



Department  
for Business  
Innovation & Skills

BIS RESEARCH PAPER NUMBER 255

Empirical research on Youth  
Transitions to, and within the labour  
market

Findings of descriptive and  
econometric analyses

OCTOBER 2015

RESEARCH

Prepared by: Stefan Speckesser (IES), Jake Anders (National Institute of Economic and Social Research, NIESR), Augustin De Coulon (King's College), Richard Dorsett (NIESR), Héctor Espinoza Bustos (King's College), Laura Kirchner Sala (IES), Vahé Nafilyan (IES)

INSTITUTE FOR EMPLOYMENT STUDIES

City Gate  
185 Dyke Road  
Brighton BN3 1TL

UK

Tel: + 44 (0) 1273 763400  
Fax: + 44 (0) 1273 763401

[www.employment-studies.co.uk](http://www.employment-studies.co.uk)

In conjunction with:

National Institute of Economic and Social Research  
2 Dean Trench Street  
Smith Square  
London, SW1P 3HE  
Tel: +44 (20) 7654 1917  
[www.niesr.ac.uk](http://www.niesr.ac.uk)

King's College London  
London, SE1 9NH  
Tel: +44 (20) 7848 4530  
[augustin.decoulon@kcl.ac.uk](mailto:augustin.decoulon@kcl.ac.uk)

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

Department for Business, Innovation and Skills

1 Victoria Street

London SW1H 0ET

[www.gov.uk/bis](http://www.gov.uk/bis)

Research paper number 255

October 2015

## The Institute for Employment Studies

The Institute for Employment Studies (IES) is an independent, apolitical, international centre of research and consultancy in public employment policy and organisational human resource issues. It works closely with employers in the manufacturing, service and public sectors, government departments, agencies, and professional and employee bodies. For over 40 years the Institute has been a focus of knowledge and practical experience in employment and training policy, the operation of labour markets, and human resource planning and development. IES is a not-for-profit organisation which has over 30 multidisciplinary staff and a network of international associates. IES expertise is available to all organisations through research, consultancy, publications and the Internet. For further details please visit: [www.employment-studies.co.uk](http://www.employment-studies.co.uk)

## Acknowledgements

We would like to acknowledge the useful guidance and feedback provided by the Department for Business, Innovation & Skills on earlier versions of this research and throughout the project. We are particularly indebted to Matthew Bursnall, who managed the research project related to this report in BIS, and Sara Davies and Peter Blyth, who engaged with helpful feedback on some of the methodological issues at various stages of the research. We would also like to thank Karen Patient in IES for her help improving the presentation of findings and readability of this report. Of course, the authors have sole responsibility for the contents of this report.

# Contents

<b>Figures and Tables .....</b>	<b>7</b>
<b>1 Introduction .....</b>	<b>9</b>
1.1 Background .....	9
1.1.1 Policy background .....	9
1.1.2 Aims and objectives of the project .....	9
1.1.3 This report .....	10
<b>2 Long-term trends of young people’s labour market transitions and impact on adult labour market trajectories .....</b>	<b>11</b>
2.1 Objectives.....	11
2.2 Descriptions of pseudo-cohorts.....	11
2.2.1 Method .....	11
2.2.2 Findings.....	12
2.2.3 Gender differences .....	14
2.3 Econometric estimation of ‘scarring’ of youth unemployment.....	14
2.3.1 Method .....	14
2.3.2 Findings.....	15
2.4 Econometric estimates based on the ‘Ad hoc module’ 2009 on youth transitions.....	17
2.4.1 Data.....	17
2.4.2 Cohort descriptions.....	17
2.4.3 Observed transition from education to employment .....	19
2.4.4 Analysis of drivers of the duration of education to work transitions .....	21
2.4.5 Analysis of long-term effect of initial transitions.....	22
2.4.6 Summary of the findings .....	23

<b>3</b>	<b>Sequence analysis of transitions of young people from different cohorts.....</b>	<b>25</b>
3.1	Objectives.....	25
3.2	Method .....	25
3.3	Findings.....	26
3.4	Extension of the sequence analysis: A) NCDS and BCS .....	29
<b>3.4.1</b>	Introduction.....	29
<b>3.4.2</b>	Entering the Labour Market .....	30
<b>3.4.3</b>	Accumulating Human Capital.....	31
<b>3.4.4</b>	Summary Statistics of Long Groupings .....	34
<b>3.4.5</b>	Cross-tabulation of shorter- and longer-sequence based groupings .....	34
<b>3.4.6</b>	Conditional associations between characteristics and being classified 'potential cause for concern' .....	35
<b>4</b>	<b>A detailed analysis of activities and transitions of the 2010/11 KS4 leavers .....</b>	<b>37</b>
4.1	Objectives.....	37
4.2	Data description.....	37
<b>4.2.1</b>	Relevant samples .....	37
<b>4.2.2</b>	Representativeness of the balanced panel for full NPD samples .....	40
<b>4.2.3</b>	Categorisations of young people's activities.....	42
<b>4.2.4</b>	Most common patterns of youth transitions.....	45
<b>4.2.5</b>	Treatment of the category 'Current situation not known' .....	47
<b>4.2.6</b>	Comparison of monthly and quarterly activities.....	49
<b>4.2.7</b>	Extent of churning over the age 16 to 19 .....	51
<b>4.2.8</b>	Mobility between schools and transitions between activities post 16.....	51
<b>4.2.9</b>	Extent and duration of the NEET state.....	52
4.3	Analysis of individual transitions between states.....	53

4.3.1	Description.....	53
4.3.2	Estimates of duration models.....	54
<b>5</b>	<b>Summary and recommendations for further research .....</b>	<b>60</b>
5.1	Summary .....	60
5.2	Policy conclusions .....	61
5.3	Further research .....	61
<b>6</b>	<b>References.....</b>	<b>63</b>

# Figures and Tables

Figure 2.1	Labour market and employment participation of selected birth cohorts .....	13
Figure 2.2	Employment rates for selected birth cohorts .....	14
Figure 2.3	Effects of high youth unemployment on labour market trajectories .....	16
Figure 3.1	Sequence substitution example .....	25
Figure 3.2	Entering the labour market.....	26
Figure 3.3	Accumulating Human Capital.....	27
Figure 3.4	Potential Cause for Concern.....	28
Figure 3.5	Impact of characteristics on ‘cause for concern’.....	29
Figure 3.6	Entering the Labour Market .....	31
Figure 3.7	Accumulating Human Capital.....	32
Figure 3.8	Potential Cause for Concern.....	33
Figure 3.9	Average marginal effects of age 16 characteristics on being in a cluster that is ‘potential cause for concern’ .....	36
Figure 4.1	Proportion of ‘Current situation not known’ (September 2011 – April 2014) .....	47
Table 2.1	Age groups/cohorts covered in 2009 EU:LFS AHM	18
Table 2.2	Highest level of education of cohorts in 2009 (% of cohorts).....	18
Table 2.3	Orientation of the highest level of formal education attained (% of cohorts) .....	19
Table 2.4	Year when education was left for the last time .....	20
Table 2.5	Mean duration (months) between leaving education and starting first significant employment	21
Table 2.6	Hazard rate estimates (exit to start first significant employment).....	22
Table 2.7	Probit estimates on employment status in 2009 (All cohorts) .....	23
Table 3.1	Summary Statistics .....	34
Table 3.2	NCDS: Cross-tabulation of groupings on basis of 29 months of sequence analysis and of groupings on basis of 98 month sequence analysis .....	35

Table 3.3	BCS: Cross-tabulation of groupings on basis of 29 months of sequence analysis and of groupings on basis of 98 month sequence analysis .....	35
Table 4.1	Full NPD data combined with NCCIS information, for the cohort born between 1 September 1993 and 31 August 1994 (who passed their GCSEs in May 2011) .....	38
Table 4.2	Full NPD data combined with NCCIS information, adjusted for those with 30 and 31 months and identical activity before and after the missing months .....	39
Table 4.3	Number of good GCSEs (A*-C) in the balanced panel compared with the NPD and the fully merge NPD-NCCIS samples.....	40
Table 4.4	Region in the balanced panel compared with the NPD and the fully merge NPD-NCCIS samples .....	41
Table 4.5	Ethnicity and samples .....	42
Table 4.6	Recoding of NCCIS activities (September 2011 – April 2014).....	43
Table 4.7	Month by month activities, % (September 2011 – April 2014) .....	44
Table 4.8	Most common patterns (September 2011 – April 2014), without adjustments on the category [7] Situation unknown .....	46
Table 4.9	Most common patterns (September 2011 – April 2014), when the category 'Current situation not known' [7] is replaced by the same activity before and after.....	48
Table 4.10	Number of spells experienced by young people from 16 to 19.....	49
Table 4.11	Main activities by quarters, percentages .....	50
Table 4.12	Number of spells (using quarters) and performance at 16.....	51
Table 4.13	Number of spells (using quarters) and performance at 16 for those who experienced at least one spell as NEET .....	51
Table 4.14	Number of quarters in NEET .....	52
Table 4.15	Number of quarters in NEET and performance at GCSE .....	53
Table 4.16	Transition between states in the full sample (all transitions).....	54
Table 4.17	Transition out of education to employment, training, NEET and Unknown .....	54
Table 4.18	Transition from education to employment and training, and from training to NEET and employment .....	56



# 1 Introduction

## 1.1 Background

### 1.1.1 Policy background

In 'Building Engagement, Building Futures: Our Strategy to maximise the Participation of 16-24 year Olds in Education, Training and work' HM Government set out priorities and strategies for supporting young people, including those at risk of disengagement, to succeed in learning and careers (December 2011). Since then, a range of policies have been implemented to increase educational achievement, help young people to improve their skills, provide them with meaningful experiences to develop and demonstrate skills and to increase their labour market success. In doing this, there is a recognition that helping young people is about more than just skills acquisition; it forms the main stage of human capital investment in life trajectories and is a precondition to adult roles, including economic independence and family formation.

### 1.1.2 Aims and objectives of the project

In the context of these policies, this project researching *Youth Transitions to, and within the Labour Market* aimed to achieve:

1. A better understanding of movements and progression within the youth labour and learning sectors, answering such questions as:
  - What types of learning and employment are young people involved in;
  - How long do those episodes last and how often do people switch from one type to another;
  - Do some transitions or movements lead to better outcomes than others or lead more swiftly to a good outcome than others
  - What influences young people's movements and progression and what is the relative influence of these factors across the different groups.
  - Is it possible to identify different groups of young people who are affected in different ways and the characteristics of these different groups?
2. Answer to key policy questions including, but not restricted to:
  - What pathways and transitions are most likely to support young people into work?
  - What pathways and transitions are most likely to support young people into a better job?
  - How do different groups within the youth labour market move into work and progress?

3. Providing a linked dataset that is capable of being used in the future to answer a range of policy questions relating to how patterns in youth transitions have changed in the last 30 years.

Methodologically, we focused on econometric analyses of combined large-scale datasets, which can provide both descriptions of youth transitions (between the age of 16 and 24) and longer-term outcomes, as well as hindsight on the drivers of these transitions and how they subsequently affect longer-term labour market and life trajectories. To some extent, this research also took into account the wider view on people's biographies by linking the experiences and achievements in primary and secondary education to transitions in adolescent life and education and labour market trajectories in adult life.

### 1.1.3 This report

The research project aimed to provide research outcomes on the **magnitude of youth transitions**, the **main drivers and barriers of youth transitions** and **potential long-term outcomes** ('scarring') on people's subsequent employment trajectories. This report summarises the evidence obtained by making use of a range of individual-level data sets and methods, in particular:

- An analysis of Labour Force Survey (LFS) data creating pseudo-cohorts to examine **long-term education and labour market trends** affecting 16-to-24 year olds from 39 different birth cohorts and **subsequent employment trajectories**.
- An analysis of the 'Ad Hoc module' of 2009 from the European Labour Force Survey (EU:LFS) on 'Youth Transitions' on **individual long-term outcomes of particular youth labour market transitions**.
- **An analysis of various cohort studies** (National Child Development Study [NCDS], British Cohort Study [BCS], Youth Cohort Study [YCS], Longitudinal Study of Young People in England [LSYPE]) **comparing the episodes young people experience when making transitions into the labour market in youth and early adulthood**. This study uses sequential analysis for individual monthly panel data to describe biographies of young people until the age of 25 (based on BCS and NCDS) and until the age of 19 (for YCS and LSYPE).
- **An analysis of the transition from secondary schooling to further destinations using a recent cohort of School Leavers** National Pupil Data (NPD) merged to records of National Client Casework Information System (NCCIS) on young people's activities after the end of compulsory education including econometric models on drivers and barriers of particular transitions.

## 2 Long-term trends of young people's labour market transitions and impact on adult labour market trajectories

### 2.1 Objectives

The analysis carried out with Labour Force Survey (LFS) data for the UK had three main objectives, namely to:

- Describe education and labour market activities between 16 and 24 of cohorts born between 1959 and 1997, and to provide a picture of changing activity patterns of young people entering the labour market over the last forty years.
- Estimate econometric models (separately by gender and controlling for educational composition and macroeconomic time trends) to understand how much labour market trajectories of (pseudo-)cohorts are affected by differences in youth unemployment experienced at the time of labour market entry.
- Use the 'Ad Hoc module' of 2009 on 'Youth Transitions' for further pseudo-cohort analyses on the long-term outcomes of particular youth labour market transitions using duration and multivariate regression modelling.

### 2.2 Descriptions of pseudo-cohorts

#### 2.2.1 Method

LFS data were collected bi-annually from 1975 throughout Great Britain until 1983, then annually and since 1992 quarterly, covering all countries of the UK. We focus on pseudo-cohorts, which provide average characteristics of people belonging to *particular birth cohorts*, for example 1959 representing the cohort of people entering the labour market or post-compulsory education in 1975 (at age 16). We provide extensive descriptions of time, cohort and age effects. More particularly, we show how education and labour market activity in early life and adult employment trajectories differ for people of different birth cohorts:<sup>1</sup>

- Cross-section data of 16-year olds entering the labour market in different years.
- Longer-term labour market outcomes in adult years based on cross-sections of the population from the same birth years in later years.

---

<sup>1</sup> Synonymous to pseudo-cohorts here

## 2.2.2 Findings

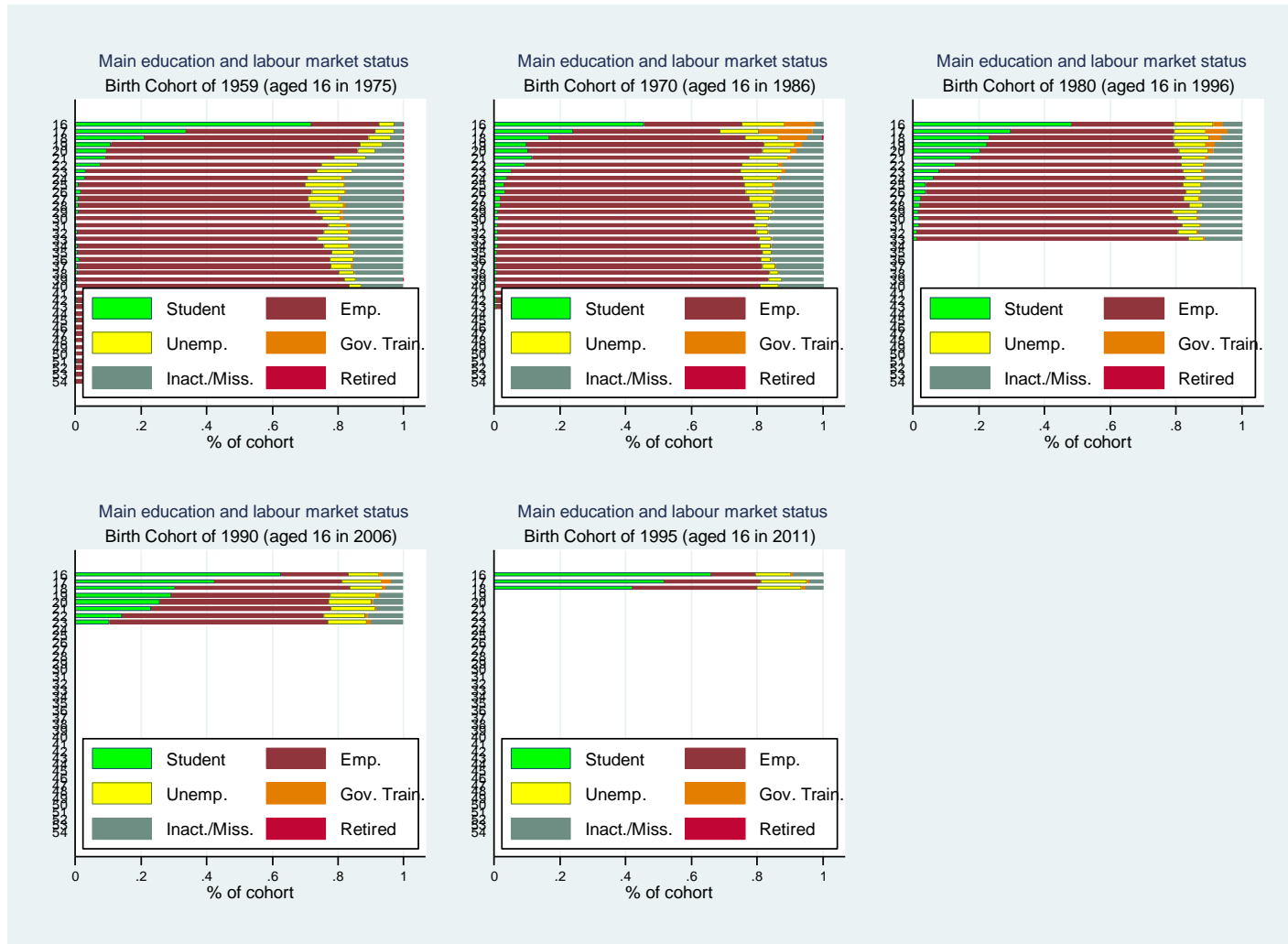
### ***Timings and levels of education attainment***

- Across cohorts, the share of individuals with higher qualifications increases and the group with qualifications below Level 2 is substantially reduced.
- Within cohorts, we observe an up-skilling process with age which differs across cohorts: a decline of the below Level 2 and semi-skilled over the life course in early cohorts, while for birth cohorts from mid-70's onwards, up-skilling affects primarily above Level 3 – this means that from about the early 20's, the composition of the low-skilled end of the labour market remains practically unchanged.
- Gender differences: a major change after the 1980s cohorts (women's share of higher education is now larger than men's and below Level 2 qualification's share substantially reduced).

