the *long-term* return to vocational qualifications. The matched data include ILR data over 2002/03-2005/06, earnings data over 2003/04-2009/10 and employment data over 1999/00-2009/10. The dataset therefore allows analysis of earnings and employment data up to seven years after completion in 2002/03 (ie the first year of ILR data).

As discussed elsewhere, there are some drawbacks with the analysis that was undertaken, due to the data that was available (or not available) in the matched dataset. The first is that, because the last year of ILR data is 2005/06 (and there is no information on higher education), the data do not allow the impact of learning that took place after that time to be taken into account. Another drawback is that the counterfactual used in the analysis, comparing completers with non-completers, is not particularly satisfactory.

Long-term impact on earnings

Notwithstanding that, the analysis shows that, apart from Level 1 qualifications, the impact of training and education on earnings does appear to vary over time (at least up to seven years) following completion. The impact of these qualifications on earnings appears to increase towards the end of the seven year period, with the impact of Level 2 and Level 3 qualifications negative initially (See Figure 5.1).

Figure 0.1: Long-term impact of education and training on earnings (RP47)

As discussed in RP47, the strong negative impact of Level 3 qualifications in the first four years following attainment reflects the inclusion of GCE A' Levels in the analysis, where completers are more likely to then progress to further or higher education and so remain as relatively low earners for a number of years.

Table 5.7 shows selected results for particular qualifications. Note that Apprenticeships have some of the largest long-term impacts on learners (as was estimated also by RP53), but that there is some erosion of the earnings premium over time.

Table 0.7: Coefficients attached to selected qualifications (RP47)

	Year 1	Year 2	Year 3	Year 7
NVQ3	0.027***	-0.005***	-0.021***	0.108***
NVQ2	-0.023	0.002	0.000	0.015
Advanced Apprenticeship	0.226***	0.199***	0.182***	0.137***

Source: RP47 Table 49 in Annex 3 and Table 18 in Chapter 5.
Note: Based on simultaneous attainment analysis on earnings effects / ***P<0.01

Long-term impact on employment

The impact of vocational qualifications on employment was also found to increase over time, with Level 3 qualifications having returns around zero in the first year after completion, but returns of around 6% after seven years (See Figure 5.2).

Long-term impact on benefit dependency

Vocational qualifications at all levels were found to reduce benefit dependency, although the impact diminishes over time for Level 2 and Level 3 qualifications (see Figure 5.3).

Figure 0.2: Long-term impact of education and training on JSA dependence (RP47)

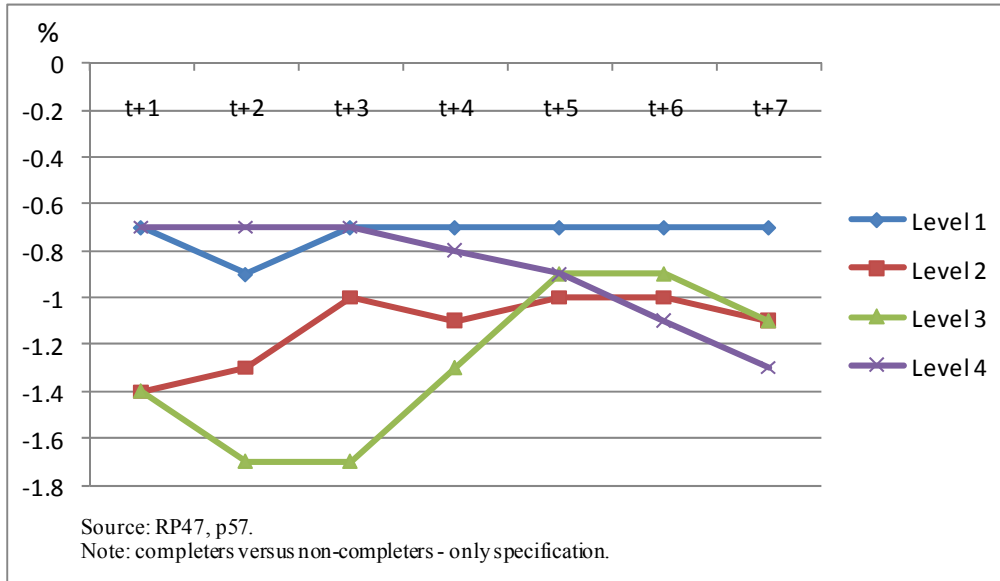
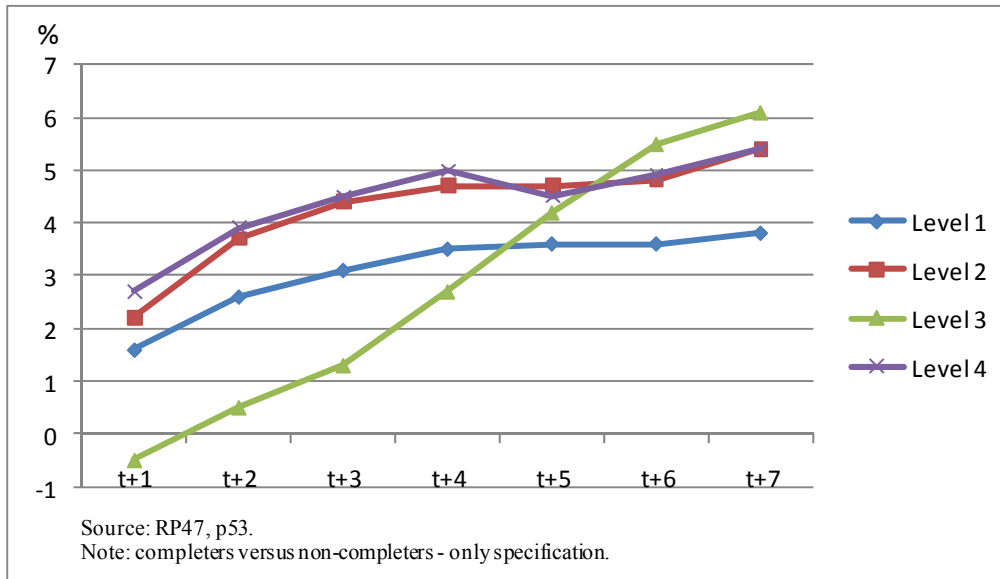


