

**DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INNOVATION  
COMMITTEE ON INDUSTRY, INNOVATION AND ENTREPRENEURSHIP**

DSTI/IND(2016)14  
For Official Use

**OECD EVALUATION FRAMEWORK FOR MODERN APPRENTICESHIPS IN SCOTLAND**

**12-13 October 2016**

*This evaluation framework sets out recommendations for evaluating Modern Apprenticeships (MAs) in Scotland. The report has been prepared by Matej Bajgar and Chiara Criscuolo (STI/SPD). It discusses the evaluation activities to carry out, the outcomes to examine, the data to use and the methods to apply. It also sets the recommendations in a broader context by introducing activities and guiding principles related to evaluating public interventions and by briefly summarising relevant literature. The report describes a long-term evaluation strategy to be followed once required links between administrative datasets are put in place, and it also outlines the steps that should be made already in the short and medium term to facilitate the data linkage and to exploit data that are more readily available. The report emphasises the crucial role of formulating an ex-ante evaluation strategy in enabling high-quality and cost-effective evaluation.*

*Please also see the background paper "Impact of apprenticeships on individuals and firms: Lessons for evaluating Modern Apprenticeships in Scotland " (DSTI/IND(2016)15) which reviews existing studies evaluating the impact of apprenticeships on individuals and firms.*

*This report is a contribution to Output Result 1.2.1.3.1 in the 2015-16 PWB of the CIIE.*

**Contacts:**

Mr. Matej Bajgar, Tel. + (33-1) 45 24 19 76, e-mail: [Matej.bajgar@oecd.org](mailto:Matej.bajgar@oecd.org)

Ms. Chiara Criscuolo Tel. + (33-1) 45 24 79 68, E-mail; [Chiara.Criscuolo@oecd.org](mailto:Chiara.Criscuolo@oecd.org)

**JT03400229**

Complete document available on OLIS in its original format

*This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.*

**OECD EVALUATION FRAMEWORK FOR  
MODERN APPRENTICESHIPS IN SCOTLAND**

## TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS .....	5
EXECUTIVE SUMMARY .....	6
1. PREFACE/INTRODUCTION .....	9
2. MODERN APPRENTICESHIPS IN SCOTLAND: BACKGROUND .....	11
History and policy drivers .....	11
Structure of Modern Apprenticeships .....	12
Stakeholders .....	12
Allocation of apprenticeship places .....	13
Rationale for public funding contribution and the co-ordination role of SDS .....	13
3. EVALUATION AIMS, ACTIVITIES AND GUIDING PRINCIPLES .....	15
Definition and aim .....	15
Evaluation activities .....	15
Monitoring inputs of the intervention .....	15
Monitoring outputs of the intervention .....	15
Evaluation of process and participant satisfaction .....	15
Impact evaluation .....	16
Benefit-cost analysis (BCA) .....	16
Guiding principles .....	16
4. EXISTING LITERATURE .....	18
Summary of literature survey .....	18
Past and Current Monitoring and Evaluation of Modern Apprenticeships .....	19
Modern Apprenticeships Statistics (quarterly, 2008/2009 onwards) .....	19
MA Outcomes Survey (2012, 2016) .....	19
MA Employer Survey (2012, 2015) .....	20
MA Employed Trainees Survey (2014) .....	20
5. SCOPE OF EVALUATION .....	21
Input monitoring .....	21
Output monitoring .....	22
Evaluation of process and participant satisfaction .....	23
Evaluation of impact on apprentices .....	23
Outcome variables .....	23
Time horizon .....	26
Evaluation of impact on employers .....	27
Outcome variables .....	28
Time horizon .....	29
Benefit-cost analysis .....	30
Macroeconomic evaluation .....	30

6. DATA FOR EVALUATION .....	32
Existing data sources.....	32
Data options for evaluating impact of MAs on individuals .....	33
Existing self-standing survey .....	34
Data linking.....	34
Dedicated survey .....	36
Data options for evaluating impact of MAs on employers .....	36
Data linking .....	36
Dedicated survey .....	37
7. IMPACT EVALUATION METHODS.....	38
Policy counterfactuals: alternative states of the world.....	38
Overall impact.....	38
Design .....	38
Evaluating impact on individuals.....	39
Control groups.....	40
Methods for identifying causal effects .....	41
Evaluating impact on firms .....	44
Measuring firms' involvement in Modern Apprenticeships .....	44
Methods for identifying causal effects .....	45
Benefit-cost analysis .....	46
Benefits .....	47
Other benefits .....	48
Costs.....	48
General equilibrium effects .....	48
8. RECOMMENDATIONS FOR EVALUATING MODERN APPRENTICESHIPS .....	50
Long-term evaluation strategy .....	50
Monitoring inputs .....	51
Monitoring outputs.....	51
Evaluation of process and participant satisfaction .....	51
Evaluation of impact on individuals using linked administrative data.....	52
Evaluation of impact on individuals using a well-being survey.....	55
Evaluation of impact on firms.....	56
Benefit-cost analysis .....	57
Medium-term evaluation strategy .....	58
Preparing data linking .....	59
Well-being survey .....	59
Annual Survey of Hours and Earnings.....	60
Annual Population Survey .....	60
9. LEVERAGING EX-ANTE EVALUATION STRATEGY .....	61
Role of ex-ante evaluation strategy.....	61
The case of the Apprenticeship Levy .....	62
RCT for efficient design .....	63
REFERENCES .....	64

## ABBREVIATIONS AND ACRONYMS

ABS	Annual Business Survey
AL	Apprenticeship Levy
APS	Annual Population Survey
ASHE	Annual Survey of Hours and Earnings
BCA	Benefit-cost analysis
CRM	Customer Relationship Monitoring
CSS	Customer Support System
CTS	Corporate Training System
DWP	Department for Work and Pensions
HESA	Higher Education Statistics Agency
IDBR	Inter-Departmental Business Register
LFS	Labour Force Survey
LP	Labour productivity
MA	Modern Apprenticeship
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PAYE	Pay-as-you-earn tax
PIAAC	Survey of Adult Skills
RCT	Randomised control trial
SCQF	Scottish Credit and Qualifications Framework
SDS	Skills Development Scotland
SQA	Scottish Qualifications Authority
SSC/Bs	Sector Skills Councils/Bodies
MFP	Multi-factor productivity
VAT	Value added tax

## EXECUTIVE SUMMARY

This evaluation framework sets out recommendations for evaluating the Modern Apprenticeships programme (MAs) in Scotland, including the evaluation activities to carry out, the outcomes to examine, the data to use and the methods to apply. The report is part of an ongoing effort by the Committee on Industry, Innovation and Entrepreneurship to support rigorous policy evaluation practices across countries. More specifically, the report assists policy makers in Scotland in generating insights about Modern Apprenticeships, while doing so it provides policy makers elsewhere with an example of an evaluation strategy in a concrete policy setting. Thus, the framework contributes to a better understanding of education and training outcomes and to strengthening the links between apprenticeships' and labour market policies, which are at the forefront of the Organisation for Economic Cooperation and Development's (OECD) agenda for inclusive growth.

Evaluation of MAs should comprise (i) input monitoring, (ii) output monitoring, (iii) evaluation of process and participant satisfaction, (iv) evaluation of impact on individuals using administrative data, (v) evaluation of impact on individuals using survey data, (vi) evaluation of impact on employers and (vii) benefit-cost analysis. Each of the first six components is important in its own right and will also provide essential inputs into the benefit-cost analysis.

### *i) Monitoring inputs*

- Collecting comprehensive information on inputs into MAs is necessary for a robust benefit-cost analysis. It should include public funding contributions; Skills Development Scotland (SDS) costs attributable to administering MAs; training contributions paid by employers; the amount of time apprentices spend in training rather than in productive work; and the administrative costs borne by employers.

### *ii) Monitoring outputs*

- SDS currently monitors the number of MA starts, MA leavers, apprentices in training and MA completion rates. Additional outputs which SDS should consider monitoring are apprenticeship duration and form of training (on-the-job vs. off-the-job vs. learning by doing). Reported statistics should be disaggregated by level, framework or framework group and key apprentice characteristics (e.g. gender, age).

### *iii) Evaluation of process and participant satisfaction*

- Process evaluation could be based on a combination of interview- and survey-based evidence. It is advised that MA Employer Survey and MA Outcomes Survey shift their emphasis away from self-reporting of impacts and towards feedback on operational aspects of MAs, information on the form of training and information on administrative costs faced by employers.

*iv) & v) Evaluation of impact on individuals*

- Evaluation of the impact of MAs on individuals should primarily rely on linked administrative and Office of National Statistics (ONS) data and investigate the effect on employment; wages; probability of being unemployed and receiving benefits; career advancement and subsequent qualifications.
- In addition, a periodically conducted well-being survey can be used to investigate the effect of MAs on outcomes not observed in administrative data, in particular job and life satisfaction.
- Evaluation of the impact on individuals should compare individuals who complete MAs with two different control groups: individuals who never started an MA (never-starters); and individuals who started but did not complete an MA (non-completers). Using both control groups will lead to the most robust results. In the baseline specification, the treatment and control group should be restricted to those individuals whose highest attempted qualification, other than MA, is of lower level than the evaluated MA.
- The analysis should be further strengthened by approaches such as using a narrower control group consisting of individuals with characteristics more similar to those of apprentices, analysing changes in outcomes before and after MAs, and using the most appropriate econometric analysis.

*vi) Evaluation of impact on employers*

- Evaluation of the impact of MAs on employers should make use of existing official firm data (e.g. the Inter-Departmental Business Register and the Annual Business Survey) which can be linked to records held by SDS. It should examine the effect of MAs on productivity, as measured by labour productivity or multi-factor productivity.
- Employers' involvement in MAs should be measured by the share of Modern Apprentices in the total employment. It is important to distinguish between the effect of apprentices during their training (which may be negative) and after the training is completed.
- The estimation should be strengthened by comparing firms employing apprentices to other firms with similar observable characteristics, by examining changes over time rather than cross-sectional comparisons and by applying appropriate econometric techniques.

*vii) Benefit-cost analysis*

- Benefit-cost analysis should produce estimates of the value that MAs generate for the Scottish economy each year. On the benefit side, the calculation should include employment and productivity increases due to MAs and, if feasible, indirect benefits of MAs such as reducing skill shortages in the economy and preventing undesirable social outcomes (e.g. criminal behaviour). On the cost side, it should include the direct training costs paid by SDS and employers, the indirect costs due to time spent in training rather than productive work and the administrative costs.

Reliable, detailed and regularly updated data is essential for successful evaluation. Whenever possible, the impact evaluation of MAs should rely on linked administrative data, which offers greater timeliness, frequency and (for at least some variables) accuracy, significantly larger sample size and lower

long-run costs than surveys. Although linking multiple external and internal datasets will not be an easy task, it is likely to bring large long-term benefits for evaluation of MAs and create a powerful tool for evaluation of other public interventions in Scotland. SDS should adjust the collection and management of its internal data so as to facilitate their linking to external administrative datasets.

In the intermediate period before the administrative data link is set up, SDS should take steps to facilitate the data linkage and future evaluation activities. It should also consider undertaking those components of the long-term evaluation strategy which do not depend on the linked administrative data, such as conducting a well-being survey and analysing outcomes observed in existing data held by the Office for National Statistics. Caution is required, however, in the case of wages, as publishing early survey-based impact estimates might overshadow later, more reliable estimates based on administrative data.

SDS should consider using experiments, implemented in collaboration with a small number of large MA employers, to test different designs of MA training (e.g. relative weight of on-the-job vs. off-the-job training), with the aim of improving the design and, ultimately, effectiveness of MAs.

Evaluation of Modern Apprenticeships in Scotland should represent an inherent part of the policy cycle. It should be viewed as a continuous process which is guided by a detailed ex-ante evaluation strategy. The results of the evaluation should be used when discussing possible reforms and changes in the system to allow for a steady improvement of policy making in the area of apprenticeship provision in Scotland.

While the evaluation of Modern Apprenticeships proposed here has the potential to generate large benefits, it will also need to overcome significant challenges. Firstly, each potential source of data has important limitations, whether the data comes from a survey (e.g. limited sample size, inaccuracies, non-response, high costs if newly collected) or administrative sources (limited set of variables, gaps in coverage, significant initial investment to gain access and clean data). Collecting detailed cost data may also be costly. It will also be necessary to address issues such as identifying suitable control groups and limiting the possibility that observed correlation between participating in MAs and examined outcomes is driven by unobserved individual characteristics. Evaluation of the impact on employers may be made particularly challenging by a reduced “signal-to-noise” ratio due to the limited share of apprentices in the total employment of most firms.

While this evaluation framework focuses on the specific case of Modern Apprenticeships in Scotland, the lessons derived in it are relevant for evaluation of policy interventions more generally. Issues discussed in this framework – relative benefits of administrative and survey data, choice of control groups, role of time horizon over which outcomes are examined and many other – are central to evaluation of many types of policy interventions. There are now a number of high-quality documents providing general guidance on evaluation; the hope of this framework is to complement these theoretical accounts by providing a hands-on example of how the principles they outline can be applied in the case of a concrete policy intervention.



## 1. INTRODUCTION

“A strong “skills” agenda will be critical. It must provide actionable answers to the challenges of inequality, productivity, growth, sustainability and well-being, and help to anticipate skills needs for the future. Continuing to build a better understanding of education outcomes at all levels and relying on those assessments for reforms is an important part of this work. We need to pave the way for more effective connections between labour market policies and apprenticeship programmes.”

OECD Secretary-General's Strategic Orientations for 2016 and Beyond (OECD, 2016)

1. This evaluation framework sets out recommendations for evaluating Modern Apprenticeships (MAs) in Scotland, including the evaluation activities to carry out, the outcomes to examine, the data to use and the methods to apply.
2. The evaluation framework aims to assist policy makers with generating insights into the outcomes, benefits and challenges of Modern Apprenticeships in Scotland and possibly also apprenticeships in other countries. By doing so, it responds to the needs for a better understanding of education and training outcomes and for more effective links between apprenticeships and labour market policies. These are at the forefront of the Organisation for Economic Cooperation and Development's (OECD) agenda for inclusive growth, as outlined in the recent strategic document quoted above (OECD, 2016).
3. While this evaluation framework focuses on the specific case of Modern Apprenticeships in Scotland, the lessons derived in it are relevant for evaluation of policy interventions more generally. Issues discussed in this framework – relative benefits of administrative and survey data, choice of control groups, role of time horizon over which outcomes are examined and many other – are central to evaluation of many types of policy interventions. There are now a number of high-quality documents providing general guidance on evaluation; the hope of this framework is to complement these theoretical accounts by providing a hands-on example of how the principles they outline can be applied in the case of a concrete policy intervention.
4. The evaluation framework has been prepared by the OECD in collaboration with Skills Development Scotland (SDS) as part of a work programme on the evaluation of industrial and innovation policies.
5. It is motivated by the strong emphasis that the Scottish Government places on Modern Apprenticeships in its strategies related to economy, skills, employment and inclusive growth,<sup>1</sup> and in particular by the recommendation made by Audit Scotland to Scottish Government, stating that the Scottish Government should “*develop appropriate outcome-based measures to assess the long-term benefits of modern apprenticeships, such as sustainable employment and increased earnings, and publish performance information against these*” and “*review the long-term benefits of modern apprenticeships and*

---

<sup>1</sup> See Scottish Government (2007); Scottish Government (2011); Scottish Government (2015).

*ensure that funding is directed to those [MAs] which demonstrate the best value for individuals, employers and the Scottish economy” (Audit Scotland, 2014).*

6. The aim of the evaluation framework is to map key issues related to evaluating Modern Apprenticeships, present potential approaches to overcoming these issues and suggest which of the available options are most appropriate in the case of Modern Apprenticeships given the availability of data today or in the near future.

7. The framework is structured as follows. Sections 2 and 3 provide background to subsequent analysis by respectively providing an overview of Modern Apprenticeships in Scotland and introducing aims, evaluation activities and guiding principles related to evaluating public interventions. Section 4 summarises implications of relevant existing literature for evaluation of MAs; its first part is based on literature reviews in companion studies by Bajgar and Criscuolo (2016a, 2016b), and its second part focuses on studies and evaluation activities undertaken by SDS. Section 5 **Error! Reference source not found.** outlines the scope of the proposed evaluation. It describes the multiple components that should together form the evaluation framework for MAs, and it discusses which inputs, outputs and outcomes the evaluation should examine and over which time horizon. Section 6 describes relevant existing data and the advantages and disadvantages of the available options for obtaining the data required for the evaluation. Section 7 discusses the methods for estimating the causal effects of MAs on individuals and firms and comparing the overall benefits and costs MAs generate; it discusses the choice of control groups to which individuals and firms participating in MAs could be compared, the approaches that can be used to overcome challenges inherent in evaluating causal effects of an intervention and guidance for benefit-cost analysis of MAs. Building on analysis in Sections 5 to 7, Section 8 lays out recommendations on how MAs should be evaluated. It first describes the long-term strategy for evaluating MAs once a link between records held by SDS and a number of administrative datasets is put in place. Secondly, it outlines the steps that should be made already in the short and medium term to facilitate the data linkage and to explore the effects of MAs using data that are more readily available. Finally, Section 9 emphasises the crucial role of formulating an ex-ante evaluation strategy in enabling high-quality and cost-effective evaluation. For Skills Development Scotland, it provides examples of ex-ante measures that can substantially facilitate future evaluation of MAs, and it illustrates how experiments administered in collaboration with a small number of large employers could produce valuable information for improving the design and structure of MAs.

## 2. MODERN APPRENTICESHIPS IN SCOTLAND: BACKGROUND

8. This section puts the subsequent analysis in a context by briefly presenting background information on Modern Apprenticeships in Scotland. It starts by discussing the history and policy drivers of MAs in Scotland and continues by discussing the way MAs are structured, their key stakeholders and how MA places are allocated. In the end, it discusses several market failures that may justify public policy intervention in the area of apprenticeships.

### History and policy drivers

9. Modern Apprenticeships were introduced in the UK in 1994 as a response to the decline of traditional apprenticeships, and the loss of skills at craft, supervisory and technical level, compared to what was happening among other major economies. They were initially available only in traditional industries such as construction or plumbing, but the range of sectors rapidly increased. MAs in Scotland were originally also subject to an age limit for funding at 24, but the limit was lifted in 2009, although funding is still prioritised for young people. MAs in Scotland have been administered by the newly-founded SDS since April 2008. They have the dual aim<sup>2</sup> to support economic growth – particularly in key sectors, and provide employment opportunities for young people.

10. For apprentices, MAs give an opportunity to train to a nationally recognised standard while in paid employment. Employers, for their part, benefit through developing a fully trained employee contributing to the success of the business. In line with SDS goals, MAs aim to ensure that

- employers are better able to recruit the right people with the right skills at the right time;
- employers have high performing, highly productive, fair and equal workplaces;
- people have the right skills and confidence to secure good work, progress in their careers and achieve their potential; and,
- there is increased equality of opportunity for all.

11. The number of MA starts has seen a rapid increase from around 11 000 in 2008/9 to slightly above 25 000 in recent years. The number of MA starts each year is based on annual targets, first established in 2011/12, and it is set out to gradually increase to 30 000 by 2020.

12. In addition to setting the overall target for MA starts,<sup>3</sup> the Scottish Government puts priority on

- Young people aged 16-24, with a particular focus on 16-19 year olds (majority of MA places should be allocated to 16-24 year olds),

---

<sup>2</sup> For more detail see the [SDS Letter of Guidance 2015/2016](#) and Scottish Government (2015).

<sup>3</sup> The government targets only apply to publicly funded MA places. In addition to publicly funded places, there are a small number of privately funded MA places (to which the targets do not apply).

- MAs in key and supporting sectors (for those aged 25+),
- Higher level frameworks (Level 3 / SCQF 6 and above),
- MAs in science, technology, engineering and math (STEM) occupations, and
- Equality and diversity (targets set for gender, disability and Care Leavers).

### Structure of Modern Apprenticeships

13. Modern Apprenticeships combine paid work with training that aims to develop transferable skills and occupational competence and leads to an industry-recognised qualification.

14. Delivery of training and assessment varies depending on the sector. In some sectors, such as Social Services and Healthcare, it is delivered entirely on the job, with occasional development sessions. In these sectors, the primary role of the training provider is to assess the apprentice's competence in the workplace. In other sectors, formal taught learning in a college or a training centre provides underpinning knowledge, while on the job training helps apprentices to build competence. Such sectors include, for example, Plumbing, Automotive and Land Based Engineering.

15. Modern Apprenticeships typically last between one and four years, but their duration depends on the sector and the occupation, the pace of the learner, amount of off-the-job taught learning and any additional learning required. Some sectors (e.g. Plumbing, Electrical Installation) also set a minimum length of time needed to develop the underpinning knowledge and competence and ability to perform the work safely.

### Stakeholders

16. A number of actors participate in operating Modern Apprenticeships.

- **Scottish Government** sets targets and priorities.
- **Scottish Apprenticeship Advisory Board** makes recommendations on the guiding principles, operational policy, systems and structures supporting MAs.
- **Sector Skills Councils/Bodies (SSC/Bs)** work with industry representatives to develop National Occupational Standards, qualifications and MA frameworks.
- **Awarding bodies** work with SSC/Bs to develop, review and certificate qualifications.
- **SQA Accreditation** approves awarding bodies, their qualifications and the qualification structures developed by SSC/Bs.
- Modern Apprenticeship Group<sup>4</sup> approves MA frameworks.

---

<sup>4</sup> Modern Apprenticeship Group is an independent group consisting of the Scottish Government, SDS, Sector Skills Councils/Bodies, SQA Accreditation, the Scottish Training Federation, the College Development Network and the Scottish Trades Union Congress.

- **Skills Development Scotland** administers MAs (e.g. commissions MA starts, contracts with training providers and employers, promotes MAs, reports on results, and manages Apprenticeships in Scotland vacancy website).
- **Training providers** are contracted by SDS to deliver an approved programme of learning and assessment as required by the employer and the individual; they include private training providers, third sector organisations, local authorities, colleges and employers.
- **Employers** offer employment, provide working and learning environment and finance training costs if these exceed the public contribution; they can be contracted directly without an intermediation of a training provider (large employers).

### Allocation of apprenticeship places

17. MA places are contracted out in an annual cycle. In the first phase SDS estimates demand for various types of MAs based on SSC/Bs' estimates; information from industry,<sup>5</sup> employers and training providers; and government information and policy direction. It then prepares its Contracting Strategy, reflecting Scottish Government's funding priorities, and publishes it on Public Contracts Scotland.

18. Subsequently, training providers and employers (in case of direct employer contracts) submit their proposals. Bids are then evaluated and SDS allocates places based on proposal quality, past performance and policy imperatives (e.g. age, sector).

### Rationale for public funding contribution and the co-ordination role of SDS

19. If left purely to market forces, apprenticeships are prone to be subject to several market failures. These market failures create the need for public intervention in apprenticeship provision.

20. Most importantly, since apprenticeships develop skills which are transferable within a sector, nothing guarantees that an apprentice does not move to a different employer after the end of the training. Consequently, employers taking on (Modern) Apprentices face the risk that they invest in training, but another employer will benefit from its fruit. This could discourage employers from providing apprenticeship opportunities and lead to a socially sub-optimal number of apprenticeship places. This market failure is known in the literature as the "poaching externality",<sup>6</sup> and is not specific to apprenticeships – it also afflicts other forms of training leading to general and transferable skills.

21. The public contribution to the costs of apprenticeship training aims to correct for this market failure by compensating the training employers for the "positive externality" they generate. The public funding is a contribution towards the cost of training and/or assessment, and it is paid directly to the contracted training provider (or employer in case of a direct employer contract). The contribution should be commensurate with the size of the externality, and it will, thus, vary with the level of training provided and the industry supported. It also varies by age group, level and type of framework.

22. In addition to the "poaching" externality, apprenticeships are characterised by significant commitment problems and asymmetric information. First, individuals will only want to enter apprenticeships if they know they will obtain high-quality training, but once they enter, it is in their employers' interests to save on the training costs. Second, if generally recognised qualifications are not awarded, apprentices will not be able to provide proof of the skills they acquired to other potential

<sup>5</sup> Information from industry is gathered through Regional Skills Assessments and Skills Investment Plans.

<sup>6</sup> See, for example, Stevens (1996); Katz and Ziderman (1990); and Brunello and De Paola (2004).

employers and their incentives to enter apprenticeships will, thus, be significantly reduced. Both of these issues call for public involvement to provide independent control of training quality and, by doing so, enable MAs and their certifications to serve as a signal of a certain well-defined set of transferable skills.

23. Finally, firms may be unwilling to employ young workers with little work experience. The fact that MA funding prioritises young people helps to overcome this issue by partly compensating firms for providing the young workers with initial work experience.

### 3. EVALUATION AIMS, ACTIVITIES AND GUIDING PRINCIPLES

24. This section provides an introduction to public policy evaluation. It provides a definition and aim of evaluation; introduces the main components, or activities, of which evaluation consists; suggests a number of principles that should guide high-quality evaluation and lists a few information sources providing further guidance on evaluating public policies.

#### **Definition and aim**

25. Evaluation can be defined as an objective process of understanding how a policy or other intervention was implemented, what effects it had, for whom, how and why (HM Treasury, 2011). Evaluations aim to inform policy making by producing evidence suggesting whether particular interventions should be extended, limited or abandoned, and how they could be improved.

#### **Evaluation activities**

26. Understanding the full benefits and costs of the evaluated intervention requires combining five different evaluation activities. When considering these activities, it is important to emphasize that outputs of an intervention (such as the number of individuals trained) are not equivalent to its impact (i.e. its effects on the outcomes of interest, such as wages or probability of employment). While monitoring outputs is clearly important, outcomes are what should matter most for policy decisions in medium and long term.

#### ***Monitoring inputs of the intervention***

27. In order to understand how an intervention works and to be able to eventually compare its benefits to its costs, it is important to keep detailed records of the resources that flow into the intervention. These include both the direct expenditure on the intervention, such as training fees paid to training providers, the costs of administering the intervention and the costs for both employers and apprentices to participate in MAs (e.g. opportunity costs; administrative costs<sup>7</sup>; reporting etc.).

#### ***Monitoring outputs of the intervention***

28. Monitoring outputs of an intervention involves measuring the quantity and characteristics of the services or products delivered by the intervention. Obtaining reliable information about outputs of an intervention is an important first step towards understanding how the intervention operates and what its effects are. As emphasized above, outputs of an intervention are not equivalent to its impact.

#### ***Evaluation of process and participant satisfaction***

29. This type of evaluation consists of monitoring the way an intervention is administered and the satisfaction of participating individuals and organisations. It helps to ensure that the intervention efficiently

---

<sup>7</sup> These signify administrative costs (in terms of time or money) incurred by firms participating in Modern Apprenticeships, e.g. due to filling in paper and online forms or extra management and supervision needed when employing apprentices.

uses inputs in the production of outputs and that the outputs are aligned with the desired outcomes. Process evaluation can rely on surveys among employers and former apprentices, and it should also draw on interviews with all the relevant stakeholders.

### *Impact evaluation*

30. Quantitative impact evaluation measures the changes in outcomes of interest that are attributable to the causal impact of an intervention. This requires comparing the outcomes with the intervention to what would have happened without any intervention or with alternative interventions. Impact evaluation addresses the central questions of what the effects of an intervention are, how effects of different types or components of the intervention compare and which mechanisms drive the effects.

### *Benefit-cost analysis (BCA)*

31. The purpose of benefit-cost analysis is to establish, in monetary terms, the net value that an intervention creates after subtracting its costs from its benefits. Such knowledge helps policy makers to determine which interventions bring the highest return to investment and, thus, represent the best use of tax-payers' money. Benefit-cost analysis builds on the results of the other evaluation activities, in particular input monitoring, output monitoring and impact evaluation. It also requires several additional assumptions and analysis. These may include, in particular, assumptions underlying methodologies that convert non-market outcomes (e.g. well-being or health) into monetary terms and choice of the discount rate for long-term benefits of MAs.

### **Guiding principles**

32. It is possible to identify several principles that the evaluation of Modern Apprenticeships should follow.

- **Evaluation as a process.** Evaluation should not be seen as an event which takes place at one point in time. Instead, it should be viewed as a continuous process which starts with a detailed ex-ante evaluation strategy and where process and impact evaluation informs policy and design changes, impact of which is then again evaluated.
- **Use of appropriate methods.** The evaluation should apply measurement and estimation methods that are most appropriate to the questions posed by the evaluation given the nature of Modern Apprenticeships and the available data.
- **Resource adequacy.** The evaluation should be supported by resources which are commensurate with the potential gains from the evaluation (e.g. due to redesigning, expanding or reducing certain MA frameworks as a result of the evaluation).
- **Openness.** Results of the evaluation should be accessible by policy makers in other parts of the public administration, academics and the wider public.
- **Transparency.** A detailed and clear description of the evaluation methodology should be made publicly accessible and communicated together with the evaluation results.
- **Sustainability.** The nature of training contained in Modern Apprenticeships, the set of people undertaking the apprenticeships and requirements in the labour market are continuously changing. Some or all components of the evaluation framework should, thus, allow being



repeated at different points in time and, by doing so, lead to information about how the effects of Modern Apprenticeships evolve over time.

- **Exploiting synergies with other evaluations.** The evaluation should exploit potential synergies with other present or future evaluations of other interventions related to education or labour markets, especially in terms of producing, linking and cleaning data and communicating results.

### Sources of information on good practice

33. A number of publications and information sources provide useful information about principles and methods that are useful when evaluating interventions. These include, among others, the following:

- HM Treasury: [The Magenta Book](#)
- European Commission: [Counterfactual Impact Evaluation](#)
- [OECD: Evaluation of Industrial Policy](#)
- World Bank: [Impact Evaluation in Practice](#)
- Better Evaluation: [Better Evaluation website](#)
- Saskatchewan Ministry of Education: [Excerpts from Review of Evaluation Frameworks](#)

#### 4. EXISTING LITERATURE

34. This section provides an overview of relevant literature. It first discusses main lessons from a literature survey carried out to support this evaluation framework, and it then briefly describes the main evaluation-related studies and activities undertaken by SDS in the past.

##### **Summary of literature survey**

35. Two studies reviewing literature evaluating impact of apprenticeships accompany this report. Bajgar and Criscuolo (2016a) provides a detailed, technical review of studies that estimate the impact of apprenticeships on individuals and firms. It discusses their methodology, counterfactual, data and results. It also provides an overview of methodology and results of studies evaluating other interventions which in some ways resemble apprenticeships – education, active labour market programmes and private on-the-job training. The literature summary by Bajgar and Criscuolo (2016b) summarizes the previous literature review in a more condensed form, focusing on implications of the existing studies for the evaluation framework for Modern Apprenticeships in Scotland. The main lessons emerging from the review of the literature are summarised below.

- Evaluating apprenticeships is a complex task that requires addressing a number of issues. Some are technical in nature, but others are ultimately related to a political decision about what the aims of the evaluation are.
- The effects of apprenticeships can change substantially over the time passed since the end of the training period. It is, therefore, important to analyse not only the short-term effects of Modern Apprenticeships but also their medium- and long-term impacts.
- Rigorous evaluation depends on the availability of suitable data that can be accessed, are of good quality and offer good coverage across individuals and over time. Continued work at creating, maintaining and linking datasets is, thus, essential for evaluation.
- To the extent that the choice of measured outcomes on individuals aims at reflecting the emphasis in the existing literature, an evaluation framework for Modern Apprenticeships should primarily focus on wages and the probability of being employed. Similarly, as secondary outcomes, it should also consider subsequent education and training and the likelihood of finding a standard, as opposed to non-standard, employment.
- To the extent that the choice of measured outcomes goes beyond the emphasis in the existing literature, the outcomes should include productivity. The impact on productivity can be either derived from the estimated impact on individual wages or directly estimated using firm-level data. If feasible, the latter approach is preferable but challenging, as witnessed by the scarcity of existing studies estimating the impact of apprenticeships on firm productivity.
- Evaluations that take into account only observed individual characteristics may incorrectly ascribe the effect of unobserved characteristics, such as ability or motivation, to the effect of an apprenticeship. It is, therefore, important to also employ methods that try to control for

unobserved individual characteristics, although such methods are sometimes context-specific and are subject to data availability.

- It is important to clearly define the counterfactual relative to which the impact of Modern Apprenticeship is evaluated. The potential comparison groups include (i) all individuals who do not do an apprenticeship, (ii) more specific groups of individuals, such as individuals who, instead of an apprenticeship, obtain non work-based vocational training at a college and (iii) individuals who start an apprenticeship but do not complete it. The evidence coming from comparing apprentices to the different “control groups” might provide different and complementary answers.

### **Past and Current Monitoring and Evaluation of Modern Apprenticeships**

36. The studies and activities related to monitoring and evaluation of Modern Apprenticeships that SDS has carried out in the past include, in particular, MA Statistics, MA Outcomes Survey, MA Employer Survey and MA Employed Trainees Survey.

#### ***Modern Apprenticeships Statistics (quarterly, 2008/2009 onwards)***

37. Its Corporate Training System (CTS) provides SDS with information on all Modern Apprentices, including information on demographic characteristics, dates of starting and leaving apprenticeship, framework and achievement. Since 2008/2009, SDS has published summary statistics based on CTS on a quarterly and annual basis and made them openly accessible on its website. SDS statistics are prepared in line with the principles set out in the Code of Practice for Official Statistics, and SDS is currently undergoing a review of their standard statistical processes with a view to becoming official statistics producers.

#### ***MA Outcomes Survey (2012, 2016)***

38. MA Outcomes Survey is a telephone survey of 2000 Modern Apprentices in Scotland who had left their apprenticeship 6 months before the survey. It was first conducted in the latter part of 2012 (Skills Development Scotland, 2013), and a new wave was collected in the first half of 2016, with analysis of the survey responses currently under way, and results forthcoming.

39. The survey aims to cover apprentices of all levels and ages, and it also includes non-completers, who represented about 17% of the sample in 2012. The areas covered include

- qualifications and employment prior to entering MA;
- satisfaction with MA;
- labour market status 6 months after MA;
- career progression after MA;
- reasons for leaving MA (non-completers);
- self-reported impact of MA on skills and abilities and on personal development; and
- self-reported utilisation of skills acquired during MA in current employment.

***MA Employer Survey (2012, 2015)***

40. MA Employer Survey is based on a telephone survey of approximately 2500 employers who had at least one employee leave an MA in the three years prior to the survey. It was conducted in 2012 and again in 2015 (Skills Development Scotland 2013b; Skills Development Scotland 2015). In 2012, the telephone survey results were complemented with information from the UK Commission for Employment and Skills Employer Perspectives Survey (Shury et al., 2012) covering views of employers who do not participate in MAs. The areas covered include

- employer profiles;
- involvement with MAs;
- motivation for participation;
- satisfaction with MAs;
- provision of on-the-job training;
- perceived benefits of offering MAs;
- reasons for non-completion; and
- future plans regarding MAs.

***MA Employed Trainees Survey (2014)***

41. MA Employed Trainees Survey, conducted in 2014, is based on an on-line survey of approximately 1500 Modern Apprentices currently in training. The areas covered include

- motivation for starting an MA;
- perceptions of the programme
- wage rates and hours; and
- expected destinations.

## 5. SCOPE OF EVALUATION

42. This section discusses what the evaluation activities outlined in Section 0 should cover in the context of Modern Apprenticeships. As apprenticeships have impacts primarily on apprentices and on their employers, the section separately discusses the scope of evaluating the impact on apprentices and on host firms. In the end, it also briefly discusses the possibility of incorporating the microeconomic evaluation results in a macroeconomic model of the Scottish economy.

### Input monitoring

43. Comprehensive monitoring of resources invested into MAs by the government, employers and apprentices is an essential ingredient in a benefit-cost analysis of Modern Apprenticeships. Multiple types of information are needed. To allow performing the benefit-cost analysis separately for each MA level and framework group, the information should be collected separately for each MA type, to the extent that cost considerations allow it.

- **Public funding contribution** paid by SDS to training providers forms part of the direct costs of the training. This information is available and held by SDS.
- **Employers' contribution** towards the costs of training represents the other part of the direct training costs. Knowing the employers' contribution is valuable not only for evaluation purposes, but also for setting future public contribution rates.<sup>8</sup> In 2015/2016, SDS asked training providers if they requested training contribution from employers. The request was non-mandatory, and the information was filled for about two thirds of all "training provider-occupational grouping" combinations. Employers' contribution was asked in only about one in five training provider-occupational grouping combinations for which the information was provided. Obtaining reliable information about employers' contributions is important for enabling future benefit-cost analysis of Modern Apprenticeships as a whole and, in particular, of individual MA framework groups.
- **Reduced time spent in productive work** compared to regular employment, which is due to the training component of MAs, represents the indirect cost of the apprenticeships. In principle, this information can be either calculated from information on the time apprentices spend in productive work – by comparing it with hours worked by regular employees in similar positions – or obtained from information on the time apprentices spend in on-the-job and off-the-job training – by assuming that the combined time apprentices spend in productive work *and* training is the same as the time worked by regular employees in similar positions. However, a complete decomposition into time spent in (i) productive work, (ii) on-the-job training and (iii) off-the-job training would provide the most reliable basis for estimating the indirect costs of output forgone due to time spent in training. The MA Employer Survey 2012 recorded whether employers provided apprentices with on-the-job

---

<sup>8</sup> In England, evidence found that employers' contributions to the direct costs of further education were limited and lower than was expected when the policy was designed (Banks et al., 2010; Department for Business, Innovation and Skills, 2014).

and off-the-job training but did not ask about the number of hours. The MA Employer Survey 2015 collected information on the number of hours spent in “formal on-the-job training”<sup>9</sup> but no information related to off-the-job training. Neither wave of the MA Employer Survey asked directly about the amount of time apprentices spent in productive work,<sup>10</sup> and the MA Outcomes survey does not feature any information on time allocation among the various components of MAs. It would be useful to collect information on the time spent in productive work, on-the-job training and off-the-job training in the future rounds of one or both of the surveys, although collection of this information may be complicated by its dispersion across employer and training provider and across multiple members of employer’s staff, and it may be costly to collect this information from a sufficiently large sample of employers. In the absence of reliable survey data, an alternative approach would be to use the number of training hours proposed in the description of each MA framework.

- **Administrative costs of MAs incurred by SDS** (and potentially by other public agencies) also need to be included in the cost calculation. Information on these costs for SDS is available in its accounts and published in its annual reports and financial statements, although certain assumptions may need to be made to apportion certain components of SDS’ costs between Modern Apprenticeships and other activities in the SDS portfolio.
- **Administrative costs of MAs borne by employers** should also be taken into account. Currently, no systematic information on such costs exists. It would be beneficial to collect it in future wave(s) of the MA Employer Survey.

### Output monitoring

44. Several key outputs should be monitored as part of the evaluation of Modern Apprenticeships. To allow tracking whether the outputs meet objectives set by the Scottish Government, they need to be measured separately by age, apprenticeship level and MA framework, although the actual granularity of the measurement will need to take into account the cost implications of producing robust statistics with increasingly detailed decompositions. The outputs are the following.

- **Number of MA starts** is the output used for setting policy targets by the Scottish Government. It can be readily calculated from CTS records held by SDS and is regularly published in the Quarterly/Annual Modern Apprenticeship Statistics.
- **Completion rates** are an important output to track, as high non-completion could signal mismatch between, on one side, the content of MAs and the employment characteristics during MAs, and, on the other side, the expectations, needs or ability of apprentices and employers. They, too, can be readily calculated from CTS records held by SDS and are regularly published in the Modern Apprenticeship Statistics.
- **Apprenticeship duration** is another important indicator to be monitored, as high prevalence of very short or long actual duration relative to the expected duration for the given MA

---

<sup>9</sup> The survey describes “formal on-the-job training” as activities that staff would recognise as training and not the sort of learning by experience which could take place all the time.

<sup>10</sup> For example, Dionisius et al. (2009) use survey data on German and Swiss apprenticeships that contains such information.

framework can signal issues to be addressed.<sup>11</sup> The distribution of apprenticeship duration can be obtained from CTS records held by SDS, but it depends on accurate recording by training providers and it is not currently published in the Modern Apprenticeship Statistics.

- **Composition of MA training** in terms of off-the-job training, on-the-job training and learning by doing during productive work is an important indicator for understanding the actual structure of MA training, and an important factor for understanding potential heterogeneity of impacts across different framework groups. As described above in the discussion on measuring the reduced time in productive work due to training, time spent in on-the-job training is recorded in MA Employer Survey 2015, but time spent in off-the-job training and productive work is currently not available. As suggested above, it would be useful to include this information in future rounds of the MA Employer Survey, MA Outcomes Survey or both these surveys, although the challenges discussed in the section on input monitoring (dispersion of information across multiple people and organisations, cost considerations) also apply here.

### **Evaluation of process and participant satisfaction**

45. The MA Outcomes Survey and MA Employer Survey are the key formal sources of information on how efficiently and smoothly MAs are administered by SDS, as seen, respectively, from the perspective of apprentices and employers. In the absence of other impact evaluation, the past waves of the surveys have attempted to inquire into the impact of MAs, largely through self-reported impacts by apprentices and employers. As impact evaluation should, ideally, be undertaken using existing administrative data, the surveys could instead be used to understand how administration (by SDS and contractors) and content of MAs could be improved and to collect information on the form of MA training (e.g. on-the-job vs. off-the-job vs. learning by doing). As discussed above, the level of decomposition at which this information is collected will need to reflect cost implications of producing robust statistics at higher levels of decomposition.

46. This information should be suitably complemented by qualitative information collected from all stakeholders involved in the administration of MAs (e.g. SDS, SAAB, SSC/Bs).

### **Evaluation of impact on apprentices**

47. Since “providing opportunities for young people” is one of the principal aims of Modern Apprenticeships, examining their impact on individuals should be at the core of the evaluation framework.

48. This subsection discusses outcomes that could potentially be examined when evaluation the impact of MAs on individuals and the time horizon over which the outcomes should be analysed.

### ***Outcome variables***

49. Part of the purpose of Modern Apprenticeships is to equip apprentices with skills and develop competences that will lead to sustained and productive employment. Measuring the impact of MAs on outcomes that capture such skills or employment should, therefore, be at the heart of the evaluation framework. The following paragraphs present such potential outcomes, discussing for each outcome its rationale and suitability for evaluation. The discussion moves from the outcomes which are most closely

---

<sup>11</sup> The expected length of MAs varies greatly, with some frameworks (e.g. in engineering) taking much longer to complete than others (e.g. social care). The expected end date may also depend on specific individual and firm circumstances.

related to the ultimate goals of MAs towards those which are somewhat less directly related but may be easier to measure and use in estimation, and which are likely to be correlated to a lesser or greater extent, with the ultimate goals of the MAs.

- **Skills.** The purpose of Modern Apprenticeships, as of any other training, is to acquire skills. It is, therefore, important, to understand the impact of MAs on skills. Directly evaluating the impact on skills would also have the benefit that it would allow separating the effect of general skills (cognitive and non-cognitive) acquired through MAs from the effect of the positive signal about ability and motivation that an individual sends to prospective employers by completing an MA. Directly measuring skills is, however, challenging, for several reasons. Firstly, skills are not observed in administrative data. Consequently, their measurement would require collecting new data, which is costly and methodologically challenging, although the data collection could build on the methodology and infrastructure of existing surveys, such as the OECD Survey of Adult Skills (PIAAC).<sup>12</sup> Secondly, the content of Modern Apprenticeships strongly varies across frameworks. The set of measured skills would, therefore, either have to be framework-specific, limiting the comparability of results across frameworks, or include only rather general skills, overlooking a large part of the occupation-specific learning that takes place during apprenticeships. For these reasons, skills should be highlighted as a desirable but challenging outcome of interest for a sustained, repeated evaluation of Modern Apprenticeships. In any case, a periodic exercise comparing the skills of Modern Apprentices and a control group before and after an apprenticeship would provide useful insights into, in particular, the more general (cognitive and non-cognitive) skills acquired through MAs. Such evidence would be more generalizable, internationally comparable and externally valid than the information implicit in completing a particular vocational qualification, and it would be a valuable complement to the core evaluation activities. Such exercise could reduce its costs and improve data quality and comparability by building on existing methodologies, such as PIAAC.
- **Productivity.** From a societal perspective, a fundamental aim of the apprenticeships is to increase the productivity of workers. Unfortunately, measuring productivity of individual workers in a comparable way is typically possible only within a specific set of workers undertaking a clearly defined and well-measurable task. Furthermore, the required information may exist in records of individual employers, but no comparable information of this kind exists for a large, representative set of employers. Direct measurement of worker-level productivity for evaluation of Modern Apprenticeships would consequently be very challenging. Instead (as discussed in more detail below), productivity can be proxied by wages or measured with employer-level data.
- **Probability of being employed.** The fact that a person is employed means there is demand for the person's skills. Employment is also an important source of income, and it has a strong effect on individual well-being above and beyond the effect of the increased income. It is relatively straightforward to measure whether an individual is employed, and employment status can be obtained from administrative datasets. For these reasons, probability of being employed is one of two outcomes most commonly analysed by studies evaluating the impact of apprenticeships (see Section 4), and it should represent a core part of the evaluation framework for Modern Apprenticeships. However, it needs to be noted that if employment

---

<sup>12</sup>

The OECD Survey of Adult Skills (PIAAC) measures adults' proficiency in key information-processing skills – literacy, numeracy and problem solving in technology-rich environments – and gathers information and data on how adults use their skills at home, at work and in the wider community. For more information, see <http://www.oecd.org/skills/piaac>.



rates of both apprenticeship completers and individuals with similar ability and other characteristics who did not complete an apprenticeship are very high (close to 1), there may be no observed effect of MAs on the probability of employment, but this should then not be interpreted as a sign of a low demand for the skills of (former) apprentices. Instead, the analysis of the impact on wages will be more important informative in such case.

- **Wage.** Expectations of higher future wages are clearly an important factor that motivates individuals to enter apprenticeships. Furthermore, wages are also important from a social perspective as a proxy for individual productivity. Under certain assumptions, the proportional increase in individual wages should equal the proportional increase in the value added generated by each worker. Such assumptions are rather strong and unlikely to strictly hold in practice. But more generally, it is likely that increases in individual productivity will be reflected in wages, although to an unknown extent. Together with the probability of being employed, wages are among the two outcomes that have been most studied by evaluations of apprenticeships (see Section 4), and they should represent an important component of the evaluation framework for Modern Apprenticeships.
- **Probability of being unemployed, proxied by the probability of receiving benefits.** Being unemployed is another measure of a demand for a person's skills. It is closely related to being employed, but an important difference is that individuals who are not observed as employed may be self-employed or economically inactive. Unemployment can be measured from administrative data on receiving related benefits. In addition, by examining multiple benefits (e.g. Jobseeker's Allowance, Incapacity Benefit, Income Support), the analysis may paint a richer picture than simply observing that certain individuals are not employed. For these reasons, it would be valuable to examine the probability of being unemployed and receiving benefits as additional outcomes for evaluation of MAs.
- **Standard/non-standard employment.** Although multiple changes of occupation or employer need not imply a negative outcome, most workers are likely to prefer a permanent contract to a series of short-term contracts.<sup>13</sup> One option is to observe the type of contract directly, although this would require survey data. Alternatively, the type of contract can be proxied by observed employment duration. Longer employment duration would additionally indicate a better match between an employer and an employee. Overall, the type of contract, observed directly or proxied by employment duration, is likely to represent a useful additional outcome variable for evaluation of Modern Apprenticeships, especially if the potential effect of MAs on employment in general is limited by a relatively high baseline probability of being employed among individuals who tend to participate in MAs.
- **Career progression.** It would also be valuable to understand if Modern Apprenticeships help individuals to advance in their career through, for example, obtaining a promotion or being charged with tasks requiring higher responsibilities. Such measures are not observed in administrative data, but they can be observed in existing or new surveys. While wages, and in particular their growth over time, can be interpreted as a proxy for career advancement, analysis of information on promotions or responsibility attached to jobs would paint a richer picture of the impact of MAs on young people's careers. For this reason, direct measures of career progression should be considered for the evaluation framework.

---

<sup>13</sup>

Short term contracts may be a norm in certain industries (e.g. creative industries) and there has been a general structural shift away from a "job for life". However, evaluation using this outcome would not be affected by these factors, because sectoral characteristics and structural changes over time would be controlled for in a regression framework.

50. In addition to the above outcomes, which are directly related to skill acquisition or employment, several other important individual-level outcomes could be affected by participating in Modern Apprenticeships.

- **Subsequent education.** The learning through a combination of work and training that Modern Apprenticeships offer could motivate apprentices to continue with higher or further education – including another MA – even after they complete their apprenticeship. It would be valuable to analyse how MAs influence subsequent education decisions, in particular for lower-level MAs and for individuals who left school at the age of 16. The required information can, in principle, be obtained from administrative records on education achievements in further and higher education; although experience in England indicates linking to higher education records may be problematic. In addition to the intrinsic interest in the effect of MAs on subsequent education, collecting such information is also essential for estimating the effects of MAs on other outcomes, as it will be needed for separating the direct effects of MAs from the effects of other subsequent qualifications. Given that the relevant data should in any case be obtained, analysing subsequent education as an outcome of Modern Apprenticeships is likely to provide interesting information for little additional costs, and it should be included in the evaluation framework.
- **Subjective well-being.** Measures of subjective well-being, such as self-reported satisfaction with job, or with life in general, are gaining an increasing relevance in policy debates. Measuring the effect of completing a Modern Apprenticeship on subjective well-being is interesting and relevant, but it would need to be collected from a survey. Furthermore, such survey would need to take into account the difficulties inherent in measuring subjective well-being and allow for a sufficient sample size. And finally, it would be necessary to have such measures prior to participation in a MA since there may be correlation between self-reported life satisfaction and motivation to undertake a MA.
- **Undesirable outcomes.** Through provision of skills, motivation and improvement of career prospects, Modern Apprenticeships could reduce the risk of undesirable outcomes such as criminal behaviour, drug abuse or teenage pregnancy. As reducing such outcomes is an important policy goal, measuring how they are affected by apprenticeships is, in principle, desirable. Unfortunately, as these outcomes tend to be relatively rare, collecting the necessary information in a survey would be costly because it would require a very large sample size or an oversampling of high-risk individuals. At the same time, while some of these outcomes are recorded in administrative data, their sensitive nature makes such data hard to access. Despite these caveats, including some of these outcomes in a potential survey (e.g. parenthood) or matching relevant administrative records (e.g. criminal records), if possible, would provide valuable additional information on potential social benefits of Modern Apprenticeships.

### *Time horizon*

51. In addition to the choice of outcome variables, the choice of the time between the end of an apprenticeship and the time(s) when the outcomes are measured is a key decision in setting up the evaluation framework. Several considerations need to be taken into account.

52. Firstly, the summary of related literature in Section 4 highlights that the estimated effects of an apprenticeship change with time passed since the end of the apprenticeship. In particular, the effect on employment tends to be substantial shortly after training but fades out over the first few years, while the effect on wages may grow over time. Not only the size of the effects, but also the mechanisms behind them, may vary between different time horizons. In particular, potential high probability of employment

immediately after the end of an apprenticeship is likely to be driven by apprentices staying in the firms where they were trained, while employment effects later on would, to a larger extent, reflect acquired skills.

53. Secondly, evaluating the long-term impact of MAs faces an important trade-off. On the one hand, long-term effects can strongly influence potential benefit-cost calculations. On the other hand, as time passes, new factors influence the measured outcomes and it becomes increasingly difficult to precisely distinguish the role of the apprenticeship from all other factors, especially if sample size is not huge.

54. Thirdly, the time horizons over which effects of Modern Apprenticeships can be estimated depend on the data used. Administrative datasets have the advantage of being more flexible, allowing evaluation at multiple different time horizons, and their large size makes it possible to more accurately estimate effects even a long time after an apprenticeship. On the contrary, a survey conducted a long time after the end of the apprenticeships may encounter significant challenges when trying to contact individuals.

55. Overall, these factors suggest that, rather than measuring the effects of Modern Apprenticeships at any one time, the effects should be measured at multiple time periods. Such approach will provide a more accurate quantitative assessment of the impacts and a better understanding of the mechanisms behind them. It is particularly suitable when large administrative datasets are used.

56. With survey data, it may be necessary to determine a particular period at which to measure the outcomes. 2-4 years after the end of the apprenticeship seems like an appropriate time horizon. It allows for a sufficient distance from the end of the apprenticeships and (at the longer end) makes it possible to observe labour market outcomes also for individuals who undertake additional training after the end of apprenticeship (e.g. Higher National Certificate in engineering), and, at the same time, it does not make the link between MAs and the outcomes too loose to be estimated with reasonable precision.

### **Evaluation of impact on employers**

57. Alongside providing opportunities for young people, “supporting economic growth” is another principal aim of Modern Apprenticeships. Evaluating their impact on employers, and especially on employers’ productivity, is the most direct way of examining to what extent MAs are fulfilling this aim. For this reason, evaluation of the impact on employers should form another core part of the evaluation framework for MAs.

58. This subsection discusses outcomes that could potentially be examined when evaluating the impact of MAs on firms and the time horizon over which the outcomes should be analysed.

59. As a word of caution, it needs to be mentioned that estimation of the effect of MAs on employers faces an important challenge due to the fact that, especially for large employers, apprentices represent only a small fraction of the workforce. This means that even if the effect of training Modern Apprentices is positive and large *relative to the number of apprentices trained*, this effect may be hard to estimate with sufficient precision, given all the other factors that affect firm performance. The difficulty of directly estimating the impact of apprenticeships on firm performance is also witnessed by the fact that only one study has done so in the past.<sup>14</sup> The existing challenges should not discourage the analysis of firm-level effects of MAs, but they should be kept in mind, as the precision of the firm-level analysis will be an important factor in determining, for instance, if productivity estimates to be included in benefit-cost analysis should be based on individual-level wage estimation or firm-level productivity estimation.

---

<sup>14</sup> See survey by Bajgar and Criscuolo (2016b) for more detail.

### *Outcome variables*

60. Evaluating the impact of Modern Apprenticeships on productivity is crucial for understanding how they contribute to the economy. The previous subsection explains that when examining the impact on individuals, estimating the effect on productivity is difficult, and, consequently, it is necessary to rely on wages as a proxy for productivity. Estimating the effect of MAs on firms would offer an opportunity to estimate the productivity outcomes directly. Relevance of such analysis is further supported by the fact that, from survey evidence,<sup>15</sup> productivity is among the most commonly cited benefits of Modern Apprenticeship for firms. Three potential outcome variables should be considered.

- **Labour productivity (LP)** reflects the output a firm can produce with a given amount of labour. It is calculated as sales or value added divided by the number of employees, where value added is obtained preferably from sales by subtracting the costs of production inputs (but not of labour and capital) but can also be closely approximated by summing pre-tax profits and total labour costs. Use of value added is preferable but also more demanding in terms of data. Information to calculate labour productivity can come from existing firm-level data, such as the business register (only turnover based measures and sometimes based on imputed values especially for smaller businesses) or enterprise surveys (such as ABS). Alternatively, the information could be gathered through a newly collected survey, but that would be costly and methodologically challenging.
- **Multi-factor productivity (MFP)** captures the output a firm can produce with a given amount of inputs (e.g. labour, capital, total intermediate goods and services and energy). Unlike LP, MFP takes into account also the amount of physical capital and intermediates used by each firm. There are multiple methods to estimate or calculate it using information on sales or value added, employment, physical capital and intermediates. Examining the impact on MFP rather than LP, therefore, better accounts for the possibility that employing apprentices goes hand in hand with acquiring more physical capital. In such case, the increased labour productivity due to the increased capital intensity would be incorrectly ascribed to the effects of former apprentices if LP was used to measure productivity, but not if MFP was used. The disadvantage of examining MFP is that it requires information on physical capital and intermediates employed by each firm, which may not be available in some datasets also because capital tends to be less reliably measured than sales and employment.

61. Potential effects of apprentices on firm-level productivity can be driven by individual-level productivity of the apprentices but also by a number of other factors that are also of interest in their own right. Examining these factors as additional outcomes would provide valuable insights into the mechanisms driving the overall productivity effects.

- **Employee turnover.** If training their apprentices allows a firm to obtain employees who have skills relevant to the firm and who fit well with the firm's culture, it may benefit the firm through reducing employee turnover. Lower turnover, in turn, may lead to saving on training costs and to improved staff morale. Indeed, increased staff retention is among the most commonly cited benefits of Modern Apprenticeship for firms based on survey evidence.<sup>16</sup> For these reasons, employee turnover, measured as an average job duration or a share of

<sup>15</sup> See Lerman, Eyster, and Chambers (2009); Department for Business, Innovation and Skills (2014); and Skills Development Scotland (2015).

<sup>16</sup> See footnote **Error! Bookmark not defined..**

employees leaving over a period of time would be an interesting additional outcome variable to examine. Unfortunately, measuring employee turnover on firm level would require either matched employer-employee data or survey information. Therefore, at least in the short run, the only feasible way to evaluate the effect of MAs on employee turnover may be measuring average job duration using individual-level data, as proposed in the previous subsection, until matched employer-employee data, coming, for example, from combining PAYE records with IDBR, become available.

- **Innovation.** Involvement in Modern Apprenticeships may also increase firm innovation, either thanks to ideas coming from the apprentices or from the training process itself. There are a number of ways to measure innovation, both technological and non-technological, for instance, as the number of newly developed or improved products and services or as patenting activity.<sup>17</sup> Patenting is the only innovation measure that is likely to be available from administrative sources (e.g. patent databases such as Patstat or UK-IPO) dataset, but the potential innovations related to apprenticeships are likely to be less radical and non-technological in nature, and thus, not captured by patents. Broader innovation data are available in innovation surveys, such as the Community Innovation Survey (7<sup>th</sup> wave collected in 2011). However, the CIS captures only a sample of firms, only a small number of which employ apprentices. For these reasons, measuring the effect of apprentices on innovation might require collecting new firm-level survey information or expanding, on an ad hoc basis, the sample covered by the Community Innovation Survey in Scotland.
- **Better working environment.** The knowledge-sharing and co-operation involved in training Modern Apprentices could improve the general atmosphere and level of collaboration in a firm. Indeed, improved staff morale is among the most commonly cited benefits of Modern Apprenticeship for firms based on survey evidence.<sup>18</sup> However, as in the case of employee turnover and innovation, measuring the impact on firm-level morale and atmosphere would require use of survey data, such as the Workplace Employment Relations Study (WERS), which is collected only approximately every seven years, with the last one conducted in 2011.

### *Time horizon*

62. When evaluating firm-level effects of apprenticeships, the key distinction is between the effects during the training period and the effects of employing former apprentices, internally or externally trained.<sup>19</sup> The effects during the training period can be negative in many cases, and they depend on the share of time apprentices spend doing productive work and on the relative productivity of apprentices compared to regular workers. Understanding these effects is valuable primarily for understanding the net costs that training apprentices may carry, with possible implications for public subsidies to the training. On the contrary, the longer-term effects of training and employing Modern Apprentices depend on the apprentices' skills, and on other factors such as staff retention and working environment. Examining these effects is central to understanding the medium and long term effects of Modern Apprenticeships.

<sup>17</sup> See the Oslo Manual (OECD and Eurostat, 2005) for more detail.

<sup>18</sup> See footnote **Error! Bookmark not defined..**

<sup>19</sup> From the perspective of a firm deciding whether to participate in MAs, the effects of internally trained apprentices are crucial. However, from the perspective of the entire economy, the effects of both internally and externally trained apprentices on their current employers should be considered: MA completers contribute to the economy irrespectively of whether they changed employer after the end of their training.

63. As the key explanatory variable, the evaluation of longer-term effects of MAs on employers should use the current number of former Modern Apprentices as a share of all employees. In an ideal world, the evaluation would separately analyse effects of apprentices with different length of time since the end of their training. However, finite sample size and the fact that the number of apprentices most firms employ is rather small mean that such analysis is unlikely to be feasible.

### Benefit-cost analysis

64. Evaluation of Modern Apprenticeships should ultimately lead to robust benefit-cost analysis, which will express, in monetary terms, the value that MAs generate for the Scottish economy. Costs are likely to be relatively easily expressed in monetary terms. However, at least some of the benefits are likely to be more difficult to monetise whether because they are avoided costs which are difficult to quantify (e.g. avoided costs associated with reduced delinquency) or because they are non-pecuniary in nature (e.g. increased life satisfaction). For this reason, in many cases, the estimates of benefits are likely to be "lower bound", and, to the extent that this is likely to be the case, this should be noted. If monetisation is not possible, then quantitative (but non-monetised) or qualitative evidence can be provided separately.

65. Several measures – all of them based on the concept of the net present value (NPV) – can provide useful information about the value generated by MAs and should be included in the evaluation framework.

- **NPV per MA start** = discounted benefits of undertaking an MA (completed or not), aggregated over the working life of an apprentice, minus the corresponding costs to government, employer and apprentice. This measure capture the value generated for the Scottish economy by each MA start.
- **Total NPV** = total NPV of all MA starts in a given year. This measure captures the total value generated by MAs for the Scottish economy in each year. If it is greater than zero, then MAs make a positive contribution to the economy.
- **NPV per £ of government spending** = Total NPV divided by government funding for Modern Apprenticeships. This measure is suitable as a basis for comparing different uses of government funds.
- **NPV per £ of investment** = Total NPV divided by the total costs to government, employers and apprentices. This measure is suitable as a basis for comparing the overall net returns to different education interventions.

### Macroeconomic evaluation

66. All elements of the evaluation framework which have been discussed so far are microeconomic in nature. Even the benefit-cost analysis, as discussed in this report, arrives at the aggregate impact of Modern Apprenticeships by simply adding up the estimated microeconomic effects (taking into account the potential for any double-counting). This is in line with most impact evaluations in general, and virtually all existing evaluations of apprenticeships. Such an approach also has the important advantage of being simpler and more transparent than approaches relying on complex macroeconomic modelling.

67. However, several mechanisms are missed by the microeconomic analysis. In particular,

- by increasing productivity of workers in one part of the skill spectrum, MAs may also affect employment and wages of other workers, and this effect may differ between high-, medium-

and low-skill workers based on their substitutability or complementarity with Modern Apprentices;

- employing Modern Apprentices may result in displacement of workers not participating in MAs;
- increased productivity in some sectors due to MAs may affect performance of other sectors through supplier buyer relationships;
- increased productivity and wages due to MAs may stimulate aggregate demand and aggregate investment; and,
- if MAs increase firm productivity through raising individual productivity of apprentices or through other channels, this may also affect these firms' demand for labour and ultimately wages of other workers.

68. Taking into account some of these mechanisms, especially the first two, by incorporating the results of the microeconomic analysis in a macroeconomic model could eventually improve understanding of the aggregate and distributional effects of Modern Apprenticeships. It is, nevertheless, important that if such analysis is undertaken, the model employed is not treated as a black box, and it is made clear which mechanisms drive any differences between the model's results and the conclusions from a more straightforward aggregation of the microeconomic results.

69. Finally, any macroeconomic evaluation of MAs should not be viewed as a substitute of microeconomic evaluation but as its complement. Indeed, a report by Fraser of Allander Institute (2016) argues that the most promising approach for macroeconomic evaluation of MAs is what it calls, micro-to-macro“ approach, which embeds the microeconomic results in an appropriately adjusted general equilibrium model of the Scottish economy. The microeconomic outcomes proposed in this evaluation framework – wage and employment premia, firm-level productivity effects and evaluation of the impacts over time passed from the end of the apprenticeship – are cited by the report as key inputs for the micro-to-macro approach it recommends.

## 6. DATA FOR EVALUATION

70. Measuring the impact of Modern Apprenticeships on the outcomes discussed in the previous subsection depends on the availability of suitable data. This section first presents existing datasets that should be considered for the evaluation, and it then discusses data options for evaluation of MAs, separately for individuals and firms. The principal options are to link existing, mainly administrative, datasets to SDS records and to conduct a new dedicated survey.

### Existing data sources

71. The datasets that might potentially be relevant for evaluating Modern Apprenticeships are the following.

#### *SDS administrative data*

- **Corporate Training System (CTS)** is an internal dataset collected by SDS. It contains information on all Modern Apprentices, collected as each apprentice is registered, proceeds with the MA and eventually leaves it.
- **Customer Support System (CSS)** is a dataset administered by SDS. It contains information on the activity of the 16-19-year-old cohort.
- **Customer Relationship Monitoring (CRM)** is an internal dataset collected by SDS. It contains information on all employers engaged with SDS.

#### *Comprehensive administrative data*

- **P45 employment data** are owned by HM Revenue and Customs (HMRC). They contain information on employment start and end dates and on whether individuals are claiming benefits.
- **P14 earnings data** are owned by HM Revenue and Customs (HMRC). They contain information on employment start and end dates, earnings and hours worked.
- **Department for Work and Pensions data (DWP)** contain detailed information on benefits claimed by each individual.
- **Work and Pensions Longitudinal Study** combines P45, P14 and DWP data. Importantly, the dataset only covers individuals who are on DWP benefits and employment programmes.

#### *Individual surveys*

- **Understanding Society** is a longitudinal survey of about 40 thousand UK households.
- **Labour Force Survey (LFS)** contains quarterly information on demographics, education and labour-market outcomes of 100 thousand individuals living in UK. It includes a variable



describing if an individual has completed an apprenticeship, but it does not specifically refer to Modern Apprenticeships. Importantly, it cannot be linked to SDS individual records.

- **Annual Population Survey (APS)** is obtained by merging waves one and five of the LFS quarters with annual boosts for England, Scotland and Wales. It contains about 340 thousand individuals in each year, although it contains fewer variables than the standard LFS. It also has a panel element: the households at selected addresses are interviewed annually over four waves and then leave the survey and are replaced by other households.
- **Annual Survey of Hours and Earnings (ASHE)** is a survey filled in by employers, containing job and earnings information on approximately 140-185 thousand individuals in each year.
- Education data
- **Pupil Census** contains mostly demographic information on all pupils in publicly funded schools in Scotland.
- **SQA Attainment Data** contains information on entry and attainment of SQA qualifications.
- **Higher Education Statistics Agency (HESA) Student Record** is an administrative dataset of all students in higher education in Scotland.

#### *Firm-level information*

- **Inter-Departmental Business Register (IDBR)** is a comprehensive list of UK businesses based mainly on Value Added Tax (VAT) and Pay As You Earn (PAYE) systems, both administered by HMRC, but also on information from Companies House, Dun and Bradstreet and business surveys. The information covered includes industry, employment, turnover and country of ownership.
- **Annual Business Survey (ABS)** is the largest survey of UK businesses in terms of the combined number of respondents and variables it covers.
- **Community Innovation Survey (CIS)** samples about 28 thousand UK enterprises with at least 10 employees. It provides detailed information on the level of innovation activity, and it also touches on aspects of the wider innovation process, such as the introduction of new management techniques.
- **Workplace Employment Relations Study (WERS)** combines information from employers, employee representatives and employees related to employment relations, skills and characteristics of workplaces. It takes place approximately every 6-8 years, with the last wave undertaken in 2011.

#### **Data options for evaluating impact of MAs on individuals**

72. Estimating the impact of MAs requires knowledge, for the same individuals, of their involvement in MAs, their subsequent outcomes (e.g. wages) and as much information as possible about their personal characteristics and background. This information either needs to appear in the same dataset or it can be distributed over multiple datasets linked together by a common individual identifier (e.g. National

Insurance number, Scottish Candidate Number) or by personal details such as name, date of birth and address.

73. There are, in principle, three ways of obtaining knowledge of all three required types of information (MA information, outcomes, individual characteristics) for some group of former apprentices: using existing self-standing survey(s), linking multiple datasets – these can include either administrative data or surveys, possibly extended by oversampling Modern Apprentices in Scotland –and collecting a new survey dedicated specifically to the evaluation of MAs.

#### *Existing self-standing survey*

74. The first option is to rely on a single existing self-standing survey. This option has the obvious advantage that it is quick and simple, as it uses information which is contained in a single existing dataset. Largest surveys containing all essential types of information include the Labour Force Survey, the Annual Population Survey and Understanding Society. Unfortunately, as Modern Apprentices in Scotland who started their MA between 2008 and 2012<sup>20</sup> represent only about 0.1% of the UK population, they are likely to represent only a few dozen individuals in Understanding Society and about a hundred in each wave of LFS, likely not enough for any meaningful analysis.

75. These apprentices are likely to account for several hundred observations in **Annual Population Survey (APS)**, which makes it the only existing self-standing survey which could, in principle, be used for examining individual-level outcomes of MAs. Using APS would have the advantages of being fast and bearing low short-run costs, but these advantages have to be weighted against a number of disadvantages, several of which are related to the fact that, like LFS, APS cannot be matched individual records held by SDS. First, while a sample of several hundred individuals allows meaningful aggregate analysis, it is insufficient for analysing the effects of MA separately by MA levels, framework groups or apprentices' characteristics. Second, the limited sample size also requires combining information on apprentices who completed MA in different years, making it hard to analyse how the effects evolve over time passed since the end of the training. Third, APS does not contain information on individuals who started but did not complete their MA, preventing use of MA non-completers as a control group.<sup>21</sup> Fourth, it also does not allow some other, more refined, ways of constructing a control group, such as individuals who were proposed taking up an MA but did not do so. Finally, wages are missing for a large number of observations in APS, and, where they are available, the figures are less reliable than in the case of HMRC administrative data or the Annual Survey of Hours and Earnings. For all these reasons, relying on APS seems like a sub-optimal option for evaluating the impact of MAs on individuals and it should be avoided if other data options turn out to be feasible.

#### *Data linking*

76. An alternative to observing all key information in a single dataset is to link multiple datasets. With this approach, information on participation in MAs can come directly from the **CTS** records held by SDS, providing all key facts about each apprentice's training. In addition, the **CSS** records, also held by SDS, can provide details of individuals who did not start an MA and can, thus, serve as a control group.

77. Outcomes variables, in particular information on employment and wages, can come from two possible sources. The first source consists of the administrative data held by **HMRC**. Linking these data to

<sup>20</sup> SDS began administering MAs in April 2008. Given that a typical length of an apprenticeship is 1-4 years, apprentices entering MAs after 2012 are highly unlikely to have left MAs early enough to allow examination of medium-to-long-term outcomes in the present or in near future.

<sup>21</sup> See Section 0 for discussion of control groups.

CTS and CSS would provide information on the entirety of the apprenticeship population from recent years as well as a large control group, overcoming all of the drawbacks related to the use of APS. Using HMRC data has the strong advantages of providing a large sample size and reliable information on wages and employments. On the other hand, it lacks information about other outcomes of interest such as job satisfaction and career progression. Importantly, linking HMRC data to SDS records represents a considerable administrative and logistical task which will require sustained co-ordinated effort by SDS, Scottish Government and HMRC and will take some time to complete.

78. An alternative source of wage information is the **Annual Survey of Hours and Earnings (ASHE)**. The number of Modern Apprentices starting in 2008-2012 appearing in ASHE is likely to be in low hundreds per year, with the possibility of combining data for multiple years. Consequently, use of ASHE would suffer from the same challenges due to limited sample size as using APS. In addition, by its nature, it cannot be used for estimating of the impact of MAs on employment. On the other hand, its linkage to SDS records would overcome the limitations of APS related to control-group construction. Furthermore, as wages in ASHE are reported by employers rather than employees, it is substantially more complete and reliable than in the case of APS. Moreover, ASHE includes information on whether an employee has a permanent contract, the employee's job title and job description<sup>22</sup> and an indicator of whether an employee has "direct supervisory or managerial control of any other employee(s)." This information captures relevant outcomes related to career advancement which cannot be observed in HMRC data. Overall, a link between SDS records and ASHE, if feasible, may represent a valuable complement to the HMRC data, allowing evaluation of outcomes which cannot be observed in the HMRC data and of wages before the HMRC link is established. It may also represent the preferable alternative for evaluating the wage effects of MAs in case establishing the HMRC link proves not possible. The use of ASHE could be further strengthened in two ways. Firstly, SDS could collaborate with ONS to oversample Modern Apprentices and a suitably selected control group<sup>23</sup> during the ASHE survey collection. Secondly, ASHE could be linked to the Annual Business Survey (ABS), which would provide a number of additional firm-level controls for the analysis.

79. A number of other datasets would ideally also be linked to the CTS and CSS records. Firstly, these include the individual information from the Pupil Census, SQA Attainment Data and the HESA Student Record. This would make it possible for the analysis to take into account richer information on individuals' background and characteristics, and it would allow examining the impact of MAs on subsequent education. Secondly, the records would ideally be linked to firm-level information in IDBR for employers where apprentices did MA (CTS) and for current employers in a given year (CTS and CSS).

80. Crucially, linking any two datasets requires that, in both datasets, records corresponding to the same person can be identified and matched. This can be done either on the basis of a unique personal identifier or based on personal details such as name, date of birth and address. Matching based on unique identifiers tends to be easier to carry out, more complete and more reliable. For these reasons, it is preferred whenever available. Importantly, this means that, to facilitate subsequent data linkage, it is highly advisable to collect key individual identifiers in all administrative records.<sup>24</sup>

81. Currently, CTS records include the National Insurance (NI) number, which makes it possible to link them to the HMRC and ASHE data. The CSS records include the Scottish Candidate Number (SCN),

<sup>22</sup> The job title and job description are recorded as words rather than codes. However, analysis could possibly be conducted relying on presence of key words such as "manager", "supervisor" or "senior".

<sup>23</sup> See more on control groups in Section 0 below.

<sup>24</sup> The importance of ex-ante evaluation strategy and collection of personal identifiers is further emphasized in Sections 0 and **Error! Reference source not found.**

which can be used for linking them to the education datasets. As apprentices recorded in CTS also appear in CSS, establishing a robust link between CTS and CSS would mediate a link between CTS and the education datasets using the SCN available in CSS. No unique identifier currently provides for linking CSS records (of individuals who have not done an MA) with HMRC and ASHE data. Such link would, therefore, need to be based on personal details (e.g. name, date of birth and address). Finally, the individual data can, in principle, be linked to the employer information in IDBR using either IDBR number contained in CTS (for linking to MA employers) and PAYE number contained in HMRC data (for linking to current employers).

82. Importantly, the initial investment into data linking, in particular with large administrative datasets, will not only enable robust evaluation of Modern Apprenticeships, but it will create a powerful analytical tool for evaluation of other policy interventions of SDS and broader Scottish public sector.

### *Dedicated survey*

83. Conducting a new survey specifically for impact evaluation of Modern Apprenticeships is the remaining data option. Its advantage is that it provides greatest flexibility in terms of gathering information on any outcomes that are of interest. On the other hand, it is important to stress that a dedicated survey represents is the most costly option, especially in the long run,<sup>25</sup> and it suffers from some of the same limitations as the existing surveys – limited sample size preventing very disaggregated analysis and less reliable wage information as compared to HMRC and ASHE data.

84. The listed strengths and weaknesses of newly collected survey data suggests that a dedicated survey would be best used not as a substitute for other data sources, notably linked administrative records, but as a complementary source of information covering outcomes which cannot be observed in the administrative data.

### **Data options for evaluating impact of MAs on employers**

85. Similar to the case of individuals, evaluating the impact of MAs on employers requires three types of information, possibly longitudinal: on their involvement in MAs, on outcomes of interest and on additional firm characteristics. No existing firm-level survey contains information on both firms' involvement in MAs and the key outcome of interest – firm productivity – and, as a result, the evaluation needs to rely either on data linking – which could involve oversampling of Scottish firms involved in MAs and the relevant control group – or on a newly collected dedicated employer survey, which is nevertheless the more costly option.

### *Data linking*

86. The most promising datasets for analysing the impact of MAs on employers are IDBR and ABS. IDBR has the advantage of covering the universe of UK firms. Only a subset of small firms is asked to complete ABS in each year, but ABS covers the universe of larger employers and it provides more detailed, reliable and up-to-date information than IDBR. Due to their different strengths, both datasets should ideally be used for the analysis.

87. Neither IDBR nor ABS contains information on employers' involvement in MAs, so these datasets have to be linked to the CTS records. An experimental link between CTS and IDBR has already been established in Scotland. CTS currently does not include any firm-level identifiers, so the link needs to

---

<sup>25</sup> Data linking requires initial investment but allows high-quality low-cost evaluation in the future. In contrast, a survey carries significant costs each time a new wave is collected.

be based on employer's name and post code. Once CTS is linked to IDBR, IDBR number can be used for further linking with ABS. In the future, collecting IDBR, PAYE, VAT and Company Registration numbers in CTS would allow easier, more complete and more reliable linking to the employer data.<sup>26</sup>

### *Dedicated survey*

88. Collecting information required for accurately measuring productivity is relatively complicated, and such information can be obtained from IDBR and ABS data. At the same time, as discussed in Section -, observing the channels through which MAs can impact productivity – employee turnover, innovation, working environment or staff motivation – would likely require either an ad-hoc sampling top up of existing surveys or a dedicated survey. As with individuals, in the case of employers a new survey should be seen as a complement rather than a substitute to links with existing data.

89. However, as discussed below in Section 0, measuring the impact of MAs using firm-level data with sufficient accuracy is likely to be challenging because former Modern Apprentices represent only a small part of employees especially for large firms and their effect may, therefore, be hard to separate from all other determinants of firm performance. For this reason, if resource constraints necessitate a choice between conducting a dedicated survey among individuals and among employers, the former should be given priority or possibly the employer survey might be more focused on SMEs, where both the potential costs and benefits of MAs might be more easily evident and which are harder to observe in ABS.

---

<sup>26</sup> The fact that UK firm-level data are collected on multiple levels of aggregation (e.g. enterprise, plant, “reporting unit”) means that multiple identifiers are needed to maximise the number of linked employers.

## 7. IMPACT EVALUATION METHODS

90. Evaluating the impact(s) of an intervention involves comparing multiple states of the world – one with the intervention as it currently stands and one (or more) without any intervention or with the intervention differently designed. The fundamental problem of evaluation is that actually observing multiple different states of the world is not possible. As a result, the alternative state of the world needs to be simulated with the use of units – the “control group” – which, for one reason or another, are not subject to the intervention.

91. The major evaluation challenge stems from the fact that the reasons why some units (e.g. individuals, firms) are generally not selected into the evaluated intervention randomly – so called “selection problem”. Outcomes for the “treated” and “control” units may, therefore, systematically differ not only because of the intervention itself, but also because of their characteristics and other “third factors”.

92. This section starts by briefly discussing what the “alternative states of the world” actually represent in the case of Modern Apprenticeships. The largest part of the section discusses control-group construction and the “selection problem” in the specific context of the impact of MAs on individuals and firms and describes methods that can be used to overcome it. Finally, the section considers some key issues related to performing benefit-cost analysis of MAs.

### **Policy counterfactuals: alternative states of the world**

93. But, first all, to ensure that the evaluation framework for Modern Apprenticeships produces results that are relevant to policy makers and likely to be correctly interpreted, it is important to be clear about the alternative states of the world to which Modern Apprenticeships, as they currently stand, are compared.

### ***Overall impact***

94. The principal aim of evaluating Modern Apprenticeships is to understand their overall impact on individuals and firms and see if, in the long term, they generate value that sufficiently exceeds their costs. Therefore, the alternative state of the world that should be the primary focus when evaluating Modern Apprentices is one where Modern Apprenticeships in the present form of apprenticeships administered by a public agency do not exist. Even in this scenario, it is likely that some apprenticeships opportunities would be organised by the private sector. However, the market failures inherent in apprenticeships, the intensive government participation in apprenticeship systems in most developed countries and the relatively low numbers of apprentices prior to establishment of Modern Apprentices all suggest that the provision of apprenticeship training in the absence of a public intervention would be rather limited.

### ***Design***

95. Evaluating the overall impact of Modern Apprenticeships is key for making crucial high-level decisions about their role in the Scottish economy. However, in the short to medium term, even more practically relevant lessons can be learnt from examining the way Modern Apprenticeships are designed.

Design options that are, or have been, considered by SDS and other key stakeholders, include, in particular, the following:

- **Training contribution rates.** SDS currently subsidises training of Modern Apprentices. The contributions vary by framework, apprenticeship level and age of the apprentice. Understanding to what extent the level at which the contribution rates are set increases the number and quality of apprenticeships places offered would inform decisions about whether (or which) contribution rates should be cut or increased.
- **Form of training.** Modern Apprentices learn while doing their job (“learning by doing”), while attending internal training organised by their employer or while attending external training outside the firm. The mix of on and off the job training is defined in each MA framework, based on the needs of employers and their sector. Understanding the relative efficacy of these components could help employers understand the benefits of different types of training and enable them to redesign MAs frameworks to increase their impact.
- **Content of training.** MA training can contain both material that is highly specific to a particular job and material that is much more generally applicable. Understanding the impact of different types of learning content would be useful for maximising the positive impact of MAs on apprentices’ careers.
- **Duration of training.** Duration of MAs varies widely across MA frameworks, depending on the time needed to develop the underpinning knowledge and competences in each sector and occupation. Identifying an appropriate duration for each MA framework would help ensuring that apprentices are given the time they need to develop full occupational competences and at the same time do not remain in training beyond the level at which the additional returns to training become very small.

96. Evaluating the (sector-specific) effect of these features in the design of Modern Apprenticeships would produce valuable information that might allow for increasing the impact of the apprenticeships while optimising the use of financial resources.

### Evaluating impact on individuals

97. Evaluating the impact of Modern Apprenticeships on individuals means asking how the examined outcomes for individuals participating in Modern Apprenticeships would be different, had the individuals not participated.

98. The central challenge of any impact evaluation is that it is not possible to observe the same individual at the same point in life both having and not having completed an apprenticeship. Consequently, it is necessary to find a *control group* of individuals who were initially as similar as possible to the apprentices (the *treatment group*) but did not complete an apprenticeship. Even if a control group is established, there is a risk that outcomes for the control group differ from those for apprentices not only because of the apprenticeships but also for other reasons. If that is the case, interpreting observed differences between the two groups as the causal effects of apprenticeships will be incorrect. A number of econometric techniques can be used to try to overcome this issue.

99. This section discusses the choice of a control group and the different methods that can be used to obtain causal estimates of the impact of Modern Apprentices on individuals.

*Control groups*

100. Generally speaking, there are two types of control groups to which Modern Apprenticeship completers can be compared – never-starters and non-completers. Using each of these control groups leads to different challenges.

*Never-starters*

101. The first option is to use individuals who never started an apprenticeship. The main challenge when using never-starters as a control group is that the decision to start an apprenticeship is not random and may be related to individual characteristics (e.g. skills, motivation, socio-economic background, local employment opportunities...) that may affect the examined outcomes (e.g. wages, employment...). As a result, the group of those starting an apprenticeship might be a “self-selected” group of individuals who are very different – already before the MAs – from non-starters, which would invalidate the use of the comparison between the two groups for estimating the causal impact of MAs.

102. One way to mitigate this issue, discussed in more detail below, would be to try to select amongst non-starters those that could have potentially started an apprenticeship, such as individuals who are observed in CSS records to being suggested to start an MA by a career advisor.

*Non-completers*

103. An alternative is to focus on individuals who started an apprenticeship but did not complete it. The issue of selection into starting an apprenticeship does not apply to non-completers, because they themselves started an apprenticeship. Furthermore, if benefits of apprenticeships differ across individuals, it is particularly important for policy purposes to understand the effects of apprenticeships on those individuals who are likely to participate, rather than on, for instance, strongly academically minded individuals who do not consider apprenticeships as an option. The advantage of relying on non-completers is that they indeed are such individuals.

104. However, relying on non-completers leads to two other challenges. Firstly, some of the factors determining who completes or does not complete an apprenticeship are also not random and may be related to the examined outcomes. On one hand, some individuals may drop out from their apprenticeships due to a lack of skill or motivation, and these factors can also lead to poorer subsequent labour market outcomes. On the other hand, some individuals may leave their apprenticeship because they obtained a more attractive job or educational offer. This new opportunity itself, or their qualities that allowed them to gain the opportunity, may lead to subsequent superior labour market outcomes for these individuals. In the former case, the non-randomness in completing apprenticeships would lead to an overestimation of the effects of apprenticeships; in the latter case, it would lead to their underestimation.

105. This issue can be mitigated by using information on the reasons for non-completion, which are recorded in CTS. One option is to construct the control group of only those non-completers, who dropped out for a seemingly unrelated reason (e.g. moving to a different city). Alternatively, two control groups can be obtained, consisting of individuals whose reasons for non-completion are likely to be either positively or negatively related to later labour-market outcomes. One of these control groups is likely to lead to an overestimation of the effect, while the other to an underestimation. Together, these estimates would provide a lower and upper bound for the true impact of MAs.

106. The second challenge related to using non-completers as a control group is that even individuals who do not complete their apprenticeship may gain skills and other useful assets from their participation. The difference between completers and non-completers, therefore, captures only the additional benefit of completion and certification rather than the full benefit of participating in an apprenticeship from start to



end. The different impacts of starting an apprenticeship, of undertaking most of the training and of being certified as an MA completer could be separately measured by comparing MA completers with never-starters, with non-completers leaving early in their apprenticeship and non-completers leaving after a long time in training.

### *Benefits of using both control groups*

107. Importantly, the issues related to each of the two control groups discussed above are different. To the extent that data allow it, it is, therefore, desirable that both control groups are used. If results turn out to be similar for both control groups, it will be possible to consider them as more reliable than if only one control group was used.

108. Two qualifications are needed. First, even use of two control groups by no means guarantees infallible results. Second, in the case of conducting a new survey, relying on two different control groups will increase the required sample size and consequently also costs (although this is not necessarily the case with existing administrative or survey data).

### *Accounting for subsequent qualifications*

109. The specification of the control group and the treatment group also needs to take into account any other qualifications that individuals start (unless subsequent qualifications are the examined outcome). The reason is that the other qualifications are also likely to affect the examined outcomes, and if they are not accounted for, their effects will be incorrectly taken for the effects of doing or not doing an apprenticeship. This is the case whether the control group consists of never-starters or non-completers.

110. There two fundamental ways of accounting for other qualifications.

111. The first approach is to restrict the treatment and control groups to those individuals whose highest attempted qualification, other than MA, is of lower level than the evaluated MA.<sup>27</sup> The advantage of this approach is that it relies on more homogeneous treatment and control groups.

112. The second approach is to keep also individuals with higher attempted qualifications, but include these qualifications as control variables in a regression analysis.<sup>28</sup> The advantage of this approach is that it provides richer information and allows directly comparing the effects of MAs to the effects of other qualifications.

113. As each approach has a distinct advantage, both approaches should ideally be employed. However, the former approach, based on restricting the treatment and control groups is likely to provide more robust estimates of the causal effects of MAs and should, therefore, be preferred in the baseline specification.

### *Methods for identifying causal effects*

114. A number of existing methods can fully or partly overcome the challenges related to using apprenticeship never-starters or non-completers as a control group. These methods differ in their ability to overcome estimation challenges, their data requirements and their suitability for evaluating Modern Apprenticeships as a whole or for evaluating their individual design features.

---

<sup>27</sup> This approach is taken in the evaluations of further education undertaken by the UK Department for Business, Innovation and Skills, e.g. Bibby et al. (2014).

<sup>28</sup> See more on regression analysis in the subsection “Methods for identifying causal effects” below.

*Randomised control trial*

115. A randomised control trial (RCT) would randomly assign some individuals into a treatment group and others into a control group or an alternative treatment group. By doing so it would overcome issues due to the fact that individuals not starting or not completing an MA systematically differ from MA completers. Evaluation of the overall impact of Modern Apprenticeships based on an RCT would award some apprentices an MA place while putting others to a control group which would not participate in the apprenticeships.

116. This strategy faces at least two challenges.

117. Firstly, administering an RCT requires a degree of centralised control over the evaluated intervention, while MAs operate in a rather decentralised fashion, with a key role being played by a large number of training providers that match apprentices with employers. This issue could be partly alleviated by organising the experiment in collaboration with a limited number of large employers, but this would restrict the range of MA frameworks and firm types covered by the evaluation.

118. Secondly, denying some interested individuals access to a publicly subsidised programme tends to be ethically and politically acceptable only when the demand for the programme significantly exceeds the number of places and some individuals would have to be denied access in any case. However, this is not the case for many MA frameworks. As a result of these challenges, the case for using an RCT for evaluation of the overall effect is weaker in the case of Modern Apprenticeships than in the case of many other public programmes.

119. Even if a randomised experiment is not used to evaluate the overall impact of MAs, it could be suitable for comparing different design features of the apprenticeships (e.g. general vs. occupation-specific content, on-the-job vs. off-the-job training, training length) and, by doing so, generating policy-relevant information on how the design of Modern Apprenticeships could be improved. In this case, an RCT is less politically and ethically problematic, because it is *ex-ante* unknown which of the design alternatives is more advantageous and the experiment, thus, does not provide some individuals with (*ex-ante*) better treatment than others. Furthermore, a similar experiment could be undertaken in collaboration with only one, or several, large employers, hence overcoming the challenges posed by the decentralised nature of MAs. Focusing on a limited number of large employers would also limit the overall transaction costs imposed on the participating employers, while ensuring their interest in the experiment thanks to their high stake in finding the most effective apprenticeship training design.<sup>29</sup>

*Instrumental variables*

120. Instrumental variable techniques would rely on factors that make some individuals more (or less) likely to enter an apprenticeship but do not directly affect the examined outcomes (e.g. wage or employment). Such variables could be related, for instance, to the availability of apprenticeship places in a given region or to the intensity with which apprenticeships are promoted in different regions or at different schools. Unfortunately, no variable that would be suitable for this purpose has been identified in the available data.

<sup>29</sup>

Bandiera, Barankay, and Rasul (2011) discuss studies describing randomised economic experiments conducted in cooperation with firms. The experiments analyse effectiveness of a wide range of practices including monetary and non-monetary incentives, worker monitoring, organisation of the workplace and provision of feedback. Bandiera, Barankay, and Rasul (2011) also provide some general guidance related to designing and conducting experiments in collaboration with firms.

### *Matching*

121. Matching techniques would construct the control group of a subset of the never-starters or the non-completers consisting of individuals who are particularly similar to apprentices in their characteristics observed prior to an apprenticeship. Characteristics potentially suitable for matching in the context of MAs include, for example, gender, age, socio-economic background, previous education, region and, if available, previous labour market history. Based on similar characteristics, apprenticeship completers can be matched to individuals whose observed characteristics made them initially similarly likely to start/complete an apprenticeship (propensity-score matching) or whose initial observable characteristics are themselves similar to those of apprentices (coarsened exact matching).<sup>30</sup>

### *More specific control group*

122. Even when the control group consists of a matched sample of people with similar observed characteristics, many individuals in the control group may significantly differ from apprentices in their unobserved characteristics, such as motivation, ambitions, persistence and skills. These characteristics are likely to be related to the examined outcomes and could, therefore, influence the impact estimates.

123. When never-starters are used as a control group, this issue could be mitigated by further restricting the control group to individuals who have either considered or been proposed to enter an apprenticeship. In particular, the CSS records contain information on whether a person was suggested starting an MA. Restricting the analysis to individuals who received such advice and subsequently either did or did not start an apprenticeship would contribute to the robustness of the results.

124. When non-completers serve as a control group, one issue is that even individuals who do not complete their apprenticeship are likely to benefit somewhat from starting an apprenticeship and pursuing it for some time. A partial solution to this issue is to construct the control group only of those non-completers who dropped out within a short time into their apprenticeship.

125. Another issue related to non-completers is that the factors leading to dropping out may be directly related to the examined later outcomes. This issue may be mitigated by restricting the control group to those non-completers who left their apprenticeship prematurely for “random” reasons that are unlikely to be directly related to the examined outcomes for the given individual. Such reasons can include, for example, moving to a different area (for reasons unrelated to employment opportunities), short-term health problems and or employer terminating involvement in MAs.

126. If no reasons can be meaningfully considered to be random or the number of individuals dropping out for such reasons is too small, an alternative approach, equally using information on reasons for dropping out, is to split all non-completers into two control groups according to whether their reason for dropping out is likely to be positively or negatively related to the examined outcomes (e.g. „received a more attractive offer“, as opposed to „poor attendance“). The effects of apprenticeships can then be estimated separately using each of the two control groups. With one of the control groups leading to an underestimation and the other to an overestimation of the impact, the two control groups together will provide a range within which the true impact is likely to lie.

<sup>30</sup>

Coarsened exact matching (CEM) is a more robust way to construct a control group that is actually similar to the treatment group than the propensity-score matching (PSM). However, CEM requires a very large sample size, and its successful application would, therefore, require comprehensive administrative data. CEM is described by Iacus, King, and Porro (2011) and has been used in the context of apprenticeship evaluation by Bibby et al. (2014).

127. This approach would be strengthened by redefining the reasons for non-completion available to training providers when filling in CTS records. The proposed answers could be redefined in a way that makes it clearer whether each answer is (a) likely unrelated to the examined outcomes, (b) likely positively related to the examined outcomes or (c) likely negatively related to the examined outcomes.

### *Regression*

128. A careful construction of a control group, as discussed in the previous paragraphs, is essential for a successful evaluation. Once an appropriate control group is established, applying regression analysis – as opposed to simply comparing means observed in the treatment and control groups – can significantly improve the analysis. In particular, finite sample size implies that matching techniques can take into account only a limited number of key individual characteristics. The regression framework allows taking into account additional observed individual and employer characteristics that may differ between the treatment and control groups and at the same time be related to the measured outcomes. Such characteristics may include, for example, gender, age, socio-economic background, previous education, region and size and industry of MA employer.

### *Panel techniques*

129. Although matching and regression techniques can limit the potential bias in impact estimates, they are able to account only for those individual characteristics that can be observed in data. Unfortunately, many important factors, such as motivation, talent or persistence are hard to observe in data. Consequently, it would be useful to compare apprentices and control group in terms of *changes* in examined outcomes (e.g. wages) over time. Importantly, this approach can be applied only to apprentices with work experience (i.e. observed wages) prior to their apprenticeship. The application of this desirable approach,<sup>31</sup> therefore, depends on an existence of a sufficient number of apprentices for whom this is the case.

## **Evaluating impact on firms**

### *Measuring firms' involvement in Modern Apprenticeships*

130. An individual either works as an MA or does not. On the contrary, a firm can participate more or less intensively. It is, therefore, useful to think of firms' involvement not only as a binary outcome depending on whether a firm participates in MAs at all, as a continuous variable measuring the intensity of the involvement.

131. To establish how such continuous variable should be defined, it is necessary to consider what firms training or hiring apprentices would do if they did not train or hire them. The main alternatives are

- a) Employ a regular worker without training.
- b) Employ and train a regular worker.
- c) Train an existing regular worker.
- d) Neither employ nor train an additional worker.

---

<sup>31</sup> The approach can also be beneficially combined with matching techniques.

132. If (a) is the most likely alternative, the explanatory variables should be based on the number of apprentices relative to the total employment. If (b) or (c) is the most likely alternative, the explanatory variables should be based on the number of apprentices relative to the total number of trained workers. Finally, if (d) is the most likely alternative, the explanatory variable should be based on the total number of apprentices.

133. Firm-level estimation that would shed light on whether firms see apprentices as substitutes for regular workers or as additional workforce would be valuable. In the absence of such estimation, option (a) seems most likely based on existing empirical and theoretical literature (see Bajgar and Criscuolo (2016b), and the number of apprentices as a share of all employees, thus, appears as the most appropriate measure of firms' involvement in Modern Apprenticeships.

134. As discussed in Section 5, it is also important to carefully determine the time at which effects of employing apprentices are examined. During the training period, the effect of apprentices relative to regular workers can actually be negative, even if the effect of participating in Modern Apprenticeships is positive in the long-term. The benefits of having trained apprentices are likely to materialise towards the end and, in particular, after the training.

135. For this reason, the evaluation should use the number of apprentices trained over the previous several years as the explanatory variable. In an ideal world, the evaluation would examine the effect of employing *former* Modern Apprentices, trained by their current employer or, possibly, by another employer. However, such information is hard to observe unless detailed employer-employee data are available. While data available in Scotland should allow observing which firms *train* Modern Apprentices, they are unlikely to allow observing which firms employ them afterwards. In such case, a second-best strategy is to exploit the fact that many apprentices remain employed by the firm where they completed their MA and use the number of apprentices trained over the previous several years as a proxy for employing apprentices.

136. Importantly, either the resulting estimates will need to be adjusted for the fact that only some apprentices stay with their employers upon completion, or the estimation should be restricted to those apprentices for whom SDS administrative records indicate that they remained with their employer after completion of their MA.

### ***Methods for identifying causal effects***

137. As in the case of individuals, evaluating the impact of Modern Apprenticeships on firms also requires separating the causal effect of MAs from other factors that may be correlated with both firms' involvement in MAs and the examined outcomes. For instance, it is possible that firms that take on apprentices are generally better managed and invest more in their employees. Such firms could be more productive, but it would be misleading to interpret their higher productivity to be a consequence of their participation in MAs. Alternatively, some firms might train apprentices in response to underinvestment in training in previous years. These firms would be likely to have, at least initially, a lower productivity, but this would again not be a result of training apprentices. A number of methods can be used to bring the impact estimates closer to the true values. Several of the methods are similar to those relevant for evaluating the impact on individuals, but there are also some important differences.

### ***Randomised control trial***

138. As in the case of individuals, a randomised control trial can be used to make some firms participate in MAs while other, otherwise similar firms, do not participate. Rather than preventing some firms from taking on apprentices, the experiment would have to be based on providing some firms with

stronger incentives to participate in MAs than others. In particular, a randomly selected set of firms could receive a voucher entitling them to receiving a larger training contribution than the remaining firms. If providing some firms with unambiguously larger subsidies than others turned out to be politically unfeasible, an alternative would be to deliver the contribution in different ways. For example, some firms could receive a training subsidy in multiple instalments related to MA milestones, while others would receive it at the end of the training. One of these settings is likely to be more effective at motivating employers to train apprentices, but it is not *ex ante* obvious which one.

139. In either case, the RCT would bring a “double dividend” in the form of two types of valuable information. In the first stage, it would reveal to what extent the training contributions are effective at increasing firm participation in Modern Apprentices – with useful implications for whether the contribution rates should be increased or cut – or which way of paying out the contribution is preferable. In the second stage, the variation in firm participation could be used to examine the causal impact of participating in Modern Apprenticeships on the ultimate outcomes of interest, such as firm productivity.

#### *Instrumental variables*

140. In the context of impact on firms, instrumental variable techniques would draw on factors which make some firms more actively engage with Modern Apprenticeships but do not affect firm performance except through the apprenticeships. Such variables could, in principle, be related to the contribution rates or to the traditional and institutional arrangements in place for apprenticeships in different occupation or regions. However, no variables that could be used for this purpose have been identified in the case of MAs.

#### *Matching*

141. Similar to the case of individuals, matching techniques can be used to construct a control group of firms that do not engage in MAs but are similar to the firms that do in their other characteristics. The variables to be used in such matching include, in particular, firm industry and firm size, but they may also include, for example, availability of a training budget and training plans (subject to data availability).

#### *Panel techniques*

142. An important advantage of evaluating the impact on firms, compared to the case of individuals, is that the examined outcomes for firms can be observed both before and after the firm engages in MAs. This is harder to do in the case of individuals, because only some apprentices have work experience prior to their apprenticeships. Being able to observe firms before engaging with MAs makes it possible to take into account firm characteristics that cannot be observed in data but do not change over time, by comparing *change* in examined outcomes over time between firms that start taking on apprentices and those that do not.

#### *Regression*

143. As in the case of individuals, regression analysis can be used to take into account firm characteristics which could be correlated with participation in MAs and at the same time influence the measured outcomes such as productivity. Such variables can include, for instance, industry, firm age and provision of training (other than MAs).

### **Benefit-cost analysis**

144. As discussed in Section -, the key outcomes of benefit-cost analysis (BCA) are all derived from the NPV per MA start, defined as discounted benefits of undertaking an MA (completed or not),

aggregated over the working life of an apprentice, minus the corresponding costs to government, employer and apprentice (normalised to the number of MAs started).

145. Producing robust benefit-cost analysis is a challenging task, and it requires a careful consideration of which effects should be included so that the key benefits and costs of MAs are taken into account but, at the same time, no benefits and costs are double-counted<sup>32</sup> and the results are not driven by some difficult-to-measure hypothetical secondary effects.

146. In particular, benefit-cost analysis inevitably requires a set of strong assumptions which may dramatically influence their results even while using the same set of microeconomic estimates as inputs. Good evaluation practice therefore requires that

- the general BCA procedure as well as the particular assumptions made are described with maximum transparency;
- the BCA includes extensive robustness tests analysing how its results change when assumptions change;
- the key assumptions are always presented alongside the results; and
- preferably, the results are presented as a probabilistic claim regarding a range of values rather than as a single number.

147. This report does not discuss the entire procedure for the benefit-cost analysis of MAs, but it highlights some of the key benefits and costs to be considered in BCAs and discusses how they are related to the other components of the evaluation framework.

### ***Benefits***

148. The primary potential benefits of MAs for the Scottish economy are increased employment and increased productivity.

#### ***Increased employment***

149. The probability of employment is one of the primary proposed outcomes for estimation of the effects of MAs on individuals. Results of this estimation will, thus, provide the key input for calculating the contribution of MAs to the Scottish economy through increased employment.

#### ***Increased productivity***

150. Increased productivity can reflect the increase in individual productivity of the apprentices, but also the increase in overall firm productivity thanks to reduced employee turnover, more innovation, better working environment and spillovers from apprentices to other workers.

151. To establish the contribution of MAs to the Scottish Economy through increased productivity, two main approaches can be used.<sup>33</sup>

---

<sup>32</sup> This could be the case, for example, if the calculation added up individual- and firm-level productivity improvements, even though the latter are largely driven by the former.

<sup>33</sup> See Section - for related discussion.

152. One option is to draw on results from the proposed estimation of the effects of MAs on individuals' wages, based on the observation that individual wages are likely to reflect individual productivity. The second option is to estimate the effect of MAs on productivity directly, from firm-level data.

153. Each of these approaches faces significant challenges. The former approach, based on individual-level wage estimates, requires taking a judgement as to how wages and productivity are related. In particular, if the effect of MAs on wages primarily reflects the signalling value of an MA completion, the effect on productivity will be smaller than the effect on wage. If, on the contrary, MAs increase firm productivity through increased apprentice productivity and other channels (e.g. other employees learning by training apprentices, improved staff morale, reduced employee turnover), the effect on productivity will be larger than the effect on wages. The judgement regarding the relationship between productivity and wage effects is rather arbitrary, although it could be informed by existing research on the role of signalling in driving wage returns to education and, in particular, by potential analysis of the impact on MAs on the additional firm-level outcomes such as staff morale and reduced turnover.<sup>34</sup>

154. The latter approach, based on productivity estimation using firm data, is less arbitrary and, therefore, in theory, preferable. However, its feasibility depends crucially on whether data constraints and the small number of apprentices in larger firms allow sufficiently precise and reliable estimation of the effect of MAs on firm productivity.

#### *Other benefits*

155. In addition, if feasible, calculations should also include indirect benefits of MAs such as reducing skill shortages in the economy and preventing undesirable social outcomes (e.g. criminal behaviour, drug abuse or teenage pregnancy).

#### *Costs*

156. The costs of MAs to the Scottish economy should include the following components:

- **Direct costs of MA training.** These include the public funding contributions and training contributions paid by employers.
- **Indirect costs of MA training.** These include the training-induced reduction in the time apprentices spend doing productive work.
- **Administrative costs of MAs.** These costs are carried by SDS and potentially also other public agencies and by employers.

#### *General equilibrium effects*

157. The benefits and costs considered so far do not take into account general equilibrium effects. These are discussed in more detail in Section 0 and include, in particular,

<sup>34</sup>

There are also a handful of studies which estimate the effect of training on both firm-level wages and firm productivity (see Bajgar and Criscuolo (2016b) for more details), but the different nature of Modern Apprenticeship and regular on-the-job training (e.g. apprentices tend to represent a small share of all employees, apprentices are junior staff while training often involves managers) makes extrapolation from these studies highly problematic.



- increasing productivity of workers in one part of the skill spectrum affecting employment and wages of other workers at different skill levels;
- displacement of workers not participating in MAs due to employing apprentices;
- productivity spillovers across sectors through supplier-buyer relationships;
- effect of increased productivity and wages due to MAs on aggregate demand and aggregate investment; and,
- effect of increased firm productivity on the firms' demand for labour and ultimately wages of other workers they employ

158. These effects could potentially be taken into account by incorporating the results of the microeconomic analysis in a general equilibrium macroeconomic model of the Scottish economy.

## **8. RECOMMENDATIONS FOR EVALUATING MODERN APPRENTICESHIPS**

159. This section proposes a strategy for evaluating Modern Apprenticeships. It first describes how Modern Apprenticeships should be evaluated once all required data infrastructure is put in place. It then outlines the steps towards understanding the impact of Modern Apprentices that can be undertaken while the long-term data infrastructure is being set up.

160. Several elements of the evaluation of Modern Apprenticeships proposed here (e.g. ex-ante strategy, use of multiple specific control groups, analysis of impact on well-being and on employer productivity) would make it unique among existing evaluations of apprenticeships and mean that it would have the potential to provide particularly valuable insights.

161. Nevertheless, such an evaluation would also carry significant challenges.

162. All survey data suffer from limited sample size, inaccuracies and non-response. Conducting one's own survey is, in addition, costly. Administrative data are also not without challenges: the set of variables they cover may be limited, they sometimes involve important gaps and gaining access to them requires overcoming non-trivial legal hurdles, co-ordination among multiple organisations and making significant investment in terms of both time and financial resources. Collecting detailed cost data may also be costly.

163. Beyond constructing data, it will be necessary to identify researchers with the right skills for undertaking the analysis and to address a number of issues inherent in counterfactual impact estimation. These will include identifying suitable control groups and limiting the possibility that observed correlation between participating in MAs and the examined outcomes may be driven by unobserved individual characteristics and other factors.

164. Finally, evaluation of the impact on employers may be particularly challenging, because the limited share of apprentices in the total employment of most firms is likely lead to limited "signal-to-noise" ratio in the estimation.

### **Long-term evaluation strategy**

165. This long-term evaluation strategy describes how evaluation should be undertaken once datasets containing the required variables are linked and access is awarded to experts undertaking the evaluation. The evaluation should consist of seven elements:

- i. monitoring inputs;
- ii. monitoring outputs;
- iii. process evaluation;
- iv. evaluation of impact on individuals using administrative data;
- v. evaluation of impact on individuals using survey data;

- vi. evaluation of impact on employers; and
- vii. benefit-cost analysis.

### ***Monitoring inputs***

166. To support robust benefit-cost analysis, SDS should collect information needed to make an informed estimate of the total public and private costs of MAs, both direct and indirect. The required information includes

- public funding contributions;
- training contributions paid by employers;
- amount of time apprentices spend in training rather than in productive work;
- SDS costs attributable to administering MAs; and
- administrative costs to employers for participating in MAs.

### ***Monitoring outputs***

167. SDS should monitor the following outputs of MAs, disaggregated by level, framework or framework group and key apprentice characteristics (e.g. gender, age):

- number of MA starts, leavers and apprentices in training;
- completion rates;
- apprenticeship duration;
- form of training (on-the-job vs. off-the-job vs. learning by doing); and

168. To reliably monitor outputs, SDS will need to ensure that training providers supply accurate and timely information.

### ***Evaluation of process and participant satisfaction***

169. SDS should reconsider the purpose of the MA Outcomes Survey and the MA Employer Survey, because (i) part of the information collected through the surveys will be collected during the impact evaluation elements of the evaluation framework and (ii) additional data collection using similar questionnaires as in the past is likely to bring only a limited amount of new information. The surveys should refocus towards collecting

- feedback on how MAs are administered and how the administration could be improved;
- information on how much time apprentices spend in productive work, on-the-job training and off-the-job training; and
- administrative costs of MAs.

170. The latter two points are in line with the needs of input and output monitoring and could potentially support MA impact estimation disaggregated according to the nature of training in different groups of frameworks.

171. A time period of 4-5 years between two rounds of each survey may be sufficient and save resources, which can be redirected towards impact evaluation.

***Evaluation of impact on individuals using linked administrative data***

172. For evaluating the impact of MAs on individual-level outcomes that appear in administrative data, linked administrative data should be preferred to surveys. Important advantages of using administrative data include

- higher accuracy;
- dramatically larger sample size;
- lower costs (in the long run);
- a higher number of individual characteristic (provided enough datasets can be linked);
- enabling panel analysis;
- replicability over time;
- possibility of following cohorts of apprentices; and
- coverage of variables that can be used for specifying policies (because of their availability in administrative data).

***Datasets***

173. The evaluation should be based on linking a number of administrative datasets. The datasets and the available identifiers are summarized in Table 1.

**Table 1. Linkable datasets for evaluating impact on individuals**

<b>Dataset</b>	<b>Owner</b>	<b>Identifiers</b>	<b>Coverage</b>	<b>Variables</b>
Corporate Training System (CTS)	SDS	NI	MAs	MA information (start, end, completion status, reason for non-completion, framework, location etc.); demographics (age, gender, disability etc.)
Customer Support System (CSS)	SDS	SCN	16-19 year old cohort	demographics; MA proposed by SDS career advisor
P45 employment data	HMRC	NI	employed individuals	employment dates
P14 earnings data	HMRC	NI	employed individuals	earnings, hours worked
ASHE	ONS	NI	1% sample of employees	earnings, hours worked, permanent contract, (uncoded) job title, supervisory role, job tenure
DWP data	DWP	NI	benefit recipients	benefit history
Pupil Census	SG	SCN	pupils	age, gender, location, ethnicity, disability, main home language, free school meal status
SQA Attainment Data	SG	SCN	qualification holders	subject, qualification level, result
Higher Education Student Records	HESA	SCN, HESA, UCAS, ULN	higher education students	course details; prior qualifications
Inter-Departmental Business Register (IDBR)	HMRC	IDBR, PAYE	all firms	size and industry of MA employer and current employer

174. The linking consists of the following steps, which should be undertaken in parallel:

1. CTS and CSS records are linked using name, birth date and postcode.
2. The education data are linked to CTS-CSS using the Scottish Candidate Number (SCN).
3. HMRC, DWP and ASHE data are linked to CTS-CSS using the National Insurance (NI) number, or name, birth date and postcode where the National Insurance number is not available.
4. IDBR data for MA employers are linked to CTS records using IDBR number and IDBR data for current employers are linked to linked CTS-CSS-HMRC data using PAYE number.

#### *Outcome variables*

175. Evaluation using linked data should examine impact on the following outcome variables:

- Employment (HMRC)
- Wages (HMRC, or ASHE if HMRC not available)
- Receiving benefits (DWP)

- Employment duration (HMRC, ASHE)
- Permanent contract (ASHE)
- Supervisory role (ASHE)
- Subsequent started and achieved qualifications (SQA, HESA)

*Control groups and methodology*

176. As underscored in the introduction to this section, correctly estimating the causal impact of apprenticeships is a challenging task; it requires identifying suitable control groups and limiting the possibility that any observed correlation between MA participation and the outcomes is driven by unobserved individual characteristics and other factors.

177. To produce more robust results, evaluation of the impact of MAs on individuals using linked data should use two different control groups: the never-starters and the non-completers. In the baseline specification, the treatment and control group should be restricted to those individuals whose highest attempted qualification, other than MA, is of lower level than the evaluated MA.

178. The analysis should be strengthened by the following approaches, applied either in the standard specification or as robustness checks:<sup>35</sup>

- applying matching techniques (coarsened exact matching or propensity-score matching) to construct an appropriate control group;
- limiting the control group of never-starters to individuals who were proposed starting an MA by an SDS career advisor;
- limiting the control group of non-completers to those who dropped out shortly after starting their MA;
- limiting the control group of non-completers to those who dropped out for reasons unlikely to be related to their later labour-market outcomes;
- splitting the control group of non-completers based on whether individuals' reasons for non-completion are likely to be positively or negatively related to their later labour-market outcomes;
- comparing changes in outcomes before and after MAs for the treatment group and the control group (feasibility of this approach depends on the number of apprentices with work experience prior to their MA);
- applying regressions techniques to account for additional observed individual and employer characteristics which may differ between the treatment and control groups and at the same time be related to the measured outcomes; and
- controlling for higher attempted qualifications in a regressions analysis rather than restricting the sample to individuals who did not attempt such qualifications.

---

<sup>35</sup>

For more details on the proposed approaches, see Section 0.

*Time horizon*

179. The impact should be measured at multiple time periods after the end of apprenticeship rather than at any one time period. This is important to understand how the effects of MAs evolve over people's careers.

*Evaluation of impact on individuals using a well-being survey*

180. Some important outcomes do not appear in administrative data. Information on these outcomes can be collected through a well-being survey. Such survey should be seen as a complement of, rather than a substitute to, the linked, mostly administrative, data. For this reason and due to costs consideration, it may be sufficient to administer it with, for example, 5-year periodicity.

*Outcome variables*

181. Evaluation based on the well-being survey should focus, in particular, on self-reported job and life satisfaction. In designing the survey, SDS can build on existing methodology, as developed, for instance, in the OECD Better Lives Initiative.<sup>36</sup>

*Control groups and methodology*

182. As in the case of evaluation using administrative data, the survey should include two different control groups: never-starters and non-completers. If financial constraints do not allow sufficient sample size for examining two distinct control groups, the control group should consist of never-starters.

183. The analysis should be strengthened by the following approaches.

- Applying matching techniques (coarsened exact matching or propensity-score matching) when constructing the control groups
- Sampling the control group of never-starters from individuals who were proposed starting an MA by an SDS career advisor.
- Sampling the control group of non-completers from individuals who dropped out shortly after starting their MA.
- Sampling the control group of non-completers from individuals who dropped out for reasons unlikely to be related to their later labour-market outcomes (provided there is sufficient number of these individuals).
- Splitting the control group of non-completers based on whether individuals' reasons for non-completion are likely to be positively or negatively related to their later labour-market outcomes.
- Applying regressions techniques to account for additional observed individual characteristics which may differ between the treatment and control groups and at the same time be related to the measured outcomes.

---

<sup>36</sup>

See <http://www.oecd.org/statistics/better-life-initiative.htm>.

184. Although the principal aim of the well-being survey is to examine the impact of MAs on outcomes which are not observed in administrative data, information from the linked administrative data should also be matched to the survey data. This will allow, for example, testing whether observed differences in job satisfaction between the treatment and control groups are mainly due to wage differences between the two groups, or whether the differences in job satisfaction persist even when wage differences are taken into account.

#### *Time horizon*

185. Unlike the evaluation using administrative data, the survey-based evaluation requires a choice of a particular time period after the end of an apprenticeship at which the outcomes are evaluated. 2-4 years after leaving the apprenticeship is likely to be a time horizon that provides an optimal trade-off between being long enough to allow medium-term (although not long-term) effects of MAs to materialize and being short enough to allow high response rates and limit the relative strength of other factors that also influence individual outcomes and can make the effect of MAs hard to isolate with sufficient precision.

#### *Evaluation of impact on firms*

##### *Datasets*

186. Evaluation of the impact of MAs on firms should be based on datasets summarized in Table 2.

**Table 2. Datasets for evaluating impact on firms**

<b>Dataset</b>	<b>Owner</b>	<b>Identifiers</b>	<b>Coverage</b>	<b>Variables</b>
Corporate Training System (CTS)	SDS	employer name & address	MAs	MA information (start, end, completion status, reason for non-completion, framework, location etc.); demographics (age, gender, disability etc.)
Customer Relationship Monitoring (CRM)	SDS	CRN	employers engaged with SDS	name, address, industry, employment
Inter-Departmental Business Register (IDBR)	ONS	Local unit reference number; Reporting unit and enterprise reference number	all firms	name, address, industry, employment, turnover
Annual Business Survey (ABS)	ONS	Reporting Unit reference number	sample of firms	name, address, industry, employment, turnover, wages, capital, R&D
Community Innovation Survey (CIS)	ONS	Reporting Unit reference number	sample of firms	innovation indicators



*Outcome variables*

187. Evaluation of the impact on firms should focus, in particular, on the following outcomes:

- Labour productivity
- Total factor productivity

*Methodology*

188. The key explanatory variable, measuring the involvement of each firm in MAs, should be based on the number of apprentices as a share of all employees. Importantly, it should take into account only those apprentices who have already completed their training, not those currently in training. Given that it is not possible to directly observe the number of MA completers working for each firm, the explanatory variable should be based on the number of former Modern Apprentices that can be expected to work for each firm given

- the number of apprentices the firm trained in each of the previous years
- the apprentice retention rates observed in individual administrative (CTS) and survey (MA Outcomes Survey, well-being survey) data.

189. To establish the causal impact of MAs on firms, the following approaches should be used:

- Using matching techniques (coarsened exact matching or propensity-score matching) to compare firms with high involvement in MAs to firms with low or now involvement in MAs.
- Comparing change in examined outcomes over time between firms that start taking on apprentices and firms that do not.
- Applying econometric techniques to account for additional observed firm characteristics which may differ between firms with different intensity of involvement in MAs and at the same time be related to the measured firm performance.

*Benefit-cost analysis**Measures of economic returns*

190. Building on the microeconomic estimates of the effects of MAs on individuals and firms, the benefit-cost analysis should aim to establish

- net present value per MA start;
- total net present value of MAs;
- net present value per £ of government spending; and
- net present value per £ of overall investment.

*Benefits and costs*

191. The primary benefits for the Scottish economy that should be included in the analysis are

- increase in employment; and increase in productivity (through individual productivity of the apprentices and firm-level productivity effects due to reduced employee turnover, more innovation, better working environment and spillovers from apprentices to other workers).

192. In addition, if feasible, the calculations should also include indirect benefits of MAs such as reducing skill shortages in the economy and preventing undesirable social outcomes (e.g. criminal behaviour, drug abuse or teenage pregnancy).

193. The costs to be considered include primarily

- direct costs of MA training (public funding contributions and training contributions paid by employers);
- indirect costs of MA training (training-induced reduction in the time apprentices spend doing productive work);
- administrative costs of MAs (carried by SDS and potentially also by other public agencies); and
- administrative costs of MAs (borne by employers).

*Presenting results*

194. When the results of the benefit-cost analysis are presented, it should be ensured that

- the general procedure and the assumptions used are described with maximum transparency;
- the analysis includes extensive robustness tests analysing how its results change when assumptions change;
- the key assumptions are always presented alongside the results; and
- preferably, the results are presented as a probabilistic claim regarding a range of values rather than as a single number.

**Medium-term evaluation strategy**

195. Some important components of the long-term evaluation strategy, in particular analysis of the long-term impact on employment and wages, will only be realised after the necessary data linking is set in place. However, even before then, a number of steps can and should be undertaken to (i) facilitate the data linkage and future evaluation activities, (ii) undertake those components of the long-term evaluation strategy which do not depend on the linked administrative data, and (iii) use existing data held by the ONS to provide preliminary answers to questions which will later be answered with the use of the linked administrative data.

### *Preparing data linking*

196. Linking administrative datasets to SDS records, as recommended in the long-term evaluation strategy, is a challenging task that carries significant fixed costs but will generate large long-term benefits by dramatically cutting marginal costs and raising quality of future evaluations of not only MAs but also other Scottish public policies related to skills, education and labour markets. As an important priority, SDS should work together with other key stakeholders, such as the Scottish Government, HMRC and HESA, to establish a broad and robust data linking strategy as soon as possible.

197. In addition to activities that need to be carried out in collaboration with other stakeholders, there are a number of steps which SDS can undertake on its own and which will facilitate the data linkage and future evaluations. These include the following:

1. **Linking existing CTS and CSS records.** CTS records include the National Insurance number, which can be used for linking with HMRC data, and CSS records include the Scottish Candidate Number, which enables a match with the main educational datasets. Hence a match between CTS and CSS records would create a set of individual records where SDS already holds relatively rich information and where identifiers allow high-quality link to both HMRC and educational data. Importantly, this would facilitate the use of individuals who sought career advice from SDS but did not participate in Modern Apprenticeships as a control group for evaluation of MAs, as recommended in the long-term evaluation strategy.
2. **Collecting both National Insurance number and Scottish Candidate Number for all individuals benefiting from SDS-administered services.** Collection of SCN in the records on Modern Apprentices and of NI number in the records on individuals seeking career advice from SDS (so that both identifiers would be available for both groups) would, firstly, facilitate integration of information recorded in the past separately in CTS and CSS and, secondly, allow easier match of information collected by SDS in the future with HMRC and education data.
3. **Collecting IDBR, PAYE and VAT numbers for MA employers.** CTS records currently include employers' names and post codes. Additionally collecting IDBR, PAYE, VAT and Company Registration numbers would facilitate linking SDS data to existing firm-level datasets such as IDBR and ABS.
4. **Allowing employer look-up in a list when filling CTS.** Adding functionality to CTS that would allow looking employer up in a pre-set list is a low-cost ex-ante measure that will dramatically reduce costs and increase precision of analysing SDS records on employers and linking them to external datasets.
5. **Ensuring complete and correct information in CTS.** This includes, for instance, making it clear that the "reason for non-completion" field should not be filled based on the destination unless the destination is actually the reason for non-completion.

### *Well-being survey*

198. The well-being survey is one of the components of the long-term evaluation strategy. Its aim is to evaluate the impact of MAs on outcomes which are not captured in administrative data – most notably, subjective well-being. Although, in the long term, the survey responses should also be linked to variables obtained from administrative data, the survey can be planned, administered and analysed even before the administrative data linkage is in place.

199. In addition to providing information not available in administrative data, the survey could, in principle, also be used in the short term to collect information which will eventually be obtained from the administrative data – in particular, wages. However, collecting and analysing information on wages through the well-being survey requires caution. Firstly, questions related to income may compromise telephone survey respondents' willingness to complete the survey; for this reason, if asked at all, they should be asked in the end of the questionnaire. Secondly, caution is necessary when considering publishing estimates of wage returns to MAs using survey data. If these were published earlier than the corresponding estimates using the linked administrative data, they would likely receive disproportionate attention as the first estimates of their kind for MAs in Scotland, which would not reflect their compromised reliability. Consequently, they might create incorrect perceptions of the benefits of MAs among policy makers', media and public, which would be hard to change even with later, more robust estimates. Last but not least, availability of wage information in survey data could undermine the effort to establish a more reliable and cheaper source of wage information for evaluation through the data linkage. For all these reasons, inclusion of wages in the well-being survey should be seen rather as a second-best or third-best option to draw upon if obtaining wages from existing data sources turns out not to be possible or to take an unacceptably long time.

### *Annual Survey of Hours and Earnings*

200. ASHE has some important advantages for estimating the effect of MAs on wages. Its relatively large sample size means that it is likely to contain information on up to several hundred Modern Apprentices in each year, and more if multiple years of data are combined. Furthermore, it is filled in by employers, and, consequently, it is considered to provide more reliable wage information than any other survey. Moreover, as discussed in Section 0, ASHE includes information on whether an employee has a permanent contract, the employee's job title and job description and an indicator of whether an employee has "direct supervisory or managerial control of any other employee(s)." This information captures relevant outcomes related to career advancement which cannot be observed in other data.<sup>37</sup> Finally and crucially, ASHE can be linked to CTS using the NI number.<sup>38</sup> For these reasons, it is highly desirable that SDS uses linked CTS-ASHE data for evaluating the impact of MAs on wages as well as the outcomes related to career advancement. In the medium term, SDS should consider collaborating with ONS to oversample Modern Apprentices and a suitably selected control group<sup>39</sup> during the ASHE survey collection, and linking ASHE to the Annual Business Survey (ABS) to gain a number of additional firm-level controls for the analysis.

### *Annual Population Survey*

201. An important strength of APS is that its large sample size and its use of an annual sample boost for Scotland means that it is likely to contain information on more former Modern Apprentices in Scotland than any other survey. Unlike ASHE, APS cannot be linked to the records held by SDS, but it contains variables recording whether an individual holds an MA qualification and which apprenticeship levels the individual completed. Hence a self-contained analysis of the effect of MAs on wages and employment using APS represents another option for exploring the effect of MAs on individuals prior to establishment of the administrative data link.

---

<sup>37</sup> Information on permanent versus temporary employment is also asked in LFS but not, it seems, in APS.

<sup>38</sup> Using MA never-starters recorded in CSS as a control group would necessitate linking ASHE also to CSS. As CSS currently does not include the NI number, such linking would have to be based on personal details such as name and post code.

<sup>39</sup> See more on control groups in Section 0 below.

## 9. LEVERAGING EX-ANTE EVALUATION STRATEGY

202. Evaluation should not be seen as a one-off event, isolated from other parts of the policy process. Instead, it will bring largest benefits if it forms an integral part of a continuous policy process and if it is undertaken within an effective governance structure for programme evaluation.

203. Evaluation should represent one of the elements of a process that is aimed at improving public programmes and which involves, in particular, the following elements:

- ex-ante evaluation and ex-ante strategy for ex-post evaluation;
- design and implementation;
- data collection and linking;
- on-going monitoring of intermediate outcomes;
- ex-post evaluation;
- trialling (testing of design changes);
- policy changes/abandonment/renewal; and
- further evaluation of the implemented changes.

204. Effectively establishing this process requires good governance for programme evaluation. On one hand, evaluators should be close enough to the decision-making process so that they focus on the most relevant questions and their findings are used for future changes to the evaluated programme. At the same time, evaluators need to be able to preserve their independence.

205. The rest of this section closes the evaluation framework by emphasizing two elements listed above that are particularly important for ensuring that evaluation is well integrated within the policy process: *i)* ex-ante strategy for programme evaluation and *ii)* trialling.

206. The section first presents the advantages of having an ex-ante strategy for ex-post evaluation and listing several steps that SDS can undertake now to facilitate evaluating MAs in the future. It then illustrates how considering evaluation early in the policy cycle could be useful in the context of the Apprenticeship Levy. Finally, it discusses how trialling based on randomising experiments can be used for identifying ways to improve the design of Modern Apprenticeships.

### **Role of ex-ante strategy for ex-post evaluation**

207. Evaluation should be seen as an inherent part of the policy making process from the very beginning. Thinking early about how an envisaged policy will be evaluated – i.e. developing an ex-ante evaluation strategy – rather than considering evaluation only when a policy has been in place for some time, makes evaluation less costly, more transparent and more powerful by, for instance,

- making it possible to collect the required information through administration of the policy rather than through ex-post surveys, which are more expensive and may be less precise;
- ensuring that the key individual and firm identifiers are recorded, allowing easier, more complete and more accurate linkage to internal and external datasets;
- ensuring that, from start, data is collected or linked not only for the individuals affected by the policy but also for a suitably selected control group;
- ensuring collection of detailed contact information to enable conducting follow-up surveys; and
- where applicable, encouraging design of the policy in ways that allow evaluation approaches based on randomisation, discontinuities in programme design or instrumental variables.

208. In addition to facilitating evaluation, the process of developing an ex-ante strategy for evaluating a policy may make the policy itself more focused, as the need to determine the measured outcomes requires clarification of the policy's objectives. In this light, the effort made by SDS to establish a detailed framework for future evaluations of Modern Apprenticeships is an important step in the right direction.

209. Several further steps that SDS should undertake to facilitate future evaluations of MAs include

- linking records collected by various branches of SDS on all individuals benefiting from SDS-administered services;
- collecting all key individual and firm identifiers for each individual and employer;
- making use of pre-defined lists when collecting administrative data fields (e.g. MA employers);
- ensuring administrative data fields that are important for the evaluation (e.g. recommended career options in CSS, reason for non-completion in CTS) are correctly filled in for all individuals to whom they apply; and
- redefining pre-defined reasons for non-completion in CTS records in a way that makes it clearer whether each answer is (a) likely unrelated to the examined outcomes, (b) likely positively related to the examined outcomes or (c) likely negatively related to the examined outcomes.

### **The case of the Apprenticeship Levy**

210. From April 2017, all UK companies with a wage bill exceeding £3 million will be liable to pay an Apprenticeship Levy (AL) at a rate of 0.5% of their wage bill. Companies in England will be able to reclaim their contributions in the form of digital vouchers to use as a payment for apprenticeship training, but the voucher system will not apply in Scotland. Instead, the Scottish Government will be allocated a part of the AL funds, and it will be able to make an independent decision regarding their use.

211. It is not clear at the moment how the Scottish Government will use the AL funds. However, if the funds are used to finance Modern Apprenticeships, the resulting changes in financing MAs may open new avenues for evaluation, either of the changes themselves or, even more importantly, of MAs. In particular, they could open new options for evaluating the impact of MAs on firms based on, for example,

- random timing of the reform’s introduction across regions or industries;
- differential policy design for firms of different size or age; or
- randomly subjecting different firms to different policy design types (not necessarily clearly ranked ex-ante in terms of their benefits to firms).

212. Provided that the Apprenticeship Levy funds allocated to Scotland result in a change in the way MA training is financed, implications of the change for evaluation should be considered. Making an ex-ante evaluation strategy an important element of the new policy could greatly improve future evaluation of (changes in) MAs and the way they are financed.

### **Trialling: RCT for efficient design**

213. Even though an RCT may not be a suitable approach to evaluating the overall impact of Modern Apprenticeships, this powerful tool can be beneficially applied to comparing different ways in which MAs can be designed and administered. For example, it can be used to compare effectiveness of

- focusing on general versus occupation-specific content;
- allocating more time to on-the-job versus off-the-job training;
- training of different length; and
- different rates of or ways of paying the public funding contributions.

214. For several reasons, RCTs are more suitable for evaluating design features of MAs than for evaluating their overall impact. Importantly, to provide a full picture of the effects of MAs, evaluation of the overall impact needs to be based on the entirety of employers or on a large and representative sample, but the decentralised nature of MAs – recruitment of apprentices is fully in hands of employers or training providers – means that administering an RCT that would be sufficiently representative would be extremely difficult. On the contrary, evaluation of design features, such as the share of training that takes place off the job, is valuable even if it is undertaken only among a small subset of employers. This makes it possible to organise an RCT related to training design in collaboration with a small number of large employers. Such evaluation may be further facilitated by the fact that it is in the interest of large employers, taking on a large number of Modern Apprentices every year, to learn which changes to the training design could strengthen the beneficial impacts of the training.

215. Moreover, using an RCT to evaluate design features is substantially more likely to be acceptable politically than using it for overall evaluation of MAs. This is because rather than randomly assigning who gets an opportunity to participate in MAs, it would only require randomly assigning who gets exposed to what type of training. Since it is not clear ex-ante which type of training leads to higher benefits – learning this is the main motivation for the experiment – it would not entail providing some individuals or firms with (ex-ante) better treatment than others.

216. For these reasons, use of experiments implemented in collaboration with large employers represents a promising way of using evaluation to improve the design and, ultimately, effectiveness of Modern Apprenticeships.

## REFERENCES

- Audit Scotland (2014). "Modern Apprenticeships". [http://www.audit-scotland.gov.uk/docs/central/2014/nr\\_140313\\_modern\\_apprenticeships.pdf](http://www.audit-scotland.gov.uk/docs/central/2014/nr_140313_modern_apprenticeships.pdf).
- Bajgar, Matej, and Chiara Criscuolo (2016a), "Apprenticeships: Impact on Individuals and Firms", OECD.
- Bajgar, Matej, and Chiara Criscuolo (2016b) "Impact of Apprenticeships on Individuals and Firms: Lessons for Evaluating Modern Apprenticeships in Scotland", OECD.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2011), "Field Experiments with Firms", *Journal of Economic Perspectives* 25 (3): 63–82. doi:10.1257/jep.25.3.63.
- Banks, Christopher N. (2010), "Independent Review of Fees and Co-Funding in Further Education in England: Co-Investment in the Skills of the Future", A report to Ministers in the +Department for Business, Innovation and Skills.  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/32317/10-1025-independent-review-fees-co-funding-fe-england.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32317/10-1025-independent-review-fees-co-funding-fe-england.pdf).
- Bibby, David, Franz Buscha, Augusto Cerqua, Dave Thomson, and Peter Urwin (2014), "Estimation of the Labour Market Returns to Qualifications Gained in English Further Education", 195. BIS RESEARCH PAPER. London: Department for Business, Innovation and Skills.  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/383646/Estimation\\_of\\_the\\_labour\\_market\\_returns\\_to\\_qualifications\\_gained\\_in\\_English\\_Further\\_Education\\_-\\_Final\\_-\\_November\\_2014.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/383646/Estimation_of_the_labour_market_returns_to_qualifications_gained_in_English_Further_Education_-_Final_-_November_2014.pdf).
- Brunello, Giorgio, and Maria De Paola (2004), "Market Failures and the under-Provision of Training", 1286. CESifo Working Paper. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=608944](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=608944).
- Department for Business, Innovation and Skills (2014), "Apprenticeships Evaluation: Employers", 204. BIS Research Paper. London.  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/387595/bis-14-1207-Apprenticeships-Evaluation-Employers-December-2014.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/387595/bis-14-1207-Apprenticeships-Evaluation-Employers-December-2014.pdf).
- Dionisius, Regina, Samuel Muehlemann, Harald Pfeifer, Günter Walden, Felix Wenzelmann, and Stefan C Wolter (2009), "Costs and Benefits of Apprenticeship Training. A Comparison of Germany and Switzerland", *Applied Economics Quarterly* 55 (1): 7–37. doi:10.3790/aeq.55.1.7.
- Fraser of Allander Institute (2016), "Evaluating the Impact of the Modern Apprenticeship Programme", Report for Skills Development Scotland.
- HM Treasury (2011), *The Magenta Book: Guidance for Evaluation*.
- Iacus, Stefano M., Gary King, and Giuseppe Porro (2011), "Multivariate Matching Methods That Are Monotonic Imbalance Bounding"? *Journal of the American Statistical Association* 106 (493): 345–61. doi:10.1198/jasa.2011.tm09599.



- Katz, Eliakim, and Adrian Ziderman (1990), "Investment in General Training: The Role of Information and Labour Mobility", *The Economic Journal* 100 (403): 1147–58. doi:10.2307/2233964.
- Lerman, Robert, Lauren Eyster, and Kate Chambers (2009), "The Benefits and Challenges of Registered Apprenticeship: The Sponsors' Perspective", *Urban Institute (NJ1)*.  
<http://eric.ed.gov/?id=ED508268>.
- OECD, and Eurostat (2005), *Oslo Manual*. Paris: Organisation for Economic Co-operation and Development. <http://www.oecd-ilibrary.org/content/book/9789264013100-en>.
- Scottish Government (2007), "Skills for Scotland: A Lifelong Skills Strategy",  
<http://www.gov.scot/resource/doc/197204/0052752.pdf>.
- Scottish Government (2011), "Putting Learners at the Centre: Delivering Our Ambitions for Post-16 Education", <http://www.gov.scot/Resource/Doc/357943/0120971.pdf>.
- Scottish Government (2015), "Scotland's Economic Strategy",  
<http://www.gov.scot/Publications/2015/03/5984>.
- Shury, Jan, David Vivian, Katie Gore, Kamilla Huckle, and Vicki Belt (2012), "UK Commission's Employer Perspectives Survey 2012", Evidence Report 79. UK Commission for Employment and Skills. <https://www.researchonline.org.uk/sds/search/download.do?ref=B29091>.
- Skills Development Scotland (2013a), "Modern Apprenticeships Outcomes Survey 2012",
- Skills Development Scotland (2013b), "Modern Apprenticeship Employer Survey 2012".
- Skills Development Scotland (2015), "Modern Apprenticeship Employer Survey 2015".
- Stevens, Margaret (1996), "Transferable Training and Poaching Externalities, In *Acquiring Skills: Market Failures, Their Symptoms and Policy Responses*, edited by Alison L. Booth and Dennis J Snower. Cambridge, UK: Cambridge University Press.  
<http://www.cambridge.org/us/academic/subjects/economics/labour-economics/acquiring-skills-market-failures-their-symptoms-and-policy-responses?format=PB>.