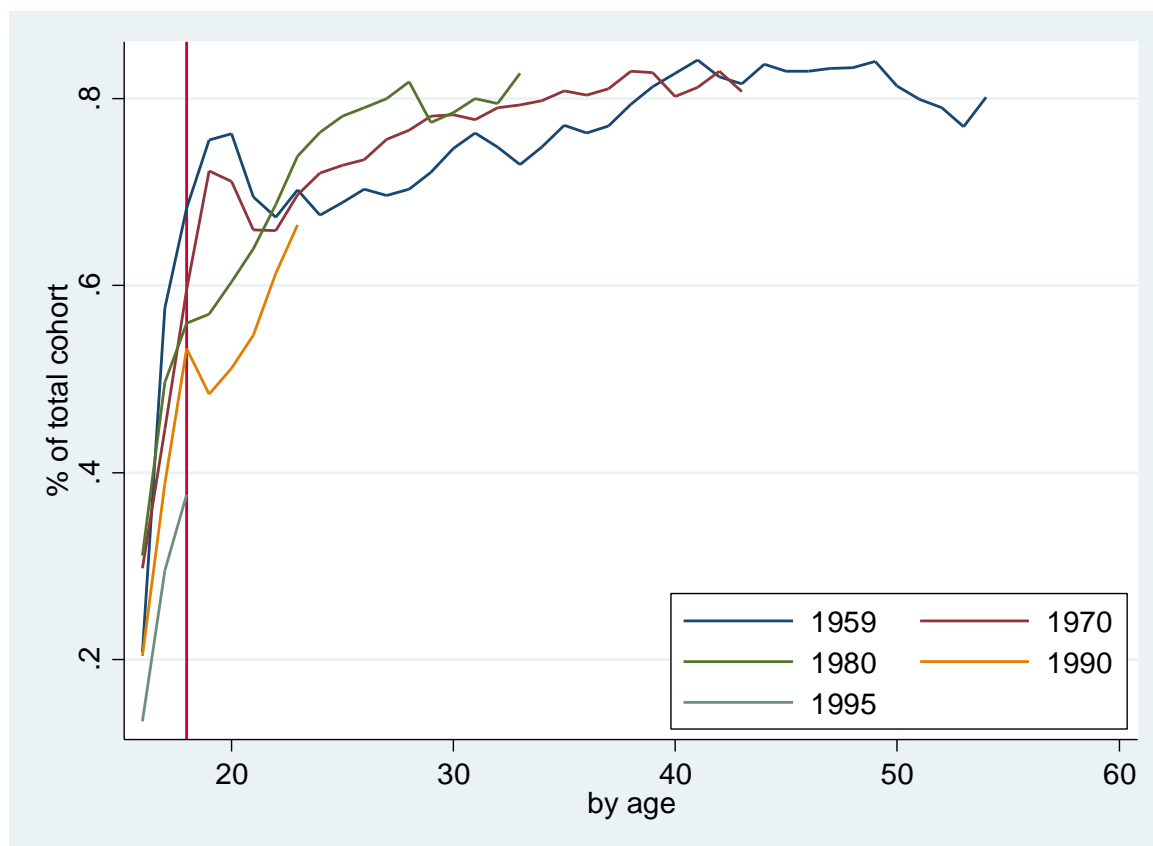
### ***Employment and unemployment***

- We observe an extension of participation in education, indicating that young people are better educated when entering the labour market (Figure 2.1).
- From 1986 (when 1970s cohort is aged 16) youth unemployment rates remain high even in 'good' economic circumstances and indicate a structural problem of the disappearance of the youth labour market providing entry level jobs and career progression.
- With progressing age, we observe a clear pattern of employment rate convergence in life course trajectories across cohorts despite very different initial unemployment experiences (Figure 2.2).

**Figure 2.1 Labour market and employment participation of selected birth cohorts**



Source: LFS and own calculations

**Figure 2.2 Employment rates for selected birth cohorts**

Source: LFS and own calculations

### 2.2.3 Gender differences

- The recessions of the early 1980s and 1990s affected both genders differently. We observe an increase in the male unemployment rate, whereas women reacted by leaving the labour market and so are not captured in the unemployment rate.
- The high share of inactive women during their twenties and thirties because of childbearing and family care (Cohort of 1959) decreases for later cohorts (1970).
- Government training participation decreases for later cohorts, especially for women.

## 2.3 Econometric estimation of ‘scarring’ of youth unemployment

### 2.3.1 Method

One of the key research questions of this project was to understand whether the experience of unemployment in early age has long-lasting effects on labour market trajectories in adult lives. Estimating such labour market ‘scarring’ resulting from youth unemployment is not possible without further assumptions because a simultaneous estimation of cohort (high/low initial unemployment), trajectory (life-course/age) and macroeconomic trend effects is not possible. Time is the sum of cohort and age. Hence, an econometric model estimating all three would suffer from perfectly correlated variables

and it would not be possible to estimate cohort effects and the life-trajectory when controlling for macroeconomic trends.

Therefore, we estimate effects of scarring in relation to *several* cohorts of young people entering the labour market in difficult economic conditions,<sup>1</sup> which are to some extent exogenous to the labour market entry of particular cohorts. For the practical estimation of the model, we pool all pseudo-cohorts for all years with available information on education into a merged panel data set. The properties of the data set are unbalanced (birth cohort 1997 is represented with only one year, while 1972 is represented with 26 years, etc.) and the overall structure of the panel is further reduced as education variables have only been consistently available from 1987. Nonetheless, the regression is based on a reasonably large data set of almost 700 observations.

### 2.3.2 Findings

We find significant effects of initially bad labour market conditions on subsequent life course trajectories and confirm the ‘scarring’ of youth unemployment found in microeconomic estimates benefitting from much richer NCDS data (Gregg 2001). We show the scarring effect by plotting the different employment trajectories predicted in the models – similar to a Mincer (1975) model earnings profile – in relation to the age of people (with confidence intervals and separately for men and women) (Figure 2.3).

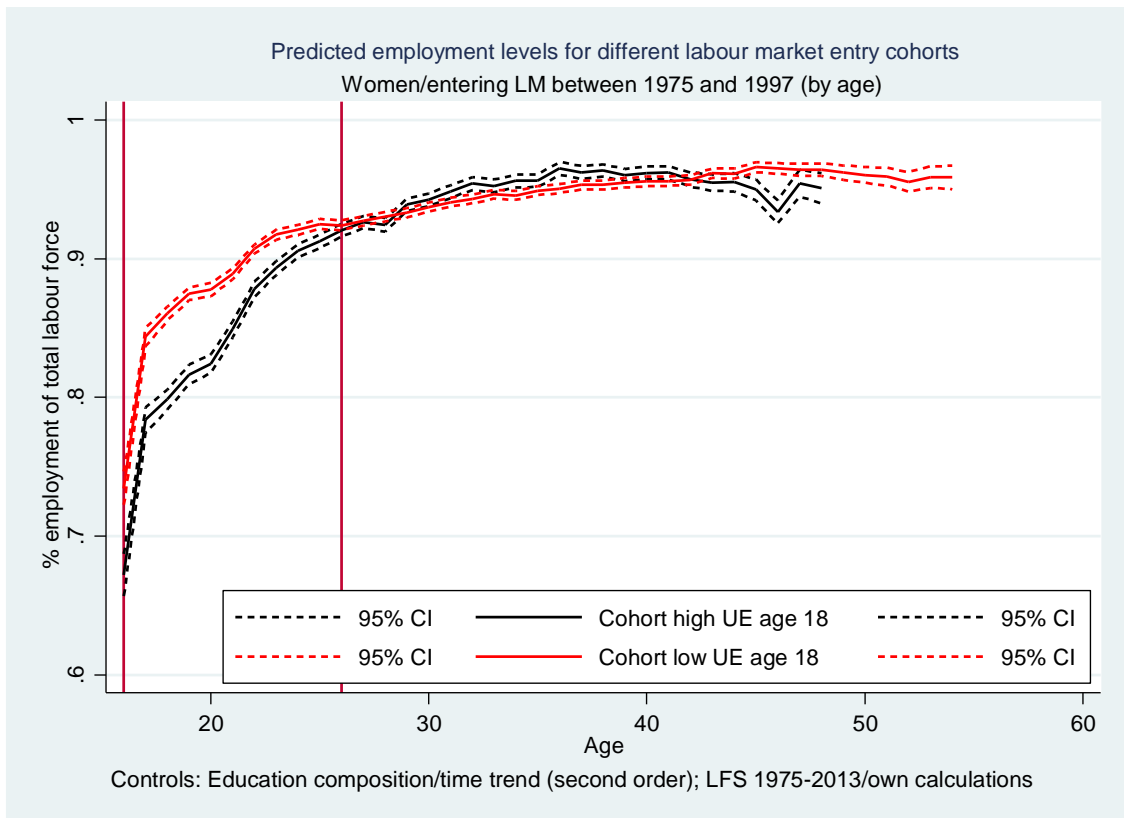
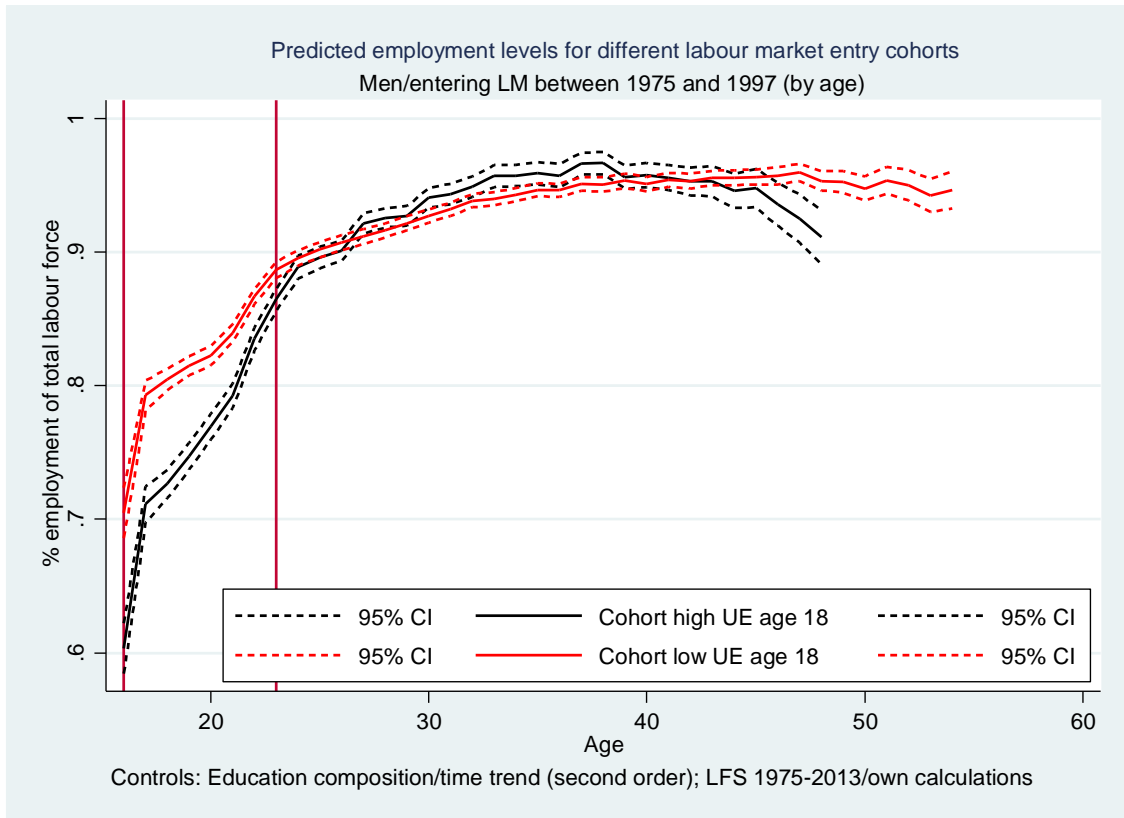
Findings of labour market scarring effects:

- Trajectories differ in *early* adult life (Male  $\leq 22$ /Female  $\leq 26$ ).
- By mid-/end-20, trajectories converge, youth unemployment differences become irrelevant for trajectory.
- Effects less substantial for men (as opposed to Gregg 2001), but this could result from not including inactivity (which Gregg can do because of more informative labour market status information).

---

<sup>1</sup> These are modelled as periods of high youth unemployment (i.e. unemployment ratios of 18 year olds at or above the 75 percentile of the 1975-2013 unemployment ratio for this population).

**Figure 2.3 Effects of high youth unemployment on labour market trajectories**



Source: LFS and own calculations



## 2.4 Econometric estimates based on the 'Ad hoc module' 2009 on youth transitions

### 2.4.1 Data

The UK Labour Force Survey forms part of the annual European Union Labour Force Survey (EU:LFS) covering the population in private households in currently 33 European countries with a sample size of about 4,000,000 individuals (2008).<sup>1</sup> These data provide principal employment indicators on the EU labour market and include further 'Ad hoc modules' since 1999 (AHM in the following), covering different topic areas, including occupational health, working time and life-long- learning.

Young people's transitions into the labour market have been covered with two AHMs:

- 2000: Transition from school to working life
- 2009: Entry of young people into the labour market.

The following analysis uses the AHM of 2009 available from the UK Data Archive (GN 33246). The early AHM on youth transitions (from the year 2000) is currently not available from the archive and would require a data share agreement with Eurostat.

### 2.4.2 Cohort descriptions

The 2009 AHM collects information for everyone under the age of 35 at the time of the interview. This corresponds to birth cohorts between 1975 and 1994, leaving compulsory education at KS4 between 1991 and 2010. As shown in Table 2.1 below, these cohorts enter the labour market in different conditions of the economic cycle and currently represent about 16 million people in the UK. Since the EU:LFS as available from the archive only offers age in Bands as shown in the table, a specific year of the cohorts when turning 16 cannot be derived.

---

<sup>1</sup> <http://www.gesis.org/en/services/data-analysis/official-microdata/european-microdata/eu-lfs/about-the-eu-lfs/>

**Table 2.1 Age groups/cohorts covered in 2009 EU:LFS AHM**

Age	Birth cohorts	Year when leaving KS4	Economic cycle	Freq.	Per cent
15-19	1990-94	2006-2010	Expanding, then contracting	3,876,285	24
20-24	1985-89	2001-2005	Expanding	4,201,910	26
25-29	1980-84	1996-2000	Expanding	4,181,353	26
30-34	1975-79	1991-1995	Contracting, then expanding	3,804,846	24
<b>Total</b>				<b>16,064,395</b>	<b>100</b>
<i>Base</i>				<i>21,926</i>	

Source: EU:LFS AHM 2009

A description of levels of education of the cohorts shows the usual age-related differences in highest education levels (Table 2.2). Based on ISCED-Codes, which can be approximately compared to Levels used in national statistics, there is a clear tendency of older cohorts to have completed education and achieved higher level qualifications, in particular academic qualifications.