Figure 0.3: Long-term impact of education and training on employment (RP47)



Conclusions

As discussed above, the main analysis in RP53 in effect estimates the average returns over a fairly long period of time, but says nothing about how those returns may change post-achievement. Some of the results based on analysis using the BCS do suggest that some indication of returns over time could be established, but the robustness of the results would need to be investigated further.

RP47 finds that the impact of vocational qualifications on earnings and employment is positive and increases over time (up to 7 years post-completion), and that achieving them

reduces benefit dependency, it should be noted that this analysis into the persistence of returns is experimental and there are methodological issues (such as those mentioned above) that need to be overcome and warrant further research.

5.5 In Which Circumstances is it Preferable to Use Which Evidence?

The three papers reviewed in this report do not allow for a straightforward decision to be made on which is the best dataset to use to estimate the returns to qualifications, even for particular situations. In terms of which data and estimation approaches to use in considering the rates of return to FE and training, this varies by the purpose of the estimates and the comparison one wants to make (i.e. FE learner compared to who?). There is no hard and fast rule however, the nature of the two main datasets, the linked ILR-HMRC-DWP data and the LFS, goes much of the way in determining which is most appropriate.

The LFS allows control for more of the individual-level characteristics of individuals such as other training/learning. From the analysis undertaken to date, the LFS proves more useful in digging down into the factors behind different outcomes for different individuals though there are shortcomings associated with this data in measuring the outcomes and in capturing accurate details of the learning programme which led to a particular qualification – the administrative data does provide a wealth of such detail.

More important than deciding which dataset to use in which circumstance is the need to clearly define the counterfactual and ensure that the comparison is the most useful and relevant for a particular purpose. This decision too can help in deciding which dataset is most useful in carrying out any analysis on the returns to qualifications.

The use of the LFS data is important to ensure that the analysis found in earlier studies (e.g. McIntosh 2007, 2009) can be replicated and results compared over time. The methodological approach applied to the matched ILR-HMRC/DWP data need to be investigated further to ensure that selection effects and other potential sources of bias are treated appropriately. Some of the comparisons (e.g. 'before and after' and 'completers v non-completers') are inappropriate (or at least need further work) in order to consider factors that affect the returns which are not accounted for in a satisfactory manner in the current specifications.

The matched data allow analysis, by Level (including basic skills), funding stream, type of course, and subject, in a degree of detail that captures better than LFS-based analysis the heterogeneity in vocational qualification provision.

In any case, having both survey and administrative data provides opportunities to explore numerous facets and features of the returns provided by FE qualifications. The results of the three papers are promising in that they do not generally contradict each other in terms of the overall picture.

6 Conclusions and Recommendations

6.1 Introduction

This chapter brings together the findings of the earlier chapters to draw out the conclusions and recommendations of the review of the three research papers (RPs): Section 6.2 discusses what the three RPs add to the existing evidence base; Section 6.3 discusses where it may be better to use the findings of one RP over another; Section 6.4 looks at the implications for BIS' NPV model, and; Section 6.5 makes broad recommendations for future research.

6.2 How does the research add to the existing evidence base?

The findings of recent analysis of the LFS/BCS (RP53) are broadly consistent with previous studies

The existing evidence base on the returns to vocational qualifications draws principally on cross-section analysis of the Labour Force Survey (LFS) and longitudinal surveys such as the British Cohort Survey (BCS). RP53 applies these well-established techniques to recent survey data to derive updated estimates of returns to intermediate and low level qualifications, and, because the methods used are similar to those previously used, the results are the most readily comparable, of the three papers, with earlier work. The differences between results from RP53 and those of other studies (e.g. McIntosh 2007) are partially attributed to the use of hourly wages rather than weekly wages. The findings of RP53 are, however, broadly consistent with previous studies that applied similar analysis:

- There continue to be positive employment and earnings returns for the vast majority of qualifications gained in adulthood.
- The returns to vocational qualifications are higher at Level 3 than Level 2, and Apprenticeships have the highest overall wage and employment returns.
- Typically, the wage returns to vocational qualifications are found to be higher for men than women.
- Returns are higher for younger learners (aged under 25 years).
- There is significant variation in returns by sector and occupation of employment: for example, there are higher returns to qualifications found in lower-skilled occupations (e.g. men working in skilled trades and machine operative occupations and for females working in personal services).

