# **Youth in Transition: Longitudinal Comparisons of Youth Transitions in the UK using Cohort and Synthetic Cohort Data**

Contents

[Youth in Transition: Longitudinal Comparisons of Youth Transitions in the UK using Cohort and Synthetic Cohort Data 1](#_Toc144117907)

[Chapter One: The National Childhood Development Survey: Youth Transitions in Education and Employment 6](#_Toc144117908)

[Introduction to Chapter One 6](#_Toc144117909)

[Literature Review: NCDS Timeframe and Context 7](#_Toc144117910)

[Story of transitions for NCDS youth 8](#_Toc144117911)

[Structural Barriers to successful transitions – the role of social-class and sex 12](#_Toc144117912)

[The role of social theory 14](#_Toc144117913)

[Data and Methods 17](#_Toc144117914)

[Introduction to the NCDS data 17](#_Toc144117915)

[Introduction to measures for subsequent analysis 19](#_Toc144117916)

[Descriptive Statistics 30](#_Toc144117917)

[Modelling Main Economic Activity 31](#_Toc144117918)

[Discussion and Conclusion 44](#_Toc144117919)

[Sensitivity Analysis of Independent Variables 46](#_Toc144117920)

[Testing Measures of Parental Social Class 46](#_Toc144117921)

[Discussion and Conclusions 53](#_Toc144117922)

[Missing Data in the NCDS 53](#_Toc144117923)

[Missing Data 53](#_Toc144117924)

[Multiple Imputation by Chained Equations 56](#_Toc144117925)

[Discussion and Conclusions 63](#_Toc144117926)

[Discussion and Conclusions for Chapter One 64](#_Toc144117927)

[Chapter Two: The British Cohort Study: Youth Transitions in Education and Employment 67](#_Toc144117928)

[Introduction to Chapter Two 67](#_Toc144117929)

[Literature Review: BCS Timeframe and Context 67](#_Toc144117930)

[Story of transitions for BCS youth 68](#_Toc144117931)

[Structural Barriers to successful transitions – the role of social-class and sex 71](#_Toc144117932)

[Data and Methods 72](#_Toc144117933)

[Introduction to the BCS data 73](#_Toc144117934)

[Descriptive Statistics 77](#_Toc144117935)

[Data Citation 88](#_Toc144117936)

[Bibliography 88](#_Toc144117937)

List of Tables

[Table 1. 1 Sweeps Included for Analysis 17](#_Toc137904670)

[Table 1. 2 Participation in the NCDS from birth to 23 years 18](#_Toc137904671)

[Table 1. 3 Frequency Statistics for Economic Activity 19](#_Toc137904672)

[Table 1. 4 Educational Attainment Count Variable by Economic Activity 22](#_Toc137904673)

[Table 1. 5 RGSC Class Schema 26](#_Toc137904674)

[Table 1. 6 NS-SEC Class Schema 27](#_Toc137904675)

[Table 1. 7 Examples of Occupations from Analytical NS-SEC 28](#_Toc137904676)

[Table 1. 8 Descriptive Statistics for Economic Activity 30](#_Toc137904677)

[Table 1. 9 Goodness-of-fit summaries for explanatory variables and Economic Activity 31](#_Toc137904678)

[Table 1. 10 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity 32](#_Toc137904679)

[Table 1. 11 Mlogit of Economic Activity 36](#_Toc137904680)

[Table 1. 12 Sensitivity analyses of alternative measures of parental social stratification 47](#_Toc137904681)

[Table 1. 13 Missing data patterns for NCDS 55](#_Toc137904682)

[Table 1. 14 Comparison of CCA NS-SEC vs Imputed NS-SEC 58](#_Toc137904683)

[Table A. 1 Goodness-of-fit summaries for explanatory variables and Economic Activity (CAMSIS) 66](#_Toc137904684)

[Table A. 2 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity (CAMSIS) 66](#_Toc137904685)

[Table A. 3 Goodness-of-fit summaries for explanatory variables and Economic Activity (RGSC) 67](#_Toc137904686)

[Table A. 4 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity (RGSC) 67](#_Toc137904687)

[Table A. 5 Average marginal effects on the probability of Economic Activity 69](#_Toc137904688)

List of Figures

[Figure 1. 1 Predicted Probabilities of Economic Activity by NS-SEC 40](#_Toc137904689)

[Figure 1. 2 Predicted Probabilities of Economic Activity by Sex 41](#_Toc137904690)

[Figure 1. 3 Predicted Probabilities of Economic Activity by Educational Attainment 42](#_Toc137904691)

[Figure 1. 4 Predicted Probabilities of Economic Activity by Housing Tenure 43](#_Toc137904692)

List of Abbreviations

BCS British Cohort Survey

CCA Complete Case Analysis

CSE Certificate of Secondary Education

MI Multiple Imputation

NCDS National Childhood Development Study

NEET Not in Education, Employment, or Training

NS-SEC National Statistics Socio-economic classification

NVQ National Vocational Qualification

O-Level Ordinary Level

RGSC Registrar General’s Social Classes

TOPs Training Opportunities Scheme

UKHLS United Kingdom Household Longitudinal Study

# **Chapter One: The National Childhood Development Survey: Youth Transitions in Education and Employment**

"Let us be frank about it: most of our people have never had it so good". – Harold Macmillan (1957)

# **Introduction to Chapter One**

Transitions in education and employment after the end of compulsory schooling in the UK are important sociological phenomena. This chapter analyses said phenomena. It does this by examining what happened to the NCDS cohort between the time they left school at 16, in 1974. The focus of this chapter is on the way structural inequalities like social class and gender shape young people’s choices and transitional experiences. The NCDS study provided a unique resource to examine the transitions from education to employment for young people leaving education in the 1970s.

There is a wealth of previous research on school-to-work transitions using the NCDS (Machin and Vignoles, 2005; Schoon, 2007; Schoon, Martin and Ross, 2007; Dex, Ward and Joshi, 2008; Martin, Schoon and Ross, 2008). This present chapter will attempt to revise this topic employing modern data analysis techniques to replicate historic analyses of youth transitions.

The data used comes from survey interviews at age 23, whereby individuals filled out a monthly retrospective economic activity calendar, between the ages of 16-23. Individuals within the NCDS were interviewed at many intervals, and so previous sweeps will be used to supplement analysis. This chapter will focus on at those individuals that entered employment, post-schooling education, continued school, training & apprenticeships, and unemployment & out of the labour force. Such data provides an analysis of the transition from mandatory schooling. Given the focus on transitions, the interest within the analysis is looking at what accounts for differences between people in their transitions and what are the essential features of the NCDS cohorts’ transitions.

## **Literature Review: NCDS Timeframe and Context**

This section provides an overview of the literature within the field of youth transitions. This literature review focuses on existing research outlining the school-to-work transition and the examination of the impact of structural inequalities has upon that transition. A major focus is placed upon the role of social class and gender. A broader focus will also look at the nature of choice and opportunity for the NCDS cohort and how this is impacted by structural inequalities. Within this review, major transition trajectories have been identified. Trajectories into further education, apprenticeships, employment, and unemployment are identified. Each of these trajectories appears to have an element of structural inequalities influencing the outcomes of individuals. These empirical findings are expounded upon with a review of the theoretical literature to provide a holistic overview of the school-to-work transition during the NCDS period.

At the time of the NCDS cohort, young people were in full time education until the age of 16. At this age, individuals were typically expected to sit some form of examination. Most common for this time was the O-level or Ordinary level (Pearson qualifications, 2023b). After this period of examination and the end of mandatory schooling individuals are open to a level of choice in what and where they wish to take their lives next. Some choose to stay within education and attend a sixth-form college and go on to a university, others went straight into full-time employment, others went to training, and some experienced forms of unemployment. This section reflects the diversity of choice and opportunity open to (some) individuals by exploring the literature in these areas.

### **Story of transitions for NCDS youth**

Education as a system regulates the individual by implementing age-graded barriers and hierarchical and time-related credentials. The labour market through its regulatory function determines who is gainfully employed and who is unemployed which influences employment trajectories (Mayer, 2004: 164). The structure and hierarchy of occupations determines social position via segmentation and segregation which is in part determined by previous systems of employment and education. The outcome of a child post-mandatory schooling impacts their life chances across their ‘life course’. Functionally, the study of youth transitions is the study of the life course, systems of education, occupation, and labour markets constitute life domains onto which the life course manifests (Mayer, 2009). Mayer argues that these life domains are organised around three major life stages: the phase of education and preparation to work, the phase of active employment, and the phase of postretirement years (Mayer and Schoepflin, 2022: 196). The study of school-to-work transitions constitutes two of these life domains. These life stages vary in timing and sequencing depending on socio-historical context (Shanahan, 2000) which needs to be acknowledged for adequate sociological inquiry. Hence a desire to provide a through context of the NCDS cohorts socio-historical context. The following section provides this socio-historical context in the form of empirical research on existing youth transitions literature within the NCDS.

Part of this context forms an understanding of what exactly a ‘transition’ looked like for NCDS youth. Within theoretical youth transition literature, the notion that in the mid-20th century, transitions were smoother, and more simplified has taken root. During the time of the NCDS the range of choice has been argued to have been narrower comparative to later cohorts, thus owing to a more homogenised pathway (Goodwin and O’Connor, 2005: 4). This section problematises this notion by adding additional context. Whilst it is true that the NCDS cohort had large homogenous clusters as they relate to transitioning out of mandatory education (Martin et al 2008), the story is slightly more complicated. The delineation between school and employment is not a strict binary – with many children engaging in what is known as the youth labour market (Bynner, 2012) whilst still in education. The choices that children make on what they wish to do after mandatory education is influenced by labour market restructuring and recession – as well as structural factors such as family background. Whilst the pathways the NCDS cohort may have been able to choose from were narrow, this is not synonymous with smooth, nor straight-forward. For example, individuals may have faced a seemingly homogenous experience of after a period of schooling ending up in a period of employment, but that does not indicate the relative smoothness of getting from A to B – some individuals may have experienced periods of unemployment during this time, others jumping from job to job etc. The school-to-work transition for NCDS youth is complex and an appreciation of choice and opportunity – and the subsequent restriction of choice and opportunity based upon structural factors needs to be considered.