**Table 2.2 Highest level of education of cohorts in 2009 (% of cohorts)**

	1975-79	1980-84	1985-89	1990-94	Total
No formal or below ISCED-2	0.3	0.4	0.4	0.8	0.3
ISCED-2 [~Levels 1 and 2]	10.0	9.2	8.6	12.6	10.0
ISCED-3 [~Level 3]	46.4	50.8	65.3	65.2	46.4
ISCED-4 [~Post-secondary/HE]	41.7	37.9	24.0	1.4	41.7
Not applicable	1.6	1.8	1.6	20.0	1.6
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<i>Base</i>	<i>5,528</i>	<i>5,234</i>	<i>5,102</i>	<i>6,062</i>	<i>21,926</i>

Source: EU:LFS AHM 2009

A further variable describes the orientation of the highest level of formal education, i.e. whether vocational or general (Table 2.3). For this variable, later cohorts show predominantly general education, but this could also be an age effects in that work-based education for the younger cohorts is currently undertaken or in subsequent employment spells. For those cohorts more likely to have completed most of their education investment (1975-79, 1984-84 and 1985-89), the table shows higher level general education, but only small increases in the proportion of people with vocational education obtained in schools or in the workplace.

**Table 2.3 Orientation of the highest level of formal education attained (% of cohorts)**

	1975-79	1980-84	1985-89	1990-94	Total
General education	21.3	23.6	42.0	57.0	35.9
Vocational education mainly (or solely) school based	3.1	3.8	3.6	1.9	3.1
Vocational education mainly work place based	14.6	15.3	15.4	5.7	12.8
Vocational education, with no distinction possible between 2 and 3	2.9	2.6	1.5	0.4	1.8
Not applicable	43.6	39.9	25.9	22.2	32.8
Missing	14.6	14.8	11.6	12.9	13.5
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<i>Base</i>	5,528	5,234	5,102	6,062	21,926

Source: EU:LFS AHM 2009

### 2.4.3 Observed transition from education to employment

Table 2.4 below shows the year when the education system was left for the last time. As a consequence, the table shows fewer transitions than were actually undertaken as people, who had left the education system initially to employment, but subsequently re-started an education status would not show a leaving date. The AHM 2009 therefore also reports many more first significant work experiences (of jobs of more than three months of duration) than people with an observed leaving date from the education system. However, the analysis of the duration of the school-to-work transition is only possible for people showing an exit from education and a subsequent entry into employment. In this, the outcome reported in this data set is biased as those not entering stable employment are more likely to return to education and therefore, will not show a date when education was left for the last time.

Apart from this, Table 2.4 shows plausible transition patterns for the different cohorts. Most people from the 1975-1979 birth cohort show education leaving dates in the 1990's. These transitions are related to the final leaving of education and are therefore for the great majority of people related to the formal level of education observed in 2009 shown above. In the subsequent analysis, the level of education will therefore have to be controlled for.

**Table 2.4 Year when education was left for the last time**

	<b>1975-79</b>	<b>1980-84</b>	<b>1985-89</b>	<b>1990-94</b>	<b>Total</b>
Before 1990	1.2	0.3	0.1	0.0	0.4
1990	2.3	0.3	0.0	0.0	0.6
1991	5.0	0.1	0.1	0.0	1.2
1992	6.0	0.1	0.0	0.0	1.4
1993	7.1	0.2	0.1	0.0	1.8
1994	7.3	0.5	0.1	0.0	1.9
1995	7.1	1.7	0.1	0.0	2.2
1996	6.8	5.5	0.1	0.0	3.1
1997	6.0	6.1	0.1	0.0	3.0
1998	5.5	8.1	0.2	0.0	3.5
1999	5.0	8.3	0.5	0.0	3.5
2000	5.0	7.4	2.1	0.0	3.7
2001	3.8	6.5	4.7	0.0	3.8
2002	2.9	6.0	6.1	0.0	3.9
2003	1.9	6.0	8.0	0.0	4.1
2004	1.6	6.0	8.6	0.2	4.2
2005	1.2	5.2	8.1	0.9	4.0
2006	1.1	4.1	7.1	3.6	4.0
2007	1.0	2.9	7.7	5.1	4.2
2008	0.9	2.4	6.7	6.6	4.2
2009	0.6	1.0	4.7	3.5	2.5
Not applicable	20.7	21.5	34.9	79.9	38.9
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<i>Base</i>	<i>5,528</i>	<i>5,234</i>	<i>5,102</i>	<i>6,062</i>	<i>21,926</i>

Source: EU:LFS AHM 2009

For people, who left the education system and did not return, the data set offers information about the duration it took (in months) from leaving until starting the first employment of a duration of more than three months ('First significant employment', Table 2.5). This description shows that the mean duration was higher for the 1975-1979 than for any of the subsequent cohorts (7.6 months compared to somewhere between 5 and 6 months).

**Table 2.5 Mean duration (months) between leaving education and starting first significant employment**

	Mean	Std. Dev.	Min	Max	N	Base
1975-79	7.61	21.05	0	193	755,966.78	1,139
1980-84	5.29	13.30	0	134	807,009.49	1,063
1985-89	5.85	14.04	0	194	675,670.06	857
1990-94	5.39	16.10	0	216	135237.4	190

Source: EU:LFS AHM 2009

#### 2.4.4 Analysis of drivers of the duration of education to work transitions

For those people, who left the education system and did not return subsequently to employment, the duration of the education-to-work transition can be analysed as well as the outcome of the duration of the initial school-to-work transition on later life. We do this in the following based on the time in months from leaving education for the last time and starting the first job of a duration of more than three months.

To do this, we reorganise the data set into a monthly panel and estimate a simple discrete time proportional hazard rate model explaining the duration of the initial transition by a set of covariates. We use a discrete logistic (proportional odds) model as shown in Jenkins (2006), with a simple binary dependent variable of entering the first significant employment by individual  $i$ 's in any of the spell months (month 1,...,  $T_i-1$ ), showing the value of one for the last month (month  $T_i$ ). The functional form for the baseline is the logged survival time until  $t$  as the only time-varying covariate. The estimation additionally includes further characteristics (education achievement, age, gender, cohort and work experience when people were in education).

The results in Table 2.6 below show the significantly higher probability of people leaving from education to the first significant employment when:

- education was primarily work-place based
- people had work experience in while being in education
- education of at least level 2 (ISCED-3) was achieved.

There are significant cohort effects indicating generally reduced transition rates for the 1975-79 cohorts. There is also a negative gender effect, indicating that women are less likely to make successful transitions.

**Table 2.6 Hazard rate estimates (exit to start first significant employment)**

	Coefficient	Std. Err.	z	P>z
<b>Education (Base: General)</b>				
Vocational education mainly (or solely) school based	0.111	0.110	1.010	0.314
Vocational education mainly work place based	0.178	0.050	3.600	0.000
Vocational education, no distinction possible between 2, 3	-0.095	0.122	-0.780	0.436
Not applicable	-0.071	0.716	-0.100	0.921
<b>Work experience while in education (Base: No work or work less than 1 month per year)</b>				
Work (only) as part of educational programme	0.444	0.076	5.850	0.000
Work while studying but outside educational programmes	0.211	0.046	4.570	0.000
Work (only) during an interruption of studies	0.147	0.097	1.520	0.129
<b>Cohort (Base 1975-179)</b>				
1980-84	-0.117	0.056	-2.090	0.036
1985-89	-0.385	0.079	-4.900	0.000
1990-94	-0.513	0.122	-4.190	0.000
<b>Sex (Base: Male)</b>				
Female	-0.161	0.038	-4.210	0.000
<b>Qualification (Base: ISCED-2 and below)</b>				
ISCED-3	0.556	0.321	1.730	0.083
ISCED-4	0.520	0.786	0.660	0.508
Logged elapsed survival time	-0.618	0.018	-	0.000
			34.070	
Time trend	0.252	0.081	3.100	0.002
Time trend ^2	-0.016	0.006	-2.650	0.008
Time trend ^3	0.000	0.000	2.750	0.006
Cons	-2.666	0.468	-5.690	0.000
Number of observations	17294			
LR chi2(17)	1902.85			
Log likelihood	-6811.0017			
Probability > chi2	0			

Source: EU:LFS AHM 2009, own calculations

#### 2.4.5 Analysis of long-term effect of initial transitions

Finally, we also estimate whether the duration of the initial transition from education to employment has an effect on people's long-term labour market outcomes as observed in 2009. We use a Probit model explaining the observed outcome (employment/unemployment) at the time of the EU:LFS survey by a set of individual characteristics (age, education, etc.) and the duration of the initial labour market transition experienced. The results (coefficients) are shown in Table 2.7.

**Table 2.7 Probit estimates on employment status in 2009 (All cohorts)**

	Coefficient	Std. Err.	z	P> z
Duration initial transition	-0.009	0.004	-2.420	0.015
Duration initial transition ^2	0.000	0.000	2.220	0.027
<b>Education (Base: General)</b>				
Vocational education mainly/solely school based	-0.002	0.155	-0.010	0.991
Vocational education mainly work place based	0.111	0.071	1.560	0.118
Vocational education no distinction possible	0.211	0.183	1.160	0.247
Not applicable	1.436	0.903	1.590	0.112
Work experience while in education (Base: No work or work less than 1 month per year)				
Work (only) as part of educational programme	-0.115	0.114	-1.010	0.315
Work while studying but outside educational programmes	0.173	0.074	2.350	0.019
Work (only) during an interruption of studies	0.290	0.179	1.620	0.105
<b>Cohort (Base 1975-179)</b>				
1980-84	0.065	0.073	0.900	0.370
1985-89	-0.015	0.075	-0.200	0.844
1990-94	-0.032	0.124	-0.250	0.800
<b>Sex (Base: Male)</b>				
Female	-0.549	0.060	-9.220	0.000
<b>Qualification (Base: ISCED-2 and below)</b>				
ISCED-3	0.719	0.394	1.820	0.068
ISCED-4	-0.015	0.985	-0.020	0.988
_cons	0.300	0.398	0.750	0.451
Number of observations	2891			
Log likelihood	-1194.4561			
LR chi2(22)	210.32			
Probability > chi2	0.0809			
Pseudo R2	0.1016			

Source: EU:LFS AHM 2009, own calculations

### 2.4.6 Summary of the findings

This analysis shows a significant impact of the duration of the initial transition on adult employment while controlling for the effect of further significant variables (work experience, education and gender). The effect is significant when modelled in a quadratic specification. Based on derived marginal effects (not shown in the table), **the employment rate (as percentage of the overall cohort) decreases by about 0.2 percentage points for every month of duration of the initial transition.** Controlling for vocational education or general education, a key driver to explain the duration of the initial transition, does not show a significant impact on employment rates.

**This is clear evidence that the duration of the initial transition has an important impact on later labour market outcomes ('scarring').**



# 3 Sequence analysis of transitions of young people from different cohorts

## 3.1 Objectives

Previous studies have found that early labour market experiences can shape outcomes much later in life. For a minority of young people, outcomes soon after school-leaving age appear potentially problematic.

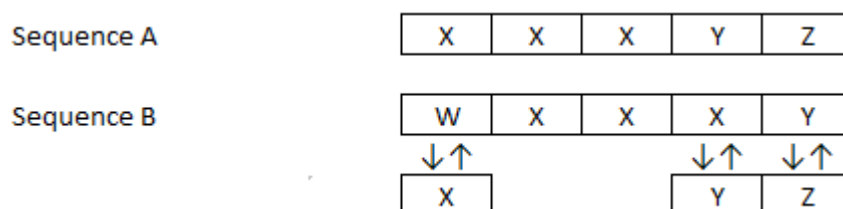
In order to shed new light on this matter, we used sequence analysis on data from four cohort studies spanning 32 years of birth:

- NCDS (born 1958)
- BCS (born 1970)
- YCS Cohort 8 (born 1979/80)
- LSYPE (born 1989/90).

## 3.2 Method

We analysed individuals' monthly labour market states from the first September after compulsory education for a period of 29 months. Sequence analysis was used to measure the similarity between individuals' labour market histories over this period. We distinguish between four potential states – education, employment, unemployment and other NEET. Similarity between two people's histories is calculated on the basis of the number of substitutions required to transform one person's sequence to be the same as that of the other person. Each type of transformation is assigned a 'cost' that reflects how common a transition between these two states is at this point in time. The similarity measure is the total cost of all substitutions. In the example below, similarity between sequence A and sequence B is calculated as the cost of substituting W for X at time 1 plus the cost of substituting Y for X at time 4 and the cost of substituting Z for Y at time 5.

**Figure 3.1 Sequence substitution example**



We then performed cluster analysis using the results of the sequence analysis, in order to group together individuals with similar sequences. We chose our preferred number of

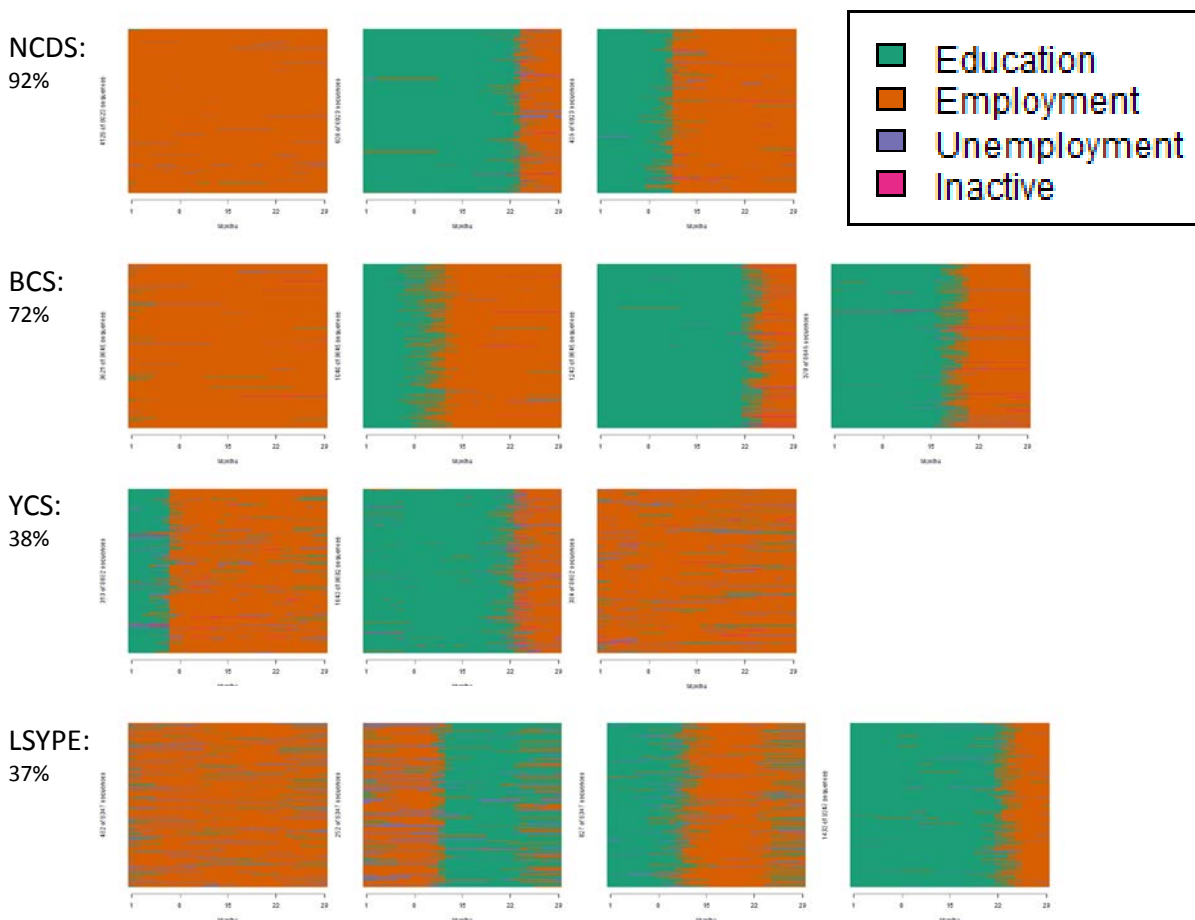
clusters largely based on diagnostic statistics, but in some cases adjusted slightly to ensure the clusters were meaningful. In all four cohorts we settled on seven cluster solutions. In addition, we collected the clusters into broader groupings of ‘Entering the Labour Market’, ‘Accumulating Human Capital’ and ‘Potential Cause for Concern’ for summary and modelling purposes.

Graphically, we use colour-coding to represent individuals' histories. Doing so provides a ready visualisation of the experiences typifying each cluster.

### 3.3 Findings

**The ‘Entering the Labour Market’ grouping has diminished significantly between the cohorts, from around 90% in the earliest to around 40% in the most recent. In addition, for those who do still follow this route, there are indications that it has become a less stable path with increased short spells of unemployment than in early cohorts.** The charts depict the case where individuals are either in employment throughout the period considered or enter employment straight after education. The exception is the second LSYPE cluster which is more ambiguous (note that it is rather small).

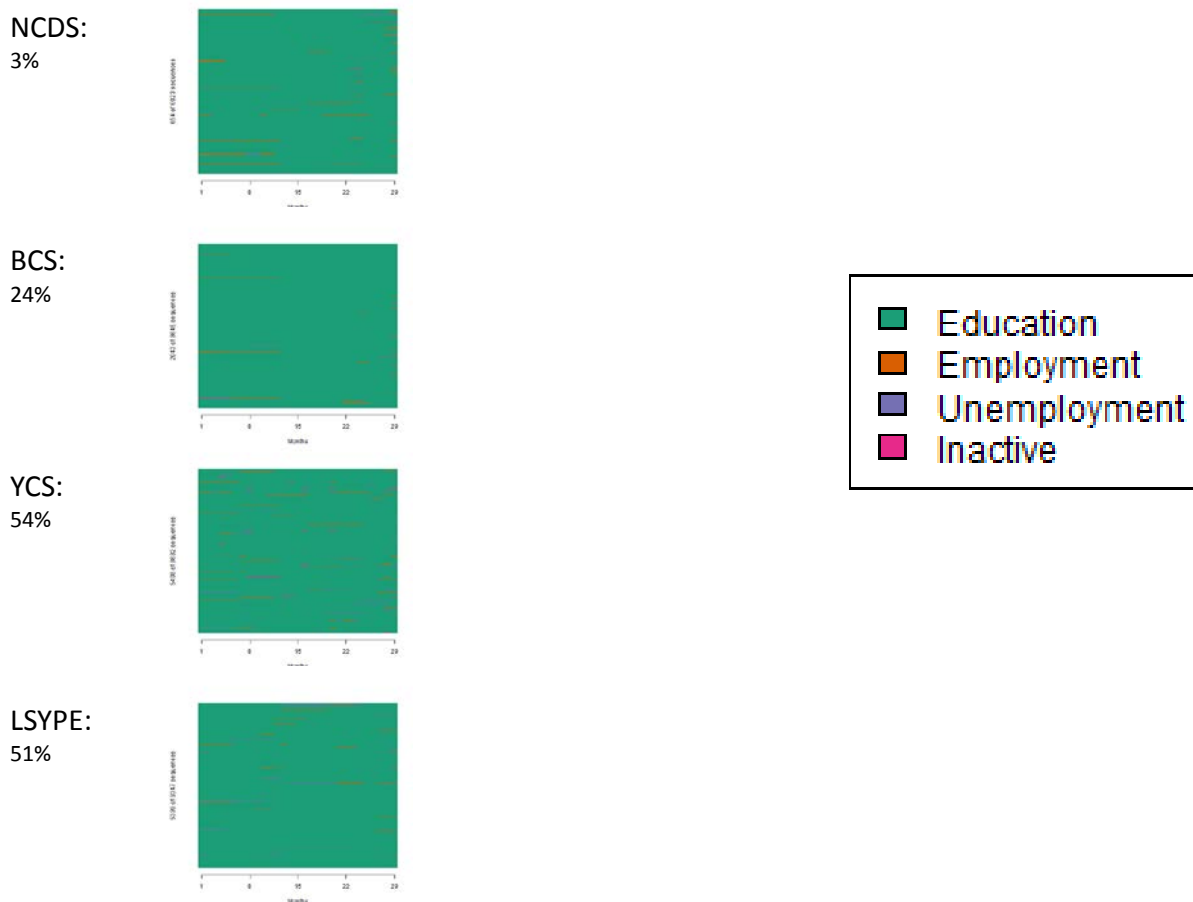
**Figure 3.2 Entering the labour market**



Source: Cohort studies, NIESR

By contrast, **the size of the ‘Accumulating Human Capital’ grouping has grown significantly across the cohorts, from 3% in the earliest, to around 50% in the most recent.** However, in other ways this transition appears to have remained similar during this period, in that the grouping remains highly homogenous: this remains a group in which young people concentrate solely on education over this period.

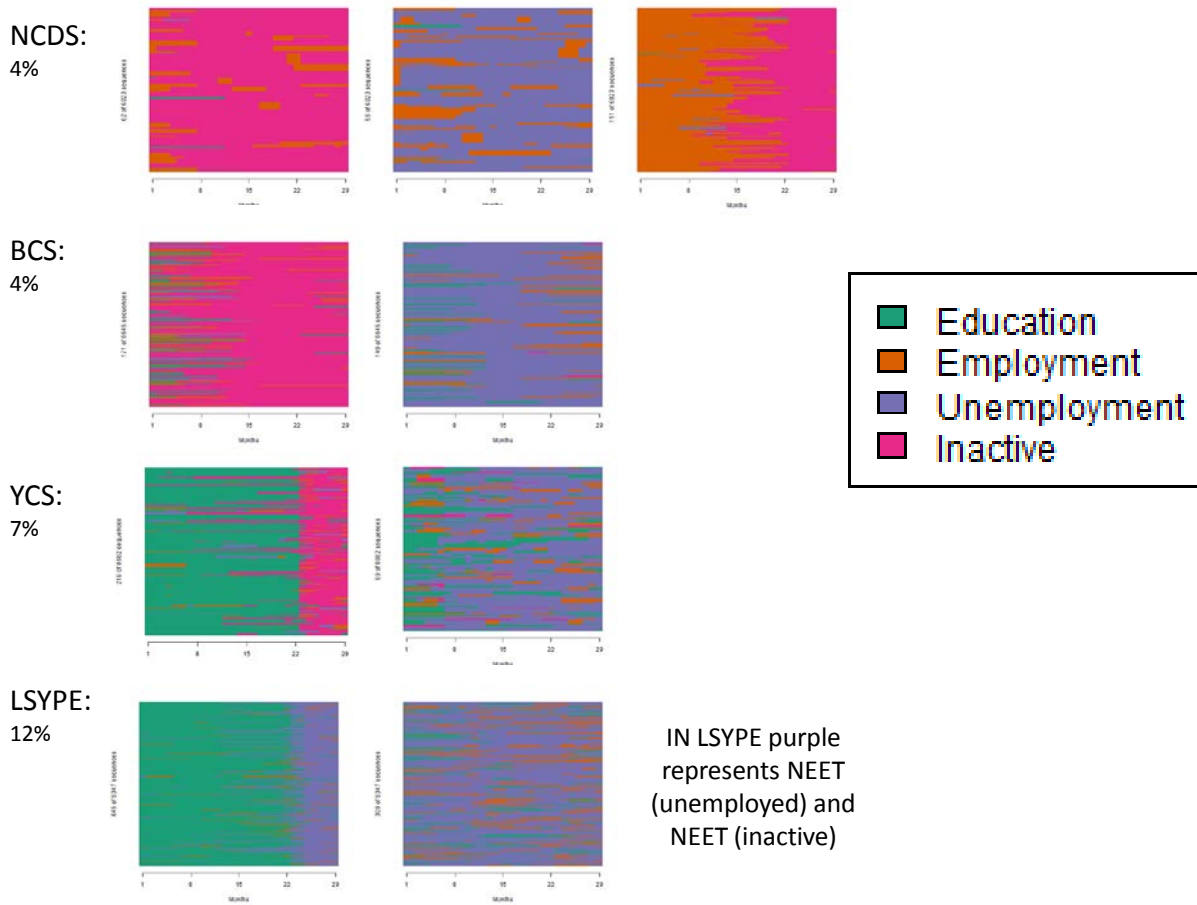
**Figure 3.3 Accumulating Human Capital**



Source: Cohort studies, NIESR

Finally, **the size of the ‘Potential Cause for Concern’ group has also grown, from 4% in the earliest cohort to 12% in the most recent.** In addition, we see a change in the behaviour of those who go straight from education into extended inactivity (predominantly young women, especially in earlier cohorts). Those in later cohorts are much more likely to receive two additional years of education.

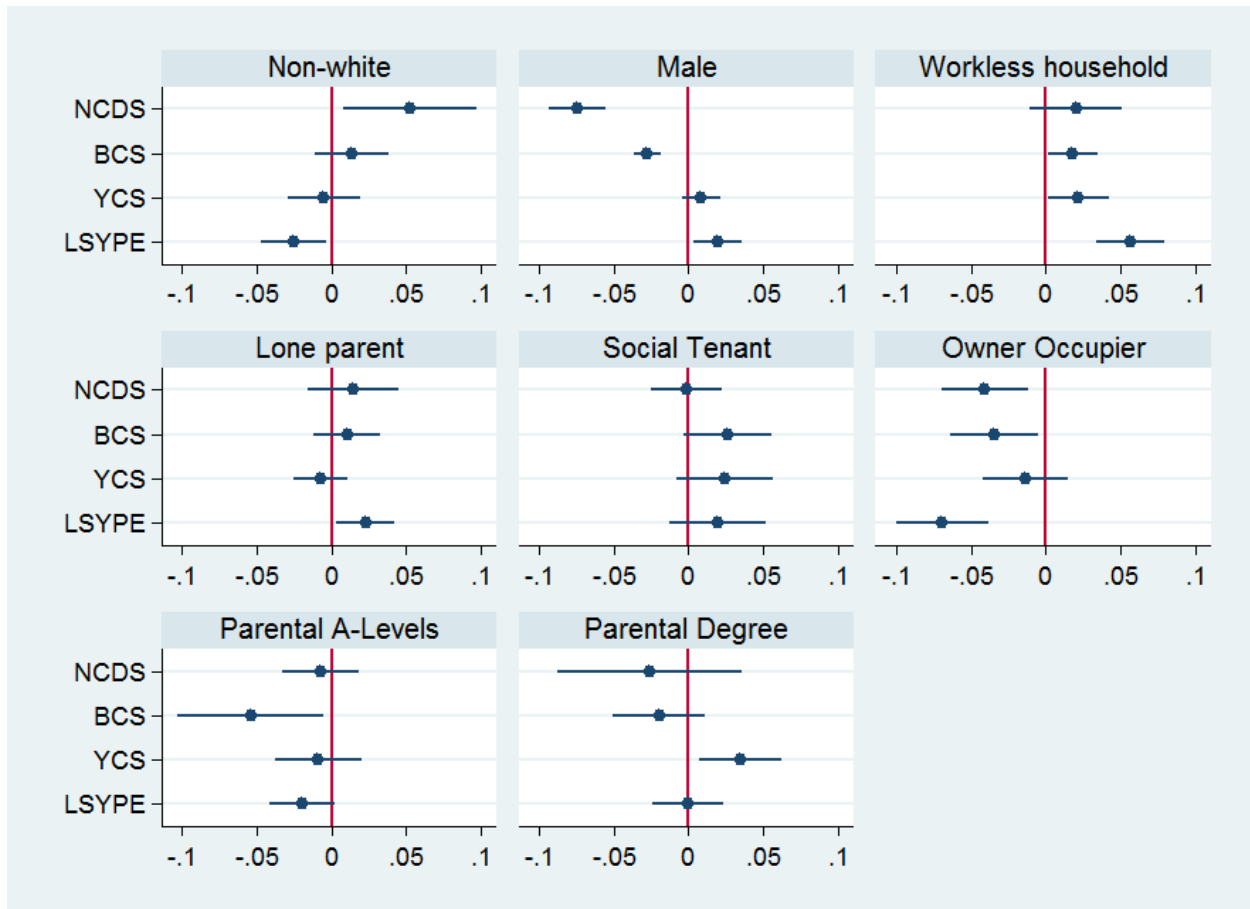
**Figure 3.4 Potential Cause for Concern**



Source: Cohort studies, NIESR

**The composition of the ‘Potential Cause for Concern’ group has also changed over time: individuals who are female and individuals of non-white ethnicity have gone from being more likely to be at risk of a difficult transition, to being less likely.** Cumulative disadvantage from various indicators of SES continue to play a large role in predicting difficult transitions.

The chart below provides more detail. For a range of age-16 characteristics, it reports the average marginal effect on the probability of an individual being in the ‘Potential Cause for Concern’ category (arms around the estimates show 95% confidence intervals). This is shown for each of the four cohorts, thereby highlighting changes over time in the composition of that group.

**Figure 3.5 Impact of characteristics on ‘potential cause for concern’**

Source: Cohort studies, NIESR

### 3.4 Extension of the sequence analysis: A) NCDS and BCS

#### 3.4.1 Introduction

In order to shed light on how meaningful the analyses of sequences of young people’s labour market states between 16 and 19 are, we also carried out an extended analysis of sequences up to age 25. This is only possible for the two datasets where the data are available: the NCDS and the BCS. We carry out sequence and cluster analysis on the same basis as was done for the shorter timeframe analyses, except that it is over a 98-month period, rather than a 29-month period. This time we use 14-cluster (rather than 7-cluster) solutions, reflecting the greater heterogeneity possible within longer sequences. Again, our choice of a 14 cluster solution is primarily on the basis of average silhouette distances.

We once again aggregate these clusters into our three broad groupings: Entering the Labour Market, Accumulating Human Capital and Potential Cause for Concern. We have sought to maintain some consistency with the definitions used with our shorter-run analysis. Most notably, clusters in which individuals remain in education until only just past the endpoint of the previous analysis (29 months) are placed into the Accumulating Human Capital Group, despite spending quite a while in the labour market over the whole period

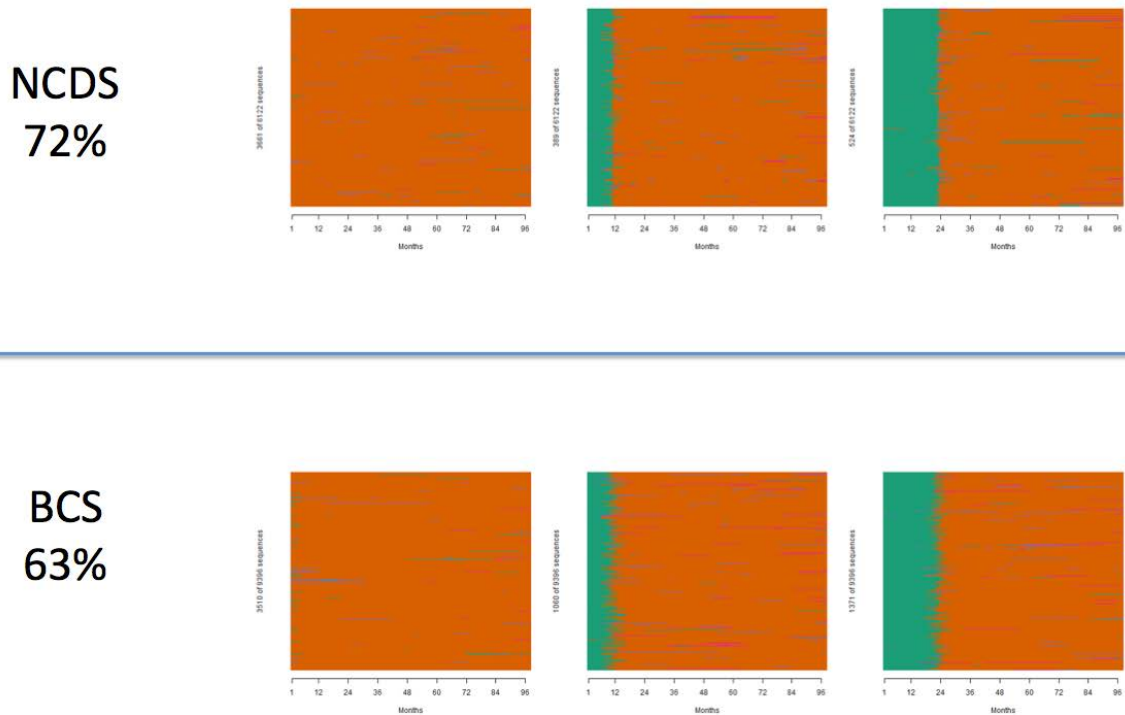
that we analyse. This is with the aim of allowing comparability between the shorter-run and longer-run analysis.

One particular challenge with conducting extended sequence analysis on the NCDS is the quality of the monthly activity data available particularly once we extend to age 25. The NCDS appears to have a rather systematic problem with gaps between different spells, which results in the loss of a substantial number individuals from our analysis, reducing the sample size from 8,372 to 6,122. This loss seems concentrated among individuals in the 'Entering the Labour Market' group, and we suspect that this is responsible for inflating the size of the 'Accumulating Human Capital' grouping compared to that estimated in the shorter analysis. Consequently, there is a concern about the ability of the NCDS to support the longer-run analysis. The BCS analysis does not suffer from the same problem; extending to age 25 reduces the sample size only marginally (from 9,518 to 9,419). In view of this, we feel more confident about the BCS results.