- However, the route of acquisition of qualification has become much less important in determining the earnings returns (previous estimates, and theory, suggested that the workplace route gave greater returns).

The use of cross-sectional analysis of the LFS does not really allow for analysis in any detail beyond Level of attainment and qualification type (as recalled by the respondent) and even the relatively large sample sizes in the LFS soon become quite small when disaggregated by type of qualification and the year in which it was attained.

Because they rely on survey responses, the LFS and BCS are both subject to errors arising from the fact that the individual may not report correctly the qualifications that they are working towards or have attained.

The BCS longitudinal survey dataset suffers from the problems of small sample sizes and sample attrition over time.

The matched administrative dataset provides a rich source of information that allows some of the drawbacks of earlier analysis to be overcome

The other two papers reviewed (RP47 and RP48), made use of a recently-developed experimental ‘matched’ database that combines administrative data from the Individualised Learner Record (ILR) with data from Her Majesty’s Revenue and Customs (HMRC) and data from the Department for Work and Pensions (DWP).

Where these linked databases add value is in the considerable amount of detail they provide about the type of FE in which the individual has engaged and the possibility via the HMRC and DWP databases of assessing the extent to which gaining a qualification at a given level via one route or another in a particular subject provides some relative long-term benefit to the individual. There do though remain a number of methodological and data issues to be overcome to improve the analyses:

- the data do not provide a complete record of the individual’s path through education (for instance, there are no data relating to the individual’s entry to and exit from higher education);
- the definition of an appropriate counterfactual proves difficult to define and construct;
- the research papers do not report fully on the potential biases introduced via the data matching and cleaning processes, and the consequent implications for the findings
- conditions in the external labour market are not always controlled for.

It is encouraging that the results of the analysis using the new linked datasets do not, in general, contradict the existing evidence base, finding that higher levels of attainment are generally associated with higher returns in terms of wages and employment.

However, the magnitude of those returns can vary considerably between studies, with, for example, all RPs confirming earlier findings that there are positive wage returns associated with completion of Apprenticeship training, but the magnitude of the estimates differing across studies and even within studies depending on model specification or counterfactual used.

Table 6.1 summarises the relative strengths and weaknesses of the analysis using the two alternative datasets: cross-section and longitudinal survey data; versus the matched administrative dataset.

The analysis using the matched administrative data is able to overcome some of the weaknesses of the survey-based analysis – the key strengths of the matched administrative data are: larger sample size; reduced recall bias; and the much richer level of detail by which to categorise the provision of vocational qualifications.

Table 0.1: Strengths and Weaknesses of the Alternative Datasets

	Survey data	Matched administrative data
Sample size	Small	Large
Sample attrition	Yes	No
Control variables for characteristics of individuals	Some Longitudinal surveys allow control for time-invariant individual characteristics	Few
Recall bias	Yes	No
‘Categorisation’ of qualifications	Less detail	Greater detail (eg subject)
Counterfactual	Outcomes of those holding various (and no) other qualifications Outcomes before attainment	Outcomes available only for those that enroll to undertake FE training. Alternative counterfactuals: ‘non-completers’; simultaneous completers at adjacent levels of qualification; before attainment
Source(s): BIS Research Papers 47,48 and 53.		

6.3 Where is it better to use one set of results over another?

As discussed in Section 5.5 In Which Circumstances is it Preferable to Use Which Evidence?, the three papers reviewed in this report do not allow for a straightforward decision to be made on which is the best dataset to use to estimate the returns to qualifications, even for particular situations.

Merits of the variables in LFS versus administrative data

On the one hand, the LFS allows control for more of the individual-level characteristics of individuals such as other training/learning. On the other hand, the larger, linked, administrative datasets allow analysis at a much finer grain of detail, such as subject.

The relevance and interpretation of the counterfactual used

Recall that a key analytical challenge to estimating the economic returns to vocational qualifications is how to define the comparison group (counterfactual). This challenge is faced by all researchers, and neither of the existing datasets, the survey and the matched administrative data, are sufficient to provide the ideal counterfactual. The choice of counterfactual can be severely limited by data and so the researchers develop alternative ways of defining the counterfactual (see Table 6.1). Consequently, perhaps more important than deciding which dataset to use in which circumstance is the need to clearly define the counterfactual and to ensure that the comparison is: the most useful and relevant for a particular purpose (see Gambin et al, 2011).

