DfE Update

NPD User Group – 7 July 2021

1. Introduction (Gary Connell)

2. DfE data sharing service (Gwen McGill)

3. Update on iteration 1 of LEO in the ONS SRS (Clare Baker)

4. LEO - Showcase and key analytical questions for DfE (Oliver Anderson)

Introduction

What is the DfE data sharing service and what have we done since the last time we met?

DfE Data Sharing Service and DSAP

Wider than NPD and covers <u>all</u> sharing of personal level data by DfE and its executive agencies (including ESFA, STA and TRA) to any other external organisation.

DSAP governs:

- All shares of individual level data - identifiable data (instant identifiers and unique identifiers), pseudo-anonymised data (meaningless identifiers) and synthetic data
- Data directly shared with organisations (direct transfer) or through Secure Environment Providers (ONS SRS, UKDS)
- Data shared as part of a DfE contract
- Data shared as part of a DfE Research Project
- Data shared with third parties that is to be part of a digital service

By default, all shares of pseudo-anonymised data will be done through the Office for National Statistics Secure Research Service (ONS SRS). Only where evidence can be provided to demonstrate that the SRS is not suitable will an exemption be considered.

What have we done since we last met?

1. **EU Exit:**
   a) Updated all of our data sharing agreements

2. COVID pandemic:
   a) ONS have significantly improved its SRS infrastructure & capabilities
   b) DfE and ONS have met the demand for homeworking researchers
   c) DfE have shared real-time data with other public bodies to support the national effort
   d) DfE has restricted access to attainment data at school level to support govt wishes not to hold schools to account during the pandemic

3. **LEO pilot, DfE has:**
   a) implemented a successful LEO pilot using DfE kit
   b) carried out a successful trial of sharing LEO data through ONS SRS under Digital Economy Act

4. Taking advice from ICO & LEO lessons learnt, DfE :
   a) Has updated its data sharing service vision
   b) Has cleared the NPD request and destruction backlog
   c) Is making some fundamental changes to its data sharing service (more use of DPIAs instead of complex GDPR Qs)
   d) Is updating all of our data sharing agreement templates
   e) Has moved to quarterly publications of external data shares
   f) Has improved the user experience of our <u>Find & Explore</u> digital tool with plans for a more integrated online digital application service

DfE Data Sharing Service

What is our vision? What are we aiming to achieve?

Data Sharing Service – Updated Vision

Where data <u>can</u> be shared through the ONS Secure Research Service under Digital Economy Act, project approval will be managed through the ONS Research Accreditation Service (RAS).

Where data <u>cannot</u> be shared through the ONS Secure Research Service under Digital Economy Act, project approval will be governed by the DfE Data Sharing Service.

This two-tier system will speed up access to DfE data and reduce burdens on researchers and DfE case workers.

Data Sharing Service - roadmap

1. Establish onward data sharing of LEO data through ONS SRS under DEA

2. By early September 2021, all requests for DfE personal data will need to use our new application form and will be required to provide a Data Protection Impact Assessment (DPIA) for their project

3. Build on the success of LEO to provide Growing Up in England (GUIE) to researchers through ONS SRS under DEA

4. Provide other key DfE data assets (e.g. NPD, ILR) to researchers through ONS SRS under DEA

5. Develop a new digital service for an online application system (similar to ONS RAS) for requests for DfE data that are not governed through ONS

Data Sharing Service - benefits

Where data can be shared through ONS under DEA:

1. Much quicker end to end service for researchers to get access to DfE data
2. No GDPR questions and no need for a DPIA as data will be functionally anonymous
3. Simple on-line application service using RAS
4. No need for DBS basic disclosure certificates
5. No need to sign a data sharing agreement

Where data cannot be shared through ONS under DEA:

1. Making better use of DPIAs should streamline the process and provide researchers much quicker access to DfE data

Data Sharing Service – DfE governance - new Application Form & Requirement for a DPIA

What do people apply for?