### 3.4.2 Entering the Labour Market

Shown first are the Entering the Labour Market groups. In both cohorts it is notable that the overall proportion of the cohort in this group is now somewhat lower than we predicted on the basis of the shorter groupings. **In the NCDS, our short run grouping estimated that 93% of individuals were 'entering the labour market', whereas here we estimate a figure of 72%. In the BCS, our short run grouping estimated that 66% of individuals were 'Entering the Labour Market', whereas here we estimate the lower figure of 63%.** As noted above, there are concerns over data quality with the NCDS. However, at least part of the reason behind the differences between the shorter- and longer-run results appears to reflect the fact that **some young people with seemingly promising starts in the labour market, nevertheless go on to either unemployment or inactivity, and so are regarded as a Potential Cause for Concern in the longer-run.** We discuss this further later in this note.

Figure 3.6 Entering the Labour Market



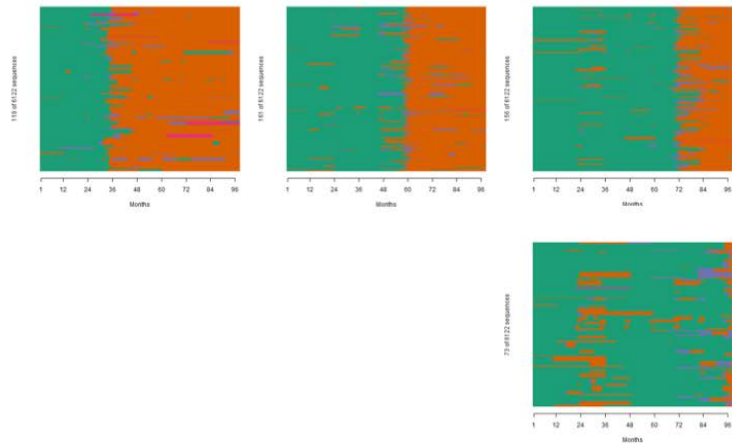
Source: Cohort studies, NIESR

### 3.4.3 Accumulating Human Capital

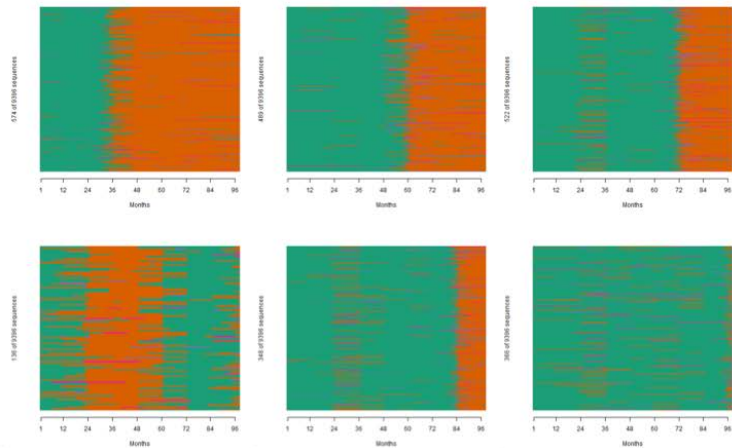
**The size of the ‘Accumulating Human Capital’ group in the NCDS, has risen from 3% in the short-run analysis to 15% in this longer-term analysis.** We cannot see a good explanation for this from graphical inspection of the groups deemed to be ‘Accumulating Human Capital’. Instead, we suspect this is due to the compositional change in the overall sample caused by losing a large number of individuals from the extended analysis due to spells of missing data. For the BCS, the size of the ‘Accumulating Human Capital’ group has also risen very slightly. This could be due to individuals with a brief time in employment subsequently returning to education.

Figure 3.7 Accumulating Human Capital

NCDS  
15%



BCS  
26%



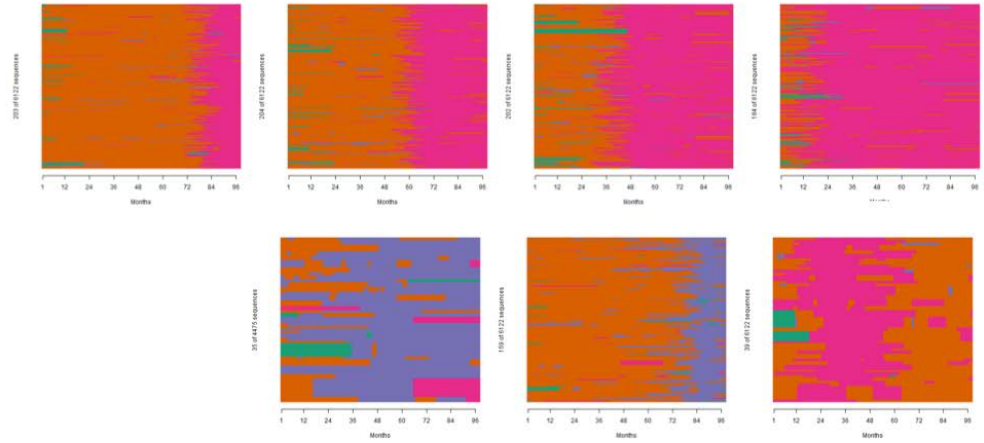
Source: Cohort studies, NIESR

Lastly, we turn to the ‘Potential Cause for Concern’ grouping. As already noted, we find that in both cohorts more individuals are in clusters with trajectories that we identify as a Potential Cause for Concern. Several are trajectories that graphical inspection reveals would not have been identified as Potential Cause for Concern within the first 29 months. Overall, **the proportion of individuals categorised as ‘Potential Cause for Concern’ increases from 5% to 13% in the NCDS and from 4% to 11% in the BCS.** Thus, despite the potential problems with the extended NCDS analysis, the size of these groupings remains similar in these two cohorts.

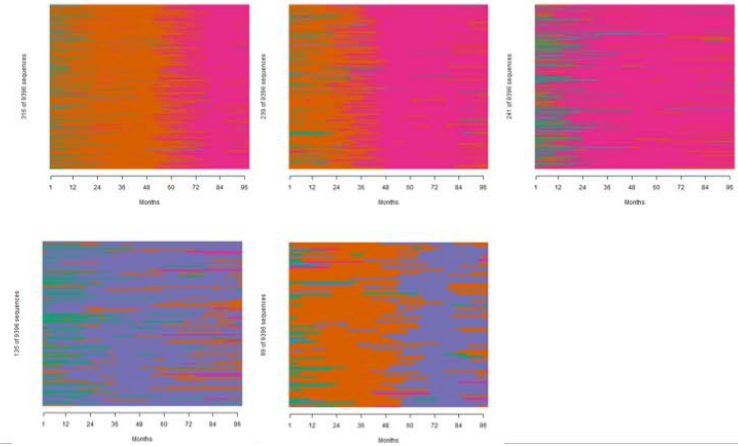


**Figure 3.8 Potential Cause for Concern**

**NCDS**  
**13%**



**BCS**  
**11%**



Source: Cohort studies, NIESR

### 3.4.4 Summary Statistics of Long Groupings

**Table 3.1 Summary Statistics**

Group	NCDS				BCS			
	ELM	AHC	PCC	Overall	ELM	AHC	PCC	Overall
<i>N</i>	4,574	509	1,039	6,122	5,947	2,448	1,024	9,419
<i>Group size (row %)</i>	0.72	0.15	0.13	1	0.63	0.26	0.11	1
<i>Male (%)</i>	0.62	0.57	0.16	0.54	0.53	0.5	0.23	0.49
<i>Non-White (%)</i>	0.02	0.02	0.02	0.02	0.02	0.04	0.03	0.02
<i>Single parent family (%)</i>	0.06	0.05	0.08	0.06	0.03	0.04	0.04	0.04
<i>Parent has A Levels (%)</i>	0.16	0.32	0.09	0.16	0.03	0.1	0.02	0.05
<i>Parent has a degree (%)</i>	0.02	0.15	0.01	0.03	0.02	0.15	0.01	0.05
<i>Home owner occupied (%)</i>	0.37	0.58	0.25	0.37	0.28	0.45	0.16	0.31
<i>Home socially rented (%)</i>	0.28	0.09	0.37	0.28	0.05	0.03	0.09	0.05
<i>Living in workless household (%)</i>	0.04	0.03	0.06	0.04	0.03	0.03	0.08	0.03

Notes: ELM = Entering the Labour Market; AHC = Accumulating Human Capital; PCC = Potential Cause for Concern. In NCDS inverse probability weights based on analyst's attrition modelling applied; in BCS attrition was too small across this period for this to be viable/necessary.

Source: Cohort studies, NIESR

The mean characteristics of individuals within each of the cluster groupings are relatively stable whether we define the groupings on the basis of 29- or 98-month sequence analysis. There are some slightly larger differences in the characteristics of the Potential Cause for Concern groupings, as would be expected given that it is now considerably larger. For example, the longer-run Potential Cause for Concern groupings are more female dominated than their shorter-run analogues. We suspect this is due to the longer-run grouping being more likely to include young women for whom not entering the labour market and instead becoming homemakers is more of a choice than a concern.

### 3.4.5 Cross-tabulation of shorter- and longer-sequence based groupings

In order to learn more about the relationship between the two sets of categorisations, we cross-tabulate the groupings in to which individuals are placed in the shorter- (29 month) and longer-term (98 month) analyses. Considering first the NCDS, we see that a large majority of individuals in the short-term groupings remain in the same grouping on the basis of the extended sequence analysis, with some movement from Entering the Labour Market into the Potential Cause for Concern, but very little movement out of the other two groupings (except into the missing category). We should also note that Potential Cause for Concern category grows primarily from individuals that were previously characterised as being Entering the Labour Market and very few from the Accumulating Human Capital

grouping. In the BCS, the picture is much the same, except for the much reduced size of the missing category, as discussed in the introduction to this section.

**Table 3.2 NCDS: Cross-tabulation of groupings on basis of 29 months of sequence analysis and of groupings on basis of 98 month sequence analysis**

Groupings	Long Groupings				Total (freq.)
	ELM	AHC	PCC	Missing	
ELM	63.5	0.7	11.3	24.6	7,110
AHC	2.5	54.2	0.8	42.5	852
PCC	7.1	0.3	57.4	35.3	394
Missing	81.3	0.0	18.8	0.0	16
<b>Total</b>	<b>54.6</b>	<b>6.1</b>	<b>12.4</b>	<b>26.9</b>	<b>8,372</b>

Notes: ELM = Entering the Labour Market; AHC = Accumulating Human Capital; PCC = Potential Cause for Concern. Reporting row proportions, except for the final (total) column, which reports frequencies.

Source: Cohort studies, NIESR

**Table 3.3 BCS: Cross-tabulation of groupings on basis of 29 months of sequence analysis and of groupings on basis of 98 month sequence analysis**

Groupings	Long Groupings				Total (freq.)
	ELM	AHC	PCC	Missing	
ELM	84.5	4.5	10.1	0.9	6,867
AHC	2.9	92.8	2.8	1.6	2,282
PCC	21.7	5.4	72.1	0.8	369
<b>Total</b>	<b>62.5</b>	<b>25.7</b>	<b>10.8</b>	<b>1.0</b>	<b>9,518</b>

Notes: ELM = Entering the Labour Market; AHC = Accumulating Human Capital; PCC = Potential Cause for Concern. Reporting row proportions, except for the final (total) column, which reports frequencies.

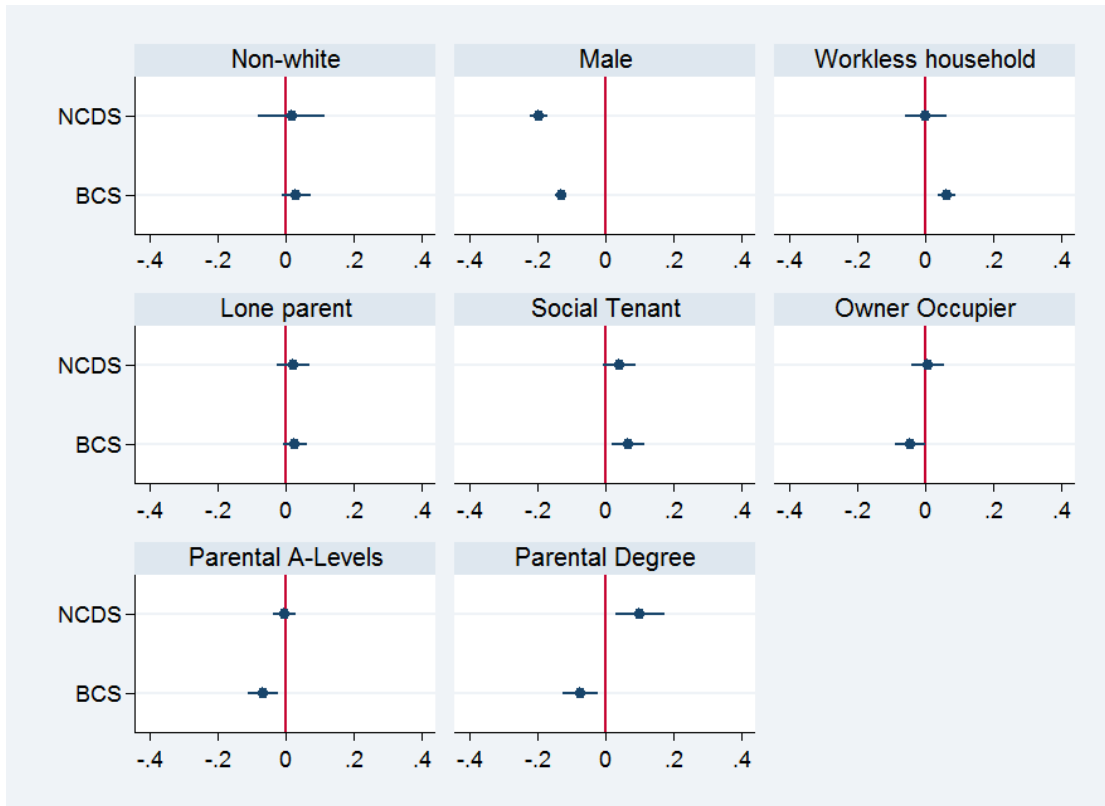
Source: Cohort studies, NIESR

What do we learn from this? Those who are potential cause for concern in 29 month analysis are likely still to be considered a potential cause for concern on the basis of the longer analysis. However, as one might expect, the longer analysis also picks up an additional number of cases that we deem to be potential cause for concern, on the basis of their trajectories post-29 months. However, we next explore whether this changes the risks of various observable characteristics associated with being a potential cause for concern.

### 3.4.6 Conditional associations between characteristics and being classified 'potential cause for concern'

Reassuringly, we find a fairly similar pattern in the effects of probability of being potential cause for concern in this analysis as we did in the 29-month analysis, although there are unexpected or surprisingly insignificant results associated with a few characteristics for the NCDS. Nevertheless, we conclude that this suggests that while the sequence and cluster analyses themselves do not necessarily pick up all the individuals who are of potential cause for concern in this shorter timeframe, our shorter-run analysis nevertheless identifies the observable groups that are likely to be at greater risk.

**Figure 3.9 Average marginal effects of age 16 characteristics on being in a cluster that is 'potential cause for concern'**



Source: Cohort studies, NIESR

# 4 A detailed analysis of activities and transitions of the 2010/11 KS4 leavers

## 4.1 Objectives

Analysis of transition from schooling to further destinations using National Pupil Data (NPD) merged to records of National Client Casework Information System (NCCIS). The data that was merged consisted of full NPD cohort of pupils (N=636,986) with valid Key Stage 4 (i.e. all children born between 01 September 1993 and 31 August 1994 in last year of compulsory education). Research consisted of:

- Assessment of quality of merge / imputations when NCCIS missing / understanding balanced panel properties and representativeness of the full NPD samples / coding of consistent activities.
- Analysis of transition patterns from KS4 to further activities (most popular spells / aggregation to quarters with dominant activities / extent of churning over the age 16 to 19 / duration of NEET

## 4.2 Data description

### 4.2.1 Relevant samples

The used data is made first and foremost of the full cohort of pupils in the National Pupil Dataset (NPD) sample (N=636,986) with valid results at their Key Stage 4 tests (GCSEs). This is the full census of all children attending a school in their last year of compulsory education. This is the cohort born between September 1, 1993 and August 31, 1994.

Early on in the project life, the research team decided to concentrate on the period between September 2011 (the first month of the academic year directly following completion of compulsory education) and the most recent month available at the time of writing (April 2014).

Starting from this initial period, we show in Table 4.1, the number and related percentages provided by the NCCIS children records successfully merged with the NPD (together with the sample that couldn't be merged, N=99,327).

**Table 4.1 Full NPD data combined with NCCIS information, for the cohort born between 1 September 1993 and 31 August 1994 (who passed their GCSEs in May 2011)**

NCCIS data available for:	Freq.	%
Less than 5 months	19,329	3.03
5 to 9 months	13,855	2.18
10 to 14 months	27,859	4.37
15 to 19 months	30,488	4.79
20 to 24 months	57,006	8.95
25 to 29 months	121,668	19.1
30 months	54,414	8.54
31 months	92,599	14.54
32 months (balanced sample)	120,441	18.91
No NCCIS data	99,327	15.59
Full NPD Sample	636,986	100

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

We note that approximately 99,000 pupils' records (16%) do not appear in the NCCIS at all. These are NCCIS observations that couldn't be merged with young people in the NPD sample. The merge 'success' rate is consequently around 84% of the full sample of children in their last year of compulsory school, which is high.

Table 4.1 also shows the number counts and percentage for pupils with valid NPD entries but missing NCCIS information on particular months between September 2011 and April 2014 (as fraction of the whole period of 32 months). Around 19% have valid NCCIS information in each month (18.91%). And further 23% have either 30 or 31 months of NCCIS information.

Given the very large number involved with NPD-NCCIS data, we proposed to start our investigations on a sub-sample of the data that has full or near full NCCIS information (i.e. a balanced panel). We will, however, first assess whether this sub-sample is representative of the whole NPD sample. For this process, we can fortunately make use of socio-economic variables available in the NPD (see Section 2, below).

But before proceeding to this comparative analysis, we make some initial and limited adjustments to the data.

We observed, indeed, that in Table 4.1 a large number of pupils (23%) have information for 30 and 31 months out of the maximum of 32 for the whole period (September 2011 to April 2014). We proposed to fill the gaps wherever possible for this large group of children for which the observation window was near complete.

We proceed as follows: we first restrict this correction to pupils for which the activity observed before and after the missing months is identical. For those with 31 months available (i.e. one month missing), we fill the gap when their missing month of activity was preceded and followed by the same activity. We do not adjust the missing month for pupils for which the activity changes before and after the missing month. The assumption is that

if only one month is missing and the young people was attending say, FE, the month before and again FE the month after, then it is very likely she was studying in FE during the month with a missing value. But we cannot infer reliably the missing activity when it is missing and the activity has changed before and after the missing month.

For those with 30 months available (i.e. two months with a missing information), we filled the missing months for those young people for which two consecutive months were missing but the activity the month directly preceding the missing spell and following the missing spell was identical, following the same logic. And similarly, we do not fill the gaps when the activity has changed in the months preceding the missing two-months spells and the month directly following it.

We do not pursue this exercise for young people with more than two months of missing information. The assumption here is that it becomes more likely that three subsequent months with missing information may indicate a real break, even in the case where the activity is the same before and after the missing spell.

**Table 4.2 Full NPD data combined with NCCIS information, adjusted for those with 30 and 31 months and identical activity before and after the missing months**

NCCIS data available for:	Freq.	%
Less than 5 months	19,329	3.03
5 to 9 months	13,855	2.18
10 to 14 months	27,859	4.37
15 to 19 months	30,488	4.79
20 to 24 months	57,006	8.95
25 to 29 months	121,668	19.10
30 months	30,685	4.82
31 months	60,316	9.47
32 months (balanced sample)	176,453	27.70
No NCCIS data	99,327	15.59
Full NPD sample	636,986	100

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

Table 4.2 show the outcome of this adjustment. We observe that the 'balanced' panel increases significantly to 176,453 individuals when filling the gaps of those with 30 and 31 months of NCCIS information. This increase from 120,441 to 176,453 has been achieved by filling the gap for 32,283 individuals with 31 months of NCCIS data in the original sample (ie. in Table 4.1), and 23,729 observations with 30 months of NCCIS data. We end up with a 'balanced' sample of 176, 453 pupils. This total represents nearly 28% of the whole NPD sample.

Another sub-sample of reference can be used in the Section that follows. The sample of successfully merged NPD-NCCIS observations. It is simply the full NPD sample to which we removed the unmerged observations (99,327). We will name this sample the 'full NPD-NCCIS merged' sample. It includes 547,713 observations, of which 176,453 have 32 months of NCCIS data and 371,260 have less than 32 months of NCCIS data. The balanced sample (176,453 observations) accounts for more than 33% of this latter sample.

## 4.2.2 Representativeness of the balanced panel for full NPD samples

In this section, we conduct investigations to determine how the balanced panel compare with the whole NPD data with regards to three key aspects: the pupils' performance at national tests when aged 7, 11 and 16, the geographical distribution, and the ethnic composition.

We start with a comparison of performance during **compulsory education**. We also include the 'fully merged NPD-NCCIS sample' (that includes records with less than 32 months of NCCIS information) for comparisons.

**Table 4.3 Number of good GCSEs (A\*-C) in the balanced panel compared with the NPD and the fully merge NPD-NCCIS samples**

Number of GCSEs at A*-C	All NPD		Full merged NPD-NCCIS		Balanced Panel	
	Freq.	Per cent	Freq.	Per cent	Freq.	Per cent
0	122,617	19.25	98,444	17.97	28,490	16.15
1	49,034	7.7	43,849	8.01	13,709	7.77
2	39,156	6.15	36,234	6.62	11,654	6.6
3	36,243	5.69	33,899	6.19	10,905	6.18
4	34,997	5.49	32,645	5.96	10,493	5.95
<b>% with less than 5 good GCSEs</b>	<b>44.28</b>		<b>44.75</b>		<b>42.65</b>	
5	36,209	5.68	33,480	6.11	10,892	6.17
6	39,387	6.18	36,340	6.63	11,964	6.78
7	45,404	7.13	41,188	7.52	13,750	7.79
8	55,212	8.67	48,338	8.83	16,592	9.4
9	70,334	11.04	56,791	10.37	19,366	10.98
10	66,449	10.43	50,766	9.27	16,653	9.44
11	31,505	4.95	26,450	4.83	8,840	5.01
12	8,683	1.36	7,685	1.4	2,627	1.49
13	1,490	0.23	1,356	0.25	440	0.25
14	242	0.04	225	0.04	72	0.04
15	22	0	21	0	6	0
16	2	0	2	0	0	0
<b>Total</b>	<b>636,986</b>	<b>100</b>	<b>547,713</b>	<b>100</b>	<b>176,453</b>	<b>100</b>

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

We decided to show the proportion with less than 5 GCSEs to focus on low performance where transitions are potentially challenging. We note that the proportions of the sub-samples appear very similar when comparing the balanced sample with the other samples. For those who achieved 5 and more GCSEs, again the proportion are remarkably similar across the samples. Given the very large numbers involved it is not too surprising. It is indeed expected that a sample of 176,453 pupils will be broadly representative of the full population of 636,986.



To investigate further the representativeness of the balanced panel, we then turn our attention to the **geographic distribution** of our merged sample. For this aim, we produce similar tables differentiated by large regions in England. We only show the percentages not achieving the target (i.e. 5 good GCSEs) in the different regions and across the three sub-samples. We note that the balanced panel appears less representative at the level of regions. In particular, West Midland and the North-West have lower low achievers in the balanced panel than in the full NPD sample. It is in the West Midland that the difference is the highest (48.2% versus 41.8% for respectively the full NDP and the balanced sub-samples). It is not entirely clear why some regional differences appear. It appears that it is mostly the proportion with no GCSE and 1 GCSE that are lower in the balanced panel for West Midland compared to the full NPD sample (tables not shown).

To investigate further this question, we looked at results from tests taken at a younger age (KS3 at 14, KS2 at 11 and KS 1 at 7). See Tables A4 to A6 in the Appendix. We observe the lower percentage for the performance in the West Midland also appears in national tests taken earlier than at 16. However, this divergence in one or two regions appears in a context where in the other seven regions the samples are remarkably similar across sub-samples. Those tables appear to confirm that local authorities do record with unequal success their 16-19 years olds activities (see, Dexter, 2014 and Mirza Davis, 2014).

**Table 4.4 Region in the balanced panel compared with the NPD and the fully merge NPD-NCCIS samples**

Region	% achieving less than 5 GCSE (A*-C)		
	All NPD	Full Merged NPD-NCCIS	Balanced Panel
London	41.27	40.48	39.51
West Midland	48.17	48.12	41.81
North West	45.51	45.34	41.22
Yorkshire	50.37	50.43	47.84
North East	50.27	50.18	46.23
South West	41.63	42.65	42.64
South East	39.98	41.99	40.81
East of England	41.74	42.45	41.52
East Midland	46.08	45.98	45.54
<b>Total (%)</b>	<b>44.28</b>	<b>44.75</b>	<b>42.65</b>

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

We next investigate the **ethnic diversity** of our sub-samples, which provides another perspective to assess how representative the balanced panel is of the full population.

We investigate how the balanced panel compares with the larger sub-sample of all NPD merged with NCCIS pupils record (N=547,713). As we are using the ethnicity information recorded in the NCCIS data, we cannot compare with the full NPD sample. We present only the proportions with less than 5 good GCSE for the different ethnic groups, but produce two tables in the Appendix (Appendix table A1 and A2) showing the full distribution of good (A\*-C) GCSE obtained.

**Table 4.5 Ethnicity and samples**

Ethnicity	Fully merged NPD- NCCIS (n=547,713)	Balanced Panel (n=176,453)
<b>% with less than 5 good GCSEs (A*-C):</b>		
White	44.71	42.3
Indian	27.38	29.24
Pakistani	51.89	53.29
Bangladeshi	44.32	43.06
Black Caribbean	55.3	54.14
Black African	45.31	46.55
Mixed	43.78	43.6
Other	45.46	43.15
Total	44.75	42.65

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

The percentages of under-performing teenagers appear very similar in the two samples. The overall percentage is slightly lower in the balanced panel but not significantly so. The proportion of three ethnic groups appears to be a bit higher in the balanced panel (mainly Indians, Pakistani and Black Africans), whereas the proportion of White is slightly lower. We perform more investigations on early performance in Key stages tests at age 7, 11 and 14. The results overall confirm that the balanced sample appears reasonably representative of the full samples. Tables are provided in the Appendix on Tables A9 to A11.

Overall, the different investigations conducted in this section suggest that **the sample appears representative of the full sample with respect to school performance, ethnicity and reasonably regionally representative as well.** The main conclusion we draw therefore is that **we can proceed with an analysis relying on the fully merged sample only.**

### 4.2.3 Categorisations of young people's activities

This section briefly discusses the choice of relevant monthly status from the long list recorded in the NCCIS.

The raw data provides 45 categories, in total, from September 2011 to April 2014. However, some new categories are created and old ones deleted (e.g. the 2014/15 data included some new codes, details are included in the appendix on Tables A12 to A16). Therefore, we recode and reduce the size of the categories, and we propose to use the most recent split as advised by the 'NCCIS Management Information Requirement 2014-15'.<sup>1</sup>

<sup>1</sup> which are the following: 'Education' [1], 'Employment' [2], 'Training' [3], 'Re-engagement activities' [4], 'Not in Employment Education or Training (NEET)' [5], 'Other' [6] and 'Current situation not known' [7].

Table 4.6 below shows the full range of activities reported in NCCIS data, and indicates how the observed categories in NCCIS were finally recoded.

**Table 4.6 Recoding of NCCIS activities (September 2011 – April 2014)**

<b>NCCIS coding</b>	<b>Recoding</b>
[110] - Registered (School/educational establishment), [120] - Educated at home, [140] - Not registered school/educational establishment, [150] - Current Situation not known; [210] - School Sixth Form; [220] - Sixth Form College; [230] - Further Education; [240] Higher Education; [250] - Part time Education; [260] - Gap Year students; [270] - Other education; [280] - Independent Specialist Provider	[1] Education
[310] - Apprenticeships; [320] - Employment with accredited training; [330] - Employment without training; [340] - Employment with non-accredited training; [350] - Temporary employment; [360] -Part time employment; [380] - Self-employment; [381] - Self Employment combined with part time study; [550] - Working not for reward combined with part time study	[2] Employment
[410] - EFA/SFA funded work based learning; [430] - Other training; [440] - Training delivered through the work programme; [450] – Traineeship; [460] - Supported Internship	[3] Training
[530] - Re-engagement provision	[4] Re-engagement activities
[510] - Personal Development Opportunity (allowance/wage); [520] - Other Personal Development Opportunities; [540] - Working not for reward; [610] – Not yet ready for work or learning; [619] - Seeking employment, education or training; [620] - Not available to LM Young carers; [630] - Not available to LM Teenage parents; [640] - Not available to LM Illness; [650] - Not available to LM Pregnancy; [660] - Not available to LM on religious grounds; [670] - Not available to LM those Unlikely economically active; [680] - Not available to LM Other reason	[5] NEET
[710] - Custody; [720] - Refugee/Asylum seekers who have not yet been granted citizenship; [130] - Custodial sentence	[6] Other
[810] - Current situation not known; [820] - Cannot be contacted; [830] - Refused to disclose activity	[7] Current situation not known

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

We propose percentages (numbers given in the Appendix, A17) during the 32 months period directly following completion of compulsory education. We see that the vast majority of our sample (around 90%) is in education at the beginning of the observation window, and that this percentage declines regularly over the period to end up at 73% in April 2014. The decrease in twenty percentage points is commensurate with an increase in employment of similar amplitude. Whereas a very small proportion of teenagers work in the initial months following their completion of compulsory education, the proportion increases steadily to reach approximately 22% in April 2014. The percentages in other categories are always much lower, and sometimes near zero.

**Table 4.7 Month by month activities, % (September 2011 – April 2014)**

Month	[1] Education	[2] Employment	[3] Training	[4] Re-engage	[5] NEET	[6] Other	[7] Not known
Sep-11	82.8	2.7	1.0	0.0	1.3	0.0	12.1
Oct-11	90.3	3.9	1.2	0.0	1.5	0.0	3.0
Nov-11	91.9	4.2	1.4	0.0	1.6	0.0	0.8
Dec-11	91.9	4.4	1.5	0.0	1.6	0.0	0.6
Jan-12	91.2	5.0	1.6	0.0	1.7	0.0	0.5
Feb-12	90.9	5.1	1.7	0.0	1.8	0.0	0.5
Mar-12	90.5	5.3	1.7	0.0	1.9	0.0	0.6
Apr-12	90.2	5.5	1.7	0.0	2.0	0.0	0.6
May-12	89.8	5.8	1.8	0.0	2.0	0.0	0.6
Jun-12	89.4	6.0	1.8	0.0	2.1	0.0	0.6
Jul-12	87.8	6.3	1.9	0.0	2.5	0.0	1.5
Aug-12	86.7	6.9	1.9	0.0	3.1	0.0	1.5
Sep-12	85.2	7.3	1.7	0.0	2.5	0.0	3.3
Oct-12	86.1	7.7	1.7	0.0	2.3	0.0	2.2
Nov-12	85.8	8.3	1.8	0.0	2.3	0.0	1.6
Dec-12	85.7	8.7	1.8	0.0	2.3	0.0	1.4
Jan-13	85.8	9.0	1.9	0.0	2.3	0.0	1.1
Feb-13	85.6	9.2	1.9	0.0	2.4	0.0	1.0
Mar-13	85.4	9.4	1.9	0.0	2.4	0.0	0.9
Apr-13	85.1	9.7	1.9	0.1	2.4	0.0	0.9
May-13	84.8	10.2	1.7	0.1	2.4	0.0	0.8
Jun-13	84.4	10.6	1.7	0.1	2.4	0.0	0.8
Jul-13	80.7	11.2	1.6	0.1	2.6	0.0	3.8
Aug-13	79.8	11.8	1.5	0.1	2.8	0.0	3.9
Sep-13	69.5	13.3	1.4	0.1	2.5	0.0	13.1
Oct-13	70.5	15.2	1.4	0.1	2.5	0.0	10.2
Nov-13	71.0	18.2	1.5	0.1	2.5	0.0	6.7
Dec-13	73.1	19.7	1.5	0.1	2.4	0.0	3.2
Jan-14	73.3	20.7	1.5	0.1	2.3	0.0	2.0
Feb-14	72.7	20.9	1.5	0.1	2.3	0.0	2.4
Mar-14	72.7	21.2	1.5	0.1	2.3	0.0	2.2
Apr-14	72.6	21.5	1.4	0.1	2.3	0.0	2.1

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

Reengagement activities only appear as a non-zero figure from May 2013 onwards when it was introduced in the data set. But the cell never exceeds 250 cases and should probably merge into another category without loss of generality. Similarly, 'the other' [6] category never exceeds 100 individuals and should be merged with another category. We thus proposed to move those two marginal categories into the 'unknown' category' [7].

#### 4.2.4 Most common patterns of youth transitions

We aggregated the 32 months of the spell into a single variable taking the categories 1 to 7 values for each of the 32 months.

Even in the sub-sample of the balanced panel, we find widely different sets of patterns of moves across categories, and the total number of different patterns is as large as 16,851. We cannot, for obvious reasons, reproduce all the different patterns but Table 4.7 shows the most common ones. We show all patterns followed by at least 500 young people. They account for 70% of our balanced panel. We can see in the Table 4.8 that up to 33% of the sample is coded as continuously in education during the period (category '1'), but a few other patterns could be thought of as following the same path. Indeed, the category 'Current situation unknown' [7] appear often preceded and followed by periods in full-time education [1].

One solution to this problem would be to treat the '7' as we did with the missing month. We discuss this option in section 5.

**Table 4.8 Most common patterns (September 2011 – April 2014), without adjustments on the category [7] Situation unknown**

Pattern	Freq.	Per cent	Cum.
11111111111111111111111111111111	58,709	33.27	33.27
111111111111111111111111111711111111	11,982	6.79	40.06
71111111111111111111111111111111	8,453	4.79	44.85
1111111111111111111111111117711111111	5,738	3.25	48.11
2222222222222222222222222222222222	3,148	1.78	49.89
711111111111111111111111111711111111	3,083	1.75	51.64
1111111111171111111111111711111111	2,352	1.33	52.97
11111111111111111111111111122222222	2,330	1.32	54.29
11111111111111111111111111177222222	2,022	1.15	55.44
11111111111711111111111117711111111	1,925	1.09	56.53
77111111111111111111111111111111	1,924	1.09	57.62
111111111111111111111111111777111111	1,754	0.99	58.61
11111111111711111111111111111111	1,692	0.96	59.57
7111111111111111111111111117711111111	1,372	0.78	60.35
7222222222222222222222222222222222	1,235	0.70	61.05
11111111111111111111111111172222222	1,209	0.69	61.73
7111111111171111111111111711111111	1,128	0.64	62.37
111111111111111111111111111711111111	973	0.55	62.92
711111111111111111111111111777111111	953	0.54	63.46
111111111111111111111111111777711111	925	0.52	63.99
71111111111711111111111117711111111	881	0.50	64.49
1111111111112222222222222222222222	855	0.48	64.97
11111111111111111111111111177777777	843	0.48	65.45
71111111111711111111111111111111	840	0.48	65.93
111111111111111111111111111777222222	816	0.46	66.39
11111111111771111111111177711111111	751	0.43	66.81
11111111111771111111111177111111111	695	0.39	67.21
1111111111177111111111117711111111	633	0.36	67.57
711111111111111111111111177771111111	622	0.35	67.92
1111111111111111111111111777722222	619	0.35	68.27
71111111111111111111111772222222	582	0.33	68.60
1111111111111111111111111222222222	556	0.32	68.92
11111111111711111111111772222222	555	0.31	69.23
7111111111111111111111177777777	555	0.31	69.54

Notes: Each figures in the activities column represents a month where the codes are: 'Education' [1], 'Employment' [2], 'Training' [3], 'Re-engagement activities' [4], 'Not in Employment Education or Training (NEET)' [5], 'Other' [6] and 'Current situation not known' [7]

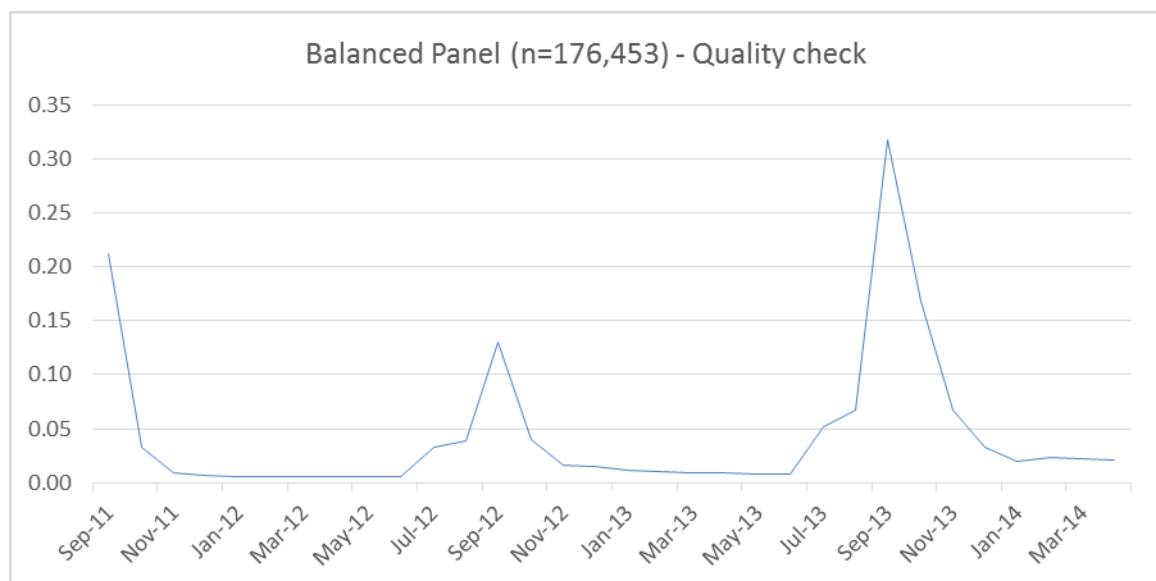
Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

Table 4.8 shows that the second most popular choice is to enter full-time employment. In this regard, there are a few interesting paths worth mentioning. The entry into full-time employment is made either straight after school (probably around 3% accumulating similar paths), or after completed a few more years of full-time education. The other categories do not appear very prevalent (most notably the NEET, Training, and Reengagement activities). This is not unexpected. The proportion of young people in situation of NEET is known to be underestimated in the NCCIS data (see Mirza-Davies, 2014). But overall the table gives us a good idea of the wide variety of patterns observed in the data. There is an obvious need to reduce the number of patterns to be able to conduct meaningful investigations. We, therefore, propose in Section 5 to first address the presence of the ‘unknown’ category 7 in the middle of two spells of the same activity.

#### 4.2.5 Treatment of the category ‘Current situation not known’

The ‘Current situation not known’ is coded with ‘7’. This code appears very often in the early months of the academic year (i.e. August, September, October). Figure 4.1 shows the proportion of the sample in each month between September 2011 and April 2014 with situation unknown. One possibility to address this problem would be to operate a similar correction to situation with missing values when the activity remains the same in the month just preceding and just following the missing month (1 or 2).

**Figure 4.1** Proportion of ‘Current situation not known’ (September 2011 – April 2014)



Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

We undertake this correction and produce the resulting Table 4.9. The adjustment allows a collapse a number of patterns and increases the size of the pattern of continuous full-time education [1]. Now around 56% of the balanced panel is observed continuously in full-time education, and approximately 2.5% are constantly in employment [2]. It is also interesting to see that only 22 patterns now account for 76% of the sample. There are still 14,800 different spells and 11,928 that are unique to one young person. The periodicity (month) combined with the 7 categories possible allows for a very large variety of spells.

**Table 4.9 Most common patterns (September 2011 – April 2014), when the category ‘Current situation not known’ [7] is replaced by the same activity before and after**

Activities from month 1 until 32	Frequency	Percentages	Cumulative percentage
11	97,591	55.3	55.3
7111111111111111111111111111111111111111	8,091	4.6	59.9
11111111111111111111111111177711111111111	3,400	1.9	61.8
1111111111111111111111111117722222222222	3,311	1.9	63.7
222222222222222222222222222222222222222	3,174	1.8	65.5
7711111111111111111111111111111111111111	2,553	1.4	66.9
11111111111111111111111111122222222222	2,516	1.4	68.4
111111111111111111111111111722222222222	1,794	1.0	69.4
11111111111111111111111111177777777777	1,537	0.9	70.3
111111111111111111111111111777222222222	1,481	0.8	71.1
722222222222222222222222222222222222222	1,251	0.7	71.8
111111111111111111111111111777711111111111	1,151	0.7	72.5
111111111111222222222222222222222222222	945	0.5	73.0
111111111111111111111111111777111111111111	925	0.5	73.5
1111111111111111111111111117777222222222	842	0.5	74.0
711111111111111111111111111777711111111111	655	0.4	74.4
11111111111111111111111111122222222222	571	0.3	74.7
11111111111111111111111111122222222222	565	0.3	75.0
11111111111111111111111111122222222222	564	0.3	75.3
111111111112222222222222222222222222222	528	0.3	75.6
11111111111111111111111111122222222222	517	0.3	75.9
111111111111111111111111111777	514	0.3	76.2
<b>Total</b>	<b>134,476</b>		
<b>Total balanced panel</b>	<b>176,453</b>		

Notes: Each figures in the activities column represents a month where the codes are: ‘Education’ [1], ‘Employment’ [2], ‘Training’ [3], ‘Re-engagement activities’ [4], ‘Not in Employment Education or Training (NEET) [5]’, ‘Other’ [6] and ‘Current situation not known’ [7]

While there are a lot of commonalities for 76% of the sample who either study or work either continuously or with small breaks where their situations is unknown, the remaining 24% of the sample is experiencing very complex and individualised transitions. Table 4.10 shows the number of spells experienced by young people and the related number counts.

The maximum number of spells experienced is 13. Unsurprisingly given what we now know, around 57% experience only one continuous spell. It is interesting to see, however, that a further 32% of the sample experience only two or three spells, which brings the total to 90% who experience less than 3 spells during the period.



We also calculated the average number of spells, which is consequently very low at 1.85. Around 9% of the sample experience between 4 and 6 spells, with the remaining 0.7% (1,158) experiencing between 7 and 13 spells. So it is a very small number of individuals that experience a large variety of spells. We consider this result as support for a reduction in the time period (i.e. the month) used in the analysis.

**Table 4.10 Number of spells experienced by young people from 16 to 19**

Number of spells	Frequency	Percentages	Cumulative percentage
1	101,033	57.26	57.26
2	29,655	16.81	74.06
3	28,849	16.35	90.41
4	9,903	5.61	96.03
5	4,061	2.3	98.33
6	1,794	1.02	99.34
7	683	0.39	99.73
8	292	0.17	99.9
9	123	0.07	99.97
10	32	0.02	99.98
11	14	0.01	99.99
12	10	0.01	100
13	4	0	100
<b>Total</b>	<b>176,453</b>	<b>100</b>	

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

We propose therefore to reduce the periodicity from a monthly to a three months period (quarters). We think this is probably a meaningful way to detect common experiences and to address the vast diversity of moves in the data for young people who do not pursue either education and employment in continuous manner. We proceed to this adjustment together with the moves of categories 'Re-engagement' [4] and 'Other' [6] into the 'Unknown' category in the Section that follows.

#### 4.2.6 Comparison of monthly and quarterly activities

In this section we briefly outline how we proceed to generate quarters. The first task is to decide the criteria to allocate one activity per quarter. As each one is made of three months, a natural process is to simply take the activity which appears the most in each quarter. We therefore simply allocate the activity which is recorded for 2 or 3 months in each quarter. By doing so, we are therefore allocating to each quarter the activity that has been exerted for the majority of time during each quarter. This reduces the number of periods to 11, with the last period being made of two months only (March and April 2014). The first period includes the month of September, October and November 2011, the second one December 2011 to February 2012, and so on until the end of the window of observation (i.e. April 2014). We list in the Appendix on Table A18 the months used to build the eleven quarters.

The resulting percentages as cross-sections are given in Table 4.11. We observe, similarly to the monthly analysis, the percentages following education start at around 90% just after completion of compulsory education and decline regularly to reach around 72% at the end of the period. This decrease is commensurate with an increase in employment over the same period. We also observe low figures for training, NEET and unknown. There is a clear spike in the 'unknown' category after completion of A-level during quarter 9 which spans over September-November 2013. This could potentially be accounted for by young people who fail to gain a place in HE and are considering their options. Indeed, there is a sharp increase in employment (and to a lesser extent in education) in quarter 10 that follows (December 2013 and January, February 2014).

**Table 4.11 Main activities by quarters, percentages**

Period (months, year in brackets)	Education	Employment	Training	NEET	Not known	3 different activities within the same quarter	Total
1 (9 - 11, 2011)	90.3	3.9	1.2	1.4	3.0	0.2	529,359
2	91.2	5.0	1.6	1.6	0.5	0.1	529,359
3	90.2	5.5	1.7	2.0	0.6	0.1	529,359
4	87.8	6.3	1.9	2.5	1.4	0.1	529,359
5	86.0	7.7	1.7	2.2	2.2	0.1	529,359
6	85.8	9.0	1.9	2.3	1.1	0.1	529,359
7	85.1	9.8	1.8	2.3	0.9	0.1	529,359
8	80.7	11.2	1.6	2.5	3.8	0.2	529,359
9 (9-11, 2013)	70.5	15.2	1.4	2.4	10.3	0.1	529,359
10 (12-2, 2014)	73.3	20.7	1.5	2.3	2.1	0.0	529,359
11 (3-4, 2014)	72.4	21.1	1.4	2.1	2.1	1.0	352,906

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

#### 4.2.7 Extent of churning over the age 16 to 19

**Table 4.12** Number of spells (using quarters) and performance at 16

Total number of spells	Number of good GCSEs at grades A*- C					Total
	0	1-3	4-6	7-9	10+	
1	13.5	18.2	17.8	30.9	19.6	111,822
2	19.0	26.7	23.1	23.3	7.8	25,530
3	16.8	21.6	20.5	26.8	14.3	29,142
4	35.0	29.2	17.4	13.4	5.0	5,876
5	34.0	25.0	14.1	17.5	9.5	3,094
6	49.9	26.6	12.6	7.6	3.3	749
7	60.7	22.9	11.4	5.0	0.0	201
8	-	-	-	-	-	32
9	-	-	-	-	-	7
<b>Total</b>						<b>176,453</b>

Notes: - means figures are too low to convey meaningful statistical inference

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

**Table 4.13** Number of spells (using quarters) and performance at 16 for those who experienced at least one spell as NEET

Total number of spells	Number of good GCSEs at grades A* to C					Total
	0	1-3	4-6	7-9	10+	
1	126	27	11	1	0	165
2	612	389	208	162	47	1,418
3	455	270	134	96	28	983
4	357	173	53	28	10	621
5	175	77	36	13	2	303
6	89	31	5	5	1	131
7	32	9	5	1	0	47
8	5	0	2	0	0	7
9	1	0	0	0	0	1
<b>Total</b>	<b>1,852</b>	<b>976</b>	<b>454</b>	<b>306</b>	<b>88</b>	<b>3,676</b>

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

#### 4.2.8 Mobility between schools and transitions between activities post 16

Mobility post-16 could be correlated with mobility during compulsory education. One obvious way to investigate mobility over the period age 6 and 16 is to check changes of school between KS1 and KS2 first and then changes of school between KS3 and KS4, as the majority of pupils will be expected to attend the same primary school for KS1 and KS2 and the same secondary school for KS3 and KS4.

## 4.2.9 Extent and duration of the NEET state

**Table 4.14** Number of quarters in NEET

Number of quarters as NEET (x)	Freq.	Per cent	Prop. 'NEET' at least x periods
0	161,425	91.48	100.00
1	5,410	3.07	9.32
2	3,423	1.94	6.24
3	2,307	1.31	4.30
4	1,353	0.77	2.99
5	811	0.46	2.22
6	552	0.31	1.76
7	479	0.27	1.45
8	257	0.15	1.18
9	161	0.09	1.03
10	110	0.06	0.94
11	165	0.09	0.88
<b>Total</b>	<b>176,453</b>	<b>100</b>	

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

The table tells us that more than 9% of the balanced sample has been NEET at any of the quarters during the period. This underlines the importance of not relying on cross-sectional analysis as it tends to underplay the significance of the status. Only 6.2% have been NEET for two quarters, and 4.3% for three quarters. A very small number (165) of young people have been NEET the whole period. The expected average stay as NEET is a bit more than a quarter (1.2), equivalent to approximately 6.5 weeks. But this average hides wide variations with a significant number of young people in NEET for a very long time (e.g. 3% for a year).

**Table 4.15** Number of quarters in NEET and performance at GCSE

Number of quarters in NEET	Number of good GCSEs A*-C					Total
	0	1-3	4-6	7-9	10+	
0	22,280	31,770	30,906	48,196	28,273	161,425
1	1,752	1,669	1,071	737	181	5,410
2	1,228	1,059	610	413	113	3,423
3	975	705	363	214	50	2,307
4	686	417	164	71	15	1,353
5	461	226	91	31	2	811
6	332	141	60	17	2	552
7	310	119	35	14	1	479
8	162	66	21	7	1	257
9	111	36	11	3	0	161
10	67	33	6	4	0	110
11	126	27	11	1	0	165
<b>Total</b>	<b>28,490</b>	<b>36,268</b>	<b>33,349</b>	<b>49,708</b>	<b>28,638</b>	<b>176,453</b>

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

How do persistent NEETS perform during their (compulsory) school? Are there distinctive events during their school experience that could help us identify young people at risk of becoming NEETS?

## 4.3 Analysis of individual transitions between states

### 4.3.1 Description

In this section, we investigate further the transitions between different states. We merged the categories 4 (Reengagement) and 6 (Other) into the unknown categories (7). This reduced the categories to: Education [1], Employment [2], Training [3], NEET [5] and unknown [7].

**Table 4.16 Transition between states in the full sample (all transitions)**

	Education	Employment	Training	NEET	Unknown	Total
Education	1,585,265	24,968	5,373	10,887	33,740	1,660,233
	95.48	1.50	0.32	0.66	2.03	100.00
Employment	3,760	158,472	737	2,216	1,043	166,228
	2.26	95.33	0.44	1.33	0.63	100.00
Training	1,814	2,756	21,080	2,403	776	28,829
	6.29	9.56	73.12	8.34	2.69	100.00
NEET	3,842	5,973	3,068	24,126	1,044	38,053
	10.10	15.70	8.06	63.40	2.74	100.00
unknown	16,883	11,277	984	2,097	16,399	47,640
	35.44	23.67	2.07	4.40	34.42	100.00
<b>Total</b>	<b>1,611,564</b>	<b>203,446</b>	<b>31,242</b>	<b>41,729</b>	<b>53,002</b>	<b>1,940,983</b>
	<b>83.03</b>	<b>10.48</b>	<b>1.61</b>	<b>2.15</b>	<b>2.73</b>	<b>100.00</b>

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

This transition matrix provides the total number of transitions experienced by young people in the sample. We observe that a vast majority of young people remained in the education category throughout the period. The unit of analysis in this table is a person-quarter observation. A quarter in education is followed by another quarter in education in 95% of the time. We can interpret the other cells in the table similarly. At the other end of the mobility spectrum, we observe that the probability to remain NEET is much lower. The largest number of moves out of NEET are into employment (16%), followed by moves into education (10%) and training (8%). It is interesting to note that by far the activity which has the lower probability of being repeated in two consecutive quarters is the unknown category, which underlines the transitory nature of this category. About 35% of quarters in this category are followed by quarters in education and 24% by quarters in employment.

This table provides guidance into which transitions are the most common in our sample and over the three years following completion of compulsory education.

As a next step we propose a preliminary investigation of some of the transitions most likely to happen. We propose those first results as preliminary results only, and to exemplify the type of analysis that could be performed more comprehensively in research in long-term outcomes.

### 4.3.2 Estimates of duration models

**Table 4.17 Transition out of education to employment, training, NEET and Unknown**

	Coefficients and Standard Errors
Number of GCSE (at levels A-C)	-0.098*** (0.001)
Female	-0.071*** (0.008)

	<b>Coefficients and Standard Errors</b>
West Midlands	-0.235*** (0.016)
North East	0.353*** (0.019)
North West	-0.294*** (0.015)
Yorkshire and the Humber	-0.123*** (0.016)
South-West	-0.314*** (0.021)
South-East	-0.243*** (0.015)
East of England	-0.253*** (0.017)
East-Midlands	-0.079*** (0.016)
Indians	-0.931*** (0.033)
Pakistanis	-0.719*** (0.026)
Bangladeshis	-0.302*** (0.035)
Black Caribbeans	-0.408*** (0.039)
Black Africans	-0.802*** (0.035)
Mixed	-0.219*** (0.025)
Other	-0.401*** (0.015)
Constant	-2.729*** (0.016)
p	1.177*** (0.004)
YP-quarter cells	1569550
Transitions (failures)	69824
Number of YP	176453

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

This first table shows the estimates of a duration model. The model is estimated through a maximum likelihood method assuming the hazard function follows a Weibull distribution - a

common choice in this type of analysis. We interpret the coefficients as affecting the probability of leaving the initial status (activity) and compared to the reference category. So for example, the negative and significant coefficient for female implies that teenage girls tend to leave education to other destinations earlier than boys, everything else kept constant. Similarly, and more surprisingly, young people tend to leave education the earlier the more GCSEs they obtained upon completion of compulsory education. We later discuss this result more in depth.

We see clear regional differences. London is the base (reference) category. The North-East is the only region where transition out of education is happening later than in London. Young people living in other regions tend to leave education earlier than in London. We can observe, as well, clear differences according to ethnicity. Young people of all ethnic groups have a tendency to leave education earlier than White British (the reference category). Of course, those results apply to all exits from education to any destinations (whether employment, training, NEET or the Unknown category). It is highly likely, however, that the observed characteristics correlates differently with exits to, say, NEET than to employment.

To account for this possibility, we perform a second set of regressions where we differentiate according to the destinations. With four destinations for each of our five categories, there are up to 20 transitions that could be investigated. Given the preliminary nature of this report, we choose the ones that, we think, are the most relevant for policy.

We present the results of four transitions in Table 4.18. Given the complexity of the problem, we investigated only the first transitions out of particular activity. So we analyse the first moves out of education and not the following ones if the same young people re-enter education and leave it again to the same activity. In practice, such two repeated transitions between two identical activities for the same person do not happen very often in this data set.

**Table 4.18 Transition from education to employment and training, and from training to NEET and employment**

	<b>Education to employment Coefficients</b>	<b>Education to NEET Coefficients</b>	<b>Training to NEET Coefficients</b>	<b>Training to Employment Coefficients</b>
Nber of GCSE (at A-C)	-0.111*** (0.002)	-0.250*** (0.003)	-0.248*** (0.031)	0.175*** (0.013)
Female	-0.200*** (0.013)	0.003 (0.020)	0.094 (0.083)	-0.119 (0.077)
West Midlands	0.536*** (0.039)	-0.046 (0.055)	-0.386 (0.281)	-0.238 (0.232)
North East	0.558*** (0.049)	1.112*** (0.056)	0.800*** (0.292)	-0.266 (0.278)
North West	0.932*** (0.036)	0.924*** (0.046)	0.244 (0.272)	-0.320 (0.238)
Yorkshire, the	0.976***	0.496***	-0.471*	0.078



	<b>Education to employment Coefficients</b>	<b>Education to NEET Coefficients</b>	<b>Training to NEET Coefficients</b>	<b>Training to Employment Coefficients</b>
Humber	(0.037)	(0.049)	(0.274)	(0.224)
South West	1.361*** (0.040)	0.626*** (0.059)	0.635* (0.327)	0.022 (0.353)
South East	0.503*** (0.038)	0.560*** (0.049)	-0.094 (0.293)	0.011 (0.242)
East of England	1.081*** (0.038)	0.412*** (0.053)	0.407 (0.292)	0.057 (0.251)
East Midlands	1.156*** (0.037)	0.352*** (0.052)	-0.037 (0.291)	-0.027 (0.245)
Indians	-1.173*** (0.069)	-0.542*** (0.090)	0.158 (0.458)	-0.610 (0.415)
Pakistanis	-1.330*** (0.060)	-0.243*** (0.054)	-0.155 (0.243)	-0.645** (0.282)
Bangladeshis	-1.272*** (0.110)	-0.258** (0.100)	0.158 (0.391)	-0.959 (0.716)
Black Caribbeans	-0.903*** (0.100)	-0.177* (0.107)	0.571 (0.388)	0.801** (0.324)
Black Africans	-2.023*** (0.131)	-1.048*** (0.122)	-14.172 (832.852)	-1.185 (1.003)
Mixed	-0.475*** (0.049)	0.225*** (0.057)	0.541** (0.226)	0.005 (0.264)
Other	-0.618*** (0.028)	-0.265*** (0.039)	0.087 (0.161)	-0.133 (0.150)
Constant	-4.397*** (0.038)	-4.477*** (0.048)	-3.095*** (0.275)	-3.581*** (0.237)
p	0.070*** (0.006)	-0.011 (0.009)	0.302*** (0.033)	0.439*** (0.029)
p	1.173 (.004)	.988 (.009)	1.352 (.004)	1.551 (.045)
Transitions (failures)	23,856	9,978	599	751
Number of YP	176,453	176,453	2,150	2,150

Source: NPD KS4 2010/11-NCCIS (September 2011-April 2014)

### ***Transitions from education (columns 1 and 2)***

The vast majority of observed characteristics correlate highly significantly for transitions from education. First, academic performance (measured by number of good GCSE) improves the transition to employment. More surprisingly, good performance appears also to accelerate transitions from education to the NEET status. This could be accounted for by a gap year effect, where some well qualified young people describe themselves as NEET for a period directly following their full-time education. Young women tend to transition earlier into employment, taking shorter periods of study. The coefficient is not statistically significant for the transition to NEET. Compared to London (the reference region), young people in every region tend to leave education and take employment at a later time. This is particularly the case in the South West, East England and the East Midlands. Transitions to NEET status happen the latest in the North East and the North West. Also, young people in the West Midlands transit from education to the NEET status less frequently than in London.

Interesting results can be observed by ethnicity. All ethnic minorities enter into NEET at earlier periods, this appears particularly the case for Black Africans. The performance of Black Africans in the labour market has already been documented in the literature. The transitions from education to NEET appears to happen at a similar time for young people from Black Caribbean and White British backgrounds.

### ***Transitions from training (Columns 3 and 4)***

While the vast majority of episodes in education follow-on straight from compulsory education, this is interesting to check whether it is also the case for episodes of training.

From the time they exit compulsory education, young people who move into training subsequently move to NEET after 4.1 quarters on average, whereas the moves to employment happens on average after 4.7 quarters. 90% of the transitions from training to NEET happen within 7 quarters of leaving compulsory education. The related figure for transition from training to employment is 8 quarters (2 years). Thus, **most episodes of training happen very soon after leaving compulsory education as well.**

From the start of the training period, the average time after which transitions are made to NEET is 3.1 quarters, whereas the value is 3.8 quarters for transitions from training to employment. This is interesting as **it may indicate that young people who transition from training to employment have on average longer periods of training. So longer training appears associated with more successful transition (to employment rather to NEET).**

With regard to how observed characteristics correlate with transitions, we find many fewer statistically significant coefficients in the transition from training compared to the transition from education. This could be due to a problem of sample size. One interesting difference is the effect of the number of good GCSEs. **Young people with more GCSEs tend to leave training to enter NEET status at earlier stage. The opposite holds true for transitions from training to employment. This appears to suggest that training combined with good academic performance lead to later transitions to employment.** Also, we find that Black Caribbeans transition later than White British and Pakistanis earlier. Again we find that young people in the North East transit later to NEETs, the effect

is highly statistically significant. We do find some (weaker) evidence that the transition to NEET happen earlier in Yorkshire and later in the South West, but the coefficients barely pass the test of significance (at 90% level only).

# 5 Summary and recommendations for further research

## 5.1 Summary

This research report provided empirical estimates of the movement of young people into the labour market, the formation of employment trajectories and the long-term outcomes of successful transition to the labour market using all currently available sources of micro data for the UK. We started with an extensive analysis of the UK Labour Force Survey data available for the last 40 years, which shows:

- Long-term changes in the youth labour market comparing early and recent cohorts of young people in the UK labour market over the last forty years. The key finding is that early labour market transitions decreased as people extend the time they spend in education.
- We found structural changes affecting young people, in particular a rise of youth unemployment from the 1980's onwards.
- Extending the available analysis, we used the 'Ad Hoc module' of the UK data from the European Labour Force Survey on youth labour market transitions in 2009 to explore whether the reported transition from education to the first significant employment (of more than three months) affects adult employment in later years. We also estimate the effect of education attainment, the work experience during the time in education and the orientation of education (i.e. whether vocational or general) on the duration of the initial transition. These results indicate that work experience and education in the workplace significantly improve people's initial transitions, which has beneficial long-term outcomes when people are observed in later years.

The second major contribution of this analysis is systematic research on the biographies of young people after the end of compulsory schooling. This analysis was based on large cohort studies<sup>1</sup> and compared young people's experiences when making transitions into the labour market in youth and early adulthood, for a period of five years after the end of compulsory education.

This research confirms the evidence from the LFS analysis about the growing group of people remaining in education until longer and shows otherwise comparatively stable patterns of labour market sequences following initial transitions. The following findings can be summarised:

- The sequential analysis shows that the group of people with difficult transitions increases over time, but because YCS and LSYPE data only provide individual

---

<sup>1</sup> National Child Development Study [NCDS], British Cohort Study [BCS], Youth Cohort Study [YCS], Longitudinal Study of Young People in England [LSYPE].

employment and education outcomes until age 19, only earlier cohorts (BCS and NCDS) could be analysed in relation to such long-term outcomes (people who are now in their 40s and 50s).

- The long-term outcomes show that people from 'entry-into-labour-market' trajectories were affected by unemployment in their early mid-20s, rather than a long time after the end of education. This, however, could also be business cycle effects as labour market outcomes of both cohorts (when in their early to mid-20s) were affected by the recessions of the early 1980's and the early 1990's.
- The growing cluster of people continuing to invest into education is less affected by unemployment than people with an 'entering-the-labour market trajectory'. Following their extended education experiences, these people make successful and sustainable transitions in their mid-20s.

Finally, the analysis described extensively the transitions experienced by a recent cohort of people after obtaining GCSEs based on a merged data set of the NPD and the NCCIS. This part of the report shows a comprehensive picture of the young people experience, including transitions in and out of the NEET state. As was found in other studies, including Dorsett and Lucchino (2013), more than 90 percent of the cohort experience unproblematic transitions in the years directly after the GCSE and have not had any NEET experiences until the age of about 19.

## 5.2 Policy conclusions

Overall, the project provided a comprehensive picture of initial labour market and education experiences and pathways that help young people into sustainable employment based on all currently available data to research the long-term outcomes of particular experiences made after the end of compulsory schooling.

Unsurprisingly, good education and successful GCSEs are important for people to show long-term labour market success. In addition, differences in transitions patterns by gender and ethnic origin exist. Further, and much in line with the findings from the literature review, early labour market experience, both before and during further education, and vocational education in the workplace significantly reduce the time between leaving the education system and first significant employment. Based on the 2009 AHM analysis, reduced unemployment for young people and the time it takes to make transitions into significant employment in early adulthood also improve labour market success in later life.

This emphasises the importance of policies to help young people in having early labour market experience, for long-term labour market success.

## 5.3 Further research

In addition to identifying current patterns in transition pathways and investigating their implications for future labour market success and welfare of young people, the project also aimed to investigate the **types of learning** and employment that young people are involved in, whether they **frequently switch between statuses** and what the implications of this would be for their long-term labour market outcomes. However, since indicators for

employment and training quality are not being systematically collected for young people when making transitions to the labour market along with suitable variables on long-term labour market outcomes, the project was not able to provide BIS with such evidence and consequentially ended.

The new Centre for Vocational Education Research (CVER), which BIS created, will make research data available, which can be used in order to provide new empirical evidence on the long-term outcomes of transitions with much richer variables about the type of learning and employment people engage in. Three contributors to this report (Dr Augustin De Coulon, Dr Stefan Speckesser and Vahé Nafilyan) will extend the research in focus in this report in the near future and produce new evidence, which will be a step-change in our understanding of young people's transitions to the labour market and the impact of vocational education quality.

## 6 References

Cappellari L and Jenkins S (2004), 'Modelling Low Income Transitions, *Journal of Applied Econometrics*, 19, 593-610.

Dexter R (2014), *NEET Quarterly Brief –April to June 2014*, Statistical First Release, SFR 29 /2014, Department for Education, Office for National Statistics, accessed at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/346024/Quarterly\\_Brief\\_NEET\\_Q2\\_2014\\_FINAL.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/346024/Quarterly_Brief_NEET_Q2_2014_FINAL.pdf)

Dorsett R, Lucchino P (2013), *Visualising the school-to-work transition: an analysis using optimal matching*, NIESR Discussion Paper, 414

Gregg P (2001), 'The impact of youth unemployment on adult unemployment in the NCDS', *The Economic Journal*, 111, F626-F653

Heckman J (1981), 'Statistical Models for Discrete Panel Data', in: *Structural Analysis of Discrete Data with Econometric applications*, Manski and McFadden eds, MIT Press.

McIntosh S (2005), *Using pseudo-cohorts to track changes in the qualifications of national populations*, Department for Education and Skills Research Report

Mincer J (1974), *Schooling, Experience, and Earnings*, National Bureau of Economic Research

Miza-Davies J (2014), *NEET: Young People Not in Education, Employment or Training*, Commons Library Standard Note SN06705, accessed 1 October 2014: <http://www.parliament.uk/briefing-papers/SN06705/neet-young-people-not-in-education-employment-or-training>

© Crown copyright 2015

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

This publication available from [www.gov.uk/bis](http://www.gov.uk/bis)

Any enquiries regarding this publication should be sent to:

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

If you require this publication in an alternative format, email [enquiries@bis.gsi.gov.uk](mailto:enquiries@bis.gsi.gov.uk), or call 020 7215 5000.

**BIS/15/612**