RP53 defines the counterfactual in a similar way to much of the earlier work that utilises survey data. LFS data are used to estimate marginal returns by comparing the returns to individuals with a qualification at a particular level with returns to similar individuals with the next lowest (or no) qualification. Individuals can be identified as similar (by a limited subset of characteristics) using the control variables that are available in the survey data. This counterfactual is a well-established approach which is considered both meaningful and intuitively appealing. However, it fails to take account of differences in ability and motivation of individuals studying for different levels of qualification.

The analyses using the linked administrative datasets (in RP47 and RP48) also have particular weaknesses in their definition of the counterfactual. These datasets only include those enrolled on FE courses and so do not provide the basis for comparison with those who have not undertaken training. The researchers develop and test alternative approaches to defining the counterfactual.

RP48 uses outcomes before attainment as the counterfactual. RP47 also investigates the use of before attainment plus two other approaches to constructing a counterfactual. The authors of RP47 conclude that, no matter which counterfactual is used, there are strong positive and persistent returns to completing a qualification at a given level. There are weaknesses to the counterfactuals used, and we note that the magnitude (and in some cases the sign) of the estimates are sensitive to the choice of comparison:

1. Before and after qualification acquisition – fails to take account of factors that change over time (e.g. age effect).
2. Simultaneous completers at adjacent levels of qualification – fails to take account of differences in ability and motivation of individuals studying for different levels of qualification.
3. Non-completers – fails to fully take account of: factors that affect the likelihood of course completion which may also affect labour market outcomes (e.g. the learner may have failed to complete because of ill-health); all types of non-completion (e.g. transfer to other course).

Section 6.5 suggests enhancements to the administrative datasets that would provide a basis to improve the counterfactuals used in RP47 and RP48.

Persistence of returns

As discussed in Chapter 5, the main analysis in RP53 in effect estimates the average returns over a fairly long period of time, but says nothing about how those returns may change post-achievement. Some of the results based on analysis using the BCS do suggest that some indication of returns over time could be established, but the robustness of the results would need to be investigated further.

RP47 finds that for many vocational qualifications the impact on earnings and employment is positive and increases over time (up to 7 years post-completion), and that achieving the qualifications reduces benefit dependency. Apprenticeships have some of the largest long-term impacts on learners (as was estimated also by RP53), but there is some erosion of the earnings premium over time. It should be noted that this analysis into the persistence of returns is experimental and there are methodological issues that need to be overcome and warrant further research.

Further development and analysis of the linked datasets will prove beneficial

The use of the LFS data is important to ensure that the analysis found in earlier studies can be replicated and results compared over time, but, despite a number of shortcomings, our general belief is that further development and analysis of the linked datasets is the route that will prove most beneficial in producing practical findings because of the substantial advantages on most counts set out in Table 6.1.

In any case, having both survey and administrative data provides opportunities to explore numerous facets and features of the returns provided by FE qualifications. The results of the three papers are promising in that they do not contradict each other in terms of the overall picture.

6.4 What are the implications of the report's findings within the context of BIS' NPV model?

The NPV model³² developed for BIS by CE and IER includes estimates of wage premia and employment premia that were drawn from the analysis by McIntosh³³ and developed through consultation between BIS and IER.

RP53 provides updated estimates of marginal returns

³² Cambridge Econometrics and IER (2011) Measuring the Economic Impact of Further Education, BIS Research Paper Number 38.

³³ McIntosh, S. (2009). The Economic Value of Intermediate Vocational Education and Qualifications, UKCES, Evidence Report 2011, December 2009.

RP53 effectively updates much of that earlier analysis, and so the estimates of wage and employment premia in RP53 could be used to update those in the NPV model, at least for vocational qualifications.

Although the analyses of the administrative data suffer some drawbacks...

Our overall feeling at this stage, however, is that the analysis in both RP47 and RP48 is too experimental for us to recommend making use of these estimates in the NPV model. It would be preferable first to overcome some of the shortcomings to make further improvements to the analyses.

...they offer potential to update and further develop the NPV model in the future

When the NPV model was originally developed, much thought was given to the trade-off between complexity and manageability. There is a need for the model to be: sufficiently complex to capture the pertinent relationships and impacts within the FE system (with the associated risk that the model will not be at all user friendly or transparent); and sufficiently manageable, and easy to understand and operate (with the associated risk that it is over-simplified to the extent that it is not suitable to undertake the kind of analysis required).

The decision was made to build into the NPV model a fairly high degree of disaggregation and detail to make the model more flexible in terms of the analysis that can be undertaken. The high level of disaggregation, with 15 provision types, 7 prior qualifications, gender, and two age bands identified separately in many of the calculations, is reflected in the complexity of the model. This complexity is at the bounds of what can be effectively and transparently handled within the software Excel. The alternative would be to migrate the model to a more flexible programmable language (but this would limit the accessibility of the model to analysts at BIS). The complexity of the model is also reflected in the robustness of some of the existing assumptions, which in many cases were made based on data at higher levels of aggregation. It was the intention that the assumptions could be updated and improved over time as better evidence became available.