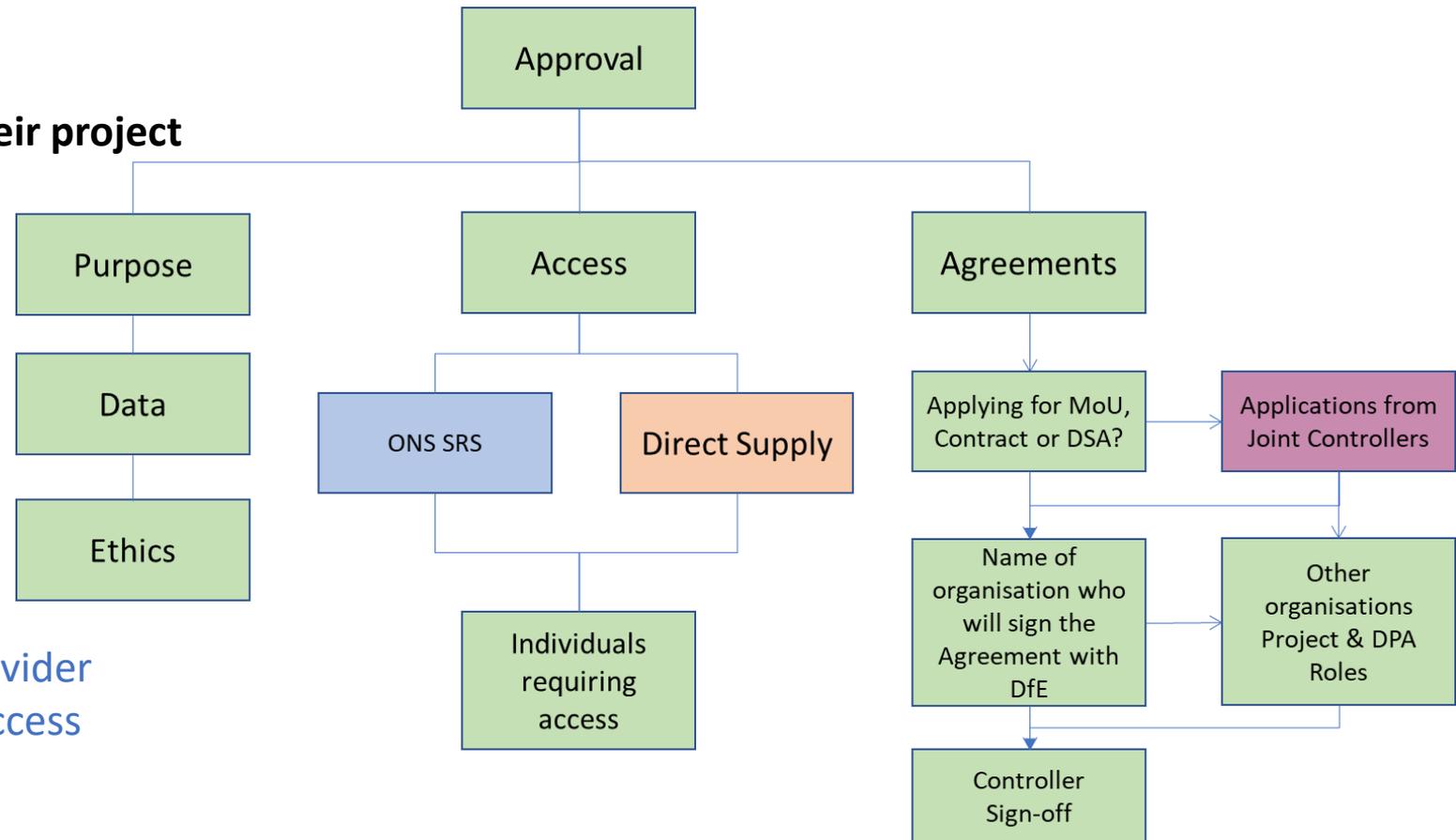1. **Approval** to use DfE personal data for their project
   - Section 1 - Purpose
   - Section 2 - Data
   - Section 3 – Ethics
   - + Evidence - DPIA

2. **Access** to DfE personal data
   - Section 4 - Direct Supply
   - Section 5 - Secure Environment Provider
   - Section 6 – Individuals Requiring Access

3. DfE **Agreement** that documents how they can use the data
   - Section 7 - Data Protection Roles

Green – all applicants;   Orange – those wanting Direct Supply;   Blue – those wanting data through ONS SRS;   Purple – applications from joint controllers.

Access to LEO through the SRS

Update on iteration 1 of LEO access through ONS SRS

LEO - Showcase and key analytical questions for DfE

An example of the potential of LEO using post 16 education and labour market activities, pathways and outcomes and some key analytical questions DfE would like answering using LEO

Post 16 education and labour market activities, pathways and outcomes – LEO showcase analysis

Research report: Post 16 education and labour market activities, pathways and outcomes

Findings in these slides taken from range of analysis focusing on the link between socioeconomic, demographic and education factors and labour market outcomes.

Contains similar analysis on socioeconomic status (free school meals eligibility and income deprivation affecting children index), special educational needs, ethnicity, gender, first language, region, school type and KS4 attainment.

Report published 27 May: [Post-16 education and labour market activities, pathways and outcomes (LEO) - GOV.UK (www.gov.uk)](#)
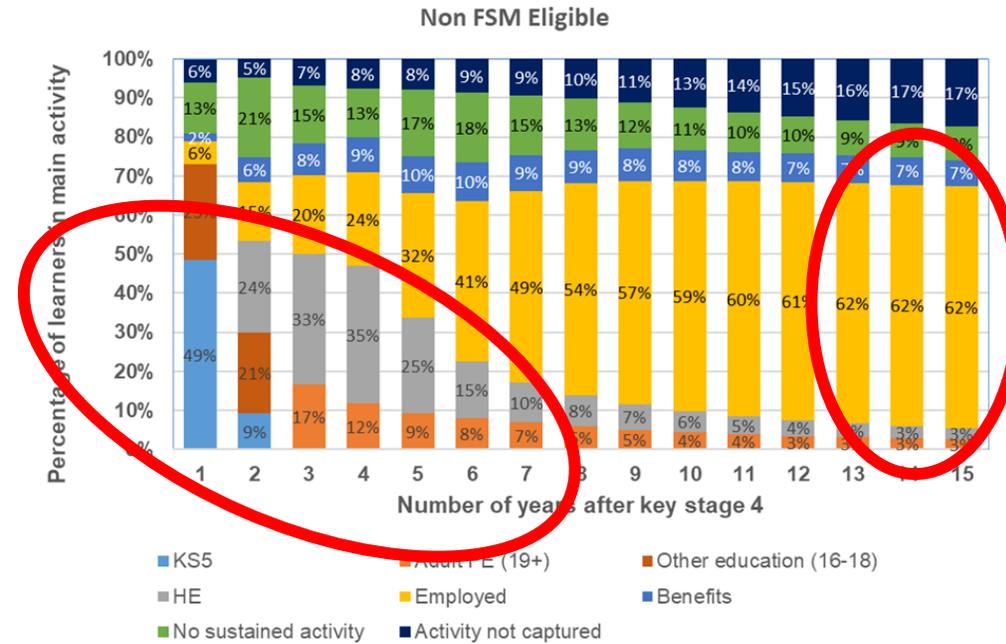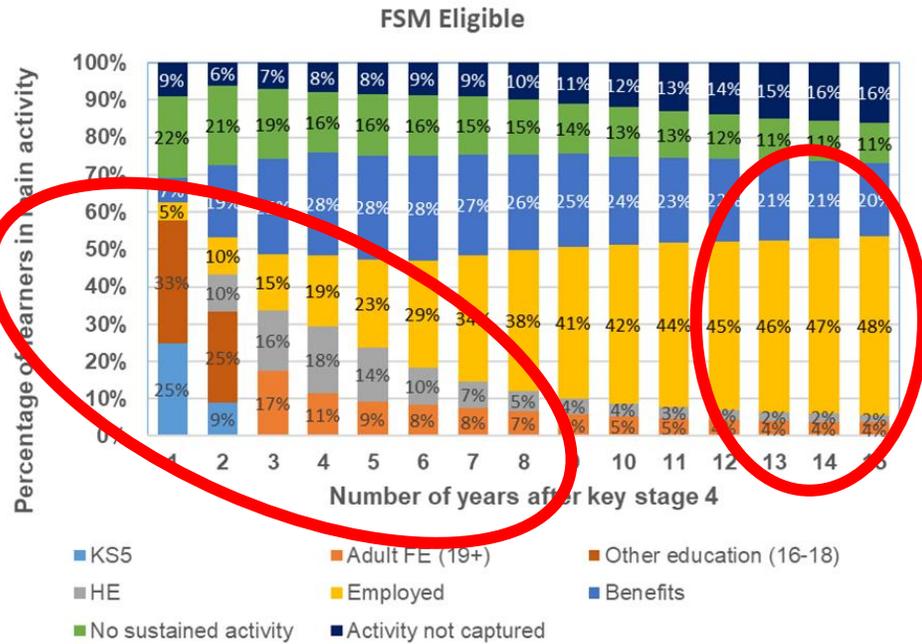
Methodology

This analysis draws upon a dataset of 3.6m individuals who completed their GCSEs between 2002 and 2007 (i.e. Key Stage 4 academic year cohorts 2001/02 to 2006/07). Their education and labour market activities can be tracked over a 10 to 15 (tax) year period. There are three main elements to the analysis:

1.   Main activities: education and labour market activities each tax year GCSEs

2.   Median earnings: trajectories over 15 years for those in employment.

3.   Labour market outcomes year 13 (age 29) by education level: graduates, non-graduates level 3 or above and level 2 or below
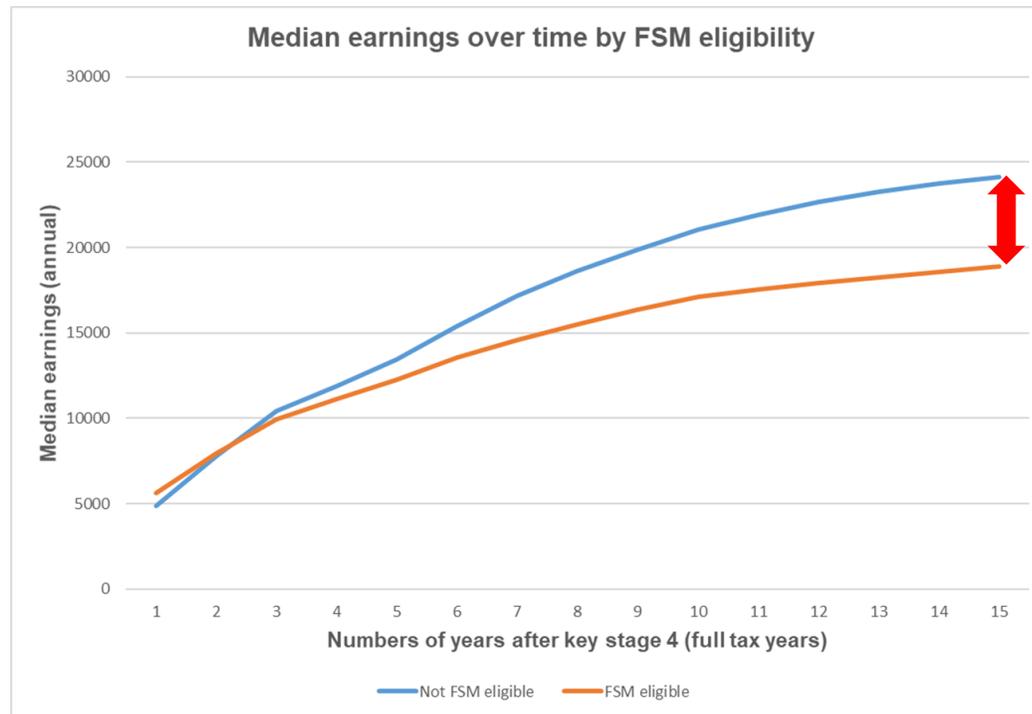
Data

Matched the following datasets:

- Key Stage 4 data from the National Pupil Database (NPD) to form sample frame.

- Socioeconomic and demographic variables taken from schools census (in NPD).

- Matched to Longitudinal Education Outcomes (LEO) data, i.e. HM Revenue and Customs employment and earnings data and Department of Work and Pensions (DWP) out of work benefits data.

- Matched to Higher Education Statistics Agency (HESA) and Individual Learner Record (ILR) data on higher and further education (respectively).

FSM Eligible

Non FSM Eligible

Median earnings over time by FSM eligibility

Around £5,000

Education achievement (by latest data)	Percentage of FSM eligible	Percentage of non-FSM
Degree completion (level 6 or above)	16%	34%
level 3 (non-graduates)	28%	45%

Key analytical questions DfE would like answering using LEO

Key question/research area 1: What is the link between location ('place'), labour mobility and education and labour market outcomes? At similar education levels do those more mobile have better outcomes?

Key question/research area 2: which groups are the most vulnerable or disadvtantaged? Which socioeconomic, demographic and education factors are associated with the poorest labour market outcomes?

Key question/research area 3: What types of education lead to 'good' labour market outcomes for disadvantaged or vulnerable groups? What investments in human capital, pathways and education factors are linked with positive labour market outcomes for those who would not be expected to do well?

Key question/research area 4: Does an academic or vocational route lead to 'better' labour market outcomes? For which groups might doing a vocational route be better VFM than doing an academic route?

Key question/research area 5: What types of education are associated with those with similar characteristics but very different labour market outcomes?

Key research/research area 6: What socioeconomic, demographic and education factors are linked with different labour market outcomes caused by Covid-19?

Key question/research area 7: is higher earnings due to an increase in productivity from education (human capital) or does the education send a signal to employers that those with high ability are more productive?

Key research/research area 8: what are the returns to key stage 5 qualifications? What are early career labour market outcomes and/or estimated lifetime earnings for A-levels and other level qualifications?

Key research/research area 9: What socioeconomic, demographic and education factors are linked with working in different sectors?

Key question/research area 10: What are the skills pipelines for key sectors? What are the characteristics of those working in key sectors? What types of skills and qualifications are needed in key sectors and are enough people studying in those areas?

If you would like to discuss please contact: oliver.anderson@education.gov.uk