RESEARCH

BIS | Department for Business Innovation & Skills

BIS RESEARCH PAPER NUMBER 56

Options Study for the Long-term Evaluation of Apprenticeships

OCTOBER 2011

Prepared for
Department for Business Innovation and Skills

By
Institute for Employment Research
University of Warwick
Coventry, UK
CV4 7AL

Lynn Gambin, Chris Hasluck, Terence Hogarth, Paul Ryan and Peter Elias

Department for Business, Innovation and Skills

1 Victoria Street

London, SW1H 0ET

www.BIS.gov.uk

BIS Research paper number 56

October 2011

Acknowledgements

The research team would like to thank Simon McKee and Stanislav Petrov at the Department for Business Innovation and Skills (BIS) for their support in carrying out the study. In addition, Matt Bursnall, David Collinge, James Davison and Ian Thomson at BIS and Anna Vignoles (Institute of Education) also provided invaluable advice and guidance on a number of data sets and data linkage.

The research team are most grateful to the participants in a workshop held at the University of Warwick on 21 January 2011 where various methodological and data issues were discussed in the context of recommending the best way forward if Apprenticeships are to be evaluated.

In addition, Jan Shury and Mark Winterbotham at IFF Research advised on sampling issues related to surveying apprentices, for which thanks are due.

Contents

Executive Summary

This study was commissioned by the Department for Business, Innovation and Skills (BIS) and undertaken by the University of Warwick Institute for Employment Research (IER). The aim of the study was to investigate the options available for a long-term evaluation of Apprenticeships. Such an evaluation would:

i. produce robust quantified estimates of the outcomes of Apprenticeships for successful trainees, in terms of progression into further and/or higher learning and improved employment outcomes; and.

ii. establish under what conditions Apprenticeships produce the best results, for whom and the relative added value of an Apprenticeship compared to other forms of learning.

The study revisited a topic previously explored in 2001 and published by the Department for Education and Skills entitled "*Feasibility Study for the Long Term Evaluation of Modern Apprenticeships*". Since then changes in the economy, the Further Education (FE) system, and within the Apprenticeship system, plus the increased number of datasets available, as well as Apprenticeship becoming the principal route for accessing vocational training, make the current review most timely.

The study has assessed how to track young people and adults into Apprenticeships, through their training (including into any subsequent learning), and then into employment, in order to gauge the extent to which their future earnings and employment status (as well as other outcomes including progression to further learning, job satisfaction, acquisition of skills and mobility between jobs) were influenced by the completion of an Apprenticeship. It has therefore considered longitudinal approaches. The study has looked at how data can be disaggregated to obtain a greater understanding of how and why Apprenticeships delivered to certain individuals (e.g. according to their socio-demographic characteristics) in particular circumstances (e.g. at different levels and in different sectors) deliver better outcomes.

The primary objective of the study was to identify the potential methods for the evaluation and assess each on its merits together with its implied data requirements. To identify the data needed, an economic model was developed that set out the relationships between potential outcomes and a number of explanatory factors. The model highlights the information requirements of the evaluation and draws attention to the need to set out definitions and measures of the variables to be included in estimating the returns to Apprenticeship. Key issues include: what constitutes an Apprenticeship; the relevant outcomes; what other factors affect returns (including characteristics of individuals and the form and quality of Apprenticeship training); the timescale over which returns should be assessed; and the group(s) to which the experiences of apprentices are to be compared.

The dependent variables in the economic model include employment status and wage levels (either in cross-section or over time) as well as other outcomes such as career progression and job satisfaction. The independent variables can be grouped as:

- the socio-demographic characteristics of the individual apprentice;
- the educational attainment of the apprentice on entry to the Apprenticeship;
- the motivations and aspirations of apprentices;
- the characteristics of the Apprenticeship (level and subject);
- the quality of the Apprenticeship (measured, for instance, with reference to the amount of on-the-job and off-the-job training delivered);
- the characteristics of the employer.

A number of data sources have been identified which shed light on the returns to Apprenticeship. These data sets include:

- administrative data including the National Pupil Database (NPD) and Individual Learner Records (ILR);
- longitudinal surveys which are general in their focus (e.g. Understanding Society);
- longitudinal surveys which have education and training as their focus (mainly Youth Cohort Survey (YCS) and the Longitudinal Survey of Young People in England (LSYPE));
- *ad hoc* or occasional inquiries.

Despite the availability of these rich sources of data, none of them completely meet the requirements of a long-term evaluation of Apprenticeship outcomes. That being so, the issue becomes how to supplement these data such that it will be possible to assess the long term impacts of Apprenticeships with respect to relative earnings, employability, occupational mobility, and overall satisfaction with working life, as well as progression to further learning, including higher education. The options set out below suggest the ways in which existing data sources might be augmented in the future in order to capture the impact of Apprenticeship. In drawing up these options two important considerations were taken into account. First, while it is possible to evaluate the impact of Apprenticeships over the whole lifecycle, at some point the signal from Apprenticeship training will become too faint to detect, hence a more feasible approach is to measure the impact of Apprenticeships over a five to ten year period after apprentices have completed their training. Second, the evaluation needs to be flexible in its approach in order to accommodate changes in the

structure and provision of vocational education and training (VET) which might emerge at some point in the future.

Option 0: Rely Upon Existing Source of Data

As detailed throughout this report there are various datasets which provide information about Apprenticeships. At the moment, unless it is possible to find a means of linking datasets which provide information about: (a) the characteristics of the apprentice before entry to their Apprenticeship (e.g. from the NPD); (b) the details of the Apprenticeship (e.g. the ILR); and (c) post-Apprenticeship labour market experiences (e.g. Her Majesty's Revenue and Customs (HMRC) and Department for Work and Pensions (DWP) databases) then it is difficult to evaluate Apprenticeship as anything other than as a homogeneous entity. Even if this approach were considered acceptable any analysis is likely to be hampered by the small number of apprentices or former apprentices in many survey datasets.

Option 1: Using Linked Administrative Data

As hinted in Option 0, one approach is to rely wholly on administrative data. This would include the NPD (for pre-Apprenticeship information), the ILR (for details of the Apprenticeship), HMRC data relating to employment and earnings, and the National Benefits Database relating to spells of unemployment/non-employment. Linking these datasets will generate evidence relating to two key measures of impact for an Apprenticeship, namely future employment and earnings. Initially this evidence would relate to the short term impacts but if data were continually linked then longer-term outcomes could be examined as well. Data access and security could be an issue as could the extent to which matching cases across the different databases is feasible. This approach would provide a relatively narrow set of data relating to the outcomes of Apprenticeship, and the factors underlying differential outcomes (such as the absence of data on quality or why a person wanted to undertake an Apprenticeship). Nevertheless, there is potential to create an administrative database which contains data which will allow the labour market experiences of apprentices to be compared with other groups. To date, much progress has been made by BIS and other Government departments in linking data. Option 1 is a most attractive proposition but there are drawbacks: (i) such a database would contain no information about attitudinal factors which might shed light on why Apprenticeships deliver better outcomes in particular circumstances; and (ii) it would contain no information on 'softer' outcomes beyond the traditional economic measures of employment and earnings. Surveys such as YCS and LSYPE contain such information.

Option2: Extending YCS and LSYPE and Linking to Administrative Data

YCS and LSYPE provide detailed information about the learning received by young apprentices relative to others who have not taken an Apprenticeship. For an evaluation of Apprenticeship, these surveys would need to be extended to encompass early labour market experiences of apprentices and others, preferably until their mid-20s to see how they have established themselves in the labour

market. The multi-cohort approach adopted in YCS is considered essential to evaluating Apprenticeships since it allows the effect of new policy to be gauged through successive surveys.

If YCS/LSYPE were repeated with larger samples, these data could be linked to administrative datasets to provide a more complete record. The most important candidates for such a linkage are NPD (for details pre-Apprenticeship), the ILR (for details of the Apprenticeship), HMRC data relating to employment and earnings, and the National Benefits Database relating to spells of unemployment/non-employment.

Option 2 provides the opportunity to comprehensively evaluate Apprenticeships. This would build upon expertise within BIS relating to undertaking YCS and LSYPE, and linking data sets.

At present, sample sizes in YCS/LSYPE are too small to fully evaluate Apprenticeships. A further drawback of YCS/LSYPE is that it does not include older apprentices (those aged 24 on entry to their Apprenticeship). Option 2 is based on using the YCS/LSYPE approach (effectively using these previous surveys as a template) and linking to appropriate administrative data. The YCS/LSYPE-type survey would cover a much larger sample of apprentices (including older apprentices) and would follow multiple cohorts over longer periods than the YCS/LSYPE. This multi-cohort longitudinal survey would aim to include details regarding individuals' attitudes, reasons for undertaking Apprenticeship and other more subjective outcomes and characteristics whilst more objective measures such as wages, employment status and institutional characteristics of the Apprenticeship programme (including Level and subject area) would be provided by linked administrative records.

Option 3: Extending Use of the Labour Force Survey (LFS)

The identification of Apprenticeship training in the LFS is cursory in that it treats it as a homogeneous entity. The LFS data could be improved if respondents were asked the subject of their Apprenticeship and the date of its completion. Adding such information would provide a valuable source of data for evaluating the outcomes of Apprenticeships. There are concerns, however, that individuals are sometimes uncertain as to whether they have completed an Apprenticeship. As a result, the LFS tends to underreport the number of people who are undertaking or who have completed an Apprenticeship. Nevertheless, this is a relatively straightforward and low-cost means of improving information about Apprenticeships.

Option 4: Maximising the Value of Longitudinal Surveys for Apprenticeships

Extending the longitudinal dimension in a YCS/LSYPE type study would provide little information about older apprentices. There are now several longitudinal surveys following cohorts born on a single date in a given year or, in the case of Understanding Society, the latest longitudinal survey in the field, all persons aged over 16 years in a sample of 40,000 households. These surveys provide a wealth of information about various aspects of a person's life at various stages in their lifecycle but face a number of problems relating to sample attrition and the small number of

apprentices – or former apprentices – they contain. An alternative is to either boost the number of apprentices in these types of survey, or run a separate survey of apprentices in parallel with, for example, Understanding Society. The ILR provides a sampling frame from which to select apprentices.

If the aim of any supplementary survey of apprentices is simply to capture information about older apprentices, then an alternative option might be to develop a variant of YCS for older apprentices to run in parallel with that survey. This would require older apprentices to provide, retrospectively, information about their pre-Apprenticeships experiences.

Option 5: The Employer Dimension

Options 1-4 gauge outcomes from Apprenticeships in terms of the impact on individuals. The role of the employer is, however, likely to be of critical importance with respect to the quality of the training provided and chances of completing an Apprenticeship. Little information is currently available about the characteristics of the employers of apprentices and there is merit in seeking to address that information gap by:

i. linking ILR records of apprentices to employers who record they have an apprentice in the National Employers Skill Survey (NESS) in order to obtain a more comprehensive picture of the characteristics of apprentices and their employers. There is also potential to link NESS to the Annual Business Inquiry (ABI) which potentially extends the scope of analysis to the relationship between Apprenticeships and organisational performance;

ii. more information could be asked in NESS – or in an additional survey which uses NESS as a sampling frame - of those employers that recruit apprentices to inquire about the number of apprentices by Framework and level of Apprenticeship delivered.

Option 6: Qualitative Research

There is an important role for qualitative research - especially in-depth analysis of particular cases - in understanding the returns to Apprenticeship; in particular, understanding why returns vary within Apprenticeship frameworks as well as between them. This is seen as an important part of an iterative research process where both quantitative and qualitative research approaches are seen as complementary. The report draws attention to the type of qualitative approaches which have done much to explain why the returns to Apprenticeships vary.

Summary of the Options

The table below summarises the information which will be produced through adopting each of the options set out above. The table provides information relating to:

i. the main outcome variables, in particular whether these potentially extend beyond the measures of the likelihood of being in employment and wage levels;

ii. the time period existing data or new data would potentially cover. Ideally, data are required which follows people through their Apprenticeship and their initial labour market experience;

iii. the age groups covered – more specifically whether data will be provided for apprentices aged 16-18 / 19-24/ and over 24 years on entry to their Apprenticeship;

iv. the level of disaggregation by level (Level 2/3/4+) and broad subject area; and

v. the potential availability of comparator data to shed light on whether Apprenticeships are, from a number of different perspectives, better or worse than some alternative.

Which Option?

Options 0 and 1 are relatively low cost options in that the data are either already collected (Option 0) or would require a limited resource to link and match various data sets (in Option 1). Option 3 is also relatively low cost since it would require nothing more than a few additional questions being added to the LFS questionnaire.

Option 1 would provide much of the information required but not all. Information on wage returns and employment returns and how these vary by age and Apprenticeship framework and level (and other factors) would be available under this option. This option would not, however, provide information on more subjective measures such as job satisfaction and progression and how these and other outcomes are affected by individuals' attitudes and reasons for undertaking an Apprenticeship. Hence, whilst there is much to recommend Option 1 as the preferred option, in many respects the preferred option is a mix of Options 1 and 2: the use of administrative data linked to a YCS/LSYPE style multi-cohort longitudinal survey which follows people into their mid-20s - since this has the capacity to provide all of the data required to fully estimate the returns to Apprenticeship, providing robust estimates of the size of those returns over time and the factors which result in there being differential outcomes.

Such a study would require large sample sizes for each cohort (starting with a total of 10,000 interviews achieved) and would need to follow individuals for at least 3 to 5 years after completion. This would be a costly exercise. Nevertheless, BIS and DfE

have considerable experience of conducting the type of survey proposed and are aware of its capacity to deliver the results needed, and have expertise in linking administrative databases. It needs to be borne in mind that YCS/LSYPE, in its present form, does not provide information in relation to older apprentices so a supplementary survey of this group would be required. The preference for a further YCS/LSYPE type survey, however, does not preclude the other options being adopted; especially so where existing data sets could be linked or where existing data collection tools could be readily amended to provide much more information about Apprenticeships. Combining this type of multi-cohort longitudinal survey with linked administrative data would allow for the survey to contain considerable detail and information not typically collected in other surveys (e.g. attitudes to training, reasons for choice of subject) as much of the detail regarding the institutional attributes of the Apprenticeship itself, prior educational attainment and other factors would be supplied by administrative data.

Summary of the Information Potentially Provided by each Option

Option	Outcome variables	Time Period	Age Groups Available	Level of disaggregation by framework	Comparator data available	Other comments
Option 0: Rely on Existing Sources of Information	Limited to economic outcomes related to wages and employment	Difficult to provide a longitudinal record for an individual much beyond their initial training.(*i.e*. after they exit the ILR)	All potentially available	Difficult to address Level or subject once person has entered labour market	Yes	Major weakness is the difficulty of following people from their training through into the labour market.
Option 1: Linked Administrative Databases	Economic outcomes related to wages and employment	Potentially provides a longitudinal record from Apprenticeship through to retirement	Available for all age groups	Allows for disaggregation by Level and subject	Yes, since education and training activities of everyone for a given cohort is available	Outcome variables limited to wage and employment
Option 2: Extending YCS / LSYPE	Wage and employment plus more qualitative measures such as job satisfaction, match of skills to current job. e*tc.* are available	At present provides information up to the end of initial training for young people and some information about very early labour market experiences	At present YCS and LSYPE contain no information about older apprentices	Allows for disaggregation by Level and subject	Yes, since activities of sample of everyone of a given age is included	Sample would need to be substantially increased to meet needs of evaluation, would also need to follow people to older age to cover first five years in labour market

Summary of the Information Potentially Provided by each Option (continued)

Option	Outcome variables	Time Period	Age Groups Available	Level of disaggregation by framework	Comparator data available	Other comments
Option 3: Extending use of LFS	Economic outcomes related to wages and employment	Cross sectional data for people of working age	Potentially all using cross-sectional analysis	Additional questions would allow disaggregation by level and subject	Yes	Additional questions added as suggested will make it a valuable resource
Option 4: Maximising Value of Longitudinal Surveys	Variety of outcome variables available	Allows for an evaluation over the lifecycle but in practice attrition rates make this difficult	All age groups included	Generally allows for disaggregation by Level and subject	Yes	Apprentice sample sizes would need to be substantially boosted.
Option 5: Employer Dimension	Variety of outcome variables potentially available through surveys such as NESS (*e.g.* alleviating skill shortages).	Data currently cross-sectional.	Current employer surveys ask about apprentices aged 16-18 / 19-24 / 24 plus	Potentially available but would require additional questions in surveys such as NESS	Yes	Potentially provides information about the characteristic of an apprentice's employer
Option 6: Qualitative Research	Allows new outcome measures to be explored	Potentially any period can be looked at	Potentially all ages can be included	Potentially can look at specific levels and/or subjects	Qualitative approaches should include some comparative dimension	Qualitative approaches allow for explorative inquiries which can be picked up later by quantitative inquiries.

1. Introduction

1.1 The Need to Evaluate Apprenticeships

Apprenticeship is the principal work-based training pathway through further education. There has been a longstanding debate in the UK, especially so in England, about the value of Apprenticeships relative to either other forms of vocational preparation or programmes which operate in other countries.[1] Compared with the attention given to higher education and the various studies which have sought to reveal the private and public returns to this form of learning, relatively little attention has been paid to Apprenticeships. Yet current policy places a strong emphasis on Apprenticeships delivering the skills the country needs. The recent White Paper – *Skills for Sustainable Growth*[2] – commented:

Our goal is to build a skills system where responsibility for quality and investment is shared between Government, employers and learners; where those using the system are in the driving seat and can select training and qualifications that are designed and valued by business, prepare them for worthwhile careers and provide a foundation for further learning.

Apprenticeships will be at the heart of this. They bring together individuals, motivated and working hard to develop themselves, employers, investing in their own success but supporting a programme with wider social, environmental and economic value, and Government providing public funding and building the prestige and reputation of the programme. (p.16)

If would-be apprentices, and employers, are to make informed decisions about their investments in human capital this needs to be supported by evidence which shows the likely returns over the long term. To date, this evidence is rather limited with respect to Apprenticeships in England, though there are a number of studies which do their best with the limited data available.[3] It is not just the returns which are of interest. From a public policy perspective there is a need to understand the processes which result in differential returns to Apprenticeships of different kinds. This is likely to relate to what happens during the Apprenticeship training period, as well as what happens afterwards, given the extent to which Apprenticeships prepare people to engage in further learning and take advantage of opportunities for career progression. Apprenticeship, as will be expanded on below, is a heterogeneous programme of vocational preparation with substantial differences in, for instance, the duration of the Apprenticeship and the volume of formal training delivered,

1 Steedman, H. (2001) Benchmarking apprenticeship: UK and continental Europe compared. CEPDP, 513. Centre for Economic Performance, London School of Economics and Political Science, London.

2 Available at: http://www.bis.gov.uk/assets/biscore/further-education-skills/docs/s/10-1274-skills-for-sustainable-growth-strategy.pdf

3 McIntosh, S. (2007) 'A Cost-Benefit Analysis of Apprenticeships and Other Vocational Qualifications', Department for Education and Skills, Research Report 834.

depending upon the Framework being considered. Even within Frameworks there is likely to be variation with respect to the delivery of the Apprenticeship. The inter- and intra-Framework variation in the delivery of Apprenticeships is likely to be important in understanding the returns eventually obtained by the apprentice, the employer, or the State. Hence any evaluation strategy needs to take into account explanatory variables and not focus solely on measuring outcomes.

The study builds upon the earlier study by Joan Payne and her colleagues which considered the options for evaluating Apprenticeships a decade ago.[4] Much has changed since then with respect to both Apprenticeships and the collection of data which makes the current review timely.

1.2 Aims and Objectives

The aims and objectives of the study are to investigate the options available for a long-term evaluation of Apprenticeships which will:

i. produce robust quantified estimates of the outcomes of Apprenticeships for successful trainees, in terms of various indicators including progression into further and/or higher learning, improved employment outcomes, and greater job satisfaction; and

ii. establish under what conditions Apprenticeships produce the best results, for whom and *why*, and the relative added value of an Apprenticeship compared to other forms of learning.

The second objective is an important one. Not only is it important to show what the returns to Apprenticeship are, but additionally there is a need to provide information which sheds light on why the returns might differ according to the characteristics of both individual apprentices and the Apprenticeships they are working towards completing. This could potentially have important implications for policy development.

With the above two objectives in mind, the study assesses how to track young people and adults into Apprenticeships, through their training (including into any subsequent learning) and then into employment, demonstrating the extent to which their future earnings and employment status are influenced by the experience of their Apprenticeship. Of key interest therefore are the outcomes for apprentices, relative to the most appropriate comparator group. This is considered in greater detail below. Bearing in mind (ii) above, the study looks at how data can be disaggregated by a number of different factors to obtain a greater understanding of how and why Apprenticeships delivered to certain individuals (e.g. by socio-demographic

4 Payne, J., Riley, R. and Coleman, N. (2001) 'Feasibility Study for the Long Term Evaluation of Modern Apprenticeships', Department for Education and Skills, Research Report 290.

characteristics) and in particular circumstances (e.g. at different levels and in different sectors) deliver better outcomes.

The primary objective of the study is to identify the potential methods for the evaluation and assess each one of them on their merits.

1.3 Structure of the Report

The report is structured as follows. Section 2 provides the overall approach to the study, while Section 3 builds upon this to specify an economic model to evaluate the impact of Apprenticeship training on an individual's progression through the labour market. The model is used to highlight the data required to evaluate the impact of Apprenticeships over the long-term. The extent to which these data are available is described in Section 4. Section 5 discusses the approach to evaluation of Apprenticeship in other countries, including information on the main sources of data available. Finally, Section 6 provides recommendations about how data gaps might be efficiently filled to allow for a much more nuanced and informed evaluation of the impact of Apprenticeships over the long-term.

2. Method

2.1 Over-Arching Approach

In order to identify the data required to evaluate Apprenticeships, the starting point was to develop an economic model, or rationale, which sets out the relationships between the dependent variables of interest and a number of explanatory ones. This model is described in Section 3. The model is used to identify what data are already available to evaluate Apprenticeships and to identify data gaps. The study, however, is not solely an exploration of economic statistics. In Grubb and Ryan's, *The Roles of Evaluation for Vocational Education and Training*, a practical guide is provided for those engaged in assessing the contribution of various training activities. As the book's subtitle implies – *Plain Talk in the Field of Dreams* – it emphasises the need for an approach that is, at one and the same time, both practical and capable of delivering meaningful and timely results to those with an interest in VET policy.[5] It emphasises the need for a well-rounded view of, in this case, Apprenticeships.

With respect to evaluation of VET interventions such as Apprenticeship, Grubb and Ryan[6] provide a number of recommendations which warrant adherence to:

i. evaluations of VET programmes should never lose sight of labour market outcomes, but in addition they should be concerned with processes leading to results;

ii. the analysis of VET programmes should try to use a variety of evaluation methods, since each of them is partial and incomplete;

iii. VET evaluations should consider a broader range of outcome measures in preference to a narrower range;

iv. evaluations should consider long run as well as short run effects of VET programmes;

v. evaluations of VET programmes should be concerned not only with their efficiency outcomes, but also their effects on equity.

These are useful guidelines for scoping the long-term evaluation of Apprenticeships by which the current study has sought to abide.

---

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

6 ibid

2.2 What is an Apprenticeship?

There is a need to clearly define the phenomenon under investigation. Apprenticeship, as a publicly funded training programme has become the principal work-based training option for many young people.[7] This is the phenomenon to be evaluated over the long-term. As a training programme it is one which has been subject to much review and reform in its modern incarnation following the introduction of Modern Apprenticeships in the mid-1990s. Given that Apprenticeships have been subject to substantial reform in the past, and might be again in the future as the programme adapts to changing needs of the labour market, any on-going long-term evaluation suggests a need for a flexible evaluative approach.

Apprenticeship in England is, in essence, a system which comprises a number of programmes (Frameworks). Whilst Frameworks contain a number of common elements they also reveal a wide variety of activity depending upon, amongst other things, the content of the Framework and the level of learning. Indeed this is considered to be one of the strengths of the Apprenticeship system in England - a system which is responsive and adaptable to labour market demand and, in so doing, meets the skills needs of both the employer and the apprentice.[8,9]

Whilst there will be occasions when there is a need to report on Apprenticeships as a whole, it needs to be borne in mind that there is substantial variety by both sector and level. A Level 3 Apprenticeship under one of the Engineering Frameworks is likely to generate significantly different returns to the apprentice compared with, for instance, a Level 2 in Social Care. The duration of the training differs, the level of employer and employee investment differs, and the value-added in each respective sector is different. Similarly, the returns which accrue to the different age groups are also likely to vary according to whether the individual is aged: 16-18 years; 19-24 years; or over 24 years (adult Apprenticeships). There is a need, therefore, to disaggregate Apprenticeships in such a way as to account for these factors.

7 It is recognised that Apprenticeships are available to older people but they have other work-based training options available to them depending upon their experience and skills.

8 House of Lords (2007) Apprenticeship: a key route to skills: Volume I - Report, House of Lords Select Committee on Economic Affairs, London: TSO.

9 Hasluck, C. et al., (2008) The Net Benefits of Training Apprentices, London: Apprenticeship Ambassadors Network.

2.3 Who or What is to be Evaluated?

If, for the purposes of tractability, the evaluation is concerned in the first instance with individual apprentices and what happens to them after they have completed their training and begin their long haul through the labour market, there are a number of outcomes which are of interest, including:

i. employability;

ii. wage levels;

iii. occupational mobility and career progression; and

iv. progression to other education and training programmes / courses.

The selection of dependent variables is, to a large extent, guided by the question: is Apprenticeship in some way better than the alternatives available?

The outcomes for apprentices can be viewed over different time periods, including taking a whole life approach. In practice, even though longitudinal data are increasingly becoming available (e.g. BCS, MCS, YCS, LSYPE, etc.) much research still concentrates on employment and earnings effects over the short-term. Nevertheless, given the increasingly rich store of longitudinal data, the study assesses the extent to which longitudinal data sets will allow a much longer term view of the impact Apprenticeship has upon those who were trained in this way.

Bearing in mind the comments above, which drew on Ryan and Grubb's evaluation guidelines, there is a need to take a more holistic approach than simply concentrating on a relatively small number of economic outcomes related to wages and employability. There is the possibility to assess the extent to which an evaluation might look at a wider set of outcomes. At best, earnings are a fairly crude indicator of skill levels – despite what human capital theory has to say - so there is a need to look at the extent to which there are alternative measures of skills available in some data sets. There is also a burgeoning literature on the determinants of job satisfaction which draws attention to the fact that where an individual's skills are poorly matched to the demands of their job, this can result in job dissatisfaction and reduced job tenure. Potentially, there is a range of variables relating to these types of issues which might be usefully incorporated in any evaluation.

The benefits or returns to Apprenticeship are not necessarily shared equally across all groups. As will be explained below, Apprenticeship covers a wide range of activities which require differing levels of both financial and intellectual inputs from employers (e.g. with respect to the ratio between work and training) and apprentices (e.g. the amount of foregone income). This suggests that the returns to Apprenticeship training may well vary according to the respective characteristics of the employer, the apprentice, and the Apprenticeship being undertaken. There is a need then, other things being equal, to identify who benefits most. This is

particularly apposite given the efforts of Government to widen participation in Apprenticeships.

It is important to devise a methodological approach which is firmly rooted in understanding the behaviour of different actors. Unlike statistical models, econometric models are predicated on behavioural relationships which can be observed and measured. Ideally, econometric models should provide some insights into why returns to different types of Apprenticeship vary, such as the existence of wage rigidities and pay structures in some sectors of the economy. Though in practice, this may be a tall order. This study has outlined an economic model designed to highlight data needs.

2.4 Evaluating Apprenticeships

There are a number of different approaches to evaluating programmes of one type or another. These are well known and the strengths and weaknesses of each approach have been subject to considerable scrutiny. The general approaches are:

- randomised control trials (RCT) – which are unlikely ever to prove tenable with respect to Apprenticeships because of associated ethical issues;[10]

- attempts to compare or match apprentices to those who have the same characteristics as apprentices except that they have made another choice with respect to VET. Typically this uses propensity score matching techniques - which can compare favourably with the results which accrue for RCT designs[11] – or to identify instrumental variables (in practice it is difficult to identify any);

- fixed effects or difference in differences approaches which tend to compare different groups before and after the 'treatment'.

The general comments made in the initial feasibility study by Joan Payne and her colleagues about each of the approaches listed above are still valid today.[12] Persistent problems encountered in carrying out such analysis include sample attrition and selection bias. Whilst there are techniques available to deal with these problems, because relatively little is known about the labour market experiences of apprentices before they start their Apprenticeship (mainly because they have no experience) or for those who do not complete, a degree of uncertainty is attached to

10 Even in those large scale experiments where RCTs have been used to investigate the effect of training programmes on wages, formidable statistical issues remain to be addressed (Zhang et al, 2008).

11 Shadish, W. R., Clark, M.H and Steiner, P. M. (2008) 'Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random to Nonrandom Assignment', Journal of the American Statistical Association, 103: 1334-1343.

12 Payne et al,ibid

how successful these measures are in practice.[13,14,15] With the emergence of longitudinal datasets which capture information from would-be apprentices before they commence their Apprenticeships, and which may provide more information on those who fail to or partially complete programmes, the use of these statistical techniques is now improved. It remains the case, as with the Payne study, that RCTs are infeasible with respect to evaluating Apprenticeship. This is addressed in more detail later on.

2.5 Why the Need for the Current Review?

Since the initial feasibility study undertaken by Payne and her colleagues,[16] much has happened in the economy, the FE system, and the Apprenticeship system to warrant the new study, including the continued expansion of Apprenticeships into non-traditional sectors and the changing demographic profile of apprentices. Much of the demographic shift can be attributed to policies directed at widening participation, particularly amongst harder to reach learners. It is also apparent that the education and learning alternatives available to people who might consider undertaking an Apprenticeship have changed or are about to change (*c.f.* the future funding of higher education). Moreover, the range of datasets available which potentially allow the returns to particular learning programmes to be assessed has also improved since the study by Payne and her colleagues. For all the reasons mentioned above, the current review is most timely.

13 Heckman, J. (1979) 'Sample selection bias as a specification error'. Econometrica, 47: 153–61.

14 LaLonde, R. (1986) 'Evaluating the econometric evaluations of training programs with experimental data'. American Economic Review, 76: 604-620.

15 Heckman, J. and Navarro-Lazano, S. (2003) 'Using matching, instrumental variables and control functions to estimate economic choice models'. NBER Working Paper No. 9497.

16 Payne et al, ibid

3. An Initial Model for Evaluating Apprenticeships

3.1 The Evaluation Problem

In evaluating any labour market/training programme, such as Apprenticeship, the 'evaluation problem' is encountered. The 'evaluation problem' arises due to the fact that the true counterfactual, what would have happened to participants/non-participants had they not participated/participated, cannot be observed. In order to truly estimate the impact of Apprenticeship on various labour market outcomes, the observed outcome needs to be compared with the outcome that would have come about had the apprentice not undergone the training.

The most robust solution to the 'evaluation problem' is the use of randomised controlled trials (RCTs). In an RCT, a pool of potential participants (based on whatever criteria of choice with, ideally, matching of individuals within the group on the basis of similar/identical characteristics) would be chosen and individuals randomly allocated to an Apprenticeship programme or not. The group not taking part in Apprenticeship training would form the control group. Such a set-up would, theoretically, eliminate self-selection problems that are encountered when people are observed retrospectively.

There are drawbacks to using RCTs, however, including costs and ethical considerations. RCTs are costly to set up and implement and require close monitoring of both control and treatment groups in order to ensure that the experiment is properly administered and to collect data. Moreover, random assignment cannot guarantee comparability of the treatment and control group. All samples, even random samples, contain potential sampling error. The statistical basis of an RCT is that the average characteristics of repeatedly drawn random samples will converge. In practice it is mainly the case that just two samples (control and treatment) are drawn and these may differ due to sampling error. This concern diminishes as sample sizes increase and sampling error is reduced. Where sample sizes can be increased the consequence may be increased cost while in other instances sample sizes are limited by the size of the underlying population of interest.

There are also ethical concerns over denying participation in Apprenticeship (or any other programme). In the case of Apprenticeships, it is unlikely that denying access to individuals for the purposes of experimental evaluation would be seen as fair and would undoubtedly cause political issues. RCTs themselves may also throw up more practical problems in that the actual experiment may alter the behaviours of both treated and control individuals. 'Randomisation bias' may, for example, arise if individuals who have been randomly denied access to Apprenticeship training

become de-motivated as a result.[17] This would result in non-apprentices having even worse labour market outcomes, not only because they did not undertake an Apprenticeship but also because of the effect that their denied participation had on their underlying behaviours.

The use of RCTs in the evaluation of Apprenticeships has been ruled out as the programme is already in full operation and to attempt to randomly assign individuals would be looked upon unfavourably from an ethical and political point of view. In practical terms, it is also difficult to visualise how individuals could be randomly assigned to the set of alternatives: (a) direct entry to the labour market; (b) academic pathway through further education; (c) neither education, employment or training; or (d) Apprenticeship. The difficulty of generating a comparison group is also enhanced due to the recent emphasis on increasing participant numbers and providing Apprenticeships for all young people wishing to undertake such training. As a programme such as Apprenticeship begins to resemble an entitlement programme, carrying out an RCT becomes even less feasible and appropriate.

A number of non-experimental approaches to evaluating programmes such as Apprenticeship exist and the choice of method depends mainly on practical concerns including the characteristics of the programme and the nature and quality of data that are available. In the absence of experimental evidence on the causal links from Apprenticeship to labour market outcomes it is necessary to make as few assumptions as possible in using a non-experimental approach to ensure that any estimated effects come close to the true effects of Apprenticeship on employment outcomes.

The discussion that follows considers the options for modelling the longer-term outcomes associated with completing an Apprenticeship when an RCT is not feasible. The discussion examines the approaches that might be adopted without consideration of whether or not the data required is currently available. Clearly, in practice any modelling is constrained by the availability of suitable data but it is illuminating to commence by considering all of the options. Whether the current data are adequate to the task of modelling long-term outcomes and, if such data are inadequate, whether, and what, steps should be taken in order to collect the required data in the future are considered in Sections 4 and 6.

17 Heckman, J. and Smith, J. (1995) 'Assessing the Case for Social Experiments' Journal of Economic Perspectives, 9(2): 85-100.

3.2 Modelling the Long Term Outcomes of Apprenticeship

3.2.1 The model specification

The basic premise to be investigated is that completion of an Apprenticeship results in some positive future outcome to the individual apprentice. In terms of modelling the longer-term outcomes of Apprenticeship, let us represent a selected outcome (e.g. earnings, employment status, etc.) as $Y_{i(t+n)}$ for individual *i* in period *t*+*n*. The period of completion is indicated by *t* and the period in which the outcome is observed is some n periods after completion thus *t*+*n*. In general terms, the model of interest can be specified as:

$$Y_{i(t+n)} = \beta_0 + \beta_1 A_{i(t)} + \beta_k X_{ik(t)} + \varepsilon_{i(t+n)} \qquad (1)$$

Where $A_{i(t)}$ is a dummy variable representing whether or not individual *i* has completed an Apprenticeship (1 has completed; 0 has not completed) in period *t*, $X_{ik(t)}$ represents a range of other factors (*k* in number) including personal characteristics of the individual. $\beta_0, \beta_1, \beta_k$ are coefficients which reflect the impact of the respective variables on the outcome, *Y*. Equation 1 indicates the additional impact of completing an Apprenticeship on the outcome of interest over and above the influence of other factors (which can be ascertained when $A_{i(t)}$ is set to zero). $\varepsilon_{i(t+n)}$ is a random error term.

This simple specification of the model highlights a number of issues that the modelling of the long term outcomes of Apprenticeship must address. These include the following questions:

- What are the other factors ($X_{ik(t)}$) that bear on the outcome and would lead to variations in outcomes and returns to Apprenticeship?
- What is the outcome of interest ($Y_{i(t+n)}$) and how is it defined?
- When is the outcome of interest to be observed (what is time '*n*')?
- How is completion of an Apprenticeship ($A_{i(t)}$=0,1) to be defined?

These issues are considered in turn in sections 3.2.2 to 3.2.4 below.

3.2.2 Explanatory Variables

Training as an apprentice is only one of many factors determining future earnings and career progression. In the model equation (1) above, $A_{i(t)}$ represents the 'pure' impact of undertaking an Apprenticeship after all the other influences on outcomes ($X_{ik(t)}$) have been taken into account. The influence of these other factors means that there will be a range of outcomes from an Apprenticeship, even within a given Apprenticeship framework. The range of outcomes (and returns) will reflect the personal characteristics of apprentices, the nature and quality of their training, the

characteristics of their jobs after completing their Apprenticeship, as well as prevailing economic and labour market conditions.

Two further considerations need to be taken into account. First, some of these other factors will be inter-related. Some Apprenticeship frameworks are, for instance, predominantly undertaken by men and others by women so that separating the impact of the framework from that of gender could be difficult. Second, some variation in the returns to Apprenticeship will be observed simply because it is impossible to take account of every factor affecting outcomes. In addition, because all explanatory factors are not readily observable or, if observable, some can be difficult to quantify or measure.

Any model of the longer-term returns to Apprenticeship will, therefore, need to include a range of explanatory variables in addition to a simple binary 'apprentice or not an apprentice' variable. Relevant explanatory variables can be considered as falling into a number of broad groups and these are set out below.

Personal characteristics: Gender, age, ethnicity and health/disability are likely to be important personal characteristics associated with different outcomes. Such variables may be important in their own right but more often reflect other factors with which they are associated. For instance, differences in the returns to men and women may be a reflection of gender differences in the Apprenticeship frameworks undertaken and the job opportunities available post-Apprenticeship. Similarly, an apprentice with a disability may face a more limited choice of Apprenticeships, have prior educational attainment that was adversely affected by their disability or poor health, and face restricted job opportunities after completion.

There is also evidence from qualitative studies of education and training that attitudes and motivation are also important factors in determining outcomes but these factors are difficult to observe and quantify. Educational attainment prior to entry to an Apprenticeship may be taken as an indicator of ability or of motivation. Similarly an apprentice's prior work history may also be used to indicate their attitude and motivation to work and employment. Again, there may be links between variables. For instance, motivation and age could be related to one another if older apprentices have a different (more mature) attitude to their training.

The Nature and Level of the Apprenticeship: Apprenticeship does not comprise a homogeneous programme of vocational education and training. There are considerable differences in terms of level, delivery and content of different frameworks. Identifying the influence of such factors would not only improve the modelling of outcomes but would also help identify the attributes of different Apprenticeships and the training associated with the best outcomes and inform future development of the programme.

At a minimum, details on the level (Level 2 or Level 3) and subject (as indicated by the Framework) would be essential covariates to include in the model. There is considerable evidence that learning at different levels leads to different returns. From a human capital point of view, this is to be expected since the investment undertaken by an apprentice at Level 3 is greater than at Level 2 and this will be

reflected (in a competitive job market) in higher future earnings for Level 3 apprentices. Both McIntosh and Dearden *et al* found the impact of vocational qualifications on the probability of being in employment was substantially greater for Level 3 qualifications than Level 2, while Dickerson and Vignoles found a positive wage return to gaining a Level 3 qualification (although the wage return to Level 2 depended critically on whether the qualification was an 'academic' or 'vocational' one).[18,19,20] These studies relate to qualifications in general, but it is plausible that these differences will also be found in regard to Apprenticeship, as found in McIntosh (2007).

In addition to the level of the qualifications attained as part of an Apprenticeship, the framework itself will be important, most critically because it will influence, even largely determine, the employment opportunities available upon completion. Apprentices training to enter occupations in the engineering or construction industry, for instance, are potentially destined for well-paid jobs with excellent career prospects. Other Apprenticeships, such as those in social care, while no less worthy, are training people to enter a career in a sector with relatively low pay and more limited career progression. Inevitably it can be expected that the financial return to the former apprentices will be greater than the returns to the latter.

Such differences are not restricted to differences between frameworks but will also occur within frameworks reflecting differences across employers of apprentices (e.g., size of employer, markets served, public or private sector). Differences in outcomes will also be found where frameworks cut across sectors. Examples of the latter include frameworks such as information technology (IT) and business administration where such apprentices can be found employed in a range of activities, both in the private and public sectors.

Framework and sector/industry are thus critical variables to include in the Apprenticeship returns model. As already observed, there is likely to be some interaction between some personal characteristics – such as gender and age – and frameworks and sectors. It is known, for instance, that female apprentices in sectors/frameworks dominated by male apprentices have a lower probability of completing their Apprenticeship (and vice versa).[21] Whether this disadvantage extends into the subsequent career is not known but is plausible.

***The Quality of Apprenticeship*:** Apprenticeships vary in quality across frameworks, often reflecting the nature and practices of the sectors in which apprentices will

18 McIntosh S. (2004).The impact of vocational qualifications on the labour market outcomes of low achieving school leavers, Centre for Economic Performance Discussion Paper 621; McIntosh, S. (2007). A Cost Benefit Analysis of Apprenticeships and Other Vocational Qualifications, Department for Education and Skills Research Paper RR834.

19 Dearden L., McIntosh S., Myck M. and Vignoles A. (2001),The Returns to Academic, Vocational and Basic Skills in Britain, DfEE Research Report, London: DfEE.

20 Dickerson A. and Vignoles A.(2007), 'The distribution and returns to qualifications in the Sector Skills Councils', Sector Skills Development Agency, ISBN: 978-0-9552029-8-8.

21 Gambin L., Hogarth T. and Hasluck C. (2010), 'Maximising Apprenticeship Completion Rates in England', Canadian Apprenticeship Journal, Issue No.4.

eventually work. Even within a specific framework there is likely to be some variation in the quality of Apprenticeships, perhaps indicating different employer attitudes to Apprenticeships, and this would be reflected in the returns to Apprenticeship. Where employers provide high quality training it can be expected that the return to an Apprenticeship will be high. It remains to be seen whether this applies mainly to the obvious 'blue ribbon' organisations or whether apprentices in smaller businesses offering high quality training also experience high returns. Some employers see Apprenticeship as a vehicle to acquire relatively cheap labour and where this is reflected in the quality of the training offered or the calibre of the apprentices hired then returns to the Apprenticeship may be lower than those obtained by other apprentices.

The 'quality' of an Apprenticeship is challenging to measure, not least because it is, to some extent, a subjective judgement. What suits one apprentice may not suit another and retrospective apprentice assessments of quality may also be conditioned by the outcomes they experienced. Nonetheless, it is important to seek objective measures of quality and to include them in the model. Possible indicators of quality are the amount of time spent in the workplace, the length of time undertaking formal learning activities off-the-job and so on.

A further dimension to the quality of Apprenticeships might relate to the quality of the training provider. The former Learning and Skills Council (LSC) provided an assessment through its Minimum Levels of Performance criteria. In general, the quality of the training provider is largely ignored in econometric analyses of Apprenticeship outcomes.

Economic and Labour Market Conditions: Macroeconomic conditions and the state of the job market will have an influence on the outcomes and returns to an Apprenticeship. There are several aspects to this, both short-term and long-term. General fluctuations in economic activity will condition the job opportunities available to apprentices, both on completion and subsequently. The opportunities facing apprentices in a recession will be more limited than for those who complete in a period of economic growth. There may even be some continuing career 'penalty' for the former who, perhaps, make sub-optimal career choices in order to gain employment. Moreover, sectors and occupations differ in their sensitivity to fluctuations in economic activity and some apprentices will be affected by such change to a greater or lesser degree. These effects could be reflected in employment stability (or otherwise) and earnings.

In addition to fluctuations in labour market conditions, the returns to Apprenticeship may also reflect longer-term trends in the demand for skills and occupations. It should be noted that the return to an investment in an Apprenticeship (by the individual or the employer) may not turn out in the event to be what was expected. Some trades (especially the traditional manual trades) are in steady decline while others are vulnerable to technological change. It is possible that the skills created by an Apprenticeship undertaken in 2011 could be largely obsolete after 30 years, and this would be reflected in the employment prospects and pay of such apprentices.

The impact of cyclical economic fluctuations and longer-term occupational change is often focussed on local labour markets. Apprentices resident in regions or areas where there is declining demand for their skills will obtain a lower return to their Apprenticeship than other apprentices located elsewhere or in growing occupations. This negative impact might be mitigated if apprentices were geographically mobile (they would go to where the returns were higher) but not all apprentices are likely to be equally mobile. Age and, particularly, gender are likely to be factors here, with women (especially those married or with children) and older ex-apprentices being least mobile and thus more exposed to variations in local labour demand. A number of indicators are available to indicate general labour market conditions including vacancy/unemployment ratios or even, simply, the local unemployment rate.

Selecting Covariates for the Model: The discussion so far has set out the many factors that could be expected to impact upon the outcomes from an Apprenticeship and thus explain variations in such returns. In an ideal world measures of all these factors would be included in the economic model. In practice what is feasible to include in the model will be determined by data that has been collected from surveys and administrative databases. The following set of explanatory variables might be regarded as a 'yardstick' against which the capacity of any particular dataset to 'explain' longer-term outcomes could be assessed or as a specification for variables to be collected in any bespoke data collection exercise:

- gender;
- age at completion and when outcomes observed;
- ethnicity;
- household status – marital status; other household members (e.g. children, dependents);
- presence of work limiting disability or health condition;
- prior educational attainment/qualifications on entry to the Apprenticeship;
- pre-Apprenticeship work history;
- indicators of an individual's motivation to work towards completing an Apprenticeship;
- the Apprenticeship – Framework; level of NVQ; completion status;
- measures of training quality – duration of training; time in formal, off-the-job training;
- nature of Apprenticeship employer – public/private/voluntary/other; sector/industry; size of employer;

- post-Apprenticeship learning;
- labour market demand – macroeconomic indicators; measures of local labour demand; measures of occupational trends.

As always, the full range of desirable covariates available for modelling purposes will be conditioned by the availability of data and the suitability of the form in which such data is collected. This availability/suitability is considered in more detail in Section 4. It is likely that data of the type referred to above will have been collected in different datasets. Where that is the case then linking such data would be one way to create the required full dataset needed to adequately explain variations in the longer-term outcomes and returns to Apprenticeship.

3.2.3 Defining and Measuring Outcomes

The hypothesis to be evaluated is that completing an Apprenticeship leads to some economic benefit to the individual in the future. The positive outcome ($Y_{i(t+n)}$) could be an indicator of:

- higher earnings;
- job retention;
- career progression and advancement;
- progression onto other learning or qualifications;
- occupational or social mobility.

The successful completion of an Apprenticeship can be expected to lead to higher lifetime earnings for the individual (that, presumably, was the basis of the decision to invest in an Apprenticeship). In some respects the process, or processes, by which such additional future earnings come about does not matter (although it would be useful to know) as future earnings can be considered as a kind of 'reduced form' of outcome that encapsulates the combined effect of several different processes or, more probably, a combination of processes. Nonetheless, to understand how such additional lifetime earnings come about it is necessary to consider the processes involved by examining evidence relating to the links in the causal chain leading to higher lifetime earnings. These intermediate outcomes would include higher pay through increased productivity (resulting from access to skilled worker wage levels), access to higher paid occupations, improved job retention and career progression. This may also increase our understanding of *why* apprenticeships deliver better outcomes in certain situations, and thus has implications for policy development.

It is possible that the whole of an Apprenticeship is not the key to higher lifetime earnings but the result of just some parts of the Apprenticeship training. Technical skills acquired through on-the-job training in the programme may result in employees becoming more productive and resulting in higher wages. Alternatively, study skills acquired through the programme may lead apprentices to undertake further

training/education and to perform better in such programmes, eventually resulting in better employment prospects and higher wages. The causal pathways between Apprenticeship participation/completion and higher lifetime earnings/reduced unemployment/career progression are undoubtedly many in number. Exploring each and every facet of these relationships is impossible under the constraints posed by existing datasets. Nevertheless, untangling the various channels through which such benefits come about is a desirable objective and considerations for how to incorporate data that could serve such a purpose is discussed (in limited fashion) in Section 4.

***Earnings*:** Various indicators of earnings can be used in evaluating Apprenticeships. Most studies of the impact of education or training on earnings consider gross hourly earnings. This standardises for variations in hours of work. Nonetheless, it is also necessary to consider whether Apprenticeships are systematically associated with differences in hours of work (higher or lower than non-apprentices). McIntosh[22] uses weekly earnings, in order to convert to annual wage increases, for the purposes of estimating lifetime benefits. From the point of view of raising productivity and maximising the impact of government investment, however the labour supply decisions of former apprentices are important. For example, hourly earnings may be greater, but if this is (at least partly) offset by a reduction in hours worked, then the impact on total productivity would be overestimated. Equally, if hours worked increase as a result of the Apprenticeship the productivity impact would be underestimated.

***Employment retention*:** Under the current definition of Apprenticeship, participants must have, in nearly all instances, employee status within a business. This means that there is no employment retention effect during the Apprenticeship itself. There is, however, the important question of whether or not the apprentice is retained by their employer upon completion and what factors are associated with such retention. In the longer term the relevant question is whether or not apprentices have more stable employment (or fewer spells of not working) than non-apprentices. If apprentices entered jobs that were more stable than non-apprentices, then higher lifetime earnings would be the result (even if there were no significant differences in wages when in employment).

There are good grounds for believing that employers are more likely to retain skilled workers over the economic cycle (particularly those in whom they have invested through training) than less skilled workers. Employment retention could be measured in several ways depending upon the data available. It could be measured, for instance, by:

22 McIntosh (2007), ibid.

- whether or not individual is in employment at some future date (say 1 year, 5 year etc.) after completing the Apprenticeship (represented by a binary (1/0) variable);[23]

- the duration of continuous employment with the same employer in some period prior to the date at which outcomes are evaluated;

- the duration of continuous employment (in different jobs) over some period prior to the date at which outcomes are evaluated;

- the cumulative duration of employment (regardless of any breaks in employment) over some period prior to the date at which outcomes are evaluated.

In an ideal world the period referred to in the above measures would be the lifetime (i.e. until retirement/death) in order to fully capture the effect of Apprenticeship training (or indeed any training/education) on all employment outcomes. As will be seen in the discussion of data in Section 4, in the real world, data constraints place restrictions on the time period to be used in evaluation approaches.

***Career progression*:** Many employers report that apprentices form the main source from which they obtain supervisors and managers. Measuring career progression could take the form of estimated probabilities of making a transition from one occupation to a 'higher' occupation with supervisory or managerial responsibilities. To produce such a measure would require longitudinal data with work histories containing 'fine grain' occupational information. Alternately, changes in employment status if recorded in the data might be used (e.g., manual/non-manual/manager etc.).

***Access to additional learning*:** Another positive outcome from Apprenticeship could be increased participation in, and completion of, further learning and acquisition of additional qualifications. These are positive outcomes in their own right but they may also reinforce or add to other effects stemming from Apprenticeship. This could be taken into account by examining the post-Apprenticeship qualifications of individuals (if such information is collected in the dataset).[24] The participation of individuals in employer provided training after completion of the Apprenticeship is another outcome of potential interest. Do apprentices tend to get jobs where employers are more inclined to continue investing in their employees' skills by providing various types of training? This provision of training by employers can be considered a factor in determining the quality of a job, which is itself an outcome of interest.

---

23 The resulting estimate of the dependent variable will then be interpreted as the probability of being in employment at that future date.

24 Analysis of such progression and the attitudes of employers towards this progression is explored in Kewin, J., Hughes,T., Fletcher, T. and Sheen, J. (2011) "The Road Less Travelled: Experiences of employers that support the progression of Advanced Apprentices to higher education" prepared by CFE for the Skills for Sustainable Communities Lifelong Learning Network. Leicester: CFE. Available at: http://www.lifelonglearningnetworks.org.uk/uploads/document/782/the-road-less-travelled.pdf

***Job Quality and Job Satisfaction*:** The economic model set out above is principally concerned with wages, occupational progression, and job security. Such outcomes may not be the only 'returns' that are valued by apprentices. Within organisational psychology, rather than economics, there has been a longstanding interest in how workers may be motivated to bring about improvements in organisational performance (*c.f.* the respective work of Herzberg, and Haslam).[25] Herzberg, for example, distinguishes between those factors which are likely to cause dissatisfaction at work, such as conflict over wage levels, and those factors which bring about satisfaction with a job, such as autonomy, control over the pace of work, etc.

Whilst it is difficult to account for indicators of job quality in a formal economic model concerned principally with wages and job security, it has become increasingly recognised that these are important outcomes for the individual in their own right and are associated with the generation of relatively high wage levels. It is important therefore that any future quantitative modelling attempts to incorporate outcomes relating to job satisfaction. These might be:

i. indicators of job satisfaction and overall satisfaction with the individual's position in life (some longitudinal studies capture these type of data);

ii. the extent to which an individual's skills are fully utilised in their current job (the SKOPE Skills Surveys asked questions along these lines);

iii. measures of work-life balance (e.g. as captured in the former DTI series of work-life balance surveys); and

iv. indicators of job quality including the physical conditions of work (e.g. Eurofound's European Working Conditions Surveys provide some comparative information about these issues).

To some extent these are second order issues to those which address wages and job security but they may well shed some light on understanding why some apprentices and some sectors of the economy obtain higher returns than others, or relative to some comparator group. Longitudinal data sets are increasingly including questions not just about job satisfaction but overall life satisfaction and relating these to the overall health of the individual. This has implications for staying in work and absenteeism, both of which will be related to the economic returns obtained by an individual over the lifecycle.

There are also statistical issues to resolve if incorporating job quality / job satisfaction into an economic model which attempts to explain any differences in wage levels or job security – especially since the direction of causality is not obvious – but it is well worth exploring how such factors can be included in any long-term evaluation of Apprenticeships.

---

25 Herzberg, F. et al. (1959) The Motivation to Work, Wiley, New York; Haslam, A. (2004) Organisational Psychology: The Social Identity Approach, Sage, London.

3.2.4 Timescales

It is a key aspect of the evaluation that the impact of Apprenticeship is to be assessed in respect of future outcomes, in particular, long-term outcomes (and ideally lifetime outcomes). Indeed, it is an essential feature of the model (especially in its fixed effects form) that observations are taken on at least two occasions. A critical question is when such observations should be taken. Many studies of the returns to learning take a short-term view and consider outcomes just one or two years after undertaking that training. On the other hand, studies of the returns to higher education tend to take a lifetime perspective. There are several possibilities for the period to be considered after completion of an Apprenticeship:

- **a short time** (e.g. 3-6 months) after completing an Apprenticeship. This would appear most appropriate to capture the immediate earnings benefits arising from the transition from 'trainee' status to 'skilled worker' status and the associated change in pay scales. This change in earnings tends to reflect part of the increased productivity of apprentices (research has suggested that employers tend to share the return from such productivity gains with the apprentice[26,27]), however effects on employment duration, job retention and progression would be virtually impossible to assess within such a short period;

- **1 to 2 years** after completion. This time interval would be most appropriate to detect the impact of productivity gains (reflected in earnings) of post-Apprenticeship work experience but again would add little to the assessment of other longer term employment outcomes such as progression and retention;

- **5 to 10 years** after completion. While ex-Apprenticeship earnings can be expected to increase directly in the short-term as the result of moving to skilled worker pay scales and increasing productivity in such jobs, in the longer-term earnings will increase as a result of occupational mobility and career progression. It could even be argued that Apprenticeship has a 'lifetime' effect, improving prospects across the whole of the apprentice's working life, much the same as has been postulated in the returns to education literature. Observing such long-term outcomes will require data to be collected in respect of a longer time period, in most instances several years. In order to achieve this period of follow-up, apprentices would need to be followed up to 29/30 years of age. An individual who takes up an Apprenticeship at age 18 years and takes two years to complete, for instance would be 30 years old 10 years after completion. In the case of employment retention, it is necessary for time to elapse in order to measure retention and

---

26 Dearden L., Reed, H. and Van Reenen, J. (2005) 'The Impact of Training on Productivity and Wages: Evidence from British Panel Data', Centre for Economic Performance Discussion Paper No 674, London School of Economics, http://cep.lse.ac.uk/pubs/download/dp0674.pdf.