The NPV model incorporates assumptions on the marginal benefit (ie relative to having a lower level qualification) of the highest qualification held, and so as long as any analysis reveals robust estimates on that basis then they could be used to update the assumptions in the model. What the matched administrative data provide is the opportunity to make estimates in more detail than previously available (from LFS) and so once robust estimates have been made, then they can be used to better populate the assumptions and parameters used in the NPV model.

Where the linked administrative datasets have the potential to add value to the NPV is to provide estimates of returns by subject of study, alongside information about Level, funding stream, and qualification type (e.g. vocational versus academic). The current NPV model does not distinguish qualifications by subject area, and adding this further dimension would substantially increase the complexity of the model. The current findings (in RP47 and RP48) by sector are not particularly robust, but, given these findings, and the theory, it may be desirable to consider ways in which to add this information to the model if more robust estimates become available in the future. One way in which to handle the additional dimension might be to use sub-models for each sector which are replicas, in

structure, of the core model, but populated with assumptions by sector. This would overcome the challenge of introducing an additional dimension in Excel, but:

- not all of the problems of complexity would be eliminated – e.g. assumptions of returns would be required by sector in combination with all the other existing dimensions (15 provision types, 7 prior qualifications, gender, and two age bands).
- the sub-model would undertake analysis for a single sector in isolation – it would likely be desirable to compare findings across sectors and so a number of sub-models would need to be set up and maintained.

Another area in which analysis of the new datasets might offer improved evidence is in the pattern of returns over time. The NPV model assumes that the returns to training and qualifications are persistent and constant over the working lifetime of each learner. This assumption was made because there was little evidence available about the persistence of returns, and to simplify the framework of the model. RP47 suggests that returns are not constant, over the first seven years post-qualification at least. The returns are found to increase over time, with negative returns in the first few years after achievement in some cases (the strong negative impact of Level 3 qualifications in the first four years following attainment reflects the inclusion of GCE A' Levels in the analysis, where completers are more likely to then progress to further or higher education and so remain as relatively low earners for a number of years). Further analyses to support these tentative conclusions would provide better evidence to further develop the NPV model to incorporate variability of returns over time. Adding the ability to vary over time the assumptions for returns would, however, add an additional layer of complexity to the model

6.5 Promising areas and methods for future research

There are limits to what survey data can tell us about returns to qualifications

Over the past 20 years there has been no shortage of analyses looking at the returns to qualifications of one kind or another. The most recent studies - based on cross-sectional and longitudinal data - were appositely summarised by McIntosh (2009). In some respects, the limits to what survey data, either cross-sectional or longitudinal, can tell about the returns to qualifications have been reached. It is possible to provide relatively robust findings relating to academic versus vocational, level, and type of qualification, whilst controlling for a variety of socio-demographic and educational factors which have some bearing on the level or rate of return.

Whilst there is merit in repeating analysis using cross-sectional and longitudinal survey data in updating findings over time, especially so given that the vocational education and training system is subject to periodic overhaul, there is also a need to progress beyond what these types of data can provide.

The matched datasets provide the most promising opportunity to further our knowledge of returns to qualifications

The use of matched and linked data as reported in RP47 and RP48 provides a considerable opportunity to progress the analysis even if, to date, there are a variety of

conceptual, econometric and statistical issues to be resolved. In looking at how to further develop the approach adopted in these research reports there are a number of issues which future research might address.

In an ideal world, if the aim is to gauge how the pathway through the FE and skills system affects labour market outcomes relating to wages, spells in and out of employment, benefit dependency, etc. then there is a wish list of data that is ideally required, and a list of methodological issues that would ideally be addressed to achieve fuller insights and more robust results. These issues include the following:

- providing a comprehensive database which allows an individual to be tracked through post-compulsory education. This will necessitate supplementing the existing databases with data which captures an individual's higher education achievements;
- there is also a need to consider how data can be obtained about the outcomes of people who are not captured by the ILR such as those who do not enter post-compulsory education. At the moment the analysis consists solely of those who have entered the further education and skills system;
- the lack of measures (to be included as control variables) to take account of ability – such as the reading, mathematics, verbal and non-verbal reasoning test scores which some longitudinal survey collect. But in many respects the key question is what value-added do qualifications obtained in the FE and skills system produce given an individual's educational attainment on entry to it;
- greater attention might be given to ensuring that there is a comprehensive record of the individual's educational attainment on entry to the further education system. The National Pupil Database may be useful in this regard if it were linked to the ILR (which has been achieved in some analyses). But alongside educational qualifications on entry, it would be useful to know something about an individual's motivation or aspirations for choosing a particular course within the system because these may well be associated with eventual success within the FE system (i.e. passing a course with commendation or merit), the predilection to continue with learning, and eventual labour market outcomes;
- an individual's passage through the FE and skills system is not included in the analyses, yet the individual's experiences whilst in the system may be a determinant of their labour market outcomes. A person who was generally satisfied with their educational experience in FE, who received fairly broad based vocational education, with plenty of opportunity to train and practice skills learnt, may well, other things being equal, fare relatively well in the labour market when compared with someone who had a less positive experience;
- the selection of an appropriate counterfactual is of critical importance since the question about any return is always that of "compared with what?". A flexible approach can be adopted here (as has been in the papers reviewed) such that the counterfactual is selected on the basis of what particular question is being addressed (Gambin et al, 2011) , but as noted above, there is a need to have data for those who are not captured in the ILR for one reason or another;