Using Optimal Matching Analysis to analyse the various routes of youth transition into work, (Martin et al, 2008) evidence states that of the NCDS cohort 96 per cent of men could be grouped into six of the largest transition ‘clusters’ compared to 90 per cent for the BCS. This suggests the NCDS cohort exhibits ‘homogenised pathways’ (Goodwin and O’Connor, 2005). For the NCDS, the predominant pattern was to leave school post-16 and move directly to employment (Schoon, 2007 :98). This is supported by sequence analysis under the likes of (Anders and Dorsett, 2017 :75), where patterns of transition amongst school leavers entering the labour market were looked at. They found that under the NCDS cohort there was a strong (91 per cent) amount of people entering into the labour market straight after mandatory schooling. This once again, supports the view that the NCDS cohort exhibited homogenous pathways of transition. The typical pathways that young people within the NCDS entered demonstrate that above all, a school-to-employment transition was dominant, though training was also an important (Schoon *et al.*, 2001) transition pathway – above that of continuing in full-time education. The NCDS cohort was however caught in a period of severe diminishing influence of apprenticeships. The number of apprenticeships in British manufacturing for example, declined from 240,400 in 1964 to 155,000 in 1979 (Blanchflower and Lynch, 1992). There was a severe gender bias when it came to apprenticeships at this time - when the NCDS cohort were 16 years old 40 per cent of male employees were apprenticed compared with only 8 per cent of females (Blanchflower and Lynch, 1992: 240).

The relative importance of employment and apprenticeship training over educational pathways suggests that the NCDS cohort experienced a pre-credentialed labour market (Bynner, 2005) post-mandatory schooling. In other words, the NCDS cohort experienced a labour market that didn’t place hard roadblocks to employment based upon educational credentials. It wasn’t until the 1980s that failing to get qualifications became a hindrance to getting work in Britain (Bynner, 2005 :377). The labour market pre-1980s was able to absorb people into large numbers of unskilled jobs (ibid). Those that did struggle to get jobs in the NCDS cohort (Bynner, 2005 :378) were significantly more likely to lead a ‘Not in Education, Employment, or Training’ (NEET) status going forward post-21 years old (Bynner 2005: 378). This is indicative of the fact that it was only in 1975 that O-levels moved from a pass/fail system to a graded one (*Pearson qualifications*, 2023).

Thus far, it has been established that the NCDS cohort exhibited a homogenous transitional experience. Historically however, the NCDS birth cohort didn’t experience a straightforward smooth school-to-work transition. Teenagers that were still in education typically engaged in what is known as the youth labour market (Bynner, 2012). Most of this work was part-time during educational studies (Dustmann *et al.*, 1996 :80). It is estimated that children aged 16 worked on average between six and nine hours a week and modal earnings in the range of £1-£2 a week (Dustmann *et al.*, 1996 :86) whilst they were still in full-time mandatory education. Within the NCDS cohort, half of 16-year-olds in 1974 had a part-time job during term time (ibid). This suggests that a straightforward delineation separating school and work is an over-simplification for the time. Children were engaging in schooling as well as employment prior to making the choice of what to do after mandatory schooling. Family income did not impact the effect of childhood part-time employment participation but parent’s unemployed status did (Dustmann, Rajah and Smith, 1997). Student employment has an adverse effect on educational decisions and choices related to continuing education post-mandatory schooling (Neyt *et al.*, 2018). This adds a layer of complexity to the aforementioned ‘homogenised pathways. Structural inequalities – in the form of family background and unemployment status – have a role to play in the choices and opportunities of youth transitions. Homogenised pathways thus are not the same as smooth transitions. Structural inequality adversely impacts the relative smoothness of an individual’s transitional experience.

Speaking as a cohort, the relative ‘smoothness’ of youth transitions from school-to-work is primarily dependent upon the relative stability of the labour market that such individuals are transitioning into. During the time of the NCDS cohort the labour market was experiencing a major period of restructuring – as well as the relative collapse of the youth labour market (Bynner, 2012). As such, the labour market at the time of school-to-work transitions for the NCDS cohort was one of instability and comparative heightened uncertainty (Leuze, 2010). The collapse of the youth labour market in the early 1980s was not a sudden affair (Bynner, 2012). In fact from between January of 1972 and January of 1977 unemployment among 16 and 17 year olds rose by 120 per cent (Maclure, 1978). Contributing factors such as the demise of heavy industry, the collapse of community networks, and the technological transformation of modes of production were all forces that the 1958 cohort were facing during their biographical lifespan (Bynner, 2012: 40). The importance of the collapse of the youth labour market relates to the introduction of uncertainty at a critical stage of development within a young person’s life (Maclure, 1978). This uncertainty has the potential to adversely impact the life domains of individuals (Mayer, 2009). These periods of instability are documented in detail with employment histories on a monthly basis (Leuze, 2010 :86), suggesting that the collapse of the labour market impacted a large number of individuals. This heightened instability during a time of transition for the youth of the NCDS presents an influencing factor into the role of choice and opportunity. In a time whereby the labour market was facing severe restructuring, on top of a collapsing youth labour market, and a major economic recession, the choices and opportunities of young people seeking to transition into the world of work would be constrained and influence the choices they made. These constraints for example, created a likelihood for disengagement with school and alienation with education (Farrall, Gray and Mike Jones, 2020).

The development of this concept of uncertainty and risk within the NCDS stands in stark contrast to the theory of ‘late modernity’ - entailing notions of risk and uncertainty in a society that provides individuals with more choice which promotes greater risk. The literature has demonstrated that the NCDS cohort experienced complicated pathways and transitions that were influenced and impacted based upon structural inequalities. The notion of ‘Late Modernity’ (Giddens *et al.*, 1991; Beck et al 1994) is based upon the idea that in the past, more concrete certainties have given way to a more fluid and dynamic notions of adult identity and its development (Bynner, 1998: 31). These past certainties gave rise to stability; these current dynamics gives rise to risk (Beck, 2014). Existing literature details the fact that the NCDS cohort didn’t have explicitly concrete certainties – it exhibited relatively homogenous pathways, but the details within these pathways were often complex and influenced by uncertainty and risk of the time. There is debate over just how fluid certainty and choice has become, Gayle et al (2009) provides a more updated version of events that appears to review and ultimately question the late modernity outlook.

This theme of constraint is evidenced in the changing influence of educational attainment during the short-term for the NCDS cohort. Whilst educational attainment – and staying on within education post-mandatory schooling - does provide a layer of protection from unemployment (Bynner, Wiggins and Parsons, 1996). Those of the NCDS cohort that stayed on within education post mandatory schooling had initial higher levels of unemployment due to exogenous shocks of rising national unemployment – due to aforementioned labour market restructuring and economic recession. Whilst experiencing short term levels of unemployment, in the long run individuals that stayed on within education had a long-term advantage income wise over their peers that didn’t stay on within education (Payne, 1987).

Looking in more detail at educational attainment within the NCDS, individuals in the UK that choose to stay on at school post-16 were a small minority, and were low by OECD standards (Micklewright, 1989 :25). Those individuals from manual backgrounds compared to their non-manual peers were less likely to stay on post-16 (Micklewright, 1989 :37), further research concurs, stating that young people from working class backgrounds were less likely than middle class peers to remain in education post-mandatory schooling (Bynner and Joshi, 2002; Schoon, 2007). There is however an incongruence in the wants and aspirations of children from working class backgrounds – of whom only 10 per cent wished to continue to further education, compared to 39 per cent of their parents (Schoon and Bynner, 2003 :25).

Whilst individuals from manual backgrounds were less likely to continue stay on within education post-16 compared to their non-manual peers, a more complicated relationship arises when looking at apprenticeships. Whilst (Schoon *et al.*, 2001) finds that young people from less privileged backgrounds are more likely to be in training or apprenticeships further research suggests that apprenticeships amongst the NCDS cohort were more likely to be offered to children of fathers that were skilled manual workers over their semi-skilled counterparts (Booth and Satchell, 1994). This suggests a fragmentation of the traditional manual/non-manual divide, with a hierarchy of skills impacting the choice and opportunity of the NCDS youth.

Within the NCDS cohort training is found to lead to full-time employment (Schoon *et al.*, 2001). Schoon and their colleagues (ibid) suggest that this is primarily because apprenticeships during the NCDS period spanned three years or longer, providing the relevant skills and development for young people to effectively transition from a period of apprenticeship training, into a stable employment. Vocational based education is generally considered to be a smoother transition from school-to-work in comparison to its academic counterpart. Whilst this short term benefit is worthwhile to consider, long-term disadvantages such as lower employment and wages impact those individuals with lower vocational education (Brunello and Rocco, 2017). Though this as a phenomenon has broken down post-NCDS (ibid).

### **Structural Barriers to successful transitions – the role of social-class and sex**

#### **Sex**

Choice and opportunity within the school-to-work transition of the NCDS youth is influenced and impacted by structural inequality factors like sex. The roles that women assume within the labour market have marked differences compared to their male peers (Dex and Bukodi, 2012). Whilst women are more likely achieve their educational aspirations than men (Cebulla and Tomaszewski, 2013 :148) and often have higher occupational aspirations compared to men at a young age (Schoon 2007; Schoon, 2022), these aspirations rarely translate to higher than average incomes and in fact, labour market segregation remains, whilst pay improvement for men continues to outpace women’s (Dolton, Joshi and Makepeace, 2002; Makepeace, Dolton and Joshi, 2004).

Due to structural changes within the British labour market that started during the NCDS birth cohort time frame, part-time work began to grow from the 1950s – though this was synonymous with low-paid jobs and the service industry. Part-time work is pervasive amongst women in the NCDS, especially when returning from giving birth (Dex *et al.*, 1998 :79). A transition (Dex and Bukodi, 2012) from full-time to part-time work is associated with downward social mobility that primarily impacts women in the workforce.

There was some slight decline in gender segregation within the labour force for the NCDS cohort due to ‘feminisation’ of industry (the growth of soft skills labour), but overall gender segregation remained consistently stable (Guinea-Martin and Elliott, 2008; Lekfuangfu and Lordan, 2022). Pay equalisation has seen a general improvement for women in the NCDS cohort (Paci *et al.*, 1995; Neuburger, Kuh and Joshi, 2009; Roantree and Vira, 2018). The growth of women in the workforce (Schoon, 2010) has promoted gender equity as education reduces the difference in earning power between men and women as well as the difference in hours of paid and domestic work seen within couples (Joshi, 2002). However equalised rates of male labour force participation has not corresponded to a matched increase in the share of women’s earnings in the household (Joshi and Davies, 1996; Joshi, Makepeace and Dolton, 2007) and there is little differential treatment of the average women in full-time employment (Makepeace *et al.*, 1999).

In addition, in the later life-course, from ages 33-42, men’s real wages rose more than women’s (Dolton, Joshi and Makepeace, 2002; Makepeace, Dolton and Joshi, 2004) and men’s wages on average grow at a faster rate than women (Dex, Ward and Joshi, 2008). Income itself is influenced by sex. The adult earnings for men are impacted based on childhood factors or parental social status, for women, traits of intellect and emotional stability impact adult earnings (Furnham and Cheng, 2013). Traits related to parental social status and educational attainment also had an impact upon personal savings and levels of investment (Furnham and Cheng, 2019).

Unequal treatment for women has decreased but has not totally disappeared (Neuburger, 2010). Women have experienced marked differences in relation to labour market outcomes in comparison to their male peers. Whilst evidence suggests that there has been improvement in the area of sex based structural inequalities, they still persist to a substantive degree.