27 Dearden L., Reed, H. and Van Reenen, J. (2000) 'Who Gains when Workers Train? Training and Corporate Productivity in a panel of British industries', Institute of Fiscal Studies, http://www.ifs.org.uk/wps/wp0004.pdf .

job stability while in the case of career progression it is likely that it takes time for apprentices to work their way up the job hierarchy;

- **long-term (e.g. 15+ years)**. This would capture more of a lifetime effect of Apprenticeship but such a lengthy timeframe is most difficult to use in practice due to data constraints, attrition, and the confounding effects of other life events such as further training/education being undertaken. Given that Modern Apprenticeships were introduced in 1994, even the first individuals to complete would only just now be 15 years post-completion thus data collection to observe the true long-term/lifetime effects of Modern Apprenticeships would be required over the next few years and into the future so that an evaluation could not be carried out at the moment. In any case, even if data were available, Apprenticeships have changed significantly since 1994 and may no longer be a good indication of the returns from present Apprenticeships.

The aim in the present study is to examine the options for evaluation of the long-term outcomes of Apprenticeship and so the preference would be to consider outcomes at least five years after completion of an Apprenticeship. But what is practical in terms of modelling future outcomes is critically dependent upon the data available. Where only cross-sectional data are available direct observation of change over time will not be possible and it may be necessary to create quasi-cohorts of ex-apprentices who completed their Apprenticeships at different times in the past (if such information is collected). Even where a longitudinal element exists it may be quite limited in terms of timescale, variables included and coverage of Apprenticeships. For instance, the Labour Force Survey (LFS) offers only a 12-month window on those who were apprentices. True longitudinal surveys tend to be small scale and while offering the prospect of tracking people over a longer period, may contain few apprentices.

3.2.5 Defining Apprenticeship and Completion

In the general economic model outlined above (equation (1)), the impact of having completed an Apprenticeship is captured by the variable *A*. In its most crude form this would be a binary variable taking the value 1 if the individual had completed their Apprenticeship and 0 otherwise. The estimated coefficient on such a variable would give an indication of the mean impact of an Apprenticeship on the selected outcome. Such a simple measure of Apprenticeship, however, takes no account of the issue of what constitutes a 'completion'. Apprenticeships consist of a number of elements including modules on key skills and the acquisition of an NVQ. While the formal definition of 'completion' may be clear (all elements have to be achieved) in practice matters may be less clear cut and there may be a variety of situations. Some apprentices will complete everything while others may complete only part of their Apprenticeship. Employers may, for instance, be indifferent to whether or not the apprentice obtains the relevant NVQ - unless such a qualification or a licence is required by law or regulation to undertake the work. Completion rates of Apprenticeships have been rising in recent years but still fall short of 100 per cent. It would be of interest to know whether completion or partial completion makes a difference to outcomes.

3.3 Econometric Considerations

3.3.1 Econometric and Statistical Approaches

A number of econometric/statistical approaches have been used in the evaluation literature, particularly in evaluations of training or active labour market programmes. Some relevant approaches were explored in the previous feasibility study.[28] Such evaluation in the econometrics literature has paid much attention to the issue of self-selection or endogeneity and associated biases. Self-selection effects treated through the use of fixed effects estimators, instrumental variables and non-parametric and semi-parametric models have been considered more recently.[29] Despite a number of advancements being made in terms of the treatment of selection effects, three general approaches are perhaps most common in the literature:

- regression methods (which include fixed effects estimators);
- methods based on propensity scores;
- matching.

While popular and well-accepted within the evaluation literature, these methods are limited and have their own shortcomings, particularly when any method is used in isolation.[30]

Regression methods involve estimation of the outcome variable (or the probability of a particular outcome) while controlling for a number of covariates (X) and Apprenticeship status (A) for both control and treatment groups. The inclusion of the X covariates is an attempt to control for all observable differences between participants and non-participants, but there always remains a risk of omitted variables, particularly those that are unobservable. Fixed effects estimation is one type of regression method that is well-suited to, and often used in, evaluations. Fixed effects or difference-in-differences estimation is discussed further in the next section.

Methods based on propensity scores present alternatives to regression estimators.[31]One such method directly replaces the covariates in a regression analysis with the estimated propensity of participation (in Apprenticeship) conditioned on individual characteristics. There are a number of techniques available to estimate the propensity scores. Another approach based on propensity scores is 'sub-classification'.[32] Sub-classification controls for differences in propensity scores in a more flexible manner than simply entering such scores directly into the

28 Payne et al, ibid
29 Imbens, G. and Wooldridge, J. (2009) 'Recent Developments in the Econometrics of Program Evaluation' Journal of Economic Literature, Vol. 47(1): 5-86.
30 ibid
31 Rosenbaum, P., and Rubin, D. (1983), 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', Biometrika, 70: 41-55.
32 Also referred to as 'substratification' or 'blocking'.

regression analysis in place of covariates. The data are partitioned into strata based on the estimated propensity scores. The data within each stratum is then analysed as if the propensity score within the stratum were constant. In effect, the data within each stratum can be considered as if it were obtained in a randomised trial as all individuals within the data in a particular block of data would be assumed to have the same probability of participation in a programme. The average treatment effect is estimated within each block of data and the overall average treatment effect can be calculated as the weighted average of the within-stratum average treatment effects.

Propensity scores are often used in the third general approach commonly found in the literature – matching. Matching methods seek to create a control group by identifying non-apprentices who have similar characteristics. Matching estimators have been used most often where the interest is in the average treatment effect for the treated and where there is a large pool of potential controls. Given a matched pair, the treatment effect within a pair is estimated as the difference in outcomes and the overall average as the average of the within-pair differences.

There are a number of approaches on which to base matching. Matching on the basis of propensity scores is one approach. Propensity score matching (PSM) creates a match based on a single index that measures the likelihood of an individual participating in an Apprenticeship (this would identify people who were very much like apprentices but who did not take up that training option). PSM methods have been shown to achieve consistent results and are more efficient than matching on all characteristics.[33] There are, however, a number of general conditions under which PSM should be ruled out. In the case of evaluating Apprenticeships, if there is an insufficient number of non-apprentices that match with apprentices in terms of their propensity scores (or probability of completing an Apprenticeship) or if there are small sample sizes, then PSM may not be an appropriate approach. If the design of the Apprenticeship programme results in almost universal participation so that there are no (or very few) non-apprentices with which to compare outcomes then PSM is not feasible. While at present this latter situation does not exist, it could be an issue facing Apprenticeships in the future.

For the evaluation of training programmes (including Apprenticeship), a combination of methods is more highly recommended than using any alone, with linear regression combined with either propensity score matching or other matching methods, being considered best practice.[34] Combining the above approaches is strongly advisable in order to achieve robust results and to avoid either over or underestimating the true effects of Apprenticeship on various outcomes.

3.3.2 Choosing a Counterfactual

In order to be certain that outcomes observed for apprentices are due to their Apprenticeship training (and not to other factors), it is necessary to know what their outcomes would have been had they not undertaken an Apprenticeship. In practice,

33 Rosenbaum and Rubin (1983), ibid
34 Imbens and Wooldridge (2009), ibid

this true counterfactual is unobservable so an alternative comparator group of non-apprentices is required. An RCT would give the closest alternative to the true counterfactual with individuals being randomly divided between participation in Apprenticeship and non-participation. The 'control group' of non-apprentices would, in an RCT, be very similar to the apprentices with the exception of their participation in the programme. As set out earlier, carrying out an RCT for the purposes of evaluating Apprenticeships is not a feasible approach. In the absence of a control group then it is necessary to identify a counterfactual or comparator group to which apprentices can be compared.

The composition of the counterfactual group will mainly depend on the questions being asked about the returns to Apprenticeship. For an assessment of the returns to Apprenticeship to be valuable, the outcomes must be viewed in relation to those achieved by others who took a different training route. The overall question is: are the returns to Apprenticeship better than, worse than or the same as the returns to something else? Deciding on this 'something else' is not straightforward. The alternative might be similar qualifications obtained through different means (i.e. other work-based training but not an Apprenticeship) or some lower level of qualification. Comparing the outcomes for apprentices to the former would indicate the relative return to an Apprenticeship while the latter approach would indicate the marginal return to training at 'the next level'. In the previous feasibility study consideration was given to what might constitute this counterfactual but no clear conclusion as to who might comprise a comparator group was reached.[35]

Within existing datasets a number of possible counterfactual groups may be considered:

- young people occupied in any VET routes other than Apprenticeships;
- young people in other VET routes and in other full-time education;
- young people in non-VET education;
- young people in work;
- NEETs;
- young people in unemployment;
- young people not in Apprenticeships (includes all of the above).

Making comparisons between young people who have undertaken Apprenticeships and each of the above groups will provide different evidence on the impact of Apprenticeships on the outcomes of interest. The choice between the above

35 Payne et al,ibid

comparators and others needs to be guided by both policy interests and the availability and quality of suitable data.

McIntosh[36] used data from the Labour Force Survey (2004/05) to estimate the net benefits and internal rates of return associated with Apprenticeship. His approach used different comparators for different levels of Apprenticeship. The returns to having completed a Level 3 Apprenticeship were compared to having a Level 2 qualification while the returns to Level 2 Apprenticeships were compared to those for Level 1 or 2 qualifications.[37] As McIntosh notes, the estimated magnitude of the estimated returns to Apprenticeships will depend upon the choice of comparator group. For instance, Dearden *et al*[38] found greater returns to Level 2 NVQs than were found by McIntosh. Dearden *et al* compared individuals with a Level 2 NVQ to individuals with no qualifications whatsoever. In McIntosh's approach the comparator group (Level 1 or 2) would be, on average, paid more than Dearden *et al*'s comparator (no qualifications) thus resulting in the lower impact of Apprenticeship found by McIntosh.

The Individualised Learner Record (ILR) (discussed further in Section 4.2) can provide a starting point for identifying a counterfactual group to compare to apprentices since it contains information about the population of FE learners. The ILR contains information about learners undertaking Apprenticeships and other programmes. Within the 2009/10 ILR, there are over 100,000 individuals undertaking an Apprenticeship or Advanced Apprenticeship. Within the Learner Responsive ILR returns, there are more than 2.7 million individuals undertaking some other programme.[39]

The largest non-Apprenticeship group in the ILR consists of those in the Entry to Employment (E2E) programme. E2E is primarily a programme for 16 to 18 year olds in England who are not participating in any form of post-16 education or training. While a sizeable proportion of young people in the ILR are in E2E programmes, it is unlikely that this group would be a suitable comparator. E2E participants are likely to differ significantly from apprentices with notable differences in terms of basic skills. The underlying characteristics of individuals are likely to differ significantly as well so that matching individuals in the two programmes would be difficult.

Given the high degree of heterogeneity across the various Apprenticeship frameworks, making comparisons *within* the Apprenticeship system itself would add much to the evaluation. Controlling for this heterogeneity is important when comparing the returns associated with Apprenticeships to any other qualifications/training or lack thereof however, looking explicitly at the varying returns stemming from this heterogeneity is also of great value. A number of different

---

36 McIntosh, ibid

37 The qualifications referred to here are the highest qualification held by an individual.

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

39 See Annex 1 for distribution of learners in non-Apprenticeship programmes.

comparisons could be made such as, the returns to Apprenticeships in the 'traditional' industries compared to newer entrants to the Apprenticeship system; comparison of young apprentices (16-18 years old) to older starters (24+ years old); comparison of Level 2 to Level 3 Apprenticeships; and so on.

Whatever the alternative pathway of individuals in the comparator (non-apprentice) group, this group needs to be comprised of individuals who are the same as the apprentices in all respects except their non-participation in Apprenticeship. In estimating the relationship set out in equation (1) the degree of comparability between apprentices (A=1) and non-apprentices (A=0) is important. Should these two groups of individuals be identical in all respects other than the fact that one has gone down the Apprenticeship route and the other has not then the estimated parameters in equation (1) would allow for calculation of the direct impact of the Apprenticeship on the outcome.

In practice it is unlikely that participants and non-participants are identical, rather it is more likely that those who have not completed Apprenticeship training differ in a number of ways from those who have. Differences between individuals are meant to be taken into account by the inclusion of a set of covariates that adjust outcomes for other factors such as individuals' demographic characteristics. This type of approach is always open to the criticism that some factor has been omitted because it is unobserved (it is not in the dataset or is not susceptible to measurement). If apprentices were to be systematically different in some unobserved way – for instance, better motivated, more ambitious or having greater ability – compared to non-apprentices then this would bias the results of any analysis (in the case of the differences mentioned, estimates would be upwardly biased resulting in the returns to Apprenticeship being overestimated). The issue for the estimation of the model is thus to reduce such unobserved differences as far as possible in order to minimise bias.

The 'selection bias' problem (that those who do not choose an Apprenticeship are different in some unobserved way from apprentices) can be partially corrected using various statistical techniques. Heckman *et al*[40] noted that there are several factors to consider in choosing a comparison group (i.e. counterfactual) in order to reduce the risk of selection bias through minimising differences between the two groups. The risk of selection bias is reduced if:

i. both groups (apprentices and non-apprentices) are placed in the same labour market;

ii. both groups have responded to the same questionnaire; and

iii. information is available on the recent labour status histories of both groups.

---

40 Heckman, J., H. Ichimura, J. Smith and P. Todd (1998) 'Characterizing selection bias using experimental data',Econometrica, 66(5):1017-1098.

The first two conditions are met through the use of existing data. Within the datasets assessed in this study, the individuals (apprentices and non-apprentices) have been drawn from the same population and with indicators of local authorities and regions (and other geographical indicators) the local labour market can be matched up or at least controlled for. Using existing datasets, it is certain that all respondents have responded to the same questionnaire (and have encountered the same interview techniques, for the most part). The only limitation here is that recent labour status histories are only available if they have been collected as part of the survey exercises already carried out. This is true of some, but not all, of the datasets considered in Section 4. The level of detail and length of these histories varies between datasets.

Identifying the comparator group within the data may use various forms of matching participants to non-participants. The matching method may simply be a crude 'rule of thumb', or may be a more sophisticated approach such as propensity score matching (where the comparison group would be a sub-group of people with a high propensity to undertake an Apprenticeship but who did not do so for some reason). Matching approaches do not guarantee that unobserved differences do not exist but such differences and associated bias may be minimised through matching.

Using 'fixed effects' estimation of the returns to Apprenticeship presents another way of accounting for differences between the two groups. This approach recognises that there may be unobserved differences between individuals who have completed Apprenticeship and those who have not, but assumes that such differences do not vary over time (e.g. apprentices are always more able or better motivated than their peers, and so on). The key to this approach is differencing as illustrated below:

$$Y_{i(t)} = \beta_0 + \beta_k X_{ik(t)} + \beta_p Z_{ip} + \varepsilon_{i(t)} \quad (2)$$

$$Y_{i(t+n)} = \beta_0 + \beta_1 A_{i(t+n)} + \beta_k X_{ik(t+n)} + \beta_p Z_{ip} + \varepsilon_{i(t+n)} \quad (3)$$

Equation (2) represents the situation at time *t* (before completing an Apprenticeship) where the outcome (say, earnings) is determined by a set of factors,*X*, which are observable and vary over time (such as age or highest qualification) and a set of factors,*Z*, some of which are observable and others unobserved, but all are time-invariant. Equation (3) represents the situation at time *t+n*, after an individual has completed an Apprenticeship (or not). In this latter period, the same factors affect outcomes but are modified depending upon whether an Apprenticeship has been completed or not (*A*).

If it is assumed that the unobserved differences amongst individuals do not change over time, then differencing equations (2) and (3) results in the following:

$$[Y_{i(t+n)} - Y_{i(t)}] = \beta_1 A_{i(t+n)} + \beta_k [X_{ik(t+n)} - X_{ik(t)}] + [\varepsilon_{i(t+n)} - \varepsilon_{i(t)}] \quad (4)$$

Differencing eliminates both the baseline value of Y (namely β_0) and the fixed effects variables (i.e. the impact of the time-invariant variables). Equation (4) suggests that the change in outcome between the two periods, t and $t+n$, is a function of whether or not an Apprenticeship has been completed (A), changes in other factors over time (X) and an error term (and since each error term is expected to have a mean of zero, the new error term would also have an expected value of zero). Thus, differencing on the assumption of fixed effects has removed potential biases from the model which can then be estimated using linear regression methods. This approach is most straightforward when variables are continuous but remains feasible even where variables are categorical. Using fixed effects estimation requires observations on the same individuals across different time periods i.e. longitudinal data. The degree to which existing data sources can meet this and other requirements is considered in more detail in Section 4.

In summary, it is difficult to empirically identify an obvious comparator group which would comprise the counterfactual in any statistical analysis. Arguably, young people upon completion of their compulsory education are faced with a number of distinct choices:

i. direct entry to the labour market;

ii. workplace based training;

iii. other vocational training;

iv. academic pathway through further education (possibly leading to higher education).

It is likely that choices made about which route to follow are, at least in part, related to the demographic, socio-economic, and educational characteristics of young people. Accordingly, it is difficult to compare like-with-like. An alternative is simply to recognise that there is a range of alternatives available to young people and to assess the relative returns to each alternative. People do have a degree of choice when selecting what to do once their compulsory education is complete. In any case, an evaluation of Apprenticeships which compares the returns to the multiple alternatives would at least indicate to people the relative costs and benefits of taking one course rather than another. A recommendation regarding the choice of counterfactual is set out in more detail in Section 6.

3.4 Conclusion

The economic model described in this section was constructed to identify the information required to evaluate the long term benefits obtained by people who have successfully completed an Apprenticeship. The model is useful in highlighting the information requirements of evaluation and draws attention to the need to set out definitions and measures of the variables to be included in estimating the returns to Apprenticeship. Key issues that require definition when setting to evaluate Apprenticeships include:

- what other factors affect returns (including characteristics of individuals, the Apprenticeship training, the employer and labour market conditions);
- the relevant outcomes;
- what constitutes an Apprenticeship (and completion);
- the timescale over which returns should be assessed; and
- the group(s) to which apprentices are to be compared.

A number of considerations for each of these key issues have been set out and discussed. The next section considers the extent to which the data specified in the over-arching framework and economic model are available.

4. An Assessment of Evaluation Data

4.1 Introduction

In Section 3, the ideal scenario for an effective evaluation of the outcomes of Apprenticeship over the long term was outlined in terms of data requirements. The feasibility of any evaluation of Apprenticeship outcomes will of course be constrained by the availability and quality of data on the outcomes of interest as well as on the characteristics of individuals and the characteristics of Apprenticeships undertaken (sector, format, level, etc.). Whilst England is rich in terms of the number of datasets sponsored by the government and carried out in a consistent manner, the applicability of such datasets to Apprenticeship evaluation is limited due to a variety of factors, including:

i. limited longitudinal data with relevant detailed information on Apprenticeship participation;

ii. limited longitudinal data on individuals' wages;

iii. insufficient sample sizes, particularly with respect to Apprenticeship details and earnings information; and

iv. the immaturity of current longitudinal surveys (i.e. respondents have not yet reached the age to become and apprentice).

The first two issues can be addressed through either:

- the introduction of a bespoke longitudinal survey with appropriate questions included to capture details on the participation of individuals in Apprenticeship including information about the individual apprentices (e.g. ages, pre-Apprenticeship activities, etc.) and their training (e.g. Framework, level, etc.) along with consistent measures of earnings, hours, and employment details; or

- revising existing longitudinal surveys which are currently in the field.

The third issue would require the sample of respondents to be increased with particular emphasis on boosting the number of apprentices included in the sample or again, for further survey(s) to be designed and administered in order to capture more of the details needed for a credible evaluation to be carried out.

The fourth issue is a much more positive issue than the others, especially in terms of evaluating Apprenticeships over the longer term. The Longitudinal Study of Young People in England (LSYPE), for example, currently contains some of the information

that is critical to evaluating Apprenticeship but is limited by the fact that, given their current ages, not all respondents have yet gone through their main educational and training experiences and entered into the labour market, thereby making it impossible to observe their labour market outcomes and the effects of Apprenticeship on these outcomes. Whilst this dataset could, in time, provide a near comprehensive set of variables needed for any evaluation, it is not yet able to do so. The main drawback of LSYPE is that due to sample attrition and the relatively small proportion of respondents participating in Apprenticeship, sample sizes are too small to provide a detailed analysis of Apprenticeships. This is a problem common to nearly all longitudinal datasets containing data on apprentices. This could, however, be mitigated by boosting sample sizes to include a greater numbers of apprentices. The original intention was for LSYPE to continue to survey people into their early 20s but a decision has been made to discontinue the survey.

The data available for the evaluation of Apprenticeships in England can be classified according to whether they are:

i. administrative databases;

ii. recurrent or longitudinal surveys with a general focus;

iii. longitudinal surveys with a training or education focus;

iv. *ad hoc* or occasional inquiries;

v. linked datasets;

vi. qualitative investigations.

The extent to which each of these respective data sources can provide information relevant to the evaluation of Apprenticeships is outlined below.

4.2 Administrative Databases

A number of administrative databases are potentially useful in examining the longer term impacts of Apprenticeship on employment and earnings. One of the overarching advantages of administrative records is their 'universal' coverage. Typically, administrative datasets contain the entire population on which they are based – e.g. the NPD has data on all children attending maintained schools in England, and HESA collects information on all higher education students. Administrative records can be used in a number of ways to estimate the impact of Apprenticeship on various individual level outcomes. They can be used on their own - though such an approach is likely to be based on a relatively narrow set of information – or in combination with other data sets (both administrative and survey data). Administrative data tends to contain information that is either absent from, or difficult to capture in, surveys. Surveys for example, often require individuals to answer retrospectively (e.g. about hours of work in an earlier period, course start/end dates, etc.) which can result in reporting errors and subsequent bias. Administrative

data should contain more accurate answers (not withstanding data processing errors). Administrative databases can also be used to improve the quality of longitudinal surveys in helping to maintain or re-establish contact with survey respondents.[41]

The main administrative datasets which are of interest for the purposes of evaluating Apprenticeships include the Individualised Learner Record (ILR), National Pupil Database (NPD), Higher Education records (HESA records), and DWP / HMRC databases.

The **Individualised Learner Record (ILR)** provides detailed information on training programmes and individual learners in the FE system. Data are recorded and provided to the Information Authority by training providers. Data in the ILR are not limited to individuals undertaking Apprenticeships; data for learners enrolled in other programmes are recorded, making comparisons between Apprenticeships and other FE routes possible. The programmes included in the ILR (2009/10 specification)[42]are:

- Advanced Apprenticeship
- Apprenticeship
- Higher Level Apprenticeship
- Entry to Employment (E2E)
- Progression Pathway to skilled work or an Apprenticeship
- Progression Pathway to first full level 2 (in the QCF)
- Progression Pathway to independent living or supported employment
- Progression Pathway to a Foundation (level 1) Diploma or GCSEs
- Diploma – Level 1 (Foundation)
- Diploma – Level 2 (Higher)
- Diploma – Level 3 (Progression)

---

41 For instance, DVLA driver records can be used to trace individuals who have been surveyed should they move or change their name without updating their survey contact details. Such use of administrative information would of course require consent from individuals.

42 A number of documents regarding the ILR are available at http://www.theia.org.uk/ilr/ilrdocuments/. The annual specification of the ILR is included amongst these.

- Diploma – Level 3 (Advanced)

There is a large set of variables recorded in the ILR with regard to the demographic profile of learners, details of programmes, and funding sources. Table 4.1 summarises some of the key relevant information contained in the ILR. The database includes: the characteristics of individual learners; the experiences/activities of learners before engaging in an Apprenticeship; details of the Apprenticeship programmes undertaken; and (limited) information on immediate post-Apprenticeship outcomes. The main advantage of the ILR is that it provides a greater level of detail about Apprenticeships (including content and structure) than the other datasets under consideration. For all apprentices, the Framework is given as is the level of study (Intermediate Apprenticeship, Advanced Apprenticeship and Higher Level Apprenticeship). Over 120 Frameworks were recorded for apprentices (at various levels) who completed in 2009/10. The Frameworks which accounted for the largest percentages of apprentices and Advanced apprentices in 2009/10 are indicated in Table 4.2. For Apprenticeship, 12 Frameworks accounted for more than 30 per cent of apprentices indicating that within many of the nearly 200 Frameworks available, participation numbers are low. In evaluating Apprenticeship many of the Frameworks will require aggregating in order to maintain sufficient sample sizes for statistically significant results to be produced.

There are many variables in the ILR critical to the evaluation of Apprenticeships. The main limitation arises with respect to information on outcomes. Information on the completion status of apprentices (and other learners) is included in the ILR but there is only one variable which provides information on what apprentices do after completion of their training. This variable, 'destination' is short term in nature referring only to the activities undertaken immediately after completion. Ideally, information about an individual's employer and job after training would be captured and individuals would be followed up over a longer period. A number of tables are presented in Annex 1 which summarise the distribution of individuals within the ILR with respect to several of the variables listed in Table 4.1.

Table 4.1: Main variables of interest in the ILR

Apprentice characteristics
- Sex
- Date of birth
- Disability/health problems
- Ethnicity
- Region
Pre-Apprenticeship Information
- Prior educational experience
- Employment status prior to start
Apprenticeship Details
- Framework
- Level
- Start and end dates (actual and predicted)
- Guided learning
- Provider and employer references
- Mode of delivery
Outcomes
- Completion status (and partial completion information)
- Destination (e.g. employment, other learning/education, etc.)

Table 4.2: Distribution of Apprentices and Advanced Apprentices by Most Common Frameworks

Framework	Apprenticeship	Advanced Apprenticeship
Customer service	12.1%	6.1%
Business administration	10.0%	9.3%
Hospitality and catering	7.5%	3.0%
Construction	7.0%	5.1%
Hairdressing	7.0%	4.7%
Engineering	4.6%	8.1%
Active leisure and learning	4.4%	3.1%
Vehicle maintenance and repair	3.4%	5.8%
Management	3.1%	3.9%
IT user	2.7%	0.5%
Accountancy	1.7%	3.7%
Sales and telesales	1.1%	0.4%
IT and telecoms professionals	0.9%	5.4%
All others	34.5%	40.9%

Source: ILR Employer Responsive returns for 2009/10

In the ILR data, learners often appear more than once as there are records for each learning aim and / or programme undertaken by each learner. Up to and including 2009/10, the ILR data has presented a number of challenges with respect to

uniquely identifying learners with identifiers being duplicated and many missing entries. In order to uniquely identify learners within the ILR there is a variable called the Unique Learner Number (ULN). The ULN should correspond to that held on the unique learner number register which is obtained from the Learner Registration Service.[43] This field became mandatory for learners receiving support from the Skills Funding Agency or YPLA from 2010/11. The 2010/11 specification sets out new validation rules for this variable which are intended to ensure that the field is completed correctly.

ILR: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Includes all apprentices
- Most detailed information about programme content (e.g. framework, hours of learning and work, delivery mode, funding arrangements, etc.)
- Detailed information on individuals' characteristics and background (e.g. age, sex, ethnicity, prior educational attainment, etc.)
- Possible to link to other datasets
- Information on other individuals in other FE programmes for potential comparisons

DISADVANTAGES

- Inputting errors
- Changes in content and coding over time
- Not originally designed for use as longitudinal database
- People drop out if they cease their course of study
- Issues around data security and confidentiality
- Limited information on outcomes – ends when individual leaves FE
- Problems with uniquely identifying individuals (especially across different providers)

POTENTIAL USE

- Potential to use as survey frame for collection of additional information from apprentices
- Link to other administrative datasets and/or survey data
- Use as standalone longitudinal database though results would be limited

43 ILR Specification 2010/11 available at http://www.theia.org.uk/NR/rdonlyres/274F58ED-3007-4189-9226-F6E9D29198F5/0/ILRSpecification201011_19Jan2011_v5.pdf

National Pupil Database (NPD) is a longitudinal database containing all children attending maintained schools in England. It links pupil/student characteristics to school and college learning aims and attainment. The database also contains individual pupil level data on attainment for some pupils in non-maintained and independent schools. The NPD contains information on the characteristics of pupils and schools including, age, gender, ethnicity, attendance and exclusions. These characteristics are matched to pupil-level data on attainment (Early Years Foundation Stage Profile, Key Stage assessments and external examinations which are collected from schools and Local Authorities). Other data on further education (including the ILR) and higher education (from HESA) have been matched to the NPD.

Using the NPD in carrying out an evaluation of Apprenticeship would require matching apprentices and the comparator group (however defined) to the database in order to obtain information on their educational attainment and performance whilst in compulsory schooling. The matching should be possible using the Unique Learner Number of respondents in the ILR and NPD, but linking to other data sets requires the use of probabilistic or fuzzy matching based on the characteristics of the individual such as their gender, postcode, etc.[44] Despite the technical difficulties in linking data, the availability of such information allows for greater control of individuals' background characteristics and supplies a proxy for 'ability' given that measures of educational attainment are available in the NPD. This would result in more of the difference in outcomes between apprentices and some counterfactual being attributed to the Apprenticeship rather than being confounded by otherwise non-observable factors such as attainment at school.

Whereas the NPD provides data about the pre-Apprenticeship educational attainment of apprentices, **HESA data** potentially provides information about the post-Apprenticeship experiences of apprentices if they have progressed into higher education (HE). At present relatively few people progress from Apprenticeships to HE. Data supplied by the LSC to the Apprenticeship Ambassadors Network reported that hardly any apprentices progressed into HE (0.1 per cent in 2006/7 and 0 per cent in 2007/8)[45] while a UVAC study from 2005 estimated that around 3 to 4 per cent of apprentices progressed into higher education.[46] Nevertheless, with the introduction of Level 4/5 Apprenticeships there is the potential for more people to progress from Apprenticeship into HE and potentially fall within the scope of HESA data collection. HESA collects data on all students in HE on courses for which the level of instruction is above that of Level 3 of the Qualifications and Curriculum Authority National Qualifications Framework (NQF), so includes sub-degree level courses such as HND. The database covers student enrolments at publicly funded

---

44 For example, see Hansen, K. and Vignoles, A. (2007)'The use of large scale data-sets in educational research', London: TLRP. Online at: http://www.bera.ac.uk/the-use-of-large-scale-data-sets-in-educational-research/

45 Reported in the Skills Commission (2009) Progression Through Apprenticeships http://www.policyconnect.org.uk/fckimages/Skills%20Commission%20-%20Progression%20through%20apprenticeships(1).pdf

46 UVAC (2005) An analysis of the progression of advanced apprentices to higher education in England. Bolton: University Vocational Awards Council.

higher education institutions (HEIs) in the UK and is available from the 1994/95 academic year onwards. Variables within the HESA student record include: age, gender, ethnicity, disability, A/AS-level points score and tariff points score, degree class, expected length of study programme, subject area, and source of tuition fees.

NPD: Advantages, Disadvantages and Potential Use

ADVANTAGES

- A longitudinal database for all children maintained schools in England which contains information about the characteristics of pupils and their educational attainment at Key Stages and external examinations
- Has been linked to other databases including the ILR and HESA data and to LSYPE survey data

DISADVANTAGES

- Access to the data is restricted by Department for Education
- Does not contain any information about Apprenticeship and later outcomes
- NPD only covers state maintained schools

POTENTIAL USE

- The particular strength of the NPD is that it provides information about the prior educational attainment of apprentices which would help to control for 'ability' and prior attainment when assessing outcomes later on.
- The prior educational attainment of apprentices is likely to be one of the factors which determines both the type of Apprenticeship they choose to work towards (by level and subject) and subsequent progression through the labour market

HESA also collects data on the destinations of leavers from HEIs. This data is collected through a survey of graduates administered approximately six months after they have left their HEI, and contains information on the activities of students after gaining a qualification from a HEI. The survey has been known as "**Destinations of Leavers from Higher Education**" (**DLHE**) since 2002/03. It was known as the "**First Destinations Supplement**" (**FDS**) from 1994/95 to 2001/02. In 2006, HESA carried out a follow up survey three years after graduation. This later survey data contains full details of individuals' higher education experience as well as wages and occupation. Variables within the DLHE include: activity (employed, unemployed, further study, etc.), location of employer, qualifications required for job, size of

employer, industry of employment, occupation, and subject and mode of further study. The data in the FDS/DLHE can be linked to student data records.

As with the NPD, the ability to link the DLHE and HESA student records to each other and to other datasets (e.g. LFS, LSYPE) might add much to the long-term evaluation of Apprenticeship in that it would provide information about any progression from Apprenticeship to HE. BIS has already linked HESA data to NPD records.

The **Department for Work and Pensions' (DWP) National Benefits Database** is a 'live' database of all benefits paid to individuals and includes personal data on these individuals. The **Work and Pensions Longitudinal Study (WPLS)** is a database which contains information about DWP administered benefits individuals have been in receipt of and HMRC data on their earnings and employment. The WPLS involves the linkage of all DWP benefit and programme participants to HMRC employment and earnings data (from P14 returns). This linking has been carried out by DWP since 1998. Such linked data have been used for a number of research purposes, particularly evaluations carried out for DWP.[47] These two databases present a rich set of variables regarding both background characteristics of individuals along with outcome measures including earnings and benefit receipts (type, amount and duration). Individuals can be tracked over time with such databases and long-term outcomes can be measured without serious concerns over attrition, recall errors (though other errors in entering data may occur), and associated biases which are encountered in carrying out longitudinal surveys.

Using the linked DWP/HMRC data alone for evaluating Apprenticeships would limit the range of outcomes which could be analysed as the WPLS only includes HMRC records for individuals who have been on DWP programmes or benefits. If the evaluation were to be limited to the individuals within this database, it would potentially misrepresent the overall effects of Apprenticeship training on outcomes as it would omit those who neither qualify nor require support from any benefits. Data on the receipt of benefits is typically difficult to capture through surveys (for a number of reasons) thus the WPLS presents an opportunity to analyse data that would otherwise be unavailable. Linking other data solely to HMRC records would avoid this problem.

Access to DWP databases is restricted due to data security and privacy concerns. Security issues are pertinent when considering using administrative records. Moreover, linking administrative datasets such as the NPD, DLHE/FDS, HRMC records, and DWP records, to survey based data requires consent to be granted by the respondents.

47 See http://research.dwp.gov.uk/asd/longitudinal_study/WPLS_Uses.pdf for a summary of uses of the WPLS.

DWP Databases: Advantages, Disadvantages and Potential Use

ADVANTAGES

- NBD contains all people – from 1999 onwards – who have been in receipt of State benefits including pensions
- WPLS links benefits data to HMRC data on an individual's employment record
- Provides a source of information on an individual's post-Apprenticeship employment record which can be analysed over time
- Consent to link responses to DWP records is found in a number of survey studies

DISADVANTAGES

- Limited to people who have been in receipt of benefits
- Use NINOs as an identifier of an individual whereas NPD and ILR use Unique Learner Number, so probabilistic or fuzzy data linking possibly required (the ILR contains NINOs but this may not be complete for all individuals)

POTENTIAL USE

- If the data can be successfully linked to other datasets which provide information about the prior educational attainment of apprentices (i.e. NPD) and details of the Apprenticeship (i.e. ILR) then there is a detailed record pre-, during, and post-Apprenticeship employment record for the individual

While the administrative databases considered here would add much detail to the evaluation of Apprenticeships, they do have their limitations. Such records are not typically designed with the express purpose of providing longitudinal data on individuals. As a result data files are often not structured in a readily accessible format for carrying out panel data analysis. The coding of records in such databases is also based on administrative requirements rather than research interests therefore coding may change over time within a database. The characteristics of administrative records do not prevent such databases from being used in research and evaluation, but depending on the dataset, the time and effort required to organise and interpret the records can be significant.

One major problem associated with using administrative data to supplement survey data is the possibility that administrative records may not be available for all survey respondents. This missing information for some respondents may result in biased estimation should the individuals with missing administrative data differ from those for whom administrative data are available. As an example, the NPD collects

information on students at maintained schools in England so that this administrative data is not available for many students in private schools. It is reasonable to assume that these two sets of students would differ in some way (e.g. socio-economic status, ethnicity, etc.) that would result in biases in analysis.

Administrative records may not be available for linking to survey data for a number of reasons. Individuals may not consent to having their data linked (though consent rates in existing datasets for which permission to link has been requested from respondents are reasonable).[48] An important consideration regarding consent is that individuals who consent to data linkage may differ significantly from those who do not give their consent. Evidence of consent bias has been found for the BHPS (Jenkins *et al*, 2004) and the Millennium Cohort Study (Tate *et al.*, 2005).

Where unique identifiers are not available within administrative data, linkage to survey data requires some other type of matching procedure, often probabilistic matching.[49] Probabilistic matching bases the match between administrative records and survey data on a number of pieces of information about individuals that are available in both datasets such as sex, date of birth, ethnicity and home postcode. The matching variables and the criteria for a valid match need to be chosen with care in order to minimise the number of incorrect matches between datasets and to minimise the number of cases for which there are no matches.[50]

48 A prerequisite for entry of individuals into the LSYPE was that they needed to provide consent for linkage of their survey records to the NPD. In the Millennium Cohort Study (MCS), 92 per cent of mothers gave consent for their child's data to be linked to the NPD and other records.

49 Other approaches to data matching and linking include 'exact matching', 'judgmental matching', and 'statistical matching' (See ONS (2004) for further information).

50 The use of probabilistic matching is explored in Jenkins et al (2004), 'Linking household survey and administrative record data: what should the matching variables be?' ISER Working Paper 2004-23. Colchester: University of Essex.

ADMINISTRATIVE DATA (GENERAL) :

Advantages, Disadvantages and Potential Use

ADVANTAGES

- Comprehensive coverage of population/sample – sample sizes not an issue
- Detailed information about particular issues (dependent on source)
- Systematic reporting
- Capture (accurately) some variables not easily obtained from surveys (e.g. benefit information)
- Most information is not self-reported thus avoiding some potential biases
- Especially useful for capturing background variables (e.g. school performance) and outcomes (e.g. wage levels)
- Possible to link to other datasets
- Cost-effective – data already collected so reduces time and expense of collecting data through surveys

DISADVANTAGES

- Inability to match or missing administrative data for some cases (e.g. due to missing identifiers or not meeting criteria for a 'good' match)
- Inputting errors, data processing errors
- Changes in coding and information recorded
- Not originally designed for use as longitudinal database – may need restructuring
- Apprenticeship not a main topic in any existing datasets (except ILR)
- Issues around data security, disclosure and confidentiality

POTENTIAL USE

- Use as standalone longitudinal data sources
- Linking to other administrative data and/or to longitudinal datasets (survey) to supplement and/or validate survey data; or to improve survey quality (e.g. improving contact with individuals, weighting survey data)

4.3 Recurrent and Longitudinal Survey Based Data Sources

4.3.1 Labour Force Survey (LFS)

The Labour Force Survey is a sample survey of households in Great Britain (GB) that is carried out on a quarterly basis. It provides information on the GB labour market and includes information on respondents' personal circumstances and labour market status.[51] The survey follows individuals for five consecutive quarters. It contains information on employment status, various indicators of income and wages, characteristics of respondents' jobs (e.g. hours, occupation, industry), qualifications (e.g. highest qualification, subject area), and activity history (e.g. redundancy, unemployment spells, status one year prior). The LFS also contains information on respondents' participation in on-the-job and off-the-job training as well as other types of learning/training/education as well as details about the format, delivery and duration of such activities. This information is important in indicating further outcomes beyond earnings and employment.

Within the LFS, Apprenticeships form one of the categories captured in a number of questions about qualifications and training. 'Trade Apprenticeships' form one possible response category to the item that records a person's highest qualification (i.e., highest qualification held and level of highest qualification held). There are also a number of survey questions addressed particularly to respondents who have completed or are currently undertaking an Apprenticeship, such as whether it is a Level 2 or a Level 3 Apprenticeship.[52]

In considering the usefulness of the LFS for the purposes of evaluating Apprenticeships, it is valuable to examine the distribution of responses to a number of variables that are of interest within the dataset.[53] The availability and consistency of variables across quarters and years will affect the feasibility of using the LFS in evaluation of Apprenticeships.

In the LFS data for January to March 2010, more than 3,500 individuals (5 per cent) indicated that their highest level of qualification was a Trade Apprenticeship. Respondents were also asked, 'Are you doing or have you completed a recognised Apprenticeship, including trade and advanced and foundation modern Apprenticeships?' Almost 7,500 respondents indicated that they had completed a recognised apprenticeship while less than 300 were currently undertaking a Modern Apprenticeship (MA).[54] Of those who had either completed or were currently undertaking a recognised Apprenticeship around 1,100 were part of the MA

51 ONS published ONS UK LFS results but fieldwork is carried out separately in GB and Northern Ireland.

52 This is achieved by asking whether the Apprenticeship is part of the Modern Apprentice initiative and whether the Apprenticeship is an 'Apprenticeship' (i.e. at Level 2), 'Advanced Apprenticeship' (i.e. at Level 3) or 'Apprenticeship plus Advanced Apprenticeship' (i.e. at Level 4).

53 Tables providing more complete details are included in Annex 1.

54 ibid

programme (with most at Level 2). More than 1,500 did not know whether or not their apprenticeship qualification was an MA.

These responses indicate one of the more major shortcomings of the LFS for use in evaluating Apprenticeships – inaccuracy in reporting Apprenticeships (as well as other qualifications). Not only are individuals subject to mistakenly recalling whether or not they had completed an Apprenticeship of any sort, but the definition of Modern Apprenticeship may not be strictly adhered to by all respondents. Inaccuracies can also occur when other individuals in a household serve as proxy respondents for others.

One of the more advantageous aspects of the LFS is its information on economic activity (employment, unemployment, inactivity) and details of jobs currently or previously held by individuals (e.g. occupation, industry, hours of work, earnings). As with any survey that asks people to give details of their activities or circumstances retrospectively, there is potential for reporting/recall errors in the LFS. That said, the breadth of outcomes covered by the LFS is particularly apposite for evaluating Apprenticeships.

The LFS has already been used to carry out cost-benefit analysis of Apprenticeship. McIntosh[55] uses data from 2004 and 2005 onwards in his analysis of the returns to Apprenticeship. The analysis estimates the returns to Level 2 and Level 3 Apprenticeships, compared to similar individuals with Level 1 or 2 qualifications and Level 2 qualifications, respectively. While acknowledging the shortcomings of obtaining information about Apprenticeship completion from the LFS, the study's findings are robust and in line with findings based on other datasets. It should be noted that McIntosh qualifies his findings with respect to the lifetime employment effects resulting from Apprenticeships (i.e. the extent to which Apprenticeship improves employability) by saying that the results reveal associations not causality. This derives from the problems of controlling for unobserved heterogeneity which inevitably arises from the use of cross-sectional datasets (such as the LFS), and the fact that people may already be in employment when they access vocational training such as Apprenticeship (thus suggesting a reverse causality).

55 Modern Apprenticeships (MA) is used to denote the formal Government funded Apprenticeship training programme.

LFS: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Large sample size of individuals who have completed an apprenticeship (but not of individuals currently undertaking them)
- Detailed information on labour market outcomes
- Information on other training after start working
- Source for comparison groups

DISADVANTAGES

- Reporting errors e.g. definition of 'apprenticeship' and 'Modern Apprenticeship' – people may say they have completed an Apprenticeship but there is no guarantee that they have done so and whether this is the Government funded programme or some other form of work-based learning
- Limited period following individual (five consecutive quarters)
- Information is not provided about subject of Apprenticeships
- Requires assumptions regarding subject area of training / education / Apprenticeship (i.e. if a person works in engineering and has completed an Apprenticeship then this will be an Engineering Apprenticeship). Other evidence suggests that people move between industries so there is no guarantee that this is a robust assumption
- Uses survey sample of addresses taken from the Postcode Address File rather than NINO or Unique Learner Number which limits the extent to which it can be linked across surveys or databases

POTENTIAL USE

- The approach adopted by McIntosh can be readily replicated over time
- There is the potential to link the LFS to other databases, such as the ILR, but it is likely that this will need to be undertaken via probabilistic or fuzzy linking given the lack of a common unique identifier

4.3.2 Longitudinal Datasets – General Considerations

Since Payne *et al*. assessed the feasibility of a long term evaluation of Apprenticeships, a number of rich longitudinal studies have been introduced and others have seen their respondents mature. In order to develop a comprehensive evaluation framework it is necessary to revisit a number of these datasets and review new studies to assess the suitability of these data for the evaluation of Apprenticeships.

Longitudinal studies present a number of advantages compared with cross-sectional data. Firstly, by following the same individuals over time, there is greater opportunity to isolate the effects of particular variables (such as Apprenticeship) on various outcomes and to investigate causal relationships. It is also possible to control for unobserved fixed effects which may influence outcomes. A longitudinal survey which follows individuals for a relatively long period of time also gives a truer lifetime perspective and permits observation of other events and actions over the life course which could influence employment outcomes. As a general point longitudinal data sets are increasingly capturing information about job and life satisfaction, and overall health and well-being (hence going beyond typical economic returns).

The main drawbacks of longitudinal surveys include attrition (i.e. loss of respondents over the course of the study), costs (in terms of time, labour and finance),and the timeliness of findings (i.e. in some instances there will be a long wait before data are provided about the employment experiences of apprentices). Some degree of attrition is inevitable in longitudinal studies. The severity of attrition depends on a number of factors including: the duration of the study (as time goes on the number of drop-outs will increase); the age of the cohort being studied (mortality and morbidity rates increase with age and will thus increase the probability of people leaving the study; alternatively younger cohorts may more mobile and more difficult to follow up); and the underlying characteristics of individuals being surveyed. The costs of longitudinal surveys are substantial (with the exception of small scale studies which are of little use for the task at hand). Designing and administering questionnaires, tracking people over time (which may require various approaches to maintaining or re-establishing contact from wave to wave) and maintaining databases have significant cost implications and require substantial time and labour inputs. It also needs to be borne in mind that a new longitudinal survey containing a single age cohort will take many years to report any employment outcomes for apprentices. If a longitudinal survey first collects data from respondents at, say, age 13 years, then it is likely to take ten to fifteen years before data are available about the early labour market experiences of these respondents.

Other issues which impinge on the usefulness of longitudinal data include the consistency of variables over time. Often, for various reasons including changes in government policy and legislative changes, the definitions for particular items captured in surveys may need to be changed in subsequent data collection sweeps. The development of new metrics may also result in survey questions and coding of responses being revised over the course of a longitudinal study. The content of questionnaires may also change in terms of questions being omitted or introduced between waves. This latter point is not necessarily negative provided the changes in

the questionnaire result in the content of the dataset being better suited for research objectives.

Most longitudinal datasets cover the range of topics relevant to evaluation of Apprenticeships. There are different designs available:

- **single cohort surveys** which follow a group of people from a given age and survey them periodically (e.g. NCDS);
- **mixed cohort surveys** which follow people of different ages from a given date and then survey them periodically (e.g. Understanding Society);
- **multiple cohort approaches** which repeatedly commence the same survey at different dates with a new sample so that there are, for example, surveys of people at, say, age 16, starting in successive years (e.g. YCS).

The benefits of tracing the same individuals over time are numerous but a single cohort design imposes some limitations, particularly in terms of the potential to examine the effects of changes in policy, funding approaches, legislation and other factors affecting the design and delivery of Apprenticeships. The cohort that is followed in a longitudinal survey will typically see all members who undertake Apprenticeship training (or other activities) do so within the same policy and administrative context. A multiple cohort design, where various age groups are followed, would add value to a longitudinal study and present more opportunities to gauge the effectiveness of particular policies, especially if they change over time. Incorporating multiple cohorts however, would significantly increase the costs of such studies.

In sections 4.3.3 to 4.3.8, a number of longitudinal datasets which are currently available and have the potential to be used in evaluation of Apprenticeships are considered.

LONGITUDINAL DATA (GENERAL):

Advantages, Disadvantages and Potential Use

ADVANTAGES

- Allows for lifetime perspective
- Control for unobserved fixed effects
- Greater ability to account for selection effects
- Observe other lifetime events/circumstances
- If multiple age cohorts are involved, policy changes over time can be observed

DISADVANTAGES

- Costly to design, implement and follow-up
- Attrition rates can be high and potentially bias any results
- Self-reported measures collected in surveys have inherent biases
- Introduction of new variables, discarding of previous variables, recoding can result in it being difficult to observe change over time
- If only one age cohort is involved this can make it difficult to observe policy changes over time
- Often small numbers of apprentices

POTENTIAL USE

- Design bespoke longitudinal study for evaluation of Apprenticeship capturing all required information
- Use existing longitudinal datasets with amendments to questionnaires to capture more information about Apprenticeships
- Boost sample of apprentices within existing longitudinal surveys so that more detailed analysis of apprentices is possible (e.g. by level and subject)
- Linking between longitudinal datasets and other data sources

4.3.3 British Household Panel Survey (BHPS) and Understanding Society

The **British Household Panel Survey (BHPS)** began in 1991 and currently contains data up to 2009. The BHPS data consists of individual-level information regarding the same individuals throughout all waves. The survey is household based with all adult (16+ years) household members being interviewed. The first wave included around 5,500 households (10,300 individuals) from Great Britain. Over time additional samples have been added in order to increase coverage to the whole of the UK and to overcome problems due to sample attrition.

The BHPS contains information on individuals' education and training, health and use of health services, labour market behaviour, and income from various sources. The BHPS also includes information on Apprenticeship, mainly in the form of response to questions about education or training. The main shortcoming of the panel in terms of being used for evaluation of Apprenticeships is its relatively small sample size, particularly when drawing on specific groups such as those who have completed an Apprenticeship. In Wave 18 (2009) only 12 individuals indicated that they had obtained an apprenticeship (trade or Modern Apprenticeship).

BHPS: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Rich information on individuals over relatively long period (18 waves)
- Covers range of age groups (not limited to one cohort)
- Labour market outcomes observed

DISADVANTAGES

- Relatively small sample and particularly small number of apprentices
- As with all longitudinal surveys, attrition and any associated bias is a potential problem

POTENTIAL USE

- Not recommended for use (on its own) in the evaluation of Apprenticeships
- Sample has been subsumed within Understanding Society so potential use set out with respect to that study

The UK household longitudinal study, **Understanding Society**, shares many of the design features of the BHPS. The BHPS sample has now been incorporated into the Understanding Society survey rather than running the two household panel surveys alongside each other. Understanding Society represents an improvement on the BHPS in that it covers a much larger sample (c. 40,000 households, c. 80,000 individuals) and has a number of features such as data linkage (with permission to link to health, education, HMRC and DWP records being sought from respondents) and an ethnic minority booster sample built into its design.[56]The first (half) wave of data from this new household panel survey was released in 2010.

In the first wave of Understanding Society, Apprenticeship is encountered in questions regarding respondents' vocational qualifications as well as the educational qualifications of their parents. The same variables are to be included in Wave 2 along with questions on qualifications gained since the first wave. The response categories for these questions include Modern Apprenticeship / Trade Apprenticeship. As with the BHPS, there are a number of variables in Understanding Society covering individuals' work, education, training and income. The variables of particular interest are summarised in Table 4.3.

Table 4.3: Main variables of interest in Understanding Society (waves 1 and 2)

Apprentice characteristics and background
- Sex
- Date of birth
- Disability/health problems
- Ethnicity
- Region
- Household composition
- Parents' vocational qualifications
Apprenticeship Details
- Vocational qualifications
Outcomes
- Employment status
- Income, wages
- Housing tenure
- Well-being

A significant limitation of this dataset at the moment is that given that around 2,200 individuals in wave 1 are aged 16 to 18 years the number of 16 to 18 year old apprentices in the sample is likely to be relatively small. Assuming that the sample is representative of the population in general and that around 6 per cent of 16 to 18 year olds were in work-based learning in 2009[57] it can be estimated that around 150 of the 16 to 18 year olds in the Understanding Society sample in Wave 1 are

56 See http://www.understandingsociety.org.uk/ for further information on Understanding Society.

57 See SFR June 2010, Table 1.

engaged in some form of WBL. Given that not all of these will be undertaking Apprenticeships, the number of apprentices will be even lower. In subsequent waves the number of apprentices will undoubtedly increase as individuals in this age group enter Apprenticeships. As members of households included in the sample reach age 16 they enter the survey which will increase the number of 16 to 18 year olds and in turn the number of apprentices. The resulting total number of apprentices amongst this age group however will still be insubstantial thus hindering robust analysis. Similarly, for 19 to 23 year olds, whose participation rate in Apprenticeships is lower than 16 to 18 year olds, there is an even smaller sample of apprentices contained in Wave 1. Given around 3,000 19 to 23 year olds in the first wave, an indicative estimate of the number of these undertaking an Apprenticeship would be around 100. There is potential to remedy this by boosting the sample of apprentices.

While the information specifically on Apprenticeship in the first two waves of Understanding Society is relatively limited, it is the immaturity of this survey that is most important in considering it for use in evaluating Apprenticeships. It may be possible to incorporate further details in the survey's questionnaires so that missing information regarding Apprenticeship, such as level and Framework, may be added. Should this be possible, Understanding Society could form a near comprehensive data set for use in the long term evaluation of Apprenticeships, provided that sample sizes are sufficient.

Understanding Society: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Much larger sample than BHPS
- New survey – content can be influenced
- Data linkage incorporated
- Ethnic minority booster sample included
- Multiple cohorts, renewing cohort - As young people in household reach age 16 they are added to survey thus providing a continuous age 16-18 cohort which will be followed throughout the panel – will allow for observation over time across individuals and consider programme/institutional setting affecting the cohorts at different times

DISADVANTAGES

- Does not currently hold detailed information on Apprenticeship e.g. level and Framework.
- Self-reported measures of income, etc.
- Small 16-18 cohort in first wave (approx. 2,200, and fewer than 150 apprentices)

POTENTIAL USE

- Incorporate additional survey items to ascertain details of Apprenticeship
- Linking the survey data to administrative records in order to fulfil further information requirements for an evaluation of Apprenticeships (e.g. NPD, HESA, HRMC, DWP records)
- Numerous cohorts can be followed and the youngest age group is renewed over time as additional individuals enter the sample when they reach age 16, so potentially provides a high-level moving picture, reflecting policy changes over time.

4.3.4 National Child Development Study (NCDS)

The NCDS is an ongoing longitudinal study of all children born in Britain (England, Scotland and Wales) between the 3rd and the 9th of March, 1958, and is thus a single cohort dataset. The NCDS provides data on a broad spectrum of variables including physical well-being, health history, social attitudes and awareness, family structure and background, earnings and labour market status, living conditions, parental educational aspirations and child abilities.

Starting with the Perinatal Mortality Survey (PMS) conducted in 1958, there have been an additional eight sweeps surveying this birth cohort: in 1965 (NCDS1, aged 7); 1969 (NCDS2, aged 11); 1974 (NCDS3, aged 16); 1981 (NCDS4, aged 23); 1991 (NCDS5, aged 33); 1999/2000 (NCDS6, aged 41/42), 2004 (NCDS7, aged 46) and 2008 (NCDS8, age 50). In the first three waves, information was collected through surveying parents, head and class teachers, school health visitors and, in the later waves, the young people themselves were interviewed and also sat aptitude tests. The timing of each wave of the survey along with sample sizes and respondents' ages are summarised in Table 4.4.

Table 4.4: NCDS Survey Structure and Sample Sizes

Sweep	Year	Age	Sample size (cross-sectional - including boost)	Sample size (longitudinal)
PMS	1958	0	17,416	17,416
NCDS1	1965	7	15,425	15,051
NCDS2	1969	11	15,337	14,757
NCDS3	1974	16	14,647	13,917
NCDS4	1981	23	12,537	12,044
NCDS5	1991	33	11,407	10,986
NCDS6	1999-2000	42	11,419	10,979
NCDS7	2004	46	9,531	9,175
NCDS8	2008-2009	50	9,790	9,408

The NCDS is not, however, an ideal data source for evaluating Apprenticeships. Perhaps the major limitations are attrition between the first and final waves, and the fact that the cohort will have completed their initial vocational education and training before the introduction of Modern Apprenticeships in 1994.Attempts were made in the January 2003 release of the series to help correct for some attrition anomalies. There is a tendency for this attrition to be concentrated amongst those individuals displaying lower ability and educational qualifications.[58] The argument put forward in Dearden *et al* (2002) suggest that this attrition may be accepted as exogenous and that it does not necessarily bias results to the extent that it depends on observable

58 Dearden.L, Machin, S. and Reed, H. (1997) 'Intergenerational Mobility in Britain',The Economic Journal, 107: 47-66.

characteristics only which are related to ability and background and have been controlled in the analysis.[59]

Respondents in NCDS were aged 36 years when Modern Apprenticeships were first introduced in 1994, by which time they would have completed their initial vocational education and training. Though apprenticeships were in existence before the introduction of Modern Apprenticeships, it is the latter which marks the beginning of a Government funded Apprenticeship training programme. Hence NCDS is of limited value with respect to this inquiry but it has much to offer in its approach to capturing information about Apprenticeships.

The information relevant to Apprenticeship provided in the NCDS covers all apprenticeships obtained by individuals.[60]The specific information recorded at each wave varies. NCDS4 (age 23) recorded the date of completion and qualifications obtained through each apprenticeship that an individual reported. In NCDS6 (age 41/42), respondents were asked questions about: the main subject of each Apprenticeship taken in the past; place of study; whether full-time or part-time Apprenticeship; the year in which they passed the apprenticeship or the year in which the qualification was awarded; and whether currently doing a Modern Apprenticeship. NCDS7 (age 46) and NCDS8 (age 50) contain some information of the apprenticeship qualifications obtained, but most of the variables about apprenticeship have small sample sizes which are likely to result in estimates related to apprenticeships not being statistically significant.

Regarding employment, the NCDS contains extensive information about work history and work experience. Information about current job, employment status, regular income in the current job, gross and net pay of the job, and the period covered for gross or net pay are available in each wave. Duration of unemployment is also available in NCDS4 and NCDS5 (ages 23 and 33 years, respectively). The sample sizes for the employment variables are reasonably large for statistical analysis.

The variables related to Apprenticeship, and employment, jobs and income, which are of interest for the purpose of evaluating Apprenticeships are shown in Table 4.5. The numbers of valid responses for the Apprenticeship variables are noticeably smaller than for the employment, jobs and income variables. The NCDS, with the degree of attrition of respondents typical of such a long time period, exhibits the expected problem of small sample sizes (and associated lack of variation for some variables), particularly for indicators of Apprenticeship. Where MA qualifications are distinguished from other apprenticeships in NCDS, sample sizes are negligible.

Table 4.5: NCDS variables related to Apprenticeship and outcomes

Apprenticeship
- Whether completed an Apprenticeship

59 Dearden, L., Ferri, J. and Meghir, C. (2002) 'The Effect of School Quality on Educational Attainment and Wages.'The Review of Economics and Statistics, 84(1); 1-20.

60 Up to nine separate apprenticeships per individual are observed in the data.

- Number of Apprenticeships obtained by individual - Start/end month and year - Any Modern Apprenticeships - Whether currently undertaking Apprenticeship - Main subject - Location of study (e.g. college, etc.) - Studied full-time or part-time
Outcomes
- Current employment and economic status - Net and gross pay - Hours (regular and in reference period) - Pay period - Duration of unemployment

Note: the specific information and variable definitions varies according to wave but overall the information referred to in the table is available in waves 4 to 8.

Given the current age of NCDS cohort members, it is possible to carry out a rudimentary assessment of differences in outcomes between apprentices and non-apprentices. Across waves 4 to 8, apprentices are found to have greater employment rates than non-apprentices. In NCDS4, 87 per cent of apprentices and 71 per cent of non-apprentices were employed. The percentage of apprentices who are unemployed is also found to be lower than that for non-apprentices in waves 4, 6 and 7 of the NCDS. It needs to be borne in mind that at NCDS4, when respondents were aged 23, they would most likely have completed any apprenticeship training they had been undertaking. At early waves any difference in employment rates between apprentices and non-apprentices might have been explained by the fact that, by definition, apprentices need to be employed in order to receive their training. By Wave 4, both apprentices and non-apprentices (even those going on to HE) will have completed their initial vocational education and training in most instances.

Within the NCDS, median weekly income is found to be higher for apprentices than non-apprentices at ages 23, 42 and 46 (waves 4, 6 and 7). With respect to unemployment rates, if attention is focused on the longest single period a person records as being unemployed, the average was 2 months for apprentices and 4.5 months for non-apprentices. Amongst all individuals, the maximum longest spell was considerably higher for non-apprentices (87 months) than for apprentices (51 months). Overall, the data available in NCDS indicates that apprentices fare better than non-apprentices on a number of outcome measures.

NCDS: Advantages, Disadvantages and Potential Use
ADVANTAGES • Details of any apprenticeships completed including start/end dates, subject area • Detailed personal information • Income and work histories **DISADVANTAGES** • Due to age of cohort, does not capture people who have undertaken a Modern Apprenticeship • Attrition rates are higher in the NCDS than in a number of other large scale longitudinal surveys under consideration • Limited to single age cohort **POTENTIAL USE** • The data set has limited use with respect to the post Modern Apprenticeship period, but its design and questioning about apprenticeships provides valuable information about how to effectively design a study which provides information relevant to the evaluation of Apprenticeships

4.3.5 British Cohort Study (BCS70)

The British Cohort Study began in 1970, collecting information about babies born in the UK in a given week. The first wave, which was called the British Births Survey, was collaboratively collected by the National Birthday Trust Fund and the Royal College of Obstetricians and Gynaecologists recording the social and biological characteristics of the mother in relation to neonatal morbidity. It is also provides data comparable to the National Child Development Study (NCDS). Eight full waves of the survey have been carried out, including the first birth information data. Thus far the survey has followed the cohort since birth through to age 38 in 2008-09. The data collection exercises have been carried out in order to monitor the health, education, social and economic circumstances of the cohort members. In addition to the full cohort studies, four sub-sample surveys have been carried out.

The first sweep was carried out in 1975 when the cohort members were aged 5 years. The second sweep took place in 1980 at age 10 of the BCS70 cohort. These two sweeps were carried out by the Department of Child Health at Bristol University and were known as Child Health and Education Study (CHES). The 16-year follow

up was carried out by the International Centre for Child Studies in 1986 and named Youthscan.[61] Successive sweeps were conducted in 1996 (aged 26), 1999-2000 (aged 30), 2004-2005 (aged 34) and 2008-2009 (aged 38). The BCS70 cohort study only includes respondents from Great Britain in the follow-up sweeps apart from the birth survey which initially included people from Northern Ireland.

The BCS variables germane to Apprenticeship and employment, jobs and income are listed in Table 4.6. From the fourth sweep (1996, aged 26 years), information about cohort members' employment began to be included in the survey. Information collected includes individuals' current employment status, whether they were ever unemployed and their current pay. Information about Apprenticeships (completed or in progress) were not included in the questionnaire until sweep five (2000, aged 30). Information has been collected on the number of Apprenticeships obtained and the subject, year of qualification and delivery location of any Apprenticeships individuals have reported completing. Such information goes some way in addressing the heterogeneity in Apprenticeships. Sample sizes for these variables are potentially large enough to perform statistical or econometric analysis. In the 1999/2000 (age 30) sweep, 509 individuals reported that they had completed an Apprenticeship, with 495 of these individuals indicating details about their Apprenticeship qualification. Subsequent waves repeat the questions pertaining to Apprenticeships (and other vocational qualifications) with regard to any qualifications obtained since their last contact with the study. In the 1999/2000 wave, 283 individuals indicated that they had completed a Modern Apprenticeship.

61 Information about the BCS70 is obtained from the website of Economic and Social Data Service.

Table 4.6: BCS variables related Apprenticeship and outcomes

Apprenticeship
- Whether completed a recognised trade Apprenticeship
- Number of Apprenticeships
- Year qualification awarded
- Any Modern Apprenticeships
- Whether currently undertaking MA
- Main subject
- Location of study (e.g. college, etc.)
- Studied full-time or part-time
- Source of fees for trade Apprenticeship
Outcomes
- Net and gross pay
- Period pay covers
- Hourly and weekly pay
- Occupation in current job
- Overtime in current job
- Status (permanent, temporary, etc.) in current job
- Ever unemployed
- Number of periods unemployed
- Longest unemployed period
- Never unemployed

In addition to the core waves of the BCS data, there is also a dataset which records the employment histories for individuals between 1996 and 2004. This data combines the data which was collected in the individual waves and tabulates the histories for individuals in chronological order. This employment histories information would be useful in looking at the impact of Apprenticeships on labour market outcomes such as the duration of employment and unemployment states, changes between various labour market statuses, and occupational mobility. The employment histories are recorded from age 26 to 34. Within this age range, individuals have most likely completed their formal education and training and would be into their working lives. This range is well-suited to evaluation of Apprenticeship outcomes.

Within the 1999/2000 BCS dataset (age 30), 90.5 per cent of individuals who had obtained a recognised trade apprenticeship were employed, 1.1 per cent were unemployed and 8.5 per cent were not in employment, education or training. Within the cohort overall at that time 81.4 per cent were employed, 2.5 per cent unemployed and 16 per cent were not in employment, education or training. The median weekly income for former apprentices in the 1999/2000 wave was £384 compared to £320 for the entire sample.

Evaluating the impact of having an Apprenticeship on outcomes would require examining the outcomes before and after the period of the Apprenticeship and comparing to individuals who have not undertaken an apprenticeship, based on a 'Differences-in-Differences' approach. The applicability to Apprenticeships is limited

in that the cohort was aged 24 at the start of the MA programme and therefore would be too old to have gone through the programme (for the most part). The BCS does, however, provide good coverage of individuals over the life-course and captures the relevant ages for observing various factors before, during and after completion of an Apprenticeship or some suitable alternative. In the 1999/2000 wave of the BCS, the majority of respondents who indicated that they had completed an Apprenticeship had done so between 1988 and 1991 when they were between 18 and 21 years old.[62]

BCS70: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Decent sample sizes, even when taking apprentices only (though diminishes with level of detail)
- Information on dates and subjects of Apprenticeships
- Information on outcomes includes employment status, unemployment duration, wages, hours, occupation

DISADVANTAGES

- Single cohort study
- Diminishing sample numbers when looking at greater level of detail regarding Apprenticeships
- Cohort will be slightly too old to have participated in Modern Apprenticeships (i.e. 24 years in 1994)
- Little on job satisfaction or other 'softer' outcomes
- Attrition rates high

POTENTIAL USE

- Possible to carry out some form of analysis on returns for this cohort given that observations span from childhood to late 30s, using difference-in-differences approach

62 The breakdown of Apprenticeship completion by age and year of completion in the 1999/2000 wave is set out in detail in Annex 1.

4.3.6 Avon Longitudinal Study of Parents and Children (ALSPAC)

The **Avon Longitudinal Study of Parents and Children (ALSPAC)** is a large scale longitudinal study of children born in Avon in the early 1990s. To be included in the sample covered by the study, the children's mother had to be resident in Avon and the expected date of delivery of the children had to fall between 1 April 1991 and 31 December 1992. The study covered more than 14,000 children at the beginning.

ALSPAC has gathered a wide range of information on the children through questionnaires of the parents, physical examinations of the children themselves, health records, biological sample analysis and tests on the home environment. A number of surveys have been conducted in schools in order to collect information on schools, classes, and cohort members. Further information was collected from the previously-named Department for Children, Schools and Families (DCSF) (now Department for Education), with regards to SATS and GCSE results. In addition, ALSPAC has obtained linkages to the following datasets: National Pupil Database, Pupil Level Annual School Census data, ONS Death and Cancer Registry and NHS Strategic Tracing Service address data. Other avenues for linking are also being explored including linking to GP data and to data from DWP and Home Office.

The ALSPAC study members are currently 18/19 years of age. The Age 18 Questionnaire is currently in the field with other questionnaires for parents being started in early 2011. The children will be surveyed again in 2013/14 at age 21/22. The coverage of the ALSPAC in terms of the subjects' ages is well-suited for use in evaluating the outcomes associated with participation in and completion of Apprenticeship training. The current age of the cohort would allow for subjects to have already begun and perhaps completed an Apprenticeship. They are also still young enough that a number of respondents may undertake an Apprenticeship within the next few years. Funding for ALSPAC has not yet been confirmed beyond the Age 21/22 Questionnaire. Should further funding be secured then the study will cover the age range required to observe labour market outcomes over the short to medium term. If ALSPAC did not continue beyond age 21/22, the measurement of outcomes for evaluating apprenticeships would be inadequate.

While ALSPAC contains a wealth of information regarding individuals' backgrounds, education, family and a host of other factors, there is at present limited detail on Apprenticeship. The only question that directly refers to Apprenticeship asks young people about their plans after Year 11 with Modern Apprenticeships as a possible response. There is no level of detail about what subject area in which they would wish to take an Apprenticeship. The major concern with using ALSPAC in a long term evaluation of Apprenticeship, however, is with respect to its limited regional scope. The study only contains individuals born in Avon. The study has published information regarding the representativeness of the sample with respect to the population of Great Britain as a whole. Overall, the socio-demographic characteristics of mothers and children in the area and in ALSPAC are not dissimilar to the rest of the country. Mothers of children under one year of age in Avon in 1991 were slightly more likely to live in owner-occupied accommodation and to have a car in the household, and were less likely to have more than one person per room and

to be non-white, than were those in the rest of Great Britain. The weights and birth lengths of children in ALSPAC were very much in line with national averages.

The major doubt regarding the applicability of ALSPAC data to the evaluation of Apprenticeships is the structure of the Avon labour market. The concentration of economic activity by sector in Avon differs from the national distribution. Such a difference would affect the types of Apprenticeships available in Avon relative to the rest of the country.

ALSPAC: Advantages, Disadvantages and Potential Use
ADVANTAGES • Rich dataset with detailed information on young person and their family • Linked to administrative records **DISADVANTAGES** • Limited to young people from Avon (the labour market is quite distinct in its industrial structure – e.g. presence of aerospace manufacturers many of which have a strong demand for apprentices). • Single age cohort • Limited information on Apprenticeship **POTENTIAL USE** • Not recommended although could be used for validation/verification of other findings

4.4 Longitudinal Surveys with a Training or Education Focus

4.4.1 Youth Cohort Study (YCS)

Payne and her colleagues highlighted the **Youth Cohort Study (YCS)** as being useful in carrying out evaluation of Modern Apprenticeships.[63] The YCS began in 1985 surveying individuals who were eligible to leave school in 1983/84. The survey has now gathered data from 13 cohorts with data available for three waves the latest cohort whose members were reached school-leaving age in 2005/06. The study

63 Payne et al,ibid

follows each cohort between the ages of 16 and 20 (for most cohorts). The YCS contains perhaps the most comprehensive list of variables which are relevant to evaluating the outcomes of Apprenticeship. This dataset includes indicators of individual's labour market status over a number of years on a monthly basis and includes indicators of whether or not an individual is an apprentice. It also includes items related to pay and hours. Demographic variables are also included.

The timing of fieldwork and the ages of respondents by cohort for the YCS are summarised in Table 4.7. The design of the study, following different cohorts over time between the ages of 16 and 19/20 (with the exception of the third cohort which covered age 23) is a limitation since it will not: (a) capture people who finish their Apprenticeships after the age of 19/20; and (b) will not include people at all who commence an Apprenticeship after age 20 years.. The design of the survey therefore severely constrains the ability to examine various employment-related outcomes such as wages or unemployment status. As Section 3 indicated, the most promising period over which to evaluate the effectiveness of Apprenticeships is around five to ten years after completion; in other words, the timeframe within which the direct effects of an Apprenticeship on an individual's progression through the labour market can be gauged. YCS stops short of providing this information. Sample attrition, especially given relatively small numbers in Apprenticeships, is also a potential problem with the YCS.

Table 4.8 summarises the information contained within the questionnaires for cohort 12 that are related to Apprenticeships. It should be noted that one of the strengths of YCS – and LSYPE – is that data are collected about the reasons why an individual decided to pursue an Apprenticeship. As noted in Section 2, there is a need to understand why returns to an Apprenticeship might differ, and these questions could shed light on this. YCS collects information about intentions at age 16 and allows a comparison with outcomes at later ages.

The most comprehensive data on Apprenticeships is included from Cohort 13 of the YCS. The data on cohort 13 (eligible to leave school in 2005/06) contains much of the same information that is contained in the Longitudinal Study of Young People in England as the methodology of the YCS has been changed to allow for linking with the LSYPE. To avoid duplication, the main variables of interest (which are common to both YCS13 and LSYPE) are presented in the discussion of LSYPE below. Samples for cohort 13 and LSYPE were taken from the same academic cohort and in order to facilitate analysis of the merged dataset the questionnaires and methodologies were harmonised. From Cohort 13, onwards, the YCS no longer surveys individuals in Wales and is conducted only in England. Another change incorporated at Sweep 1 of Cohort 13 was the use of face-to-face interviews in order to address falling response rates, particularly for people with low educational attainment. The sample sizes for Cohort 13 are: Sweep 1 - 7,525; Sweep 2 - 6,297; and Sweep 3 – 5,411. A small proportion of individuals in sweeps 1 to 3 indicated that they were currently undertaking an Apprenticeship: 4.6 per cent in Sweep 1, 7.8 per cent in Sweep 2, and 6.2 per cent in Sweep 3. In each Sweep these proportions amount to less than 500 observations. Fieldwork for Sweep 4 was completed in Autumn 2010. While sample sizes, particularly those for those enrolled in an Apprenticeship, are relatively small in the YCS the main potential for using the data

comes about through merging Cohort 13 data with data on the same academic cohort covered by LSYPE.

Table 4.7: Age and Year of Sweeps by cohort, YCS

	Sweep 1		Sweep 2		Sweep 3		Sweep 4	
Cohort	**Age**	**Year**	**Age**	**Year**	**Age**	**Year**	**Age**	**Year**
1	16	1985	17	1986	18	1987	--	--
2	16	1986	17	1987	18	1988	--	--
3	16	1987	17	1988	18	1989	23	1994
4	16	1989	17	1990	18	1991	--	--
5	16	1991	17	1992	18	1993	--	--
6	16	1992	17	1993	18	1994*	18	1994
7	16	1994	18	1996	--	--	--	--
8	16	1996	18	1998	20	2000	--	--
9	16	1998	17	1999	18	2000**	18	2000
10	16	2000***	16	2000	18	2002	19	2003
11	16	2002	17	2003	18	2004	19	2005
12	16	2004	17	2005	18	2006	19	2007
13	16	2007	17	2008	18	2009	19	2010

*Cohort 6 was surveyed twice in 1994 (sweeps 3 and 4)
** Cohort 9 was surveyed twice in 2000 (sweeps 3 and 4)
*** Cohort 10 was surveyed twice in 2000 (sweeps 1 and 2)

Table 4.8: Variables related to Apprenticeship in Cohort 12 questionnaires

Apprenticeship or Job Details
- Modern Apprenticeship/Government supported training
- Level – Advanced Apprenticeship, Apprenticeship or do not know
- Activity within company
- Name of job or job being trained for
- Contract permanent or temporary
- Take home pay
- Hours worked
Employer/Business Attributes
- Number of employees
- Main business activity
- Other training received in past 4 weeks

YCS: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Multiple cohorts - continuously renewing 16-19 cohorts surveyed – can trace changes in programme over time
- Detailed information on Apprenticeship and training
- Asks why people wanted to undertake an Apprenticeship which provides valuable information about the aspirations of apprentices which might help explain why some people obtain higher returns from their Apprenticeship
- Linkage with LSYPE

DISADVANTAGES

- Ends at early age (19/20 years) – cannot observe late starters, longer term outcomes
- Excludes older people so can provide limited information for those aged over 18 years
- Attrition is a problem, particularly as attrition rates vary by gender and educational attainment. The sample size in each cohort varies. Typically, the achieved sample size reduces by around 25 per cent every year that the cohort is followed up

POTENTIAL USE

- Use with LSYPE
- Use cohort 13 with LSYPE data to follow-up to age 29/30 (though sample sizes are relatively small for apprentices)
- Provides an excellent template for the design of a survey which will fully capture the effect of completing an Apprenticeship

4.4.2 Longitudinal Study of Young People in England (LSYPE)

The **Longitudinal Study of Young People in England (LSYPE)** was flagged up as a potential future dataset to be used in evaluation of Modern Apprenticeships by Payne and her colleagues but at the time of that report, the LSYPE was still in development stages.

The LSYPE commenced in 2004 interviewing individuals in Year 9 of compulsory education (i.e. age 13/14 years) and has continued with annual surveys of these

individuals. There are currently six waves of data available (2004-2009) which follow the respondents from age 13/14 to 18/19 in 2009. Fieldwork is currently being carried out for Wave 7. The main objectives of the LSYPE are:

- to gather evidence about the transitions young people make from secondary and tertiary education or training to economic roles in early adulthood;
- to enhance the ability to monitor and evaluate the effects of existing policy and provide a strong information base for future policy development; and
- to contextualise the implementation of new policies in terms of young people's current lives.

Waves 1 to 4 of the survey included an interview with the young person (YP) along with interviews of at least one parent/guardian of the YP. Subsequent waves do not (and will not) include interviews with parents/guardians. Sample sizes and response rates for waves 1 to 6 are summarised in Table 4.9.

Table 4.9: Sample sizes and response rates, LSYPE, waves 1 to 6

Wave	Year	Age	Respondent(s)	Sample Size	Response Rate
1	2004	13/14	Young person and both parents	15,770	74%
2	2005	14/15	Young person and both parents	13,539	86%
3	2006	15/16	Young person and both parents	12,439	92%
4	2007	16/17	Young person and one parent	11,449	92%
5	2008	17/18	Young person	10,430	89%
6	2009	18/19	Young person	9,799	87%
7	2010	19/20	Young person	In progress	--

Source: DCSF (see http://www.civilservice.gov.uk/Assets/DCSF%20Conference%20-%20Making%20the%20most%20of%20Longitudinal%20Data_tcm6-37309.pdf)

The LSYPE dataset can be linked to the National Pupil Database (NPD) and other data sources such as geo-demographic data from the 2001 census but this is not included in UKDA datasets.[64] It follows that if the LSYPE could be linked to the ILR for those who individuals who go onto FE this would provide much information about

64 For further information on this data linkage see: Department for Education/NatCen (2010) LSYPE User Guide to the datasets: Wave One to Wave Six. October 2010. (http://www.esds.ac.uk/doc/5545/mrdoc/pdf/5545lsype_user_guide_wave_1_to_wave_6.pdf).

the educational attainment of people before they entered an Apprenticeship. In order to assess the feasibility of linking these datasets, one would ideally know the number of individuals likely to move onto FE after compulsory education. Around 85 per cent of the sample intend be in full-time education and training after age 16 (which is in line with other information on participation in post-16 education and training). Given a sample size of the LSYPE of around 12,000 this would equate to more than 10,000 young people linked between the LSYPE and either the ILR or HESA records.

There are a number of variables in the LSYPE dataset that are specifically concerned with Apprenticeship (summarised in Table 4.10). The dataset (and the linked YCS cohort 13) presents the greatest level of detail regarding Apprenticeship programmes outside of the ILR with survey items including the subject area, starting date, completion date, location of training provision and whether full-time or part-time. In addition, there are a number of questions which consider individuals' reasons for undertaking an Apprenticeship, alternatives they considered, and whether an Apprenticeship was their first choice of activity. Such information is vital in considering usually unobservable factors, such as motivation which influence individuals' employment outcomes. The distribution of responses to a number of the items summarised in Table 4.10 are presented in Annex 1.

Table 4.10: Summary of relevant variables in LSYPE[65]

Background Characteristics and Pre-Apprenticeship Information
- young person's family background and household information including socioeconomic status of parents, parental employment, household composition, household income - personal characteristics of young person (e.g. sex, ethnicity, disability status) - educational attainment - post-16 plans - information on schools attended
Details of Apprenticeship
- whether currently doing Apprenticeship - main subject - level - reasons for choosing Apprenticeship - satisfaction with aspects of Apprenticeship - knowledge of and intentions towards Apprenticeship (at 16) - alternatives considered when choosing Apprenticeship - payments received while training - post-Apprenticeship plans
Outcomes
- economic activity - occupation - income and benefits - attitudes to work and other issues - life satisfaction

Within Wave 6 of LSYPE, 5.7 per cent of the 9,107 valid responses to the question regarding main activity were doing an Apprenticeship. 16.1 per cent were in education while 32.6 per cent were in paid work. Amongst apprentices, just under 200 were at Level 2 and around 320 were at Level 3. These numbers, while small, would be increased with linking to YCS13. Given age of the cohort, it is also likely that later entries to Apprenticeship programmes will be observed in the future.

In determining the counterfactual for comparison with apprentices, the LSYPE (and YCS13) proves useful as it contains considerable information on young people's intentions regarding studying and work and their preferences and attitudes that are not captured in as much detail in any other study. Details related to such variables are again reported in Annex 1. In Wave 6 of LSYPE, nearly 80 per cent of apprentices indicated that this was their first choice of activity. Young people were

65 As discussed, methodological changes have been made to Cohort 13 of the YCS to allow linking with LSYPE as the respondents in both studies come from the same academic cohort (eligible to leave school in 2005/06). As such, the relevant variables indicated for LSYPE are also (for the most part) available in the YCS13 data.

also asked if they were considering other options at the same time when deciding whether or not to undertake an Apprenticeship.

Young people in the LSYPE were also asked in Wave 6 about the reasons they decided to do an Apprenticeship. More than 40 per cent of apprentices in the survey who responded to these questions 'strongly agreed' that: they wanted to do something practical rather than academic (48.2 per cent); they liked the idea of getting a job and training at the same time (47.1 per cent); and Apprenticeships is a well-recognised qualification (41.9 per cent).

The LSYPE is a promising longitudinal database which, potentially could be used in the long term evaluation of Apprenticeship should sample sizes be sufficient and specific data related to Apprenticeship be collected in future. At the present time, when wave 7 is being carried out, the respondents of the LSYPE are 20/21 years old and so the significant outcomes associated with Apprenticeship are unlikely to be visible. At age 19, many individuals, particularly those who did not enter Apprenticeship upon leaving school, are still in training or have not yet made the transition into the labour market.

Initially the plan was to follow LSYPE respondents to age 25 years (in 2015) - which would allow for a more comprehensive analysis of their initial entrance into the labour market upon completion of an Apprenticeship (or other forms of initial vocational education and training). The key issue is whether there would be a large enough sample of apprentices to allow analysis by level and subject. With around 6 per cent of respondents in Wave 6 recorded as apprentices it is unlikely that there would be a sufficiently large sample of apprentices by 2015.

LSYPE provides a research design which is appropriate to the evaluation of Apprenticeships if it contained more apprentices and adopted a multiple cohort approach. In the absence of such an approach, attempts could be made to link the responses of respondents to, for example, the WPLS, in order to obtain information about the labour market experiences of the apprentices included in the survey. But as pointed out elsewhere in this report – see Section 6 in this regard – linking datasets is far from straightforward. It is understood that LSYPE will be discontinued and will not now collect data on individuals into their mid-20s. Whilst the current sample sizes are too small to provide a comprehensive evaluation of Apprenticeships, as a research design LSYPE has much to recommend it and much to recommend its repeated use albeit with larger samples of apprentices.

LSYPE: Advantages, Disadvantages and Potential Use

ADVANTAGES

- Relevant age group
- Information from age 13
- School information
- Items on attitudes, ambitions, values, etc. (more qualitative in nature)
- Links to NPD, YCS, other
- Apprenticeship details – level, alternatives considered, first choice, subject, reasons for undertaking

DISADVANTAGES

- Only currently age 19
- Labour market outcomes not currently observed, for most part
- As with all longitudinal surveys, attrition and any associated bias is a potential problem
- Survey will not now continue to collect data from respondents into their mid-20s
- Single cohort survey

POTENTIAL USE

- Appropriate research design which could be implemented via a multiple cohort approach with much larger sample of apprentices
- Link to YCS, NPD, HMRC, DWP, HESA datasets

4.5 Other Quantitative Datasets

The data description provided to date has concentrated on datasets which provide information about the characteristics of apprentices (in comparison with other groups in the labour market), but there are other data sets which provide valuable information about Apprenticeships. These consist of *ad hoc* surveys of apprentices and surveys of employers offering Apprenticeships.

Three surveys are particularly important:

i. Survey of Prior Qualifications;
ii. Apprenticeship Pay: Survey of Earnings by Sector;
iii. National Employers Skills Survey.

4.5.1 Survey of Prior Qualifications

There have been three Surveys of Prior Qualifications:

i. amongst adult Apprenticeships;
ii. amongst those participating in the previous Train to Gain programme; and
iii. amongst adults undertaking college-based learning under the previous Adult Learner Responsive budget.

The first survey, which is most relevant to this study, was undertaken to better understand the extent to which investment in Apprenticeship provision is being directed towards up-skilling individuals with lower skill levels, especially those without attainment at Full Level 2 (equivalent to 5 GCSEs at grades A-C or NVQ Level 2).

The survey included a total of 3,000 telephone interviews with adults (aged 19 plus) who had been undertaking an Apprenticeship at Full Level 2 or Full Level 3 in November 2009. The sample of learners was drawn from the Individualised Learner Record (ILR). A similar number of interviews were undertaken for adult learners enrolled in Train-to-Gain and college-based provision.

The surveys are of interest because they provide information about the extent to which Apprenticeships add value insofar as people commence Apprenticeships with differing levels of educational attainment already achieved. As noted elsewhere in this report, the educational attainment on entry to an Apprenticeship is likely to provide valuable information which helps explain differential returns.

4.5.2 Apprenticeship Pay: Survey of Earnings by Sector

An important occasional survey of apprentices is the Apprenticeship Pay: Survey of Earnings by Sector.[66] This survey was conducted in 2005 and 2007 in England, and a UK-wide survey is now being undertaken in 2011. The survey is limited to providing information from apprentices, with employed status. The eleven sectors are:

i. Business Administration;
ii. Construction;
iii. Customer Service;
iv. Early Years Care and Education;
v. Electro-technical;
vi. Engineering Manufacturing;
vii. Hairdressing;
viii. Health and Social Care;
ix. Hospitality;
x. Motor Industry;
xi. Retail.

With up to 500 interviews in each sector the survey provides fairly robust estimates of apprentice pay. It also provides estimates of the total amount of time spent on off- and on-the-job training which can be used as an indication of Apprenticeship quality. Data drawn from other sources, such as the Labour Force Survey, allows a comparison to be made of apprentice pay with that of other groups in the labour market (such as fully experienced workers or other types of trainee). The survey also provides information about the amount of training received by the apprentice which can be used as an indicator of training quality.

4.5.3 National Employers Skill Survey

The National Employers Skills Survey is a representative survey of employers undertaken, on average, every two years. Around 75,000 establishments are surveyed to obtain information about:

- the structure of employment;
- skill shortages and skill gaps;
- training activities;
- training costs; and
- recruitment of people straight from school or college.

The survey provides an indication of the number, or percentage of establishments, which have recruited an apprentice over the past 12 months (around 8 per cent of establishments in 2009). The usefulness of the survey, from an evaluation perspective, is that it is able to provide information about the types of employer which recruit apprentices (by size and sector) alongside other information about

66 Fong, B., and Phelps. A. (2008) Apprenticeship Pay: 2007 Survey of Earnings by Sector, Department of Innovation Universities and Skills,Research Report 08-05.

human resource policies and training activities. Potentially it provides the opportunity to assess whether employers who recruit apprentices have a unique set of characteristics – which has implications for widening employer participation in Apprenticeships – or whether they are more or less likely to experience skill shortages or gaps which are likely to inhibit organisational performance.

The 2011 National Employers Skill Survey will be extended to include Scotland, Wales, and Northern Ireland in addition to England.

4.6 Linked Datasets

There is a wide range of data relating to Apprenticeship available in existing statistical resources. As the discussion in this report highlights, such sources seldom contain all of the information required for a longer-term evaluation of Apprenticeship outcomes. There may be insufficient detail in the data collected (for instance treating Apprenticeships as homogeneous) or the necessary data may not be collected at all. One response to this might be to design a bespoke data collection instrument (survey or management information system) but an alternative approach is to link, or match, existing data sources.[67]

There are several potential advantages from data matching. First, it reduces the need to collect new data. This will significantly reduce the cost of any evaluation. Regardless of whether the data source is administrative data or a survey, the cost of data collection has already been incurred and the principal cost is limited to the time and resource spent on the matching process. Second, a matched dataset will provide a much richer and comprehensive dataset than the original data sources considered individually. Where several data sources are matched the resulting matched dataset will contain data for a range of variables far beyond what could realistically be collected by a single survey.

While data matching is attractive, it needs to be borne in mind such an approach for the long-term evaluation of Apprenticeships requires a number of conditions to be met: matching cases must be feasible; the resulting matched dataset must be capable of supporting the robust analysis required; and participant consents are obtained where survey data are being linked to other data sets.

In terms of the feasibility of matching, ideally, a unique 'identifier' should be present in all data sources to allow exact matching of common cases. Unfortunately there is at present no unique identifier common to all relevant datasets. A person's National Insurance Number (NINO) could, potentially, provide such a unique identifier but in practice is not perfect (duplicate NINOs for several people, multiple NINOs to the

67 Data linking refers to the creation of an association between data in two or more data sources, although those data sources remain separate. An example of linking would be if more recent information from one source were used to up-date (say) current economic status in another data source. Data matching goes further that data linking by matching the records of cases common to the two sources and creating a new, matched dataset containing the records from both data sources.

same person or cases where a person has no NINO). In addition, many surveys do not collect the respondent's NINO, thus preventing matching to dataset that do. Common usage of the Unique Learner Number would, in time, help to facilitate matching (but, as with NINOs, would still be subject to administrative errors).[68] In the absence of a unique identifier other matching methods would need to be used. Matching would be on the basis of a key variable (such as postcode) with a follow-up comparison of further variables (such as gender, date of birth etc.) to arrive at a 'probable match'.

Even if matching is feasible, the resulting matched data may have its limitations. The matched dataset can only be a large as the smallest data source from which it is matched (thus, matching does not overcome sample size limitations). In addition, any systematic difference between matched and unmatched cases will result in a biased sample. Such a difference could result from systematic errors in the recording of variables used for matching or from differences in the propensity of respondents in source data to give consent to their data being linked.

Finally, data matching raises a number of privacy and data security issues. Source data may have restrictions on the purposes for which it is used, while data can only be linked where consent has been given for that to happen[69]. Issues of this type can often be resolved if the possibility of data linking or matching is designed into data collection in the first instance. For instance, LSYPE was designed from the outset to be linked to other data sources (NPD and Work and Pensions Longitudinal Survey, WPLS) and consent was a pre-requisite for inclusion in the survey. Understanding Society also had linkage capabilities built in from the start.

Bearing in mind the advantages and disadvantages set out above, what are the options regarding use of matched data for a longer-term evaluation of Apprenticeships? Central to any linked or matched dataset for that purpose must be the ILR, as this contains the greatest detail of the Apprenticeship undertaken (level, framework, completion etc.) and covers the population of apprentices. In fact, the ILR has already been linked to NPD since 2000/01 and has linked to HESA records in order to identify learners who complete their learning aim in one year and remain in learning in a subsequent year. The link to NPD provides valuable information about the prior educational attainment of apprentices. At present, a critical gap exists relating to the destinations and subsequent careers of apprentices (the ILR provides only limited information about immediate destination. This gap could be partly filled if the ILR were to be linked to the National Benefits Database (NBD) and HMRC tax return records (P45 and P46). These two data sources would allow the tracking of the labour market status apprentices (identified via the ILR), covering

68 The Unique Learner Number (ULN) is a 10-digit reference number used to access the Personal Learning Record of anyone over the age of 14 involved in UK education or training. Learners retain the same number throughout their lives, whatever their level of learning and wherever they choose to participate in education, training and learning. ULNs are administered by the Learning Records Service on behalf of the Department of Education and Department for Business Innovation and Skills.

69 See 'National Statistics Code of Practice: Protocol on Data Matching', National Statistical Office, 2004.

employment, pay and non-employment and have the advantage of providing comprehensive coverage of the population of apprentices.

The Department for Business, Innovation and Skills (BIS) is in the process of developing just such a matched dataset linking the ILR, the NBD and HMRC records. The matching is undertaken by DWP on behalf of BIS who only receive an anonymised individual dataset. Matches are made, initially on the basis of NINOs (providing an exact match in between a third and a half of all records) and then probability, or 'fuzzy', matching (utilising name, date of birth, gender, and postcode) is used to generate other matches. The overall level of matching achieved is around 80-90 per cent.

The ILR is not a cumulative database, having a separate dataset for each academic year so that the matched data refers to a matching between each of these annual ILRs and NBD and HMRC records.[70] Each ILR is, however, up-dated through the academic year and the matched data for each year is up-dated on a quarterly basis to add in new learners who have appeared in the ILR (even though the ILR record of Apprenticeships is up-dated on a monthly basis), and updated information on those learners who have already been matched. One limitation to the BIS matched data is that it relates only to learners in the ILR who are aged 19 years or above, although data relating to 16-18 year olds is available once the learner has turned 19. This means that the matched data has comprehensive coverage of 19 year old apprentices but potentially less complete coverage of 16-18 year old apprentices. Because the latter group have had less interaction with the labour market, it is less likely that a match for these individuals will be achieved anyway.

The BIS developmental match dataset has been analysed with an emphasis on examining outcomes for learners who have completed their learning aims (such as an Apprenticeship). The analysis undertaken to test the datasets is reported to have been satisfactory with results that make sense and are potentially useful to the Department. The resulting matched datasets are large, being the result of merging very large source datasets and contain very large amounts of information. While the number of variables in, for instance the HMRC dataset is small, the number of pieces of data is large, with the number of observations on earnings exceeding 70 million. The principle issue with both the NBD and HMRC records is in regard to the recording of dates (beginning and ends of spells of benefits and employment) where there is a degree of omission and error in the records.

To use the matched data for the purposes of a long-term evaluation of Apprenticeship would require a number of practical issues to be addressed. The matched data is made available to BIS in the form of individual records, albeit in anonymised form, while the use of HMRC records is always likely be a matter requiring additional sensitivity to data security and access issues. Use of the data for an evaluation of Apprenticeships will require appropriate permissions to be obtained from HMRC and DWP concerning who has access to the records, what they may be used for and their disposal at the end of the evaluation. BIS is of the

70 Note that NBD and HMRC databases are cumulative.

view that the most likely options here would be for the Department to undertake the analysis in-house (perhaps to a specification determined by expert advisers from outside the Department) or to commission an outside research organisation to undertake such an evaluation under strict conditions relating to data access and security (or a combination of both). The option selected will reflect the nature of the evaluation activity being undertaken: routine monitoring of apprentice outcomes might be best undertaken in-house but advanced analysis of, say, the additional financial returns to Apprenticeship might be best undertaken by specialist research organisations.

While the matching of the ILR with the NBD and HMRC records can yield extremely valuable information relating to the economic outcomes of an Apprenticeship – it would readily allow an analysis of the economic outcomes over a five to ten year period after the Apprenticeship has been completed - it is unable to address qualitative questions relating to the Apprenticeship, such as job satisfaction, access to further learning or provide evidence of the quality of Apprenticeships and relevance to post-Apprenticeship employment. These matters could be addressed by linking the ILR to existing surveys (if suitable ones are available) or by using the ILR as a sampling frame from which to derive a bespoke and linked survey of apprentices to collect data on these other matters. The former option will be limited by the limitations of existing surveys (such as small sample sizes or limited relevant data collection). The latter would allow precisely relevant data to be collected but such a survey, being new, would have cost implications.

In summary:

- matching existing data sources offers a low cost option with the potential to provide a rich, more comprehensive evaluation dataset;
- matching cases from existing data sources is a challenging task that would need to deal with issues of identification of matched cases and a range of privacy and data security issues;
- a matched dataset will only be as good as the weakest data source (e.g. the size of the smallest sample);
- the ILR needs to be central to any matching exercise as it contains detailed information on Apprenticeships and has already been successfully matched to NPD and HESA records and to NBD and HMRC records;
- the BIS development matched ILR/NBD/HMRC dataset has the capacity to generate longitudinal outcome data in terms of employment and earnings (but not qualitative matters such as job satisfaction or ex-post satisfaction with an Apprenticeship);
- The ILR could be linked to other sources (existing surveys) but this would face the same limitations (e.g. small sample size, rack of relevant data) as those other data sources;
- The ILR could be used as a sampling frame from which to launch a bespoke (and perhaps longitudinal) survey of apprentices to provide evidence of broader outcomes than just earnings and employment status.

4.7 Qualitative Approaches

The final set of data which provides important information about apprentices and their employers are those studies which adopt a more qualitative approach. The limitation of these approaches is that they provide non-representative samples of either apprentices or their employers but are able to provide a considerable amount of detail and address issues relating to: (a) the decision to either become an apprentice or take on an apprentice (if an employer); (b) the reasons for completing or not completing training;[71,72] and (c) the quality of training provided.

The benefit of the case study approach is that it allows more complex theories to be tested and developed. There are a large number of case study based analyses of apprentices but two approaches are worth mentioning:

i. the work of Fuller and Unwin which has sought to conceptualise Apprenticeships with respect to whether they are expansive or restrictive;[73,74] and
ii. the IER Net Costs of Training series of studies which have attempted to measure employer investment in Apprenticeships.[75]

Both groups of study provide an indication of the quality of Apprenticeships. The former looks at the work situation within enterprises to gauge the extent to which this facilitates the provision of learning opportunities. The implication is that organisations which have a more expansive approach which allows new forms of activity to develop in the workplace provide a higher quality learning environment for the trainee or apprentice. The latter group of studies adopt a less theoretically sophisticated approach based on measuring the level of employer investment in training apprentices. In some sectors the evidence suggests that employers regard the investment as a short-term one where the investment is recouped more or less within the formal training period, whereas in others the employer's investment will only be recouped over the medium-term (i.e. around three years) so long as the employer is able to retain the services of the apprentice.

The case study approach is one that complements the more quantitative approaches and is able to inform the construction of large-scale data collection exercises. As such, qualitative research methods have an important role to play in the evaluation of Apprenticeships, not least with respect to whether the quality of training is improving over time.

---

71 Hogarth T., Gambin, L., Hasluck, C., de Hoyos, M. and Owen, D. (2009) 'Maximising Apprenticeship Completion Rates', Coventry: Learning and Skills Council.
72 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.
73 Fuller, A. and Unwin, L. (2008) 'Towards Expansive Apprenticeships'. A Commentary for the ESRC's Teaching and Learning Programme. London: Institute of Education.
74 Fuller, A. and Unwin, L. (1998) 'Reconceptualising Apprenticeship: Exploring the relationship between work and learning', Journal of Vocational Education & Training, 50(2): 153 – 173.
75 Haslucket al, 2008; Hogarth and Hasluck, 2003.

Quantitative and qualitative approaches to evaluating Apprenticeships should not be seen as separate, discrete exercises. Each has, through an iterative process, the ability to inform the other. In particular, qualitative investigations allow for exploratory analyses which highlight the factors which might explain why the returns to one type of Apprenticeship is greater than for others, or why some groups benefit more than others. The results from these types of investigations inform quantitative approaches with respect to the issues to be addressed and measures to be included. The results from any subsequent quantitative investigation can then be investigated further through case study techniques.

QUALITATIVE INVESTIGATIONS:

Advantages, Disadvantages and Potential Use

ADVANTAGES

- Qualitative dimension would add further interest and detail to evaluation
- Allows for investigation of more detailed and specific points of interest
- Can provide much assistance with the specification of hypotheses for testing in quantitative research

DISADVANTAGES

- Typically smaller scale than other datasets considered
- Difficult to be sure how representative data are in practice unless cases drawn from a robust sampling frame – such as a representative survey

POTENTIAL USE

- Exploratory investigations which provide insights relating to issues and measures to be included into quantitative surveys, or exploring insights provided by quantitative surveys in more detail.

4.8 Conclusions

The description provided in this section indicates that there is a rich store of information available on apprentices which allow a comparison with a range of other groups in the labour market. The assessment of the various datasets updates and expands upon the discussion provided by Payne *et al*. There have been a number of developments in terms of data availability and quality over the past decade which are relevant. New survey data has come to the fore (i.e. Understanding Society) and the cohort members of other longitudinal datasets have progressed with age, making transitions between childhood and adolescence, from school to the further education, higher education or the labour market (i.e. with reference to participants in LSYPE,

ALSPAC, BCS, and NCDS). Inevitably in the face of practical constraints, other data sources have not been continued or enhanced or currently are the subject of revision with respect to whether or not they will continue in the future.[76] A summary of the advantages and disadvantages along with potential uses of the most suitable datasets considered is presented in Table 4.11.

The use of administrative data sources for research and evaluation purposes has also increased since the previous feasibility study which is a promising development as such records can, with proper use, supplement other data sources and enrich the information available. Furthermore, there are no issues with sample sizes, as these usually cover the whole population. Along with this increased usage of administrative databases comes the need for enhanced data security. Ensuring confidentiality and data security presents increased costs, both financial and opportunity costs due to labour and time use. Such costs need to be considered alongside the costs of introducing and administering a new survey designed to capture all information required to assess the returns to Apprenticeship.

Each dataset that has been considered has its own set of advantages and disadvantages. The main shortcomings found include: insufficient observations on Apprenticeships (particularly when disaggregating programmes by framework, level or some other attribute) to carry out statistically robust analysis; insufficient information about Apprenticeships in terms of programme attributes such as subject or level; and insufficient follow up periods to track individuals' progression through the labour market and relevant outcomes over the medium to long term. While separately, none of the datasets considered present a complete set of outcomes and explanatory variables over a sufficiently long period to robustly evaluate Apprenticeships, the potential to fill in gaps is there. Possible ways forward include improving some surveys through enhanced questionnaire content about Apprenticeships (at least level and subject) and boosting the coverage of Apprentices in the sample; extending the period over which the subjects of certain surveys are followed; and linking longitudinal surveys to other such datasets and/or administrative records. The options for overcoming present data limitations are presented further in Section 6.

76 BHPS has been absorbed into Understanding Society so not completely abandoned. LSYPE has been discontinued.

Table 4.11: Principal Datasets Potentially Providing Information Germane to the Evaluation of Apprenticeships

Data Set	Sample Size	Details about Apprenticeship	Research Design	Time Period	Age of individuals	Matching / Linking Potential
Administrative Databases						
National Pupil Database	Census of all students in maintained schools in England	Not relevant	Administrative Database	Ongoing	4/5 to 16 years	Has been matched to ILR and HESA data using Unique Learner Number
Individual Learner Record	Census of all participants in FE	Detailed information about level and framework, start and end dates	Administrative database updated yearly (by academic year)	Ongoing since 2002/3	All in system (no age restriction)	Has been linked to HESA and NPD via Unique Learner Number. Also can be linked to HMRC and DWP data
DWP National Benefits Database / WPLS	Census of people in receipt of benefits	Not relevant	Live administrative database	Ongoing	All in system in receipt of benefits (no age restriction)	Can be linked through NINO or probabilistic matching to other databases to provide details of employment record
Longitudinal Surveys (General)						
Labour Force Survey	c. 120,000 individuals	Treats Apprenticeship as homogeneous entity	Sample survey of households	Quarterly survey	Individuals aged 16 years or over interviewed	Via probabilistic sample – contains neither NINO nor Unique Learner Number Information
NCDS	Wave 0: c. 17,000 Most recent wave: c. 10,000	Information on subject and start and end dates; not MA	Single cohort study	Started in 1958; most current 2008/09; lapse between waves increased over time. Respondents will not have been in scope of the Modern Apprenticeship initiative.	Birth to 50/51	Not relevant
BHPS / Understanding Society	BHPS: c. 5,500 households; c. 10,300 individuals Understanding Society: c. 40,000 households; c. 80,000 individuals	Treats Apprenticeship as homogeneous entity	Longitudinal survey of households and individuals within households aged 16+	BHPS: annual; 1991 (W1) to 2009 (W18) Understanding Society: annual ; Wave 1 2009/2010	16 years and over	Understanding Society: built in linkage capabilities to health and education data and is aiming to link to HMRC and DWP records

Data Set	Sample Size	Details about Apprenticeship	Research Design	Time Period	Age of individuals	Matching / Linking Potential
BCS 70	c.16,500 in 1970; less than 10,000 in 2004 (age 34)	Year of award, subject, full/part time,	Single cohort study	1970 to 2008/09; various intervals	Birth to 38 years	No consent requested to date; next wave (age 42) to request consent.
ALSPAC	c. 14,000 at start	Only indicate of whether plan / likely to do MA post-Year 11	Single cohort longitudinal study; includes YP interviews as well as parent, teacher interviews and other examinations	1991/1992 to present	Pre-birth to 18/19 years	Many built in linkages; biological/ medical records.
Longitudinal Surveys (Training / Education Specific)						
YCS	Varies Cohort 12, Sweep 1: c. 14,000 Cohort 11, Sweep 3: c. 7,800	Detailed information on Apprenticeship by Level and Subject, plus reasons for taking an Apprenticeship	Multiple cohort, individuals surveyed over two years	Ongoing / Yearly	Individuals aged 16-19 years	Via probabilistic matching though harmonisation with LSYPE in Cohort 13 should enable same linkages as LSYPE (see below)
LSYPE	c. 16,000	Detailed information on Apprenticeship by Level and Subject, plus reasons for taking an Apprenticeship	Single cohort linked to YCS	Yearly	Follows people from 13/14 years to age 20	Permission given to link to NPD. Feasibility of matching to WPLS and other data being explored. Can be linked to other data sets using Unique Learner Number

5. Apprenticeship Evaluation in Other Countries

Although unique in many respects, the Apprenticeship system in England has some commonalities with systems of such training in other countries. Considering the approach to evaluation of Apprenticeships adopted in other countries, and indeed whether or not such evaluations are carried out in other systems, can help to provide some direction as to the most appropriate practices to adopt and on what aspects can be improved.

Given the success and popularity of the dual system in **Germany**, the prospects for the evaluation of Apprenticeship appear promising at first glance. Many of the components that would facilitate the evaluation of Apprenticeships are in situ: one half of each youth population cohort enters Apprenticeship; labour market programmes for adults have been the subject of many academic-led econometric evaluations;[77] longitudinal and linked employer-employee datasets have lately become available;[78] and a large public institute (BIBB) is responsible not only for administering the Apprenticeship system but also for analysing its functioning.[79]

The promise has not however been realised in practice. The reality in Germany is that public evaluations of labour market programmes have mostly been confined to indicators of gross outcomes for participants (notably employment rates after participation[80]). Various factors contribute to this situation. First, RCTs are essentially ruled out on grounds of equity (denial of entry to qualified individuals). Second, the sheer size of Apprenticeship makes econometric evaluation harder, by reducing the availability of a reasonably well matched comparison group (a problem that may possibly arise in England in the future given a drive toward expanding the programme and turning it into the main vocational education route for young people). Third, occupational specificity also impedes evaluation: unlike France and the Netherlands, where Apprenticeship and full time vocational education provide alternative paths to particular vocational qualifications, in Germany each occupation is typically attained by one or the other, rather than by both. Occupation-specific and gender-specific factors therefore cloud the assessment of outcomes – as in the case of the low pay of ex-apprentices in hairdressing compared to engineering in the UK and the US.[81] Fourth, the institute charged with administering and studying Apprenticeship (BIBB)

---

77 Fitzenberger, B. and Speckesser, S. (2000), 'Zurwissenschaftlichen Evaluation der aktivenArbeitsmarketpolitik in Deutschland: einÜberblick', Mitteilungenaus der Arbeitsmarkt- und Berufsforschung, Nr. 3.

78 Dustmann, C. and Schönberg, U. (2009), 'Training and union wages', Review of Economics and Statistics, 91(2): 363-76.

79 Pfeiffer, H., Schönfeld, G. and Wenzelmann, F. (2011), 'How large is the firm-specific component of German apprenticeship training?', unpublished paper, BIBB, Bonn.

80 OECD (2005), Thematic Review of Adult Training: Country Note, Germany. Paris: OECD (http://www.oecd.org/dataoecd/44/5/36341143.pdf).

81 Ryan, P. (2001), 'The school-to-work transition: a cross-national perspective', Journal of Economic Literature, 49:34-92.

has concentrated its research efforts not on broad evaluations but rather on mechanisms and finance, in response to recurrent concerns about the viability of Apprenticeship in general, and the fear of a dwindling appeal to employers in particular.[82]

More generally, there is the difference between institutions and outcomes. Apprenticeship is constituted in Germany as an institution, part of the formal education system, rather than as a labour market programme. As such, it has benefited from the post-war spread of public belief in its importance and efficacy, which both reduces the perceived urgency of formal evaluation and implicitly makes any such evaluation harder, in pointing to broader criteria than those in programme evaluation in the US, viz. educational development and youth socialisation. Thus a typical response to criticisms that much Apprenticeship is wasteful, in that occupational mobility after training means that many of the skills that apprentices have learned are left unused (as typified by the qualified baker working on the car assembly line), runs along these lines: they are young, they need to learn about life, and Apprenticeship teaches them that, so it's an effective institution.[83]

There is significant scope for evaluating Apprenticeship in **France**, as it and full time vocational education offer alternative routes to a wide range (in principle all) vocational qualifications recognised by the Ministry of Education. In France, the CEREQ's 'Generation' series of cohort surveys provides information about the activities of people who have exited the education system and initial education and training system. This series have been the basis of a major research effort, much of it descriptive (classification of pathways e.g.) rather than evaluation proper. The aim of the Generation surveys is to analyse the early labour market experiences of young people.[84] This series can be compared to the Youth Cohort Study (YCS) in the UK. The Generation series surveys run for much longer than YCS. Surveys were undertaken in 1992, 1998, 2001, and 2004. The 2004 cohort survey contained 65,000 young people who were followed up again in 2007. The survey covers all levels of post compulsory education and training so potentially provides the basis for a comparison between various groups.

A number of sources of Apprenticeship-related data exist at both the national level and within a number of provinces and territories in **Canada**. The Canadian Apprenticeship Forum overviewed existing data sources relevant for analysing the labour market outcomes of Apprenticeship in Canada since 2000[85]. The National Apprenticeship Survey (NAS) provides standardised data on individual apprentices across trades and provinces. The 2007 NAS covered the period 2002 to 2004 while

82 Beicht, U., Walden, G. andHerget, H. (2003), Kosten und Nutzen der betrieblichen Berufsausbiludung in Deutschland. Bonn: BIBB.

83 Taylor, M.E. (1981), Education and Work in the Federal Republic of Germany. London: Anglo-German Foundation.

84 CEREQ (2008) Enquête Generation 2004, CereqBref No. 248, http://www.cereq.fr/pdf/b248.pdf.

85 Canadian Apprenticeship Forum (2008)'Apprenticeship Labour Market Outcomes in Canada since 2000: An assessment of the Resources for Research'. Available at: http://www.caf-.fca.org/en/report/ApprenticeshipLabourMarket.pdf.

earlier versions of the survey covered 1989-1990 and 1994.[86] This dataset contains information on apprentices' demographic characteristics, the length of time to complete an Apprenticeship, sources of funding, and pre-Apprenticeship activities. The dataset also contains information on apprentices' labour market outcomes including employment status (i.e., permanent, temporary, seasonal), details for their most recent job, and compensation (wage/salary and benefits). A number of non-financial outcomes of interest are also included in NAS: job satisfaction (with respect to wages, job security, health and safety); details of their training experience (e.g. use of up-to-date equipment, satisfaction with teaching); and reasons for doing an Apprenticeship. The survey revealed that individuals who had completed an Apprenticeship had better labour market outcomes than discontinuers in terms of both employment and hourly wages three to five years after completion (or discontinuation).[87] Completers were more likely to be in employment, less likely to be unemployed or out of the labour force, and more likely to have permanent jobs where employed than were discontinuers. The Canada Overview Report[88]also considered labour mobility between provinces as an outcome for analysis. The NAS also set out to determine apprentices' overall satisfaction with their programmes and to link their responses to completion and retention rates.

Another national level data source in Canada is the Registered Apprenticeship Information System (RAIS) which covers apprentices and individuals who have recently become certified (similar in scope and content to the ILR). This data has been gathered annually since 1980. RAIS is more limited in scope than the NAS as it does not include post-completion outcomes but it does contain details of individuals' Apprenticeship programmes including duration of training, level of training, and sector. The National Graduate Survey (NGS) contains information on the labour market outcomes of graduates from university, college, and trade/vocational programmes. NGS studies have covered the graduating classes of 1976, 1982, 1986, 1990, 1995, 2000 and 2005.

While the NAS, RAIS and NGS datasets have been used in the research literature to a relatively great extent the scope of analysis has been rather limited (predominantly focused on completion and participation behaviour).[89]Relatively few studies have used these datasets to explore the medium to long term impacts of Apprenticeship on labour market and other outcomes for the individual.[90] Apprenticeship is often combined with other forms of training/vocational education outside of higher education in studies of the returns to education or training. An overall evaluation of Canadian

86 These earlier versions of NAS contained much less comprehensive data on apprentices than the 2007 data.

87 Menard, M., Menezes,F., Chan, C.K.Y. andWalker, M. (2008) "National Apprenticeship Survey: Canada Overview Report 2007" Ottawa: Statistics Canada. Available at: http://www.statcan.gc.ca/pub/81-598-x/81-598-x2008001-eng.pdf.

88 ibid

89 For example, Desjardins, L. andPaquin, N.(2010).Registered Apprentices: the Cohorts of 1994 and 1995, One Decade Later, Statistics Canada Catalogue no. 81-595-M No. 080.

90 Of the few exceptions are Stoll and Baignee (1997), Sweet et al (2000) and Boothby and Drewes (2010) (though Boothby and Drewes look at trade certification which includes apprenticeship).

Apprenticeship programmes with respect to the outcomes under consideration in this report is not evident in the Canadian literature.

In **Australia**, responsibility for collecting, managing, analysing, evaluating and communicating research and statistics about VET has been centralised into the National Centre for Vocational Education Research (NCVER). This is a not-for-profit company owned by state, territory and federal ministers responsible for training. NCVER collects a range of data relating to Australian vocational education and training. The main data collection instrument is a quarterly survey of all VET in Australia. In addition, NCVER has undertaken historical time series analyses (going back as far as 1963) and surveys of Apprenticeship destinations. The large amount of data available to NCVER has led to a wide range of Australian publications.

The most comprehensive dataset on which evaluation of Apprenticeship (and other forms of vocational education and training) in Australia can be based is the Longitudinal Surveys of Australian Youth (LSAY). LSAY tracks young people as they move from school into further study, work and other destinations. It uses large, nationally representative samples of young people to collect information about education and training, work, and social development. The LSAY data are collected through computer assisted telephone interviews.

Survey participants enter the study when they turn 15 years old, or as was the case in earlier studies, when they were in Year 9 of school. Individuals are contacted once a year for 10 years. Studies began in 1995 (Y95 cohort), 1998 (Y98 cohort), 2003 (Y03 cohort), 2006 (Y06 cohort) and more recently in 2009 (Y09 cohort). Since 2003, the initial survey wave has been integrated with the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA). Over 10,000 students start out in each cohort.

LSAY provides a rich source of information to help better understand young people and their transitions from school to post-school destinations, as well as exploring social outcomes, such as wellbeing. Information collected as part of LSAY covers a wide range of school and post-school topics, including: student achievement, student aspirations, school retention, social background, attitudes to school, work experiences and what students are doing when they leave school. This includes vocational and higher education, employment, job seeking activity, and satisfaction with various aspects of their lives.

LSAY is managed and funded by the Australian Government Department of Education, Employment and Workplace Relations (DEEWR), with support from state and territory governments. On 1 July 2007, the National Centre for Vocational Education Research (NCVER) was contracted to provide analytical and reporting services for LSAY. NCVER is undertaking this service for the Department in collaboration with the Australian National University's Social Policy Evaluation, Analysis and Research Centre (SPEAR).

While there is a well-developed and comprehensive system for collecting, monitoring and analysing data relating to Australian Apprenticeships, there has been only limited use of such data for longer-term evaluations of the returns to Apprenticeships.[91] The LSAY offers the best scope for undertaking such long-term evaluations but, as in the UK, the veracity of any findings is conditioned by small sample sizes and sample attrition.

The **United States** has developed during the past fifty years the world's most comprehensive and sophisticated evaluation industry. Every new federal labour market programme must nowadays make provision for its detailed evaluation; RCTs (or social experiments) are to be used wherever possible, and cost-benefit analysis features prominently. Academics and policy research institutes (e.g., MDRC, Mathematical) have evaluated a vast number of public programmes to date.

The strengths of American evaluation research can be seen in the sequence of evaluations performed for the longest-running programme for disadvantaged youth, the Job Corps. An econometric evaluation (using regression analysis and a comparison group) in the 1970s (number one) was followed in the 1990s by a RCT (number two), in which eligible individuals were assigned at random to participation or non-participation. That evaluation was subsequently improved by extending the evaluation period (i.e., the time that elapsed between initial assignment and outcome measurement) from four to seven years, and by replacing self-reporting by administrative records in the measurement of labour market outcomes (number three). All three evaluations used a quite comprehensive cost-benefit framework, including such social benefits as reduced criminal activity. The results proved sensitive to the data used. The first two evaluations inferred substantial benefits, to both participants and society, from participation – which suggested comfortingly that the choice among sophisticated evaluation methods was not an issue. The third evaluation however found smaller private benefits, and negative social benefits – which suggested the potential importance of a longer evaluation period and more reliable data on outcomes. The reputation of the Job Corps has suffered accordingly.[92]

The strengths of the American evaluation industry do not however transfer to Apprenticeship. Econometric research has suggested substantial benefits for participants in Apprenticeship, and implicitly for society, too.[93] Application of more systematic evaluation of Apprenticeship in the US has unfortunately been hampered by its small scale and occupational specificity (urban construction trades).

As the brief summary provided in this section demonstrates, England appears to be particularly well placed compared to other countries with established Apprenticeship systems with respect to the availability of data to evaluate the effects of this form of training.

---

91 A number of analyses have been carried out using the LSAY see for example, Lee (2010), McMillan et al (2005) and Marks (2009).

92 Long et al (1981); McConnell and Glazerman (2001); Schochetet al (2003).

93 Blanchflower and Lynch (1994).

6. Options for the Long Term Evaluation of Apprenticeships

6.1 Introduction

The long-term evaluation of Apprenticeships needs to satisfy two principal information requirements:

- the provision of robust, quantified estimates of the outcomes of Apprenticeships for successful trainees in terms of progression into further and/or higher learning, skills acquisition, job satisfaction, and improved employment outcomes (including, but not limited to, wages, employment duration, mobility between jobs and career progression); and
- the collation of data which will reveal under what conditions Apprenticeships produce the best results, for whom and the relative added value of an Apprenticeship compared to other forms of learning.

In deciding how this might be achieved there are a number of issues to consider, including:

- making best use of data which is already collected via surveys and administrative databases. Section 4 has indicated that there are a number of data sets which provide much of the data required to meet the two objectives set out above;
- obtaining value for money with respect to how much additional information will be obtained from commissioning new surveys, *etc.*;
- the need for timely results. There is a need for policy to be informed over the short-term as well as over the medium to long-run. Moreover, whilst it is possible to evaluate the impact of Apprenticeships over the lifecycle, at some point the signal from the Apprenticeship training will become faint, simply because other, more contemporary experiences will dominate explanations of the labour market position of an individual;
- the need for an approach which can flexibly accommodate changes in the structure and provision of VET. The training infrastructure has changed much over the last thirty years such that the comparisons one might have made between Apprenticeship and some other form of training would have been very different in 1981 than in 2011.

With these caveats in mind the options for the long-term evaluation of Apprenticeships are set out below, but before doing so it is worthwhile to summarise what information is required and how much of it is already collected.

6.2 Information Needs

Section 2 set out a basic economic model to evaluate Apprenticeships. This is summarised in Table 6.1 below. The dependent variables which are of most interest are those relating to: (a) the earnings of the former apprentice; (b) the employment of the former apprentice; and (c) progression to further learning. Other outcomes of interest include job satisfaction, mobility between jobs, career progression and skills improvements. The time period over which these variables should be measured is not discussed in detail below, other than to say now that the principal effects of Apprenticeship training are likely to be most apparent over the early stages of an individual's entry into the labour market on completion of their initial vocational education and training (IVET). It is suggested that the time period over which the effects of Apprenticeship training should be observed in the labour market should be around five to ten years. In this way there is a degree of assurance that the effects being observed are the result of Apprenticeship training. As mentioned elsewhere in the report, the effects of the Apprenticeship training will lessen to be superseded by other events in the working lives of former apprentices.

There are also other, secondary dependent variables of interest such as the extent to which apprentices experience job satisfaction or are employed in relatively good jobs (however this might be measured given the large literature devoted to the subject of job quality).

In Table 6.1 the independent variables have been divided between those where it is possible to obtain information about the individual apprentice, and contextual variables which provide some information about their work situation once in employment, such as local labour market conditions. Similarly, the sectoral composition of employment may also affect outcomes, other things being equal; for example, where women are working in sectors mainly staffed by men (and vice versa).

An indication of the extent to which data are considered essential based on the existing research evidence and the extent to which any variable or group of variables is likely to substantially affect the dependent variable has been provided in Table 6.1. Implicit in this is that data are required at different points in time:

i. prior to the Apprenticeship commencing (*e.g.* prior educational attainment);
ii. during the Apprenticeship (*e.g.* quality measures / level / subject);
iii. immediately upon completion;
iv. after an initial period in the labour market (say after five to ten years after completing the Apprenticeship).

The extent to which data are available as specified above is outlined next.

Table 6.1: Information requirements to evaluate Apprenticeships

Dependent variable	Type of information	Essential data
Earnings	*e.g.* earnings relative to some comparator group	
Employment status	*e.g.* whether in or out of work, employment history	
Access to further education and training	*i.e.* post-Apprenticeship training and education (Foundation Degrees / degrees / professional qualifications)	
Other outcomes	e.g. level of job satisfaction; skills acquired and degree of match to job; changes in occupation and progression	
Independent variables		
Socio-demographic characteristics of former apprentice	Age, gender, ethnic group, *etc.*	√
Current educational attainment	Highest level of educational attainment (if not the Apprenticeship)	√
Educational attainment on entering Apprenticeship	Qualifications already held on entry to Apprenticeship	√
Age on entry to Apprenticeship	Age	√
Year Apprenticeship completed	Date	√
Type of Apprenticeship	Information about the subject of the Apprenticeship	√
Level of Apprenticeship	Levels 2, 3, 4/5	√
Quality of Apprenticeship	*e.g.* Amount of formal on/off-the-job training, number of guided learning hours	√
Attitudes and aspirations	Why Apprenticeship was chosen by individual (e.g. purposeful choice versus lack of an alternative)	
Characteristics of Employer providing the Apprenticeship	Since apprentices are delivered via employers, the characteristics of the employer are important: size, sector, *etc.*	
Training provider characteristics	Not much is known about how the training provider affects outcomes but some information on this might be useful, such as whether an FE college or a private training provider	
Contextual variables		
Key sectoral information relevant to framework	The employment composition of a sector can affect outcomes for certain groups	
Local labour market conditions	*e.g.* local unemployment rate. Other things being equal, returns will be lower overall in labour markets with relatively weak labour demand, *etc*.	

6.3 Data Gaps

From the information provided in Table 6.1 it is apparent that there is a core or essential data set which relates to:

- **the characteristics of the apprentice / former apprentice**
  - socio-demographic characteristics of the individual currently;
  - educational attainment of individual (current and prior to the Apprenticeship).
- **details of the Apprenticeship**
  - the level of Apprenticeship completed (Level 2, 3, or 4/5);
  - the broad subject area of the Apprenticeship (e.g. engineering, social care, *etc.*);
  - the year the Apprenticeship was completed;
  - the quality of the Apprenticeship.
- **details of the individual on access to the Apprenticeship**
  - age;
  - aspirations.

In general these data are available, but the critical weakness is that it is currently difficult, though not impossible, to make comparisons across time for the same apprentice. In other words it is difficult to provide a history of individuals which contains information about: (a) their educational attainment before commencing an Apprenticeship; (b) their experience of working towards completing an Apprenticeship in a given Framework at a given level; and (c) their initial progression through the labour market. Longitudinal data sets reviewed in Section 4 provide this type of historical information but in all cases sample sizes are too small to deal with the heterogeneous nature of Apprenticeships, and in many cases not all of the data required are collected. But all is not lost. As subsequent Sections discuss there is the potential to remedy this problem by linking several data sets. Before providing this information there is a need to consider a further aspect of the evaluation process: the choice of a comparator to Apprenticeships. Implicit in the information provided above is the notion of comparing the experience of apprentices or former apprentices to some other group.

Table 6.2: Data Availability and Data Gaps

Dependent variable	Type of information
Earnings	Data generally available in surveys and databases
Employment status	Data generally available in surveys and databases
Access to further education and training	Can be determined from longitudinal surveys but the data are limited with respect to the independent variables available
Independent variables	
Socio-demographic characteristics of former apprentice	Some of these data are available in surveys and databases
Current educational attainment	Data generally available in surveys and databases
Educational attainment on entering Apprenticeship	From ILR and Survey of Prior Qualifications, but otherwise generally not available
Age on entry to Apprenticeship	From ILR but otherwise generally not available
Year Apprenticeship completed	From ILR but otherwise generally not available
Type of Apprenticeship	From ILR and some surveys but generally limited availability with reference to other independent variables
Level of Apprenticeship	From ILR and some surveys but generally limited availability with reference to other independent variables
Quality of Apprenticeship	Limited to Survey of apprentice Pay
Attitudes and aspirations	Available in selected surveys but generally limited availability with reference to other independent variables
Employer characteristics whilst training	Limited
Training provider characteristics	From ILR only
Contextual variables	
Key sectoral information relevant to framework	Generally available
Local labour market conditions	Generally available from LFS

6.4 Comparator Group(s)

The evidence presented in Section 3 suggests that while experimental methods (an RCT) provide a robust means of identifying the benefits of a programme such as Apprenticeships, the approach can be ruled out here because of the practical and ethical difficulties of randomly assigning people on such a large, national programme and because the programme has run for many years and any existing data relating to Apprenticeships is already 'contaminated' by individuals who were not randomly assigned and who self-selected (and been selected by their employer) into an Apprenticeship programme.

In the absence of an RCT it is then necessary to determine a comparison group against which the outcomes for apprentices can be compared. The initial options study (Payne *et al*) highlighted the difficulty of identifying such a comparator group and this study has also found it difficult to identify such a group. In light of this, it is suggested that the comparator group(s) should consist of the main alternatives to

Apprenticeship regardless of the differing demographic, socio-economic, and educational composition of those groups. The principal comparator groups are:

i. Level 2 and 3 vocational qualifications obtained other than by Apprenticeship;

ii. qualifications below the level at which the Apprenticeship is offered (i.e. comparing Level 3 Apprenticeships with Level 2 qualifications, and Level 2 Apprenticeships with Level 1 and other Level 2 qualifications, along the lines of the comparisons provided by McIntosh);[94]

iii. the academic pathway through further education (i.e. to A-level and equivalent);

iv. first degree level in higher education;

v. no further training or education at the end of compulsory education:

vi. those with no or limited qualifications on exit from the education system;

vii. those with qualifications on exiting the education system;

viii. those who did not choose an Apprenticeship but otherwise had the same socio-demographic and educational characteristics.

There are benefits to taking a flexible approach to the comparator group and having a number of alternative comparisons. The principal benefit is that as Apprenticeship changes the approach to its evaluation can change too. The way data are currently collected potentially allows for any of the comparator groups listed above being used in analysis. To some extent the choice of comparator group depends upon the policy issue being addressed.

6.5 Core Options

As noted at the beginning of this Section there is a need to make the most of the data which are already available. The UK is well placed compared with many other countries in the availability and volume of data which might usefully shed light on the returns to Apprenticeships. These are:

- administrative data including the NPD and ILR;
- surveys of young people's progress through post-compulsory education and into the labour market – mainly YCS and LSYPE;
- longitudinal studies which cover the lifecycle, such NCDS, BCS, Understanding Society, *etc.*

The longitudinal studies, many of which select people born on a single day in a given year and follow them for a period of time, sometimes provide multiple cohorts which show how people's transition through the labour market has changed over time. Understanding Society breaks with this tradition in that it contains people of various

[94] McIntosh, S. (2007).A *Cost Benefit Analysis of Apprenticeships and Other Vocational Qualifications*, Department for Education and Skills Research Paper RR834.

ages. It is suggested that the approach to be adopted in evaluating Apprenticeships is a multi-cohort one in order that: (a) any changes to Apprenticeships which might be made over time can be captured in the evaluation; and (b) changes in skill demand over time are taken account of as well. Such an approach is in keeping with the tradition of YCS.

Option 0: Maximise Use of Existing Data

Before suggesting a range of alternatives it is worth summarising what existing data will provide with respect to the evaluation of Apprenticeships. There are essentially three sets of data which need to be identified, which:

i. provide data on the pre-Apprenticeship and Apprenticeship period;
ii. capture the initial labour market position of apprentices; and
iii. reveal the position of the apprentice at a later date (say after five to ten years post-initial training to see how well apprentices have established themselves in the labour market).

If Apprenticeship is treated as a homogenous entity then it is possible to provide much of the information required. For example, the NPD and ILR – which are available to Government researchers – have been linked and provide information relating to (i) and (ii). In addition, YCS / LSYPE provide information which reveals something about the aspirations of apprentices. Similarly, a number of longitudinal studies provide information relating to (ii) and (iii). The sample sizes, however, in most data sets which cover (ii) and (iii) are at the margin of what will produce statistically significant results. The principal drawback in relying exclusively upon existing data sets and their current interlinking is that it is difficult to go much beyond regarding Apprenticeship as a homogeneous entity when looking at the early labour market experiences of apprentices. Thirty years or so ago it might have been acceptable to treat Apprenticeship as a near homogeneous entity given the extent to which this form of training shared a common structure across the relatively few industries which provided it. This is not true today. Apprenticeship has been purposely developed as a heterogeneous training programme which can flexibly meet the demands of employers with widely differing skill needs whilst ensuring that apprentices acquire the skills necessary to progress through the labour market. An evaluation which seeks to both measure and explain the differential returns to Apprenticeship must capture its heterogeneous nature with respect to level and subject (as outlined above with respect to the minimum data requirement). The data currently available will not satisfy the minimum data requirement. How this may be remedied is outlined in the options described below.

Option 1: Using Linked Administrative Data

Rather than relying upon survey data one approach is to rely wholly on administrative data. This would include the NPD (for pre-Apprenticeship details), the ILR (for details of the Apprenticeship), HMRC data relating to employment and earnings, and the National Benefits Database relating to spells of non-employment and benefits.

The ILR and the other datasets have almost complete coverage of people in learning and their subsequent economic and earning status (sampling and sample size is not an issue). It is also an efficient way to conduct an evaluation because the data already exists. A linked dataset of the type indicated above would be capable of generating evidence relating to two key measures of impact for an Apprenticeship, namely future employment and earnings. Initially this evidence would relate to the short term impacts but if data were continually linked then longer-term outcomes could be examined as well.

This option poses a number of challenges, particularly relating to data access and security but also relating to the extent to which matching cases across the different databases is feasible (and the extent of unmatched cases). This linking exercise has already been undertaken by BIS on a developmental basis and so represents a feasible option. Restrictions on access to data on 16-18 year olds means that the matched dataset would be most complete in regard to apprentices aged 19 years or older (although data about younger apprentices can be added retrospectively once they reach the age of 19, provided they remain in the ILR at that time).

This approach would provide a relatively narrow set of data relating to the outcomes of Apprenticeship (in terms of earnings and employment) but can draw on the ILR to explore the factors underlying differential outcomes.

Option 2: Extending YCS and Linking to Administrative Data

Both the YCS and LSYPE provide a basis for providing detailed information about the learning received by young apprentices relative to others who have not taken an Apprenticeship. From an evaluation perspective, these surveys need to be extended to encompass the early labour market experiences of apprentices and others, preferably until their mid to late 20s to see how they have established themselves in the labour market. At present YCS stops when the cohort has reached 19/20 years of age. Potentially it is possible to re-survey Cohort 9 (where respondents were aged 16 in 1998) or Cohort 10 (where respondents were aged 16 in 2000) because these cohorts would now be aged 29/30 and 27/28 respectively. There are problems of attrition when attempting to survey people so long after their last interview. Tracking young people over the early stages of their entry to the labour market poses a number of problems not least tracing them after they leave the parental home. Through the use of mixed methods in surveying, especially the use of telephone surveys using mobile phones, these types of problem are becoming more amenable to solution.[95]

Whilst it is possible to continue with the current cohort in YCS13 and LSPYE and follow them through to their mid-20s, by which time it will be possible to assess their

95 Increasingly, organisations such as the National Opinion Research Centre (NORC) are employing a mix of methods to ensure that surveys full cover the population which they are sampling. The typical approach is now to supplement random digit dialling with a sample of mobile telephone numbers. In addition, telephone surveys are further supplemented by selecting a sample of addresses, matching these to phone numbers, and then contacting them (face-to-face, if necessary).

early labour market experiences, sample sizes are likely to be too small for a detailed analysis of Apprenticeships. It is perhaps best to regard YCS/LSYPE as an ideal template which needs repeating albeit with a much larger sample of apprentices.

It should be noted that there is the potential with YCS and LYPSE to link to administrative datasets to provide a more complete record. The most important candidates for such a linkage are NPD (for details pre-Apprenticeship), the ILR (for details of the Apprenticeship), HMRC data relating to employment and earnings, and the National Benefits Database relating to spells of non-employment and benefits.

Administrative data and YCS / LSYPE potentially complement one another in a variety of ways. They also provide scope for efficiencies to be obtained; if data are already available in an administrative database this allows more time in a survey questionnaire to ask about other issues, such as those identified in Section 6.4 below. There is scope to link the NPD and ILR to surveys such as YCS and LSYPE. Much depends upon the ability to use unique identifiers. If, in principle, it were possible to link the NPD and ILR to, say, YCS, there is the potential to link to a variety of other data sets too. This relates very much to co-ordination of data collection and the need to provide unique identifiers which are carried across databases and survey datasets.

Option 2 is to use the YCS/LSYPE as template in terms of its content and methods and repeat it with a much larger sample of apprentices and to build in links to other administrative databases.

The advantage of Option 2 - a multi-cohort longitudinal approach - is that it provides information from the formative years of the apprentice (i.e. the period of their training and their entry to the labour market as captured by YCS and LSYPE), including their aspirations and reasons for undertaking an Apprenticeship, and relates this to other data: (a) about their educational attainment (from the NPD); (b) precise details of the Apprenticeship training they received (from the ILR); and (c) a detailed record of their labour market experience (via a link to HMRC databases).

Option 2 will provide the most robust means of measuring the outcomes of Apprenticeship training if the aim is to use an economic model as set out in Section 2. Using the existing cohorts covered by the YCS/LSYPE could be a starting point but doing so would not provide information about older apprentices. Instead, using these surveys as a template for a new multi-cohort longitudinal study will allow for inclusion of apprentices of various ages.

Option 3: Extending Use of the LFS

The identification of Apprenticeship training in the LFS is cursory in that it treats this form of training more or less as a homogeneous entity. It is potentially possible, through adding question to the LFS, to ask respondents what the subject of their Apprenticeship was, and the date of completion. The use of such questions, if it were possible to include them in the LFS, would need to be tested. Adding such

information would provide a valuable source of data for evaluating the outcomes of Apprenticeships. The problems associated with using the LFS as a source of information on Apprenticeships has been dealt with in Section 5. Even bearing in mind the caveats made there, there is merit in expanding the questioning on Apprenticeships in the LFS so that a comparable set of data are collected as for graduates where subject of study and level of degree are recorded.

Option 4: Maximising the Value of Longitudinal Surveys for Apprenticeships

Whilst extending the longitudinal dimension in a YCS/LSYPE type survey (i.e. following the current participants for a longer period) will provide much information about the under-19 age group doing so will provide little information about older apprentices. This is an important consideration given the existence of Adult Apprenticeships and higher level Apprenticeships where participants will be aged over 19 years on entry. There are now several longitudinal surveys following cohorts born on a single date in a given year or, in the case of Understanding Society, the latest longitudinal survey in the field, all persons aged over 16 years in 40,000 households. These surveys provide a wealth of information about various aspects of a person's life at various stages. These surveys face a number of problems relating to sample attrition especially where they are surveying people in the latter stages of their life. This is not a problem with Understanding Society since it includes people of various ages.

The principal problem with these types of survey is the small number of apprentices – or former apprentices – they contain. For instance, there are around 2,200 people aged 16-18 in the first wave of Understanding Society which suggests, other things being equal, that this includes around 150 apprentices. This is clearly too small for a robust analysis of apprentices' progression through the labour market. An alternative is to either boost the number of apprentices in these types of survey, or run a separate survey of apprentices in parallel with, for example, Understanding Society. The ILR provides a sampling frame from which to select apprentices.

If the aim of any supplementary survey of apprentices is simply to capture information about older apprentices, then an alternative option might be to develop a variant of YCS for older apprentices to run in parallel with that survey. This will need to capture data about experiences before commencing an Apprenticeship retrospectively.

6.6. Data Issues Relating to Options 1, 2, and 4

In many respects, Options 1 and 2 are the preferred options. This is discussed further below, but they do present two significant data issues:

i. linking datasets (*c.f.* Option 1).
ii. the scale and feasibility of conducting a representative survey of apprentices which can capture robust information by level and broad Framework (*c.f.* Option 2 and, also, Option 4).

6.6.1 Data Linking

In the Options set out above there is much given to the possibilities of data linking. The key advantage of data linking is that it exploits existing data. The costs involved in matching are not trivial but if routinized there are efficiency gains to be obtained.

The ILR has been linked to NPD since 2000/01 and to HESA records. Linking has been undertaken using the Unique Learner Number (ULN) plus 'fuzzy matching' based on name, date of birth, postcode etc. It is reported that up to 96 per cent of records were successfully matched. A second matching exercise has successfully linked the ILR and records from the National Benefits Database (NBD) and HMRC tax return records (P45 and P46). These three data sources allow the employment and pay of learners (identified from the ILR) to be tracked. The data was matched using National Insurance Numbers (NINO) and 'fuzzy' matching and achieved a match in 80-90 per cent of cases. There are two principal issues with this matched dataset. First, the ILR is not a cumulative database so it is necessary to carry out a separate matching exercise for each academic year. Second, the BIS matched data is only comprehensive for learners in the ILR who are aged 19 years or above, although data relating to 16-18 year olds becomes available once the learner has turned 19 (if they remain in the ILR).

Access to all matched data would require permission to be obtained from the various owners of the data (DWP, HMRC, DfE, HESA) and there will be restrictions on access, use and disposal.

While a matched NPD/ILR/NBD/HMRC database can provide information relating to economic outcomes, it cannot provide data on qualitative matters such job satisfaction, access to further learning or the quality and relevance of apprenticeships. These matters could, however, be addressed by linking administrative databases to survey data sets. It needs to be borne in mind that administrative data, such as the ILR and NPD, can only be linked to survey data sets where respondents have given their consent to have their data linked.

6.6.2 Surveying Apprentices

A preference was expressed in the options for conducting a survey of apprentices which collected information from them to a period about five years after they had completed their Apprenticeship. Essentially, sufficient observations to fill in each cell in Table 6.3 are required.

Apprenticeship Frameworks are grouped as follows on the Apprenticeships website (www.apprenticeships.org.uk):

a) Agriculture, horticulture and animal care
b) Arts, media and publishing
c) Business, administration and law
d) Construction, planning and the built environment
e) Education and training
f) Engineering and manufacturing technologies
g) Health, public services and care
h) Information and communication technology
i) Leisure, travel and tourism
j) Retail and commercial enterprise

From an investigation of the ILR to inform sampling, there are likely to be relatively few apprentices in (a) agriculture, horticulture and animal care; (b) arts media and publishing; and (c) education and training. Similarly, the health, public services and care group of Apprenticeships will comprise a more homogeneous group of Apprenticeships if defined as health and social care. On the basis of this, a suggested grouping of Frameworks is provided in Table 6.3.

There are a number of practical problems to be resolved in surveying apprentices given the current level of inflow by broad framework. If one wanted to survey apprentices at repeated stages, say, at the start of their Apprenticeship, at the end, three years after the end, and five years after the end, and achieve margins of error of ±3% at the level of broad Framework (as listed in Table 6.3) this would suggest that a large initial sample would be required. If the aim was to include an equal number of Level 2 and Level 3 Apprenticeships within each broad Framework this would indicate that by the end of the surveying period a target sample of 500 for each level would be required. Based on discussions with experts familiar with longitudinal survey research, attrition rates and non-response rates, achieving 500 interviews in each of the 8 main framework categories (4,000 interviews in total) would require a relatively large sample of individuals to begin with. Allowing for attrition rates, it would be necessary to achieve 1,250 interviews in each category (10,000 in total) in the first sweep. With a 1 in 4 response rate, it would be necessary to target 5,000 individuals in each of the 8 main framework categories (40,000 in total) in the initial sweep in order to achieve the desired sample size at the end of the surveying period. Given the population size of each new cohort entering Apprenticeships – *i.e.* 280,000 starts in 2009/10 - this suggests that the sample will comprise 14 per cent of the annual inflow (based on 2009/10 starts). Assuming a one in four response rate at the start of the survey suggests that there would be sufficient sample overall to commence such a survey. The ILR provides a sampling frame for such as survey.

The cost of such an exercise could be around £1 million for each cohort (based on four Waves using telephone interviewing, and some online surveying, with a 15 minute questionnaire). The sampling for the survey is likely to prove quite complicated given that the number of starts varies by both level and Framework.

The above, nevertheless, provides a broad brush indication of likely sample sizes required.

The value in conducting this exercise is that it will provide additional information about the returns to completing an Apprenticeship currently not provided by linked administrative data.

Table 6.3: Sampling Frame for Evaluating Apprentices

Broad Framework	Level 2			Level 3		
	Age Groups					
	16-18	19-24	25+	16-18	19-24	25+
Business Administration & Law						
Construction, Planning & the Built Environment						
Engineering & Manufacturing Technologies						
Health & Social Care						
Information & Communication Technology						
Leisure, Travel & Tourism						
Retail & Commercial Enterprise						
Other						

6.7 Additional Options

In addition to the core options listed above, two additional options are suggested relating to:

i. surveying employers about the types of apprentices they recruit and why they do so;
ii. collecting more qualitative information through the use of employer based case studies.

Option 5: The Employer Dimension

The core options see outcomes as being determined by the socio-demographic characteristics and prior educational experiences and attainment of apprentices. Much of the research literature points to the role of the employer being of critical importance with respect to the chances of completing an Apprenticeship and the quality of the training provided. There is relatively little information in current data sets about the characteristics of the employer providing the Apprenticeship apart from their size and sector. There are two options:

- to link the ILR records of apprentices to the responses of employers who record they have an apprentice in the National Employers Skill Survey (NESS). This then provides details of the employer providing the Apprenticeship alongside characteristics of the apprentice. There is also potential to link NESS to the Annual Business Inquiry (ABI) which potentially extends the scope of analysis to the relationship between Apprenticeships and organisational performance;
- more information could be asked in NESS – or in an additional survey which uses NESS as a sampling frame - of those employers that recruit apprentices to inquire about the number of Apprenticeships delivered by broad Framework and level, and the volume of training delivered. In addition, further questions could be asked of employers about why they do not recruit apprentices. The costs of achieving this goal are relatively modest given that additional telephone interviews, for a 10 to fifteen minute interview, are around £30, on average.

Option 6: Qualitative Approaches

As noted at various places in the main body of this report, qualitative research based on semi-structured interviews with employers and apprentices can yield a great deal of valuable evidence about the effectiveness of Apprenticeships in different employer settings. If the aim is to understand what works best, and for whom, then more qualitative and exploratory analyses can provide advances which further quantitative approaches can utilise in the design of questions in surveys. The need for periodic investigations into the effectiveness with which apprentices are trained and utilised in the workplace will explain much about why returns differ across sectors. It should also be noted that detailed quantitative investigations can, depending upon the

results of the analysis, provide hypotheses which are best tested and refined in the first instance through more qualitative investigations. In summary, quantitative and qualitative approaches need to be intertwined and form part of an iterative process which improves the overall understanding of what works best in different contexts.

6.8 Recommendations

All of the options presented above are worth pursuing. With respect to prioritising the options the evidence points, in order of priority, towards:

i. Option 1: Linking Administrative Datasets;
ii. Option 2: A LSYPE type survey containing a large number of apprentices and following them to around five years after they have completed their Apprenticeship.

Data collected in Option 2 could be linked to administrative databases provided that respondents to any survey grant their permission for this to happen. Whilst Option 1 would provide much of the data required, the addition of Option 2 would provide a more comprehensive coverage of the issues of interest.

Linking data provides a number of formidable challenges but through improved coordination to ensure that people have unique identifiers across databases, this should become easier. There are, in addition, a number of ethical and security issues to consider in additional to the technical ones in constructing linked databases and using them in research and evaluation. In general, these are issues with which those engaged in data linking are well aware.

As noted throughout this report, using administrative data alone will provide a narrow set of outcome variables and will not provide information about the aspirations of apprentices or their attitudes towards training. Hence the priority given to a LSYPE type survey which follows apprentices until they have spent around five years in the labour market following completion of their Apprenticeship.

Some of the other options are also attractive in that they make use of existing data sets or require relatively small amendments to existing data collection tools, such as adding questions to the LFS, for instance, is relatively straightforward. The options which are adopted depend to a large extent upon the priorities of policy makers. A set of options are provided which provide the means to look at a number of different dimensions of Apprenticeship training – the interest in which may change over time as VET policy adapts to meet the changing demands of the economy – based on making best use of the data currently available and filling data gaps as effectively as possible. Options 1 and 2 will achieve this goal but this does not preclude adopting any of the other options.

References

Beicht, U., Walden, G. and Herget, H. (2003), Kosten und Nutzen der betrieblichen Berufsausbiludung in Deutschland. Bonn: BIBB.

Blanchflower, D. and Lynch, L.M. (1994), 'Training at work: a comparison of US and British youths', pp. 233-60 of L.M. Lynch (ed.), *Training and the Private Sector*. Chicago: University of Chicago/NBER.

Boothby, D. and Drewes, T. (2010) 'Returns to Apprenticeship in Canada' Canadian Labour Market and Skills Researcher Network, Working Paper No. 70. CLSRN. Available at: http://www.clsrn.econ.ubc.ca/workingpapers/CLSRN%20Working%20Paper%20no.%2070%20-%20Boothby%20and%20Drewes.pdf.

Canadian Apprenticeship Forum (2008) 'Apprenticeship Labour Market Outcomes in Canada since 2000: An assessment of the Resources for Research'. Available at: http://www.caf-fca.org/en/report/ApprenticeshipLabourMarket.pdf. Accessed on 10 February 2011.

CEREQ (2008) Enquête Generation 2004, CereqBref No. 248, http://www.cereq.fr/pdf/b248.pdf

Dearden, L., Ferri, J. and Meghir, C. (2002) 'The Effect of School Quality on Educational Attainment and Wages',*The Review of Economics and Statistics*, 84(1); 1-20.

Dearden, L., Machin, S. and Reed, H. (1997) 'Intergenerational Mobility in Britain', *The Economic Journal*, 107: 47-66.

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

Dearden L., McIntosh S., Myck M. and Vignoles A. (2001), The Returns to Academic, Vocational and Basic Skills in Britain, DfEE Research Report, London: DfEE.

Dearden L., Reed H. and Van Reenen J. (2000), 'Who Gains when Workers Train? Training and Corporate Productivity in a panel of British industries', Institute of Fiscal Studies, http://www.ifs.org.uk/wps/wp0004.pdf.

Dearden L., Reed H. and Van Reenen, J. (2005), 'The Impact of Training on Productivity and Wages: Evidence from British Panel Data', Centre for Economic Performance Discussion Paper No 674, London School of Economics, http://cep.lse.ac.uk/pubs/download/dp0674.pdf

Department for Education/NatCen (2010), 'LSYPE User Guide to the datasets: Wave One to Wave Six' October 2010 (http://www.esds.ac.uk/doc/5545/mrdoc/pdf/5545lsype_user_guide_wave_1_to_wave_6.pdf).

Desjardins, L. and Paquin, N. (2010).*Registered Apprentices: the Cohorts of 1994 and 1995, One Decade Later*, Statistics Canada Catalogue no. 81-595-M No. 080.

Dickerson A. and Vignoles A.(2007), The distribution and returns to qualifications in the Sector Skills Councils, Sector Skills Development Agency, ISBN: 978-0-9552029-8-8.

Dustmann, C. and Schönberg, U. (2009), 'Training and union wages', *Review of Economics and Statistics*, 91(2), 363-76.

Fitzenberger, B. and Speckesser, S. (2000), 'Zur wissenschaftlichen Evaluation der aktiven Arbeitsmarktpolitik in Deutschland: ein Überblick', Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, Nr. 3.

Fong, B., and Phelps. A. (2008) *Apprenticeship Pay: 2007 Survey of Earnings by Sector*, Department of Innovation Universities and Skills Research Report 08-05.

Fuller, A. andUnwin, L. (1998) 'Reconceptualising Apprenticeship: exploring the relationship between work and learning', *Journal of Vocational Education & Training*, 50(2): 153 – 173.

Fuller, A. and Unwin, L. (2007) 'What Counts as Good Practice in Contemporary Apprenticeships?Evidence from Two Sectors in England', *Education and Training*, 49(2).

Fuller, A. and Unwin, L. (2008) 'Towards Expansive Apprenticeships'. A Commentary for the ESRC's Teaching and Learning Programme. London: Institute of Education.

Gambin L., Hogarth T. and Hasluck C. (2010), 'Maximising Apprenticeship Completion Rates in England', *Canadian Apprenticeship Journal*, No.4.

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

Hansen, K. and Vignoles, A. (2007).The use of large scale data-sets in educational research. London: TLRP. Online at: http://www.bera.ac.uk/the-use-of-large-scale-data-sets-in-educational-research/

Haslam, A. (2004) *Organisational Psychology*: *The Social Identity Approach*, Sage: London.

Hasluck, C., Hogarth, T., Briscoe, G. and Baldauf, B. (2008) 'The Net Benefit of Training Apprentices', London: Apprenticeship Ambassadors Network.

Heckman, J. (1979) 'Sample selection bias as a specification error', *Econometrica*, 47: 153–61

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1998) 'Characterizing selection bias using experimental data' *Econometrica*, 66(5):1017-1098.

Heckman, J.andNavarro-Lazano, S. 2003. 'Using matching, instrumental variables and control functions to estimate economic choice models'. NBER Working Paper No. 9497.

Heckman, J. and Smith, J. (1995) 'Assessing the Case for Social Experiments', *Journal of Economic Perspectives*, 9(2): 85-100.

Herzberg, F., Mausner, B. and Snydermann, B. (1959) *The Motivation to Work*, Wiley, New York.

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

Hogarth, T., and C. Hasluck (2003) 'The Net Costs of Training to Employers', Department for Education and Employment Research Series, Sheffield: DWP.

House of Lords (2007) 'Apprenticeship: a key route to skills: Volume I – Report', House of Lords Select Committee on Economic Affairs, London: TSO.

Imbens, G. and Wooldridge, J. (2009) 'Recent Developments in the Econometrics of Program Evaluation', *Journal of Economic Literature*, 47(1): 5–86

Jenkins *et al*(2004), 'Linking household survey and administrative record data: what should the matching variables be?' ISER Working Paper 2004-23. Colchester: University of Essex.

Jenkins, S.P., Cappellari,L. *et al* (2004) "Patterns of consent: evidence from a general household survey". ISER Working Paper 2004-27. Colchester: University of Essex.

Kewin, J., Hughes,T., Fletcher, T. and Sheen, J. (2011) "The Road Less Travelled: Experiences of employers that support the progression of Advanced Apprentices to higher education" prepared by CFE for the Skills for Sustainable Communities Lifelong Learning Network. Leicester: CFE. Available at: http://www.lifelonglearningnetworks.org.uk/uploads/document/782/the-road-less-travelled.pdf

LaLonde, R. 1986. 'Evaluating the econometric evaluations of training programs with experimental data'.*American Economic Review*, 76: 604-620.

Lee, J-S (2010) 'Returns from education: an occupational status approach' National Centre for Vocational Education Research. Adelaide: NCVER.

Long, D, Mallar, C. and Thornton, C. (1981), 'Evaluating the costs and benefits of the Job Corps', *Journal of Policy Analysis and Management*, 1: 55-76.

Marks, G. (2009) 'The occupations and earnings of young Australians: The role of education and training', Australian Council for Educational Research (ACER), Research report no.55.

McConnell, S. and Glazerman, S. (2001), 'National Job Corps Study: the Benefits and Costs of Job Corps'. Washington DC: Mathematica Policy Research Inc. (http://wdr.doleta.gov/opr/fulltext/01-jcbenefit.pdf).

McIntosh, S. (2004) 'The impact of vocational qualifications on the labour market outcomes of low achieving school leavers', Centre for Economic Performance Discussion Paper 621.

McIntosh, S. (2007) 'A Cost-Benefit Analysis of Apprenticeships and Other Vocational Qualifications', Department for Education and Skills, Research Report 834.

McMillan, J., Rothman,S. and Wernert, N. (2005) 'Non-Apprenticeship VET courses: Participation, persistence and subsequent pathways', Australian Council for Educational Research (ACER), Research report no.47.

Menard, M., Menezes,F., Chan, C.K.Y.and Walker, M. (2008) 'National Apprenticeship Survey: Canada Overview Report 2007' Ottawa: Statistics Canada. Available at: http://www.statcan.gc.ca/pub/81-598-x/81-598-x2008001-eng.pdf.

OECD (2005), Thematic Review of Adult Training: Country Note, Germany. Paris: OECD (http://www.oecd.org/dataoecd/44/5/36341143.pdf).

ONS (2004) 'Protocol on data matching' National Statistics Code of Practice. Available from: http://www.ons.gov.uk/about-statistics/user-guidance/lm-guide/methods/analysis/data-linking/index.html. London: TSO.

Payne, J., Riley, R. and Coleman, N. (2001) 'Feasibility Study for the Long Term Evaluation of Modern Apprenticeships', Department for Education and Skills, Research Report 290.

Pfeiffer, H., Schönfeld, G. and Wenzelmann, F. (2011), 'How large is the firm-specific component of German apprenticeship training?' unpublished paper, BIBB, Bonn.

Skills Commission (2009) *Progression Through Apprenticeships* http://www.policyconnect.org.uk/fckimages/Skills%20Commission%20-%20Progression%20through%20apprenticeships(1).pdf

Rosenbaum, P. and Rubin, D. (1983), 'The central role of the propensity score in observational studies for causal effect', *Biometrika*, 70(1): 41-50.

Ryan, P. (2001), 'The school-to-work transition: a cross-national perspective', *Journal of Economic Literature*, March; 49: 34-92.

Schochet, P.Z., McConnell, S. and Burghardt, J. (2003), National Job Corps Study: Findings Using Administrative Earnings Records Data, Final Report, Executive Summary pp. xvii-xxx. Princeton, NJ: Mathematica Policy Research.(http://www.doleta.gov/reports/searcheta/occ/papers/2005-06_Final_Report.pdf)

Shadish, W. R., Clark, M.H. and Steiner, P. M. (2008), 'Can Nonrandomized Experiments Yield Accurate Answers?A Randomized Experiment Comparing Random to Nonrandom Assignment', *Journal of the American Statistical Association*, 103: 1334-1343.

Steedman, H. (2001), 'Benchmarking Apprentices: UK and Continental Europe Compared', London, UK: Centre for Economic Performance, London School of Economics and Political Sciences.

Stoll, P. and Baignee, A. (1997) 'The National Apprenticed Trades Survey: An Overview of the Apprenticeship Experience', Canada: Human Resources Development Canada.

Sweet, R., Anisef, P. and Lin, Z. (2000), 'Exploring Family Antecedents of Participation in Post-Secondary Education', Ottawa: Learning and Literacy Directorate, Human Resources Development Canada.

Taylor, M.E. (1981), Education and Work in the Federal Republic of Germany. London: Anglo-German Foundation.

UVAC (2005) *An analysis of the progression of advanced apprentices to higher education in England*. Bolton: University Vocational Awards Council.

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.

Zhang, J. L., Rubin, D.B. and Mealli, F. (2008), 'Evaluating the Effects of Job Training Programs on Wages through Principal Stratification', in D.Millimet, J.Smith and E.Vytlacil (eds.) *Modelling and Evaluating Treatment Effects in Econometrics*, Vol. 21 (Amsterdam: Elsevier) 119–147.

Glossary

ALSPAC – Avon Longitudinal Study of Parents and Children

BCS70 – 1970 British Cohort Study

BIBB – Federal Institute for Vocational Education and Training (Germany)

BHPS – British Household Panel Survey

DLHE – Destinations of Leavers from Higher Education

DWP – Department for Work and Pensions

FDS – First Destinations Survey

HESA – Higher Education Statistics Agency

ILR – Individualised Learner Record

IVET – Initial Vocational Education and Training

LFS – Labour Force Survey

LSAY - Longitudinal Surveys of Australian Youth

LSYPE – Longitudinal Study of Young People in England

MA – Modern Apprenticeship

NBD – National Benefits Database

NCDS – National Child Development Study

NCVER - National Centre for Vocational Education Research (Australia)

NPD – National Pupil Database

PSM – Propensity Score Matching

PQS – Prior Qualifications Survey

RAIS - Registered Apprenticeship Information System (Canada)

ULN – Unique Learner Number

WPLS – Work and Pensions Longitudinal Study

YCS – Youth Cohort Study

ANNEX 1: Additional Tables

Table A.1: Apprenticeships starts by level and gender, 2009/10

Gender	Apprenticeship (Level 2)	Advanced Apprenticeship and HLAs (Level 3 and Level 4+)	Total
Female	50.0%	48.6%	49.6%
Male	50.0%	51.4%	50.4%
Total	100.0%	100.0%	100.0%

Source: ILR 2009/10 (BIS)
Notes: 1) Percentages are calculated using pre-rounded data
2) Based upon self-declaration by learner

Table A.2: Apprenticeship starts by level and region, 2009/10

Region	Apprenticeship (Level 2)	Advanced Apprenticeship and HLAs (Level 3 and Level 4+)	Total
East of England	8.6%	8.1%	8.5%
East Midlands	8.8%	8.7%	8.8%
London	7.4%	7.1%	7.3%
North East	6.5%	6.8%	6.6%
North West	16.9%	16.9%	16.9%
South East	14.1%	13.6%	14.0%
South West	12.4%	12.7%	12.5%
West Midlands	11.0%	12.1%	11.3%
Yorkshire and the Humber	13.2%	12.8%	13.1%
Not Known	1.0%	1.0%	1.0%
Total	**100.0%**	**100.0%**	**100.0%**

Source: ILR 2009/10 (BIS)
Notes: 1) Percentages are calculated using pre-rounded data
2) These figures are based on the geographic boundaries for regions as of May 2010.

Table A.3: Apprenticeship starts by level and ethnicity, 2009/10

Ethnicity	Apprenticeship (Level 2)	Advanced Apprenticeship and HLAs (Level 3 and Level 4+)	Total
Asian or Asian British	3.3%	3.3%	3.3%
Black or Black British	2.4%	2.0%	2.3%
Mixed	1.9%	1.7%	1.8%
White	90.8%	91.7%	91.1%
Chinese or Other Ethnic Group	0.6%	0.6%	0.6%
Not Known/Not Provided	1.1%	0.7%	1.0%
Total	**100.0%**	**100.0%**	**100.0%**

Source: ILR 2009/10 (BIS)
Notes: 1) Percentages are calculated using pre-rounded data
2) Based upon self-declaration by learner

Table A.4: Apprenticeship starts by level and LLDD, 2009/10

LLDD	Apprenticeship (Level 2)	Advanced Apprenticeship and HLAs (Level 3 and Level 4+)	Total
Learner Considers Himself Or Herself To Have A Learning Difficulty And/Or Disability And/Or Health Problem	10.3%	7.5%	9.4%
Learner Does Not Consider Himself Or Herself To Have A Learning Difficulty And/Or Disability And/Or Health Problem	88.6%	91.8%	89.6%
No Information Provided By The Learner	1.1%	0.6%	0.9%
Total	**100.0%**	**100.0%**	**100.0%**

Source: ILR 2009/10 (BIS)
Notes: 1) Percentages are calculated using pre-rounded data
2) based upon self-declaration by learner

Table A.5: Level of highest qualification held (LFS)

Level of highest qualification held (levqual8)	%
NQF Level 4 or below	29.3%
NQF Level 3	14.9%
Trade Apprenticeships	4.8%
NQF Level 2	15.9%
Below NQF Level 2	13.1%
Other qualifications	8.6%
No qualifications	13.4%
Total	100.0%

Source: LFS Jan-Mar 2010

Table A.6: Have completed or currently doing recognised Apprenticeship (LFS)

Recognised Apprenticeship	%
Yes, completed	10.2%
Yes, still doing	0.3%
Yes, completed one Apprenticeship and doing further one	0.1%
No	89.4%
Total	100.0%

Source: LFS Jan-Mar 2010

Table A.7: Whether recognised Apprenticeship is part of Modern Apprenticeships (LFS)

Whether recognised Apprenticeship part of Modern Apprenticeships	Completed Apprenticeship	Currently doing Apprenticeship	Completed one, doing further
(Foundation) Apprenticeship	9.7%	25.1%	15.6%
Advanced Apprenticeship	3.4%	11.8%	0
Apprenticeship + Advanced Apprenticeship	1.0%	2.4%	0
Not Modern Apprenticeship	65.6%	11.0%	35.6%
Don't know	20.3%	49.8%	48.9%
Total	100.0%	100.0%	100.0%

Source: LFS Jan-Mar 2010

Table A.8: Age and Year of completion of Apprenticeship (BCS70, 1999/2000 Wave)

Age	Year of Completion	N	Age	Year of Completion	N
16	1986	6	24	1994	10
17	1987	11	25	1995	5
18	1988	55	26	1996	2
19	1989	102	27	1997	5
20	1990	168	28	1998	10
21	1991	72	29	1999	4
22	1992	24	30	2000	2
23	1993	15			
Total with Apprenticeship				**491**	

Source: BCS70, 1999/2000 Wave

Note: Where an individual has reported having obtained more than 1 recognised trade Apprenticeship, the year of completion is shown only for their first Apprenticeship

Table A.9: Main activity of respondents (derived variable) (LSYPE, Wave 6)

Main current activity (derived)	n†	%
Doing an Apprenticeship	549	5.7%
Doing course at University	2,724	28.2%
In education	1,559	16.1%
In paid work	3,150	32.6%
On training course or scheme	100	1.0%
Waiting for course or job to start	313	3.2%
Looking after family and home	202	2.1%
Unemployed, looking for work	918	9.5%
Waiting for exam results or result of job application	14	0.1%
Spending part of week with employer, part at college	81	0.8%
Voluntary work	47	0.5%
Total (validresponses, non-missing)	9,107	100.0%

Source: LSYPE, Wave 6
† data have been weighted by cross-sectional weights

Table A.10: Apprenticeships by level (LSYPE, Wave 6)

Level of Apprenticeship	n†	%
Apprenticeship – Level 2	197	38.2%
Advanced Apprenticeship	320	61.8%

Source: LSYPE, Wave 6
† data have been weighted by cross-sectional weights

Table A.11: Whether or not Apprenticeship was first choice of activity (LSYPE, Wave 6)

Apprenticeship was first choice?	n†	%
Yes – was first choice	322	79.0%
No – was not first choice	86	21.0%

Source: LSYPE, Wave 6
† data have been weighted by cross-sectional weights

Table A.12: Whether respondent considered other things at same time as Apprenticeship (LSYPE, Wave 6)

Alternatives considered	n†	%
Getting a job	256	62.1%
Going to university	91	21.9%
Another Higher Education course	157	38.0%
Retaking GCSEs	56	13.6%
Different Vocational course	82	19.8%
Another training course	97	23.6%
Other	67	16.1%

Source: LSYPE, Wave 6
† data have been weighted by cross-sectional weights

Table A.13: Reasons for doing an Apprenticeship (LSYPE, Wave 6)

Reason for doing Apprenticeship	Strongly agree	Agree	Disagree	Strongly disagree
It allows me to keep my options about the future open	30.3%	63.6%	5.8%	0.3%
I wanted to do something practical rather than academic	48.2%	43.7%	8.1%	0.0%
I liked the idea of getting a job and doing training at the same time	47.1%	52.3%	0.6%	0.0%
It porivdes good pay prospects for the future	34.9%	57.3%	6.8%	1.0%
I have good career prospects on completion	38.8%	59.6%	1.7%	0.0%
It provides qualifications you need to enter certain occupations	39.8%	56.2%	4.0%	0.1%
It is a well recognised qualification	41.9%	55.6%	2.3%	0.2%

Source: LSYPE, Wave 6
† data have been weighted by cross-sectional weights

ANNEX 2: Workshop Attendees

Rod Kenyon	Apprenticeship Ambassadors Network
Jonathan Stewart	UK Department for Business, Innovation and Skills
Simon McKee	UK Department for Business, Innovation and Skills
Vikki McAuley	UK Department for Business, Innovation and Skills
Robert Clogher	Department for Education
Kurt Vogler-Ludwig	Economix Research and Consulting
Chris Hasluck	Hasluck Employment Research
David Owen	University of Warwick Institute for Employment Research
Lynn Gambin	University of Warwick Institute for Employment Research
Peter Elias	University of Warwick Institute for Employment Research
Robert Lindley	University of Warwick Institute for Employment Research
Terence Hogarth	University of Warwick Institute for Employment Research
Paul Ryan	University of Cambridge
Mark Winterbotham	IFF Research
Martin Ward	National Apprenticeship Service
Geoff Mason	National Institute of Social and Economic Research
Stephen Morris	Policy Studies Institute
Stephen McKay	University of Birmingham
Alan Felstead	University of Cardiff
Ewart Keep	University of Cardiff
Stephen McIntosh	University of Sheffield
Alison Fuller	University of Southampton

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

Any enquiries regarding this publication should be sent to:

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

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

URN 11/1312