- consideration also needs to be given to constructing a consistent time series of data relating to the measurement of employment spells and the measurement of earnings over time based on the use of HMRC data. Related to this, if estimates are to be compared between studies and over time there is a need to control for conditions in the labour market and economy more generally, since the returns to some qualifications will to a greater or lesser extent simply reflect growth in the economy.

If it were possible to make the various improvements to the database used to provide the analysis in RP47 and RP48 then the analysis will begin to provide detailed information about the subject of study, alongside information about Level, funding stream, and qualification type (e.g. vocational versus academic). By taking an approach which looks at the value-added by different types of qualification for people who otherwise share similar characteristics then much progress will be made towards answering questions such as what qualifications give the highest return for people with a given set of characteristics. Both RP47 and RP48 provide many valuable insights but, for the reasons outlined in the main body of this report, we recommend that caution be exercised in interpreting the results of these initial studies.

In looking more to the future, any data wish list can be divided between possessing information about:

- an individual's experience before entry to the FE and skills system such that there is a comprehensive record of their achievements on entry;
- their experiences in passing through the FE and skills system;
- the outcomes achieved at the point of exiting the FE and skills system;
- any further education and learning post-FE; and
- economic status over a period of, say, seven years post-FE.

Table 6.2 below contains the wish list of data. What is apparent is that much of the data is already collected but in separate surveys so there needs to be a means sought of being able to link various data sets. In this way a much more comprehensive set of data will be available with which to gauge the relative success of following different pathways within the FE and Skills system. It will also potentially explain why there is variation in outcomes within a particular pathway.

The data set out in Table 6.2 is what would be required for an approach which allows the passage through the FE system – and beyond where it leads to higher education or other learning – to be modelled at each stage and then to determine the overall outcome with respect to employment. As noted above, the linking of relevant databases is close to making this realisable. The data needed are held in various administrative databases and surveys so that the major issue is not so much about collecting more data but being able to match and link data sources. A unique identifier is required to do this ideally – though the probabilistic matching techniques used to link ILR and HMRC data appears to obtain a high degree of match – which would require being able to include the unique learner code

in surveys of learners. From a technical perspective, if the ILR is being used as a sampling frame for surveys of learners, this is not an insurmountable barrier. Respondents, however, will need to give their permission for survey responses to be linked to other databases. And where permission is not given, or where probabilistic matching is used, there is a need to consider what types of bias may be being introduced into any analysis. But the fact remains, as evidenced by RP47 and RP48, that the ability to link and match is improving all the time such that the comprehensive analysis envisaged above is close to being achieved. The move to a unique identifier to link databases which record individuals' pathways through the education, further education, and higher education systems would make this much more achievable.

Table 0.2: Ideal Data Requirements

Stage in the education system	Type of data required	Availability of data
Entry to FE and Skills system	<ul style="list-style-type: none"> Type of programme studying / working towards Subject(s) / framework of study Prior educational attainment Socio-demographic characteristics Aspirations / motivations Date of entry 	<ul style="list-style-type: none"> NPD ILR Surveys of Prior Educational Attainment Data on aspirations and motivations from ad hoc surveys
Passage through FE and Skills system	<ul style="list-style-type: none"> Indicators of training received Satisfaction with learning Satisfaction with work environment (if applicable) 	<ul style="list-style-type: none"> Principally available in survey data such as Apprenticeship Pay Survey and National Learning Satisfaction Surveys. Other ad hoc surveys; e.g. of apprentices
Situation on exit from FE & Skills system	<ul style="list-style-type: none"> Whether completed Qualification / education award obtained (and level) Date of award Employment status on exit 	<ul style="list-style-type: none"> ILR
Experience of HE	<ul style="list-style-type: none"> Type of course pursued Qualification obtained Date of award 	<ul style="list-style-type: none"> UCAS / HESA databases
Initial labour market position over 7 years post-FE	<ul style="list-style-type: none"> Employment status Year 1 - 7 (years 1 - 7 respectively) Wage (years 1 - 7 respectively) Benefit status (years 1 - 7 respectively) 	<ul style="list-style-type: none"> HMRC / DWP databases

7 References

- Aspin, D., Chapman, J., Hatton, M. and Sawano, Y. (2001), *International Handbook of Lifelong Learning*
- Cambridge Econometrics and IER (2011), *Measuring the Economic Impact of Further Education*, BIS Research Paper Number 38
- De Coulon, A. and A. Vignoles (2008) *An Analysis of the Benefit of NVQ2 Qualifications Acquired at Age 26-34*, CEE DP 106
- Gambin, L., Hasluck, C., Hogarth, T., Ryan, P. and Elias, P. (2011) *Options Study for the Long Term Evaluation of Apprenticeships*, Department of Business Innovation and Skills, Research Paper No.56
- Green, A., Wolfe, A. and Leney, T. (2000) *Convergence and divergence in European education and training systems*, London, Institute of Education.
- Green, F. and Y. Zhu (2010) *Over qualification, job dissatisfaction, and increasing dispersion in the returns to graduate education*, *Oxford Economic Papers* 62 (2010), 740–763
- Hartog, J, and H. Oosterbeek (1998) *Health, Wealth and Happiness: Why Pursue a Higher Education?* *Economics of Education Review*, 17, 3, 245-256
- Hogarth, T. et al., (2010) *Analysis of Apprenticeship Completion Rates*, Learning and Skills Council, Coventry
- Jenkins, A., A. Vignoles, A. Wolf and F. Galindo-Rueda (2002). *The Determinants and Effects of Lifelong Learning*. Centre for the Economics of Education. Discussion Paper 19. London. April.
- Jenkins, A., C. Greenwood, C. and A. Vignoles (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. London.
- McIntosh, S. (2002). *Further Analysis of the Returns to Academic and Vocational Qualifications*. DfES Research Report 370. London.
- McIntosh, S. (2009). *The Economic Value of Intermediate Vocational Education and Qualifications*, UKCES, Evidence Report 2011, December 2009
- Mincer, J. (1991) *Education and Unemployment*, NBER Working Papers # 3838
- Nickel, S. (1979) "Education and the lifetime pattern of employment", *Journal of Political Economy*, 87, 5, 117-131.
- OECD (2000) *From initial education to working life* (OECD, Paris).

Oreopoulos, P. and K. Salvanes (2011) Priceless: The Nonpecuniary Benefits of Schooling, *Journal of Economics Perspectives*, 25,1, 159-184.

Purcell, K., P. Elias, R. Davies and N. Wilton (2005). *The Class of '99. A study of the early labour market experiences of recent graduates*, Warwick Institute for Employment Research.

RM Data Solutions (2010), *Research into Measuring Adult Attainment Using the Labour Force Survey*.

Rutter, M. (1979) *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Open Books Publishing

Silles, A. (2009) The causal effect of education on health: Evidence from the United Kingdom, *Economics of Education Review*, 28, 122-128.

Smith, D.J. and S. Tomlinson (with H. Tomes, T. Hogarth and L. Bonerjee (1986) *The School Effect: A study of multi-ethnic secondary schools*, Policy Studies Institute / Hutchinson, London

Winterbotham, et al. (2000) Winterbotham, M., Adams, L. and Lorentzen-White, D. (2000), *Modern Apprenticeships: Exploring the Reasons for Non-completion in Five Sectors*. RR217, London, Department for Education and Employment.

Wolf, A., Jenkins, A. and Vignoles, A. (2006), "Certifying the workforce: economic imperative or failed social policy?", *Journal of Education Policy*, Vol.21, No. 5; pp. 535-565.

Wolf, A. (2011) *Review of vocational education: The Wolf Report*, Department for Education, DFE-00031-2011, March 2011

Woolbers, M.H. (2000) "The effects of level of education and mobility bet on employment and unemployment in the Netherlands", *European Sociological Review*, 16, 2, 185-200.

8 Glossary

ACL	adult and community learning
BCS	British Cohort Study
BIS	Department for Business, Innovation and Skills
CE	Cambridge Econometrics
DWP	Department for Work and Pensions
ESF	European Social Fund
HMRC	Her Majesty's Revenue and Customs
IER	University of Warwick Institute for Employment Research
ILR	Individualised Learner Record
IMD	Index of Multiple Deprivation
LFS	Labour Force Survey
NBD	National Benefits Database
NCDS	National Child Development Study
RP	Research Paper
VQ	vocational qualification

Appendix A: The Review Questionnaire

Table A.1 below shows the template questionnaire that was developed and used to provide the structured framework for the review and to bring consistency to the review process.

Table A.0.1: Review Questionnaire

Question	Response		
Introduction			
Date reviewed			
Name of reviewer			
BIS Research Paper reviewed (<i>place X in the relevant box</i>)		BIS Research Paper No. 53 London Economics (Sep 2011) Returns to Intermediate and Low Level Vocational Qualifications	
		BIS Research Paper, No.48 Frontier Economics / IFS (2011) Reporting on Employment and Earnings Using Experimental Matched Data	
		BIS Research Paper No. 47 London Economics (2011) The Long Term Effect of Vocational Qualifications on Labour Market Outcomes	
Or other document	Title		Author(s)
	Source		Date
	Web link		
Section A: Conceptualisation of the education and training system in England			
A1: Is consideration given to prior educational attainment?			
A2: Is consideration given to socio-demographic characteristics?			
A3: Which learning streams and levels of qualifications are covered?			

Question	Response			
A4: Is consideration given to when the learning took place? <i>(e.g. changing content of qualifications over time)</i>				
A5: Are different sectors/subjects considered, if so which?				
A6: Which routes of acquisition are covered?				
A7: Is there sufficient recognition of how people are expected to progress through a given further education pathway?				
A8: Which economic benefits are modelled? <i>(e.g. earnings, employment)</i>				
A9: Is consideration given to the timeframe over which returns are expected to be accrued?				
A10: Do you have any further comments on the way in which the report conceptualises the education and training system in England?				
Section B: Data used				
B1: Which data sets/sources are used and which data are drawn from each source?	Source		Data used	
	Source		Data used	
	Source		Data used	

Question	Response					
	Source		Data used			
	Source		Data used			
	Source		Data used			
B2: What are the dimensions of the data? <i>(e.g. cross section, panel)</i>						
B3: What are the observations? <i>(e.g. individual learners)</i>						
B4: What is the sample size?						
B5: Comment on the suitability of the sample (size, sampling error)?						
B6: Are missing data and selection bias considered a problem; how are they dealt with?						
B7: Which indicators are used to represent the variables modelled; do the indicators sufficiently capture the concept of interest?	Variable		Indicator		Suitability	
	Variable		Indicator		Suitability	
	Variable		Indicator		Suitability	
B8: Are appropriate indicators available (and used) to control for important effects? <i>(e.g. changes over time, other non-qualification training)</i>						
B9: Have the data sources been linked, if so, how, and what are the strengths and weaknesses of the method						

Question	Response
used?	
B10: What, if any, comments does the report make on the strengths and weaknesses of the data and how it is used?	
B11: Do you have any further comments on the strengths and weaknesses of the data and how it is used?	
Section C: Econometric/statistical techniques used	
C1: Please summarise the overall approach and techniques used and the rationale for using this approach	
C2: Does the model allow for appropriate specification of dependent and explanatory variables? <i>(e.g. are any obvious variables missing, why?)</i>	
C3: What, if any, matching methods are used? Why is this method preferred in this case?	
C4: Does the method establish a robust counterfactual/comparison group for the study? What counterfactual is used? Does this elicit meaningful findings? Are alternative counterfactuals potentially available?	
C5: Which analytical issues arise and are the econometric/statistical methods used able to address these?	

Question	Response
<i>(e.g. accounting for heterogeneity using fixed effects)</i>	
C6: Does the method/model yield statistically significant results in the relationships?	
C7: Are the inferences drawn from the method statistically reliable? <i>(e.g. with reference to sampling error, confidence levels, and magnitude of any effects)</i>	
C8: What, if any, comments does the report make on the strengths and weaknesses of the econometric/statistical techniques used?	
C9: Do you have any further comments on the strengths and weaknesses of the econometric/statistical techniques used?	
Section D: Evidence to support the conclusions	
D1: What are the main findings (regarding economic benefits) relating to each funding stream (broken down by level and subject of study where possible)? <i>(note page/table numbers of key results)</i>	
D2: What are the main findings (regarding economic benefits) relating to socio-economic demographic and educational characteristics of individuals? <i>(note page/table numbers of key results)</i>	

Question	Response
<p>D3: What are the main findings relating to the time-frame, persistence, long-term benefits?</p> <p><i>(note page/table numbers of key results)</i></p>	
<p>D4: Are the findings compared to results of other studies? What explanation (e.g. approach used) can be given to differences/similarities?</p>	
<p>D5: Are the report's findings warranted by the evidence provided?</p>	
<p>D6: If not, which particular conclusions require a more cautious interpretation and why?</p>	
<p>Section E: Contribution to policy issues</p>	
<p>E1: What does the report conclude about the relative returns to investing in different types of education and training?</p>	
<p>E2: What does the report conclude about the relative returns to investing in different types of learner?</p>	
<p>E3: Does the report comment on economic benefits other than earnings and employment premia (e.g. educational progression, employment security)?</p>	
<p>E4: In what other ways does the report contribute to policy issues?</p>	

Question	Response
<p>E5: In which circumstances is it relevant (and not relevant) to use these estimates of economic benefits? <i>(i.e. matched datasets versus the (cross-sectional / pseudo cohort) survey analysis)</i></p>	
<p>E6: What further work would you recommend to add value to the report's contribution?</p>	
<p>E7: What are the implications of the report's findings within the context of BIS' NPV model?</p>	
<p>Section F: Additional comments/reflections</p>	
<p>F1: Please note any additional points made in the document that are relevant to the study and not covered above</p>	
<p>F2: Please note any initial reviewer reflections on key messages for our study</p>	

© Crown copyright 2013

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. Visit www.nationalarchives.gov.uk/doc/open-government-licence, write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

This publication available from <https://www.gov.uk/government/organisations/department-for-business-innovation-skills>

Any enquiries regarding this publication should be sent to:

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

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

BIS/13/636