For women, unlike men, differences exist for those considered to be eventual full-time workers compared to their part-time counterparts. Eventual part-timers are more likely to enter working class positions regardless of their class origins and educational attainment. It appears that those part-time women from more advantaged social origins do not exploit such origins in the same way their full-time counterparts do (Bukodi *et al.*, 2017). When limiting attention of social mobility to solely full-time workers, mobility does not greatly vary by gender (Bukodi, Goldthorpe and Kuha, 2017), though research (Savage and Egerton, 1997; Savage, 2011) emphasises the impact gender has on social mobility. Part-time female workers have highly varied pathways and often combine periods of part-time employment with full-time employment and non-employment (Connolly and Gregory, 2010). Whilst broadly speaking the NCDS cohort experiences homogenous transitional pathways, some sub-groups, like female part-time workers, experience a much more complex less-smooth transition into employment.

#### **Social Class**

Class based structural inequalities impact educational attainment of NCDS youth during mandatory schooling (Galindo-Rueda, 2003; Sianesi, Dearden and Blundell, 2003; Holm and Jæger, 2011). This then consequently has an impact upon transition outcomes and later life-chances. When looking at educational attainment (Holm and Jæger, 2011 :1) it is important to consider that family background variables like social class matter (Machin and Vignoles, 2005), with the most advantaged children seeing the best returns (Sianesi, Dearden and Blundell, 2003). Variables such as parental education play a more important role in the life chances of young people than income (Feinstein, Duckworth and Sabates, 2004; Field, 2010). Early successful ability within education confers an advantage in later educational attainment and labour market experience (Dolton, Makepeace and Marcenaro‐Gutierrez, 2005). In other words, educational attainment leads to more educational attainment. Achieving while young, impacts educational attainment at later parts of the life course (Hutchison, Prosser and Wedge, 1979). As such the influence of family background on early educational attainment appears to influence later life-chances. Whilst educational inequality has declined in the NCDS cohort (Blanden and Macmillan, 2014), educational inequality still persists within high attainment.

Educational attainment translates to higher levels of wage growth in later-life. Individuals with higher educational ability experience faster wage growth compared to their peers with lower ability (Galindo-Rueda, 2003). Early successful ability is influenced however by a structural class effect. Those from working class backgrounds are less likely to be successful at early life-stages compared to their non-working class peers (Machin and Vignoles, 2005), in part due to poorer families being less likely to invest in education over their more affluent peers (Chevalier and Lanot, 2001).

Low levels of qualifications and educational attainment are related to higher propensities toward unemployment (Bynner and Parsons, 2000 :246). Propensity toward experiencing unemployment also has a social class effect, with the growth in unemployment during the NCDS cohort being attributed to the subsequent decline in the manufacturing sector that is linked to working class labour (Schoon *et al.*, 2001). Those that are unemployed also appear to hold the lowest levels of employment commitment. Unemployment is found within the NCDS cohort to have a scarring effect on potential earnings (Gregg, 2001; Bynner, 2012; Schoon, 2020) – the youth labour market thus plays an important role in establishing adult future earnings (Gregg, 2001 :628). Data suggests that a scar from early unemployment can have an estimated 12-15 per cent damaging impact on income at age 42 (Gregg and Tominey, 2005 :2). The influencing impact that social class has upon attainment and propensity toward unemployment appear to have long-term consequences for later life-chances. The impacts of social class on the transitions of youth from school-to-work are not just felt in the short term but in the long term also.

Work related training, or training on the job has been lauded as a way for those that enter the labour market with relatively low levels of education to build up necessary skills. A study by (Arulampalam and Booth, 1997) suggests the opposite is in fact the case. Work-related training seems to give a boost to the already well-educated and leave those less-educated behind. In a later study, (Arulampalam and Booth, 2001) re-affirm their findings by stating that while work-related training does improve wages, it positively affects the wages of the well-educated more so than the less-educated in the labour market. The fact that those that happen to be well-educated related to those that come from advantaged social class positions (Machin and Vignoles, 2005) demonstrates that advantage breeds advantage. Of those from less affluent backgrounds that choose to engage in work-related training, they will not see equal levels of growth associated with their affluent peers (Arulampalam and Booth, 2001).

The returns of higher education degrees – in other words, the income gained from educational attainment - appear to be substantial within the NCDS cohort. These returns whilst generally lower than undergraduate degrees also exist for higher degrees and non-degree higher education courses (Blundell *et al.*, 2000; Blundell, Dearden and Sianesi, 2001). Another study found that each successive qualification level at the National Vocational Qualification classification corresponds to a 5 per cent rise in income (Conlon, 2001).

Evidence suggests that those individuals with advantaged family background see occupational earnings increase by at least 7 per cent (Connolly, Micklewright and Nickell, 1992). Conversely, three or more months of unemployment is associated with a fall in occupational earnings by around 7 per cent (ibid). Unemployment at the youth stage provides a higher likelihood of unemployment at the adult stage of the life course (Gregg, 2001).

### **The role of social theory**

Over the course of this literature review two key themes have been expressed: the importance of looking at the life course to assess youth transitions and the role of structures in influencing choice and opportunity. These two themes will be better expounded upon within the literature of social theory to ground future analysis within a sociological tradition.

The life course is a term that seeks to do away with static ‘snapshot’ notions of sociology. Instead views the individual in a constant web of changing temporal context that is influencing the agent. The life course approach is best suited for an analysis of youth transitions using longitudinal data. It incorporates the changing processes and influences that ultimately impact an individual’s choices and opportunities when engaging in transitions during the youth stage.

The life course approach has established itself as a substantively significant research paradigm within the last few decades (Elder, 1994: 4). The term ‘life course’ is a concrete multilevel phenomenon that is defined via the social trajectories of individuals through structured pathways of given institutions that form the developmental experience of a given individual (Elder, 1994: 5). These ‘structured pathways’ are interwoven with what Elder argued were ‘age-graded trajectories’ (ibid). These trajectories took the form of work, family, and housing transitions. Such transitions are always historically and temporally located, giving them specific form, and meaning (ibid). The structured pathways that are interwoven within the life course support an analysis that focuses upon inequalities in relation to race, class, gender, and other structural aspects of social life (Bernardi et al, 2019: 1). The life course approach is implicitly linked with a study of youth transitions. Youth transitions by their very nature detail pathways of trajectories that individuals choose at certain points in their life that are ultimately influenced and dependent upon structural inequalities.

The definition that Elder gives of the principle of agency: ‘’individuals construct their own life course through the choices and actions they take within the opportunities and constraints of history and social circumstances’’ (Elder, 2003) also known as ‘’bounded agency’’ (Evans, 2007). Bounded agency is a concept that argues that agency of the individual is situational, and bounded to the circumstances of place and time (Bernardi et al 2019: 3).

By focusing upon a life course perspective, analysis can extend beyond static moments in time. This allows research to be expanded both in reference to within individual and between individual analysis. A life course approach appreciates the fact that structured pathways are temporally grounded and as such acknowledge that any youth transitions and trajectories must be understood within that embedded temporal context. The life course perspective lends itself to a study of youth transitions due to its focus upon the interdependence between life domains (ibid). This means that outcomes within one domain (say school) are interrelated with the outcomes and behaviours of other domains (say work). Finally, a life course perspective allows for insightful comparison across cohorts to study how such cohorts have responded differently to the consequences of their early transitions (Elder, 1994: 5).

The second theme of structuration argues that structural factors like social class, gender, ethnicity still play an important role in shaping the lives of individuals and are indeed determinants for the individual who is pursuing the ‘imperative of living a life of one’s own’ (Beck, 2002). Individualisation argues that in place of these ‘collective guides’ (Gayle et al, 2009: 4) individualised identities that have greater scope beyond the mere structures (Murray, 2011: 26) they inhabit are able to create complex and subjective lifestyles that deviate from the much more rigid structures detailed above (Gayle et al, 2009: 5).

If the individualisation thesis were to be correct it would demonstrate itself empirically and repeatably. However, as Gayle et al (2009) found) the thesis’ strong claim against structures is not to be born out within the data. Pathways toward transition may have certainly altered, and even in some cases become more complex, but that does not mean there is support for ‘detraditionalization’ (Gayle et al 2009). There has always been an element of navigation and choice within youth transitions. Though in the past the range of choice may have been narrower thus owing to a more homogenised pathway for those in past contexts (Goodwin and O’Connor, 2005: 4). In other words, individualisation fails to account for the still apparently strong influence of structural elements on a person.

Another critique towards structuration theory – the relevance of social class – argues that structures are important, but the specific role of social class is one that is on the decline. In other words, new structural cleavages have arisen over and above class based effects (Devine, 2017). This ‘new structuralism’ has argued that consumption-based cleavages – most important of which related to housing tenure – are more influential on outcomes compared to social class. Unfortunately for proponents of new structuralism, empirical literature within the NCDS demonstrates a persistent class effect on outcomes for young people (Micklewright, 1989).

Structuration appears to not appreciate the increasing levels of complexity that are placed upon such persons. In this then, it is best to call for a structured individualism thesis. One that recognises in a risk society, that whilst pathways are different and numerous, as empirical data (Roberts, 2003: 484) demonstrates they are still heavily influenced by the structures of society (Gayle et al, 2009: 26).Overall, the literature stresses the relevance of contextual factors that also highlights the important of individual agency (Steiner et al, 2021: 8).

A central theme that has unfolded across this literature review relates to the notion of choice and opportunity. Concepts of choice and opportunity are constrained through structural inequalities and socio-historical context. Within the life course approach the structure/agency dichotomy is replaced with an agency within structure (Diewald and Mayer, 2008: 7) theoretical modelling. The concept of agency is impacted based on different temporal foci (Hitlin and Elder, 2007: 170). Understanding the life course requires a multidimensional notion of agency (Hitlin and Kirkpatrick Johnson, 2015: 1431), as such Hitlin and Elder (2007: 171) breakdown agency into four distinctive categories: existential, identity, pragmatic, and life course. The first, existential agency, refers to all action related to a fundamental level of human freedom – linked to Giddens’ notion that one might have acted otherwise (Hitlin and Elder, 2007: 177), identity agency refers to actions that are based on personalised social behaviour, pragmatic agency refers to the expression of action based on heuristic like devices of commonality, and finally, life course agency refers to actions with long term implications based upon an internal calculation of self-control, which reflexively guides decision making (Hitlin and Elder, 2007: 182).

The life course is embedded in the individual within social structures in a way that presents mutually interdependent sub-structures that act as mechanisms that steer individuals. These social structures – in the form of structural inequalities, manifest themselves in the form of social class and gender within the NCDS cohort. The tightness of said social structures is dependent upon socio-temporal aspects that leave the room for individual decision-making or ‘agency’ (ibid). Over the course of an individual’s life course, there are times where due to social position, and other times due to socio-historical constraints, the individual experiences differing forms of social pressures and ability to impress themselves upon social structures in the forms of choice. Empirical literature reviewed thus far has illuminated such cases – with relation to class position and likelihood to enter into higher education (Micklewright, 1989), as one example.

The concept of life course and agency intersect. It highlights the socio-historical temporal constraints that are placed upon individual decision making for the future and also prompts a core methodological desire to investigate these constraints upon choice and opportunity to discover how that impacts later life decisions and outcomes. These decisions in other words are youth ‘transitions’ (Hitlin and Elder, 2007: 182). Treating the individual as an active agent in the shaping of their biographies is important as it deters a deterministic theoretical orientation whilst maintaining that some individuals will structurally have more agentic opportunities based upon power relations (Hitlin and Johnson, 2015; Schmitt, 2021). These structural influences have been identified as primarily related to social class and sex based structural inequalities.

## **Data and Methods**

The relationship between father’s socio-economic background and economic activity after mandatory schooling is examined using large-scale, nationally representative data (Bynner, 1998a) collected from the National Childhood Development Survey (NCDS) – longitudinal data, allowing the analysis of long-term processes and outcomes of individuals (Bynner and Joshi, 2007; Field, 2011). Educational attainment, housing tenure, sex, reading, and maths scores are also included in a model to understand the individual sorting into different forms of economic activity: employment, post-education schooling, school, training & apprenticeships, and unemployment & out of the labour force. NCDS data and is accessible using the UK Data Service.

Multinominal logistic regression will be used to understand the choice and opportunities of NCDS youth when it comes to economic activity post-mandatory schooling. This model will attempt to understand the role of structural inequalities in the form of social class and sex as well as other structural consumption cleavages such as housing tenure. The model also accounts for individual merit or ability in the form of reading and maths scores as well as prior attainment. After establishing the initial multinominal logistic regression model a sensitivity analysis will be conducted on various measures of social stratification surrounding class to assess the most appropriate measure to include within the model. Finally, an analysis of missing data involving multiple imputation will be conducted in order to assess the impact, if any, missing data has on the substantive findings of the multinominal logistic regression model.

### **Introduction to the NCDS data**

This work will use the National Child Development Study (University College London et al, 2023). The NCDS is a nationally representative birth cohort study – the second of its kind in the UK. It followed 17,415 participants using a cross-sectional sampling design to collect participants from birth within the same week in 1958. For this analysis, sweeps 0-4 (up to age 23) will be used. For the outcome variable, monthly data from 1974-1981 was collected. These were all collected post-hoc at age 23. The outcome variable focuses on the economic activity state of individuals in the month of September 1974, or month 201.

Table 1. 1 Sweeps Included for Analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | 1958 | 1965 | 1969 | 1974 | 1981 |
| Age | Birth | 7 | 11 | 16 | 23 |

The NCDS cohort originated in 1958 (when participants were born) and continues to present day. For analysis in this analysis only data up until age 23 (wave 4) is considered. Table 1.2 details the sample size of the NCDS. At birth in 1958 the total cohort consisted of a sample of 17,638 with 17,415 participants. By 1974 at age 16, the total cohort had increased to 18,558. This is due to the original sample being supplemented by migrants born in 1958. The number of participants at age 16 had fallen to 14,654, or 91.6 per cent of the eligible sample. This is a reduction in real participants from birth wave of 2,761. Of this reduction, 873 people died and a further 799 emigrated leaving 1,089 missing for reasons other than death or emigration. By 1981 at age 23, the total cohort was 18,558. The eligible sample after considering 960 dead and 1,196 emigrants is 16,402. In total there were 12,357 participants, or 75.3 per cent of the sample.

The sample size at age 23 is substantially smaller than that of the initial sample, this sample attrition is primarily determined to be caused by not being able to trace participants (there is also a relatively low refusal rate – 7.1 per cent at age 23) (Power and Elliott 2006: 35). Of refusal rates it is best to understand the dynamics of how the NCDS survey was conducted to appreciate the possibilities related to refusal. At age 23, this was the first time that participants themselves filled out the survey on their own without parental or guardian assistance (like at age 16) or having it done for them by their parents, teachers, and medical professionals. Data was collected by a paper and pencil-based survey. With respect to the outcome of economic activity – that was a retrospective employment history question that had to be filled out from the age of 16 every month until their current age at 23. This level of detail does make the possibility of refusal a possibility.

According to Hawkes and Plewis (Hawkes and Plewis, 2006) non-response: other (cases where there is no data for this sweep but there is for later sweeps, and ‘temporary emigrants’) makeup 10.7 per cent of non-response. The rest of the missingness – around 6.3 per cent is categories as ‘eligibility unknown’ (ibid). Eligibility unknown relates to those that either died or permanently emigrated. There is overall a substantive amount of missingness within the data used for analysis. An issue that comes with sample attrition for the NCDS is that the size of certain ethnic minority populations back when the sample was first collected were small, meaning that attrition makes analysis of ethnic minority populations extremely difficult (ibid). The nature of the level of missing data in the NCDS suggests that there is little support for the position that the data is missing completely at random (Hawkes and Plewis, 2006: 489; Silverwood *et al.*, 2021: 3). This supports the need to apply missing data techniques.

Table 1. 2 Participation in the NCDS from birth to 23 years

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Total cohort | Dead | Emigrants | Eligible sample | Participants | (% of eligible sample) |
| Birth – 1958 | 17638 | 0 | 0 | 17638 | 17415 | 98.7 |
| Age 7 – 1965 | 18016a | 821 | 475 | 16720 | 15425 | 92.3 |
| Age 11 – 1969 | 18287a | 840 | 701 | 16746 | 15337 | 91.6 |
| Age 16 – 1974 | 18558a | 873 | 799 | 16886 | 14654 | 86.8 |
| Age 23 – 1981 | 18558 | 960 | 1196 | 16402 | 12357 | 75.3 |

a the original sample was supplemented by migrants born in 1958.

### **Introduction to measures for subsequent analysis**

The following section provides an overview of key variables used for subsequent multivariate analysis. Initially variable selection was attempted by using the UCL CLOSER search platform that promotes a resource that allows for all variables within the NCDS cohort to be searched for (Parsons, 2013). On closer inspection, the CLOSER search platform did not present itself as an adequate answer to finding various variables for inclusion in analysis. Original codebooks and manuals produced via the NCDS were thus used to construct updated codebooks[[1]](#footnote-1) that gave detail not just of variable breakdown, but also of qualitative descriptions of each variable within the dataset. From this, variable selection became a much smoother process. Ultimately, economic activity, educational attainment, sex, social class, housing tenure, reading, and maths scores variables were selected.

#### **Economic Activity**

The main outcome variable of interest is the main economic activity of month 201. In other words, what were individuals doing after mandatory schooling in the month of September at age 16. The month of September was selected to allow time for children to gain their O-level results. This economic activity variable was a retrospective work history collected at age 23, participants were asked to note down each month from age 16-23 their current economic activity. This variable comes from sweep 4 (Age 23) of the NCDS. The Economic Activity of everyone was recorded retrospectively by the participant at age 23 at each month from when they turned 16 to when they turned 23. Information for the following variable comes from the data dictionary part 1 within the UK data service package of sweep 4 of the NCDS. Each month is recorded as a diary that covers one possible main activity defined as ‘Jobs’, ‘Full-time Education’, ‘Unemployment’, ‘Out of the labour force’, and ‘Fill-in-time’[[2]](#footnote-2). The monthly diary of economic activity filled out by participants was coded by a coder, resulting in unique values that fall outside of the range of these original categories.

The original economic activity variable for month 201 has 28 unique values. Five of these collapsed into an unemployment & out of labour force category, one into a full-time education post-school category, one into a school category, four into an employment category (using both Full-time and Part-time employment as well as FT+Other and PT+Other), one into missing data, and the rest into a training/apprenticeship category – this was done via a dominance approach, any combination of categories whereby training & apprenticeship were mentioned, they were given priority in coding over and above other categories – this means for example that those within the fulltime job + apprenticeship category were coded into the training & apprenticeship category over that of the employment category. For this last category a dominance approach was taken- any variation of training/apprenticeship alongside employment, education etc was taken to be training/apprenticeship. The training/apprenticeship category contains apprenticeships, like the Training Opportunities Scheme (TOPs) training courses. The NCDS codes main economic activity in a way that creates five categories: Employment, Post-schooling Education, School, Training & Apprenticeships, and Unemployment & Out of Labour Force . Main Economic Activity is determined based on if that activity is done 21 hours or more per week for Education (Full and Part-time), a full-time job of more than 30 hours, a part-time job of less than 30 hours, unemployed if respondent is actively searching for work, and out of the labour force if all else is not false.

Table 1. 3 Frequency Statistics for Economic Activity

|  |  |
| --- | --- |
|  | Frequency |
| Economic Activity in Month 201[[3]](#footnote-3) |  |
| MISSING | 86 |
| FT JOB | 4,716 |
| FT JOB+PT ED | 144 |
| FT JOB+APP | 1,842 |
| FTJ+APP+PT ED | 22 |
| FTJ+APP+DBR TC | 21 |
| FTJ+OTH TC | 1 |
| FTJ+DBR | 366 |
| FTJ+DBR TC+PTED | 4 |
| FTJ+OTH | 20 |
| FTJ+OTH TC+PTED | 1 |
| FTJ+FT NT TOPSTC | 35 |
| FTJ+FTTC+PTED | 1 |
| FTJ+LGSS | 2 |
| FTJ+LGSS+DBR TC | 1 |
| PT JOB | 37 |
| PTJ+PT ED | 2 |
| PTJ+DBR TC | 2 |
| TOPS | 1 |
| LGSS[[4]](#footnote-4) | 1 |
| FTEDPOSTSCHL | 1,046 |
| AT SCHOOL | 3,717 |
| UNEMP | 276 |
| UNEMP+PTED | 3 |
| UNEMP RULE6 | 11 |
| OLF | 164 |
| OLF+PT ED | 3 |
| PT ED | 11 |
| Total | 12,536 |

Re-coding this variable was a necessity to get at the nuance of some of the economic activity data. For example, a lot of data was coded as full-time employment – including training schemes, apprenticeships, Technical and Vocational Educational Initiative (TVEI), and TOPs schemes[[5]](#footnote-5).

Re-coding this variable translates into five categories: Employment, Post-Schooling Education, School, Training & Apprenticeship, and Unemployment & OLF. Employment was collapsed from part-time and full-time into a singular employment category due to the negligible sample size of part-time work. Post-school education refers to credit received for completion of courses not in a school environment but given by an accredited college, trade school, workshops etc. School is defined as anyone that after completion of mandatory schooling at age 16 decided to continue education at school for A-levels etc. Training & Apprenticeship is defined by any individual undertaking a training, work, or apprenticeship related scheme. Finally, Unemployment and OLF is a combined category of all unemployed and those out of the labour force. Unfortunately for sample size reasons these two categories had to be combined for statistical power though it is recognised that there is a qualitative distinction between these two categories that may impact the statistical power presented within the models. A full breakdown can be found in 1.8.

#### **Educational Attainment**

The NCDS cohort members reached the compulsory school leaving age in 1974. At this time the main educational qualifications were either the Certificate of Secondary Education (CSE) (Pearson qualifications, 2023a), introduced in 1965. The second option was the Ordinary level or O-level, introduced in 1951 (Pearson qualifications, 2023b). The O-level was understood to be more complex that CSEs, and thus fewer people achieved O-level grades. This is the best and most advanced ability measure for the age of 16 and thus makes a good measure of educational attainment for those at 16 after mandatory schooling ends.

Mandatory schooling ended for these individuals at 16 years of age. The examination taken at this age would have been the O-level examination. Researchers have advocated for the use of established education measures in order to better facilitate replication and comparison (Connelly, Gayle and Lambert, 2016). I have constructed an educational attainment measure in a binary less than five O-levels/five or more O-levels variable. Within contemporary literature on educational attainment, gaining five or more GCSEs at grades A\*-C is a standard benchmark measure used within official reporting (ibid). Another suggested alternative – especially when dealing with the UK education system over time whereby educational qualifications changed is to use a National Vocational Qualification (NVQ) measure. This would be attractive for analysis however because educational attainment as a measure is being used just after mandatory schooling the level of qualification within the NCDS model has a ceiling threshold at O-levels, or NVQ 2.

There is an argument that GCSEs and O-levels are analytically distinct concepts (Murray, 2011), and as such a like-for-like measure may not be the most attractive. Firstly, as a measure of attainment GCSEs and O-levels provide considerable barriers to entry for young people pursuing future foals (ibid). Due to this rationale, using a threshold measure for number of O-levels given the restriction of age on the amount of attainment an individual could have undertaken at this time, it appears to be the best operationalisation of the measure. For this reason, it is rationalised to prefer the five or more-measure used within GCSE based literature for O-level attainment.

This variable was constructed from two separate variables – the first was a simple binary variable of whether an individual had any O-levels, the second, on condition of the first then asks how many O-levels that person had passed. Combining these two variables together produces a single count variable that includes the number of zeros. This attainment variable was then further recoded into a binary variable of less than five O-levels and greater than five O-levels. This was done for two reasons. The first has been discussed above, as it relates to standard practice amongst GCSE research. The second, more important reason for recoding is one of practicality. Keeping O-levels as a count variable means that the n is too low for some sub-categories when moving on to modelling – as seen in table 1.4.

Table 1. 4 Educational Attainment Count Variable by Economic Activity

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Educational Attainment – Number of O-levels | | | | | | | | | | |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 or More | Total |
| Economic Activity |  |  |  |  |  |  |  |  |  |  |  |
| Employment | 3,024 | 586 | 368 | 251 | 234 | 181 | 113 | 70 | 35 | 25 | 4,887 |
| Post-Schooling Education | 161 | 116 | 110 | 100 | 98 | 145 | 132 | 79 | 46 | 66 | 1,053 |
| School | 159 | 93 | 100 | 135 | 203 | 368 | 399 | 562 | 612 | 1,081 | 3,712 |
| Training/Apprenticeships | 1,136 | 375 | 227 | 168 | 129 | 123 | 65 | 52 | 21 | 20 | 2,316 |
| Unemployment and OLF | 350 | 45 | 16 | 14 | 9 | 9 | 6 | 5 | 2 | 1 | 457 |
| Total | 4,830 | 1,215 | 821 | 668 | 673 | 826 | 715 | 768 | 716 | 1,193 | 12,425 |

#### **Sex**

Sex is a variable derived from sweep 0. It’s inclusion for analysis is because during the timeframe of the NCDS, gender dynamics played an important role in economic activity. The evolution of part-time work, the differences in populations out of the labour force, and those choosing to go on to higher education are all influenced by sex according to previous literature (Jones, 1986; Gayle et al, 2009; Duckworth and Schoon, 2012; Dorsett and Lucchino, 2013). For these reasons, sex provides a theoretically compelling case for inclusion within a model of economic activity post-mandatory schooling.

#### **Race**

Echoing the arguments for the inclusion of sex in models of analysis, the role of race also impacted individuals’ economic activity during the NCDS timeframe[[6]](#footnote-6) (Payne, 1995; Lindley, 1996; Gayle et al, 2009). Race as a variable for inclusion in this model however presents too many statistical issues to be an effective measure.

#### **Housing Tenure**

Housing tenure has been used in previous analyses regarding educational attainment and labour market outcomes (Di Salvo and Ermisch, 1997; Duta et al, 2021). For subsequent analysis, tenure is a measure of whether an individual lives in their own home or not[[7]](#footnote-7). Housing tenure enables the inclusion of a ‘consumption cleavage’ (Saunders, 2003, 2021) based variable. This form of cleavage encapsulates the ‘new structuralist’ notions that structural inequalities do in fact matter, but not necessarily older structures such as class and gender. The inclusion of housing tenure in this model allows for a more direct investigation of this sentiment. Enabling the evaluation of influence of different forms of structural inequalities.

#### **Social Stratification and Socio-Economic Background: NS-SEC, CAMSIS, RGSC**

Social stratification is the persistence of inequalities which occur or are reified across generations (Stopforth, 2020: 17). Inequalities can manifest themselves in terms of gender, ethnicity, social class etc. These social inequalities impact individuals in terms of their aspirations and desires, educational outcomes, labour market position, and destinations.

Socio-economic background is a cornerstone of social scientific research. There is no one universally agreed measure employed. There are two main schools of thought when attempting to capture socio-economic background. The first is a measure of social class which contemporarily employs occupation-based schema. The second are social stratification scales which instead rely on capturing a continuous measure.

Social class as a variable is one that has constant and consistent debate throughout sociological literature – even today whilst current schemas reign dominant, there is no universal measure of social class. Three social stratification measures of NS-SEC, RGSC, and CAMSIS will be used due to their theoretical distinctiveness and the empirical practicality of operationalisation.

A key aspect of this chapter is to find the most appropriate measure of social class to fit the given models of analysis. This is to find the most empirically useful schema to distinguish most effectively the analytical purposes in mind for this research (Bergman and Joye, 2001: 14). Multiple measures of social class are reflected upon. This following section seeks to establish the major measures of social class and weigh their common strengths and weaknesses, which may affect model parsimony.

Prestige scales, social class schemes, occupational grading all rely on rather static temporal procedures. Longer-term structural transformations of society will alter the underlying distribution of stratification over time (Lambert and Barnett, 2021: 191). Whilst the Treiman constant[[8]](#footnote-8) is oft hailed as the single most important empirical generalisation to be confirmed through social stratification research (Lambert *et al.*, 2008), and thus justifies the use of universal and semi-universal social stratification coding of occupational data, structural transformations over time (ibid) can potentially alter the underlying distribution within these universal and semi-universal coding schemas. As an example, the decline of manufacturing industry in the UK may have a qualitative impact upon transitions into the type of skilled manual occupations as well as a quantitative impact in the number of occupations available.

The NCDS provides occupational coding measures for father’s socio-economic position using a variety of measures (Gregg, 2012). The measures provided are the Registrar General Class Schema, National Statistics Socio-Economic Classification, and the Cambridge Social interaction and Stratification scale. Occupational coding measures were taken in 1969. Unfortunately, no such occupational measures were taken for mothers making it impossible to employ a dominance approach (Connelly, Gayle and Lambert, 2016). The first occupational measure gives a full six class categorisation[[9]](#footnote-9). The last measure is continuous and as such no recoding was required. The occupational coding conducting by (Gregg, 2012) was subsequently merged with the main data of sweeps 0-4 (up until age 23). The occupational coding data was then cleaned and re-coded into relevant NS-SEC and RGSC schemas, with CAMSIS being left as it is due to it being a single continuous measure.

The next three variables for subsequent analysis are all considered socio-economic variables. Two are social class schemas and one is a stratification scale. Multiple socio-economic measures are considered to see whether there are different patterns for different dimensions of social stratification both within cohorts and across them by comparing between cohort substantive findings.

There are many other social stratification measures. The rational for including the RGSC, NS-SEC, and CAMSIS are based upon theoretical diversity and empirical practicality. In terms of theoretical diversity, each of these measures is theoretically constructed using distinct analytical frameworks – two being constructed using a categorical delineation of social class, and another constructed as a continuous measure. The most straightforward explanation as to why these three measures were included specifically is that these three measures are easily able to be constructed using the NCDS. An occupational coding file is provided for Sweep Two of the NCDS which enables the construction of the full RGSC and NS-SEC social class schemas as well as CAMSIS. All three socio-economic measures are constructed using a different theoretical orientation. The RGSC’s strict hierarchical nature is slightly different from the NS-SEC that argues against a strict hierarchy, though does maintain a ‘semi-ordinal’ structure – whilst not a major difference, this is a variation comparatively. Both RGSC and NS-SEC are categorical gradational scales in comparison to the CAMSIS which is a single continuous measure. Other socio-economic measures were considered – for example the neo-Marxian social class schema developed by (Wright, 2005). This measure would provide an alternative social class schema that challenges the occupational dominance of the RGSC and NS-SEC. However, due to data limitations this schema cannot be practically constructed.

The following section provides a detailed breakdown of each chosen social stratification measure, explaining their theoretical makeup and their analytical construction.

##### **Registrar General Class Schema**

The Registrar General’s Social Class (RGSC) is one of the oldest social class measures in the UK – first used in 1911 to show variation in infant mortality according to parents’ occupation(‘Annual report for the 1911 of the registrar-general’, 1913). The measure is built upon the assumption that society is graded based upon a hierarchy of occupations(Murray 2011: 67). The schema is broken into six distinction categories and rages from unskilled manual occupation to higher level professionals (ibid). The RGSC once formed the basis of all commonly used social classifications within Britain (Szreter 1984: 523). With, alternative measures, such as the National Statistics Socio-Economic Classification have risen to prominence. The RGSC was a popular social class measure of its time (particularly around the 1980s). The RGSC first being developed in 1911 (‘Annual report for the 1911 of the registrar-general’, 1913) means that as a measure of social-stratification, it had existed for 47 years prior to the commencement of the NCDS. Compared to other social stratification measures such as NS-SEC, that were created after the NCDS, this presents an opportunity to test the substantive interpretations of different social-stratification measures constructed at different times. This temporal distinction will become an important part of the subsequent sensitivity analyses.

The RGSC rests upon a theoretical assumption that social inequality exists within society (Szreter, 1984). This social inequality is structured around a single scale of social position/status within society encapsulated within occupational categories. These occupational categories form a single uni-dimensional hierarchy across all of Britain. The original creator of the schema, Stevenson, created the model of RGSC based upon an assumption that society is comprised of an upper middle, middle, and working class (Prandy, 1999: 468). This assumption is baked into the theoretical implications of a unidimensional hierarchy mentioned above.

The Full RGSC class schema is detailed below:

Table 1. 5 RGSC Class Schema

|  |  |  |
| --- | --- | --- |
| Class | Occupations | Example Occupations |
| I | Professional Occupations | Accountant |
| II | Intermediate Occupations | Police Officer |
| IIIN | Skilled Non-Manual Occupations | Clerical Worker |
| IIIM | Skilled Manual Occupations | Butcher |
| IV | Partly Skilled Occupations | Postal Worker |
| V | Unskilled Occupations | Labourer |

The NCDS provides occupational codes taken in 1969 – these codes are SOC2000 (Gregg, 2012) and social-class categories constructed from subsequent SOC2000 codes. Amongst the social stratification variables that are provided, a full-auto, semi-auto, and verification processing variable are provided. Semi-auto processing social stratification variables are used within subsequent analysis (Gregg, 2012) as suggested.

##### **National Statistics Socio-Economic Classification**

Rose and Pevalin developed the National Statistics Socio-Economic Classification (NS-SEC) (Rose and Pevalin, 2002). The operational categories of the NS-SEC represent labour market positions, employment statuses, and employment relations.

NS-SEC was developed from the EGP perspective before it (Rose and Pevalin, 2002). The NS-SEC was developed from the EGP perspective (ibid). Central to the NS-SECs ideas on social class – and the development of social class schemas is employment relations. These employment relations are split into three distinctive formations: those that purchase labour and have authority over those they have purchased labour from, self-employed workers, and employees who sell their labour and are thus under the authority of employers (Bergman and Joye 2001: 12). It is within this differentiation of employment relations that gives rise to class-based patterns of social stratification (Williams, 2017). Like other social class schemas already mentioned, a central tendency for Goldthorpe’s study of social class rests upon an analysis of relationships – one occupational group is relational to another within the broader social class schema (Goldthorpe and Marshall, 1992).

The full NS-SEC classification schema has 14 operational categories as they relate to employment relations but can be broken down into as few as three analytical categories. This ability to break down the social class schema is attractive – particularly when using data that has limited sample sizes or complications related to multiple imputation convergence.

Table 1. 6 NS-SEC Class Schema

|  |  |
| --- | --- |
|  | Analytical Variables for NS-SEC |
| Operational Categories | **Eight Classes** |
| L1  Employers in large establishments | 1  Higher Managerial |
| L2  Higher managerial occupations |
| L3  Higher professional occupations |
| L4  Lower professional and higher technical occupations | 2  Lower managerial and professional occupations |
| L5  Lower managerial occupations |
| L6  Higher supervisory occupations |
| L7  Intermediate occupations | 3 Intermediate occupations |
| L8  Employers in small establishments | 4  Small employers and own account workers |
| L9  Own account workers |
| L10  Lower supervisory occupations | 5  Lower supervisory and technical |
| L11  Lower technical occupations |
| L12  Semi-routine occupations | 6  Semi-routine occupations |
| L13  Routine occupations | 7  Routine occupations |
| L14  Never worked and long-term unemployed | 8  Never worked and long-term unemployed |

As with the RGSC, the NCDS allows for operationalisation of the full NS-SEC class schema. This will provide the basis for comparison and sensitivity analysis of socio-economic measures within this chapter. The following analytical variables within the NS-SEC have been broken down with example occupations to aid in interpretation within subsequent models in table 1.7.

Table 1. 7 Examples of Occupations from Analytical NS-SEC

|  |  |
| --- | --- |
| Analytical Variables for NS-SEC | Example Occupations |
| 1  Higher Managerial | CEOs, senior officers, finance directors |
| 2  Lower managerial and professional occupations | Doctors, dentists, lecturers |
| 3 Intermediate occupations | Travel agents, nursing assistants, teaching assistants |
| 4  Small employers and own account workers | Taxicab drivers, farmers, brick layers |
| 5  Lower supervisory and technical | Mechanics, plumbers, bakers |
| 6  Semi-routine occupations | Sales assistants, dental nurses, housekeepers |
| 7  Routine occupations | Labourers, hairdressers, barbers |

##### **CAMSIS**

CAMSIS – similar to the Cambridge scale argues that individuals are embedded within socially moderated spaces and networks within which they engage in various social and economic interactions which are different from interactions with persons who are more distant from these networks (Stewart, Prandy and Blackburn, 1973, 1980). In other words, CAMSIS represents a social stratification scale based on measures of social distance (Prandy and Lambert, 2003). These relationship networks are ultimately hierarchical and reify themselves in reproducing hierarchical inequalities (Bergman and Joye, 2001). The Cambridge Social Interaction and Stratification scale (CAMSIS) is a subsequent evolution and development of the Cambridge scale.

By its nature, CAMSIS does not delineate between concepts of class and concepts of status (Bergman and Joye 2001: 40). Its continuous nature means that numerical values are attached to occupations, meaning the relative value of each occupational value is only meaningful in comparison to other occupations on the same scale (Connelly et al 2016: 7). This is meaningful when it comes to interpretation of the CAMSIS measure within models of analysis as the value of the coefficient is always going to be in relation to the comparison to other occupations along the CAMSIS scale. The largest difference between CAMSIS and other social stratification measures discussed is that CAMSIS does not believe in the notion that distinct groups in the form of social classes are differentiated through material and status differences. CAMSIS rejects a ‘simple structuralism’ dependent on a static structure – it proposes instead a much more dynamic constantly re-constitutive process (Bergman and Joye, 2001).

Whilst CAMSIS stands in contrast to both other social stratification measures mentioned, they do share some similarities. CAMSIS contends - as do the NS-SEC and RGSC – that occupational groups are the major mechanism by which social and economic rewards are distributed within modern societies (ibid) and as such are some of the best indicators of social stratification in society.

As with RGSC and NS-SEC, CAMSIS codes are provided within the occupational codes provided by (Gregg, 2012).

#### **Reading and Maths Scores**

Previous research has reflected upon the influence of prior educational attainment upon later life chances. The NCDS provides various prior attainment measures. Reading and maths scores are two of these measures. Reading and maths scores have a history of being used within the sociology of education (Bynner, 1998b; Patacchini and Zenou, 2006). Adding these to the model allows an element of prior individual attainment to be controlled so that structural inequalities and effects can be better understood in the totality of an individual’s life domains (Mayer, 2009).

## **Descriptive Statistics**

Table 1.8 shows the frequencies and summary statistics for the NCDS. Overall, 38.12 per cent of our sample is in employment. Whilst 30.44 per cent remain in school 8.70 per cent moved on to full-time post school education. Unemployment and being out of the labour force makes up 3.02. Finally, 19.72 per cent of the sample are in some kind of training or apprenticeship scheme.

When it comes to educational attainment, 64.43 per cent of individuals received less than five O-levels, with the remaining 35.57 per cent receiving five or more O-levels. Sex presents a relatively equal split between men (49.93 per cent) and women (50.07 per cent). When it comes to home ownership, 47.93 per cent of individuals grew up in a home owned by their parents compared to 52.07 per cent that didn’t. The NS-SEC categories all see a relatively even distribution between 10-18 per cent except for the largest category – routine occupations, at 23.63 per cent – and the smallest category – Intermediate occupations, at 2.07 per cent. RGSC is much more unevenly distributed comparative to NS-SEC, with skilled manual making up 41.48 per cent of individuals, with professionals only making up 4.27 per cent of individuals. This uneven distribution on top of their analytical differences presents some evidence to suggest that substantive findings of a sensitivity analysis could potentially find diverging findings. Finally, CAMSIS has a mean of 4.44, reading scores of 16.50 and maths scores of 17.60.

Table 1. 8 Descriptive Statistics for Economic Activity

|  |  |  |
| --- | --- | --- |
|  | n | % |
| Economic Activity |  |  |
| *Employment* | 3,017 | 38.12% |
| *Post-Schooling Education* | 689 | 8.70% |
| *School* | 2,409 | 30.44% |
| *Training/Apprenticeships* | 1,561 | 19.72% |
| *Unemployment and OLF* | 239 | 3.02% |
| Educational Attainment |  |  |
| *<5 O-Levels* | 5,100 | 64.43% |
| *5 or more O-Levels* | 2,815 | 35.57% |
| Sex |  |  |
| *Female* | 3,963 | 50.07% |
| *Male* | 3,952 | 49.93% |
| Housing Tenure |  |  |
| *Own Home* | 3,794 | 47.93% |
| *Don't Own Home* | 4,121 | 52.07% |
| NS-SEC |  |  |
| *Higher managerial, administrative and professional occupations* | 1,268 | 16.02% |
| *Lower managerial, administrative and professional occupations* | 1,121 | 14.16% |
| *Intermediate occupations* | 164 | 2.07% |
| *Small employers and own account workers* | 807 | 10.20% |
| *Lower supervisory and technical occupations* | 1,286 | 16.25% |
| *Semi-routine occupations* | 1,399 | 17.68% |
| *Routine occupations* | 1,870 | 23.63% |
| RGSC |  |  |
| *Professional* | 338 | 4.27% |
| *Managerial and Technical* | 1,637 | 20.68% |
| *Skilled non-manual* | 869 | 10.98% |
| *Skilled manual* | 3,283 | 41.48% |
| *Partly skilled* | 1,123 | 14.19% |
| *Unskilled* | 665 | 8.40% |
|  |  |  |
|  | Mean | SD |
| CAMSIS | 4.44 | 1.37 |
| Reading Score | 16.50 | 6.17 |
| Maths Score | 17.60 | 10.26 |
|  |  |  |
| n | 7915 | |
| Data Source: NCDS | | |

## **Modelling Main Economic Activity**

The main outcome variable is the main economic activity of individuals in September of 1974. This is the first month individuals were in when they received their O-level results after mandatory schooling. The first set of analyses estimate a multinominal logistic regression model. Table 1.9 details the deviance, change in deviance, change in degrees of freedom, the McFadden’s Adjusted Pseudo R2, AIC, and BIC measure to compare the null model with models of one explanatory variable. Table 1.10 details the same but through a sequential building of the null model with each subsequent independent variable added.

This model has been tested for the goodness of fit of two competing statistical models based on the ratio of their likelihoods in a likelihood-ratio test, and again with a Wald test. Both found that the hypothesis that all the coefficients associated with educational attainment, sex, tenure, NS-SEC, reading, and maths scores are simultaneously equal to 0 can be rejected at the 0.01 level.

The model output is using the reference category of full-time employment for two reasons. The first is that said category has the largest sample of all economic activity sub-categories. The second is that a contrast with full-time employment is sociologically compelling. Contrasting full-time employment with other economic activity destinations like education or apprenticeships is temporally relevant given the possible impact that increasing the mandatory school leaving age, decline in manufacturing industry, rise in part-time work may have on the economic destinations of youth.

Table 1. 9 Goodness-of-fit summaries for explanatory variables and Economic Activity

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome Variable: Economic Activity | Deviance | Deviance (from Null) | d.f. (from Null) | McFadden’s Adjusted Pseudo R2 | AIC | BIC |
| Null Model | 21656.47 | - | - | - | 21664.47 | 21692.38 |
| Null Model + Educational Attainment | 17732.61 | 3923.86 | 4 | 0.18 | 17748.61 | 17804.43 |
| Null Model + Sex | 20709.43 | 947.05 | 4 | 0.04 | 20725.43 | 20781.24 |
| Null Model + Tenure | 20920.88 | 735.59 | 4 | 0.03 | 20936.88 | 20992.70 |
| Null Model + NS-SEC | 20492.89 | 1163.58 | 24 | 0.05 | 20548.89 | 20744.23 |
| Null Model + Reading | 19584.87 | 2071.60 | 4 | 0.10 | 19600.87 | 19656.69 |
| Null Model + Maths | 19373.19 | 2283.28 | 4 | 0.11 | 19389.19 | 19445.00 |

Explanatory variables are entered sequentially in the subsequent multiple logistic model following (Gayle and Lambert, 2009) example.

Table 1. 10 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome Variable: Economic Activity | Deviance | Deviance (from Previous) | d.f. (from Previous) | McFadden’s Adjusted Pseudo R2 | AIC | BIC |
| Null Model | 21656.47 | - | - | - | 21664.47 | 21692.38 |
| Null Model + Educational Attainment | 17732.61 | 3923.86 | 4 | 0.18 | 17748.61 | 17804.43 |
| Null Model + Educational Attainment + Sex | 16787.60 | 945.01 | 4 | 0.22 | 16811.60 | 16895.32 |
| Null Model + Educational Attainment + Sex + Tenure | 16572.38 | 215.22 | 4 | 0.23 | 16604.38 | 16716.00 |
| Null Model + Educational Attainment + Sex + Tenure + NS-SEC | 16318.67 | 253.71 | 24 | 0.24 | 16398.67 | 16677.73 |
| Null Model + Educational Attainment + Sex + Tenure + NS-SEC + Reading | 16097.12 | 221.55 | 4 | 0.25 | 16185.12 | 16492.09 |
| Null Model + Educational Attainment + Sex + Tenure + NS-SEC + Reading + Maths | 16026.66 | 70.46 | 4 | 0.26 | 16122.66 | 16457.53 |

The model fit statistics demonstrate that there are normally distributed residuals and that the model is correctly specified. Table 1.10 suggests that for the full proposed model, deviance is reduced by 5,629.81 from the null. AIC and BIC statistics also suggest that the full model is the best fit model amongst those entered. Finally, the full model presents with an adjusted pseudo R2 of 0.26. In other words, the full model explains 26 per cent of the variance of economic activity, leaving 74 per cent unexplained. The following analysis with the full model is a complete case analysis with 7915 observations.

Prior to discussing the results of this model, a discussion on interpretation must be had. When dealing with multinominal logistic regression results in the form of coefficients are reported in the default Stata output as log odds. Log odds are notoriously difficult to interpret and are rarely well described in sociological studies (Gayle and Lambert, 2009). For example, for a categorical explanatory variable the coefficient associated with category effects is thought of as the effect on the log odds of moving from the reference category to the particular category of the X variable. Due to this difficulty in interpretation, some (Norton and Dowd, 2018) have advocated for the use of odds ratios over log odds. However, odds ratios have their own issues which result in an inability to compare across models and across datasets, even if they have the same model specification (ibid). Sometimes, odds ratios are unable to compare and interpret within a model itself (ibid). This issue stem from odds ratios changing if variables are added to the model, even if such additional variables are independent from the other variables. Due to these issues, both log odds and odds ratios provide an underwhelming desire to use them in the form of interpretating multinominal logistic models beyond establishing basic substantive effects of ‘higher’ and ‘lower’ (Gayle and Lambert, 2009). The popular alternative to the use of logs odds and odds ratios is the average marginal effect of an explanatory variable on the probability that equals 1 versus 0. The rationale for interpreting multinominal logistic models using average marginal effects is based upon the fact that the marginal effect is less sensitive to changes in model specification than the odds ratio, that the average marginal effect can be either positive or negative, and finally, average marginal effects for subgroups (like social class) can differ from each other leading to different implications and interpretations (Norton and Dowd, 2018).

For subsequent analysis, log odds will be presented for use of basic substantive effects, with average marginal effects being used alongside as a complement to interpretation.

The results of the multinominal logistic regression model are reported in table 1.11. It is not possible to ascertain the significance parameters of variables other than in relation to the reference category (Firth, 2003). This is known as the reference category problem. Quasi-variances were considered to overcome this reference category problem, these are detailed for NS-SEC (the only variable that can provide quasi-variance statistics) in table 1.11. Quasi-variances are reported alongside the standard outputs of log odds and average marginal effects to overcome the reference category problem. Quasi-variances are normally produced within Stata using the custom Stata command ‘qv’. Unfortunately, the qv command and subsequent graphing subcommands do not currently work with multinominal logistic regression models. Due to this, the creation of quasi-variance statistics was done via a quasi-variance calculator[[10]](#footnote-10) (Firth, 2000). Whilst this does produce the required quasi-variance statistics, there are two notable issues with this direction. The first is that by producing quasi-variance statistics outside of Stata, this necessarily breaks the workflow and increases the possibility of manual error. The second, is that the given quasi-variance calculator does not provide lower and upper bound 95% Cis for quasi variance, instead producing a singular quasi-variance statistic. Whilst not ideal, the production of quasi-variance statistics does provide a utility in interpretation of NS-SEC within the given model.

The output for post-schooling education demonstrates that individuals that received five or more O-levels had an increased higher log odds of being in post-schooling education over employment. Using average marginal effects there is a 3 per cent increased probability for an individual to be in post-schooling education over employment if they received five or more O-levels. Individuals that are men had a decreased log odds of being in post-schooling education over employment than woman, or in terms of average marginal effects, a 7 per cent decreased probability of being in post-schooling education over employment if the individual is a man. Individuals that didn’t own their own home compared to those that did have a decreased log odds of being in post-schooling education over employment. In other words, there is a 3 per cent decreased probability of being in post-schooling education over employment if the individual resided in a home that their parents did not own. Results suggest that individuals whose parents have lower NS-SEC social class occupations have a significantly lower log odds of being in post-schooling education over employment. For a full breakdown of social class across each economic activity category see figure 1.1 for the predicted probabilities. The 95 per cent quasi-variance comparison intervals suggest that there are significant differences in post-schooling education activity of children whose fathers occupy different NS-SEC categories. Whilst reading and maths scores are statistically significant and demonstrate increased scores indicate a slight increased log odds in being in post-schooling education over employment. Substantively speaking, neither reading nor maths provide an increased probability of being in post-schooling education over employment when looking at average marginal effects.

The output for school demonstrates that individuals that received five or more O-levels had an increased log odds of being in school over employment, using average marginal effects individuals with five or more O-levels saw a 44 per cent increased probability of being in school over employment. Individuals that are men over women had an increased log odds of being in school over employment, this translates into a 0 per cent increased probability of being in school over employment. Individuals that didn’t own their own home compared to those that did have a decreased log odds of being in school over employment, this translates to a 3 per cent decreased probability. Results suggest a downward gradient for individuals whose parents have lower NS-SEC occupations. Those individuals whose parents have lower NS-SEC occupations have a significantly lower log odds of being in school over employment – the greatest decrease in log odds comes from those fathers that are classified as being small employers. Quasi-variance statistics concur with this. Reading and maths scores are against statistically significant and show an increased log odds of being in school over employment – though only reading scores translate to a substantive 1 per cent increase probability of being in school over employment.

The output for training & apprenticeships demonstrates that individuals that received five or more O-levels had a decreased log odds of being in training & apprenticeships over employment, in terms of average marginal effects, this corresponds to a decreased probability of 16 per cent. Individuals that are men over women had an increased log odds of being in training & apprenticeships over employment, or a 25 per cent increased probability. Results suggest that individuals that didn’t own their own home compared to those that did have a decreased log odds of being in training & apprenticeships over employment. As this corresponds to average marginal effects, there is a 1 per cent decrease in probability of being in training & apprenticeships over employment for an individual that lives in a home that isn’t owned by their parents over people that do. Results also suggest that for individuals whose fathers were in lower managerial and routine occupations, there was a decreased log odds of being in training & apprenticeships over employment. Reading scores are not statistically significant. Maths scores are statistically significant though hold no substantive significance – there is a 0 per cent increase in probability of being in training & apprenticeships over employment with increased maths scores.

The output for unemployment & OLF demonstrates that individuals whose fathers were lower supervisory & technical occupations has a decreased log odds of being in unemployment & OLF over employment. This translates to a 1 per cent decreased probability of being in unemployment & OLF over employment if an individual’s father is part of lower supervisory & technical occupations over higher managerial occupations. Increased reading scores translate to a decreased log odds of being in unemployment & OLF over employment, though there is no substantive effect as average marginal effects demonstrate a 0 per cent increased probability of being in unemployment & OLF over employment.

Table 1. 11 Mlogit of Economic Activity

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | NS-SEC | | | Average Marginal Effects | | Quasi-variance | |
| Economic Activity | **Coef.** | **S.E** | **Sig.** | **Prob.** | **S.E** | **S.E** | **QV** |
| Employment | Ref. | (.) |  | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |
| Post-Schooling Education |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | 1.60 | (0.11) | \*\*\* | 0.03 | (0.01) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Male* | -0.43 | (0.10) | \*\*\* | -0.07 | (0.01) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Don't Own Home* | -0.71 | (0.10) | \*\*\* | -0.03 | (0.01) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) | 0.12 | 0.01 |
| *Lower Managerial* | -0.47 | (0.16) | \*\* | -0.01 | (0.01) | 0.11 | 0.01 |
| *Intermediate Occupations* | -0.66 | (0.30) | \* | -0.01 | (0.02) | 0.27 | 0.07 |
| *Small Employers* | -1.05 | (0.19) | \*\*\* | -0.04 | (0.01) | 0.15 | 0.02 |
| *Lower Supervisory & Technical Occupations* | -0.71 | (0.16) | \*\*\* | -0.02 | (0.01) | 0.11 | 0.01 |
| *Semi-Routine Occupations* | -1.05 | (0.17) | \*\*\* | -0.04 | (0.01) | 0.12 | 0.01 |
| *Routine Occupations* | -1.04 | (0.16) | \*\*\* | -0.03 | (0.01) | 0.10 | 0.01 |
| Reading Scores | 0.04 | (0.01) | \*\* | 0.00 | (0.00) | (.) | (.) |
| Maths Scores | 0.01 | (0.01) | \* | -0.00 | (0.00) | (.) | (.) |
| Intercept | -1.35 | (0.20) | \*\*\* | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |
| School |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | 2.94 | (0.09) | \*\*\* | 0.44 | (0.01) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Male* | 0.37 | (0.08) | \*\*\* | -0.00 | (0.01) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Don't Own Home* | -0.56 | (0.08) | \*\*\* | -0.03 | (0.01) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) | 0.10 | 0.01 |
| *Lower Managerial* | -0.67 | (0.14) | \*\*\* | -0.05 | (0.01) | 0.10 | 0.01 |
| *Intermediate Occupations* | -1.16 | (0.27) | \*\*\* | -0.11 | (0.03) | 0.25 | 0.06 |
| *Small Employers* | -1.21 | (0.16) | \*\*\* | -0.11 | (0.02) | 0.12 | 0.01 |
| *Lower Supervisory & Technical Occupations* | -1.06 | (0.14) | \*\*\* | -0.10 | (0.01) | 0.09 | 0.01 |
| *Semi-Routine Occupations* | -1.09 | (0.14) | \*\*\* | -0.09 | (0.02) | 0.09 | 0.01 |
| *Routine Occupations* | -1.48 | (0.14) | \*\*\* | -0.13 | (0.01) | 0.09 | 0.01 |
| Reading Scores | 0.06 | (0.01) | \*\*\* | 0.01 | (0.00) | (.) | (.) |
| Maths Scores | 0.04 | (0.01) | \*\*\* | 0.00 | (0.00) | (.) | (.) |
| Intercept | -2.19 | (0.18) | \*\*\* | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | 0.09 | (0.11) |  | -0.16 | (0.01) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Male* | 1.94 | (0.08) | \*\*\* | 0.25 | (0.01) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Don't Own Home* | -0.25 | (0.07) | \*\*\* | -0.01 | (0.01) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) | 0.12 | 0.01 |
| *Lower Managerial* | -0.50 | (0.15) | \*\* | -0.03 | (0.02) | 0.10 | 0.01 |
| *Intermediate Occupations* | -0.40 | (0.27) |  | 0.00 | (0.03) | 0.24 | 0.06 |
| *Small Employers* | -0.19 | (0.15) |  | 0.04 | (0.02) | 0.10 | 0.01 |
| *Lower Supervisory & Technical Occupations* | -0.17 | (0.14) |  | 0.03 | (0.02) | 0.08 | 0.01 |
| *Semi-Routine Occupations* | -0.26 | (0.14) |  | 0.02 | (0.02) | 0.07 | 0.01 |
| *Routine Occupations* | -0.45 | (0.14) | \*\* | 0.01 | (0.02) | 0.07 | 0.00 |
| Reading Scores | 0.01 | (0.01) |  | -0.00 | (0.00) | (.) | (.) |
| Maths Scores | 0.02 | (0.01) | \*\*\* | 0.00 | (0.00) | (.) | (.) |
| Intercept | -1.90 | (0.17) | \*\*\* | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | 0.12 | (0.28) |  | -0.02 | (0.01) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Male* | -0.01 | (0.14) |  | -0.01 | (0.00) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |
| *Don't Own Home* | 0.13 | (0.16) |  | 0.01 | (0.00) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) | 0.25 | 0.06 |
| *Lower Managerial* | -0.46 | (0.34) |  | -0.01 | (0.01) | 0.23 | 0.06 |
| *Intermediate Occupations* | -0.39 | (0.58) |  | -0.00 | (0.02) | 0.52 | 0.27 |
| *Small Employers* | -0.52 | (0.33) |  | -0.01 | (0.01) | 0.22 | 0.05 |
| *Lower Supervisory & Technical Occupations* | -0.72 | (0.32) | \* | -0.01 | (0.01) | 0.20 | 0.04 |
| *Semi-Routine Occupations* | -0.26 | (0.29) |  | 0.00 | (0.01) | 0.14 | 0.02 |
| *Routine Occupations* | -0.33 | (0.28) |  | 0.00 | (0.01) | 0.12 | 0.01 |
| Reading Scores | -0.06 | (0.02) | \*\* | -0.00 | (0.00) | (.) | (.) |
| Maths Scores | -0.02 | (0.01) |  | -0.00 | (0.00) | (.) | (.) |
| Intercept | -1.34 | (0.33) | \*\*\* | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |
| Number of observations | 7915 | | | | | | |
| McFadden’s Adjusted Pseudo R2 | 0.26 | | | | | | |
| Cox-Snell Pseudo R2 | 0.51 | | | | | | |
| Nagelkerke Pseudo R2 | 0.54 | | | | | | |
| AIC | 16122.66 | | | | | | |
| BIC | 16457.53 | | | | | | |
|  | \*\*\* p<.001, \*\* p<.01, \* p<.05 Data Source: NCDS  Note: Complete Case Analysis | | | | | | |

To understand this in a more manageable format, each variable[[11]](#footnote-11) is graphically visualised with their predicted probabilities. This allows for a more intuitive understanding of the multinominal logistic regression as well as providing a different outlook for interpretation. Graphing predicted probabilities by variable rather than looking at a table with variables grouped by outcome variable allows for each variable to have cross-outcome group trends to be compare.

Focusing on NS-SEC, figure 1.1 depicts the predicted probabilities at means of economic activity of those categories that were statistically significant. Predicted probabilities for each economic activity category are presented. There is a general upward trend for individuals to be in employment going down the class schema. The opposite can be said with respect to post-schooling education and even more sharply so for those individuals that are continuing in school. Training & apprenticeships demonstrate that those individuals whose fathers were from NS-SEC class IV – in other words skilled manual occupations – are most likely to enter into training & apprenticeship pathways. This is not entirely surprising giving the socio-historical context of the NCDS. Apprenticeship & training schemes were heavily influenced by skilled manual worker occupations (Booth and Satchell, 1994).

Chart

Description automatically generated

Figure 1. 1 Predicted Probabilities of Economic Activity by NS-SEC

Moving on to Sex, figure 1.2 demonstrates that women are more likely than men to enter in employment post-mandatory schooling. This is also true for less traditional post-schooling education and school (though comparatively the effect sizes are much smaller for school than for post-schooling education). The largest sex-based effect relates to men being more likely to enter training & apprenticeship pathways compared to women. Given that the ‘feminization’ of the labour market had only recently begun (Guinea-Martin and Elliott, 2008) under the NCDS cohort, alongside the predisposition for apprenticeship programs to be focused primarily upon skilled manual labour (Booth and Satchell, 1994) it is somewhat understandable as to why this effect size is as large as it is.

A picture containing text, diagram, line, parallel

Description automatically generated

Figure 1. 2 Predicted Probabilities of Economic Activity by Sex

Moving on to educational attainment, figure 1.3 demonstrates that effect sizes comparatively are largest when discussing attainment over other variables. Predictably, those individuals that gained five or more O-levels were more likely to enter school pathways and less likely to enter in employment pathways compared to their peers that received less than five O-levels. Those that earned less than five O-levels were more likely than their academically successful peers to enter training & apprenticeship and post-schooling education pathways. This suggests that for the NCDS cohort, even those that did not do well academically, there were a variety of pathways above and beyond entering employment straight after mandatory schooling ended. For those that did want to continue their education in some manner, pathways were available to them that were not rigidly restricted based upon academic merit at earlier life stages.

A picture containing text, diagram, line, parallel

Description automatically generated

Figure 1. 3 Predicted Probabilities of Economic Activity by Educational Attainment

Moving on to housing tenure, figure 1.4 demonstrates that whilst there are substantive findings, the effect sizes compared to other variables are the smallest. Those that grew up in households that owned their own home were less likely to enter employment straight after mandatory schooling ended and more likely to enter some form of education – either post-schooling education or school-based pathways.

A picture containing text, diagram, line, screenshot

Description automatically generated

Figure 1. 4 Predicted Probabilities of Economic Activity by Housing Tenure

### **Discussion and Conclusion**

The multinominal logistic regression model indicates that educational attainment by far has the largest single substantive impact upon an individual’s economic activity sorting post-mandatory schooling. The impact of educational attainment had the greatest level of impact on those sorting into schooling over employment. Sex was also found to have substantively significant relationships with certain economic activities – primarily training & apprenticeships. The negative gradient amongst NS-SEC indicates that for educational economic activity outcomes, there is a class-based impact upon individual activity sorting post-mandatory education. Though, this class impact is not as pronounced as educational attainment. Cognitive ability provides a small but statistically significant effect upon economic activity outcomes. The interaction between cognitive ability and educational attainment within this model presents evidence for a mediating effect.

There are several implications these findings have on previous discussions of social theory. The first relates to the discussion on individualisation and structuralism. These findings present a clear picture that structural inequalities do in fact matter. Social class, sex, and housing tenure all present clear substantive effects on the pathways that individuals choose post-mandatory schooling. Importantly however, is that different structural inequalities have varied levels of influence and effect sizes dependent upon the given pathway being discussed. For example, some of the most pronounced social class effects relate to individuals choosing to stay on in school post-mandatory schooling. With those whose parents occupy NS-SEC VII having a much lower chance of choosing this pathway comparative to say, choosing to start employment. Comparative to another structural inequality such as sex however, and the sexed effects for staying on in school are very small substantively speaking. When it comes to training & apprenticeship pathways however, there is a much more pronounced effect here – with men much more likely to choose this pathway comparative to women. These varied structural effects speak to the complex socio-historical context of the NCDS. As previously mentioned, the status of the labour market at the time of the NCDS cohort, as well as the skilled manual dominance of training & apprenticeship schemes during this time is probably the best explanation for this structural variation.

Social class and sex were not the only structural inequalities measured within this model. Housing tenure was included to assess the views of ‘new structuralism’. Firstly, the view that social class is disaggregated (Saunders, 2003) is demonstrated to be incorrect. That does not necessarily mean that the main tenant of new structuralism – that being the social restratification of advanced capitalist societies (ibid) presenting new consumption cleavages like housing tenure – is necessarily totally incorrect. Whilst this model presents clear evidence that housing tenure, when controlling for social class influences individual’s pathway selection, the effect size across pathways is substantively small. Thus, whilst it would be correct to state that consumption cleavages in the form of housing tenure, constitute a level of influence in shaping the material life chances of individuals (Saunders, 2021), there needs to be care not to overstate this influence. Social class matters – above that of housing tenure. Ultimately whilst new structuralisms arguments that social class has become disaggregated does not find empirical evidence amongst the NCDS cohort, the view that housing tenure is important in influencing pathway choice does find some support.

These findings are not without caveats. Firstly, given the socio-historical context of the NCDS cohort other structural factors such as race would have been relevant for inclusion within the model. Unfortunately, practical reasons related to how the NCDS measured race, alongside the weak statistical power of any race variable given the low sample size means this is not feasible. Alongside this, the combination of an unemployment category with an out of the labour force category qualitatively conflates two sociologically distinct concepts – the latter of which has a rich history with structural inequalities related to sex, that for reasons related to statistical power could not be analysed.

In addition to this, when constructing social class for this model, the choice was made to use NS-SEC. There is no definitive reason for choosing NS-SEC over any other social class schema. The fact that a different choice of schema could potentially influence the substantive interpretation of the findings presented here is cause for concern. Finally, another potential issue relates to missing data. The model presented here covers 7915 individuals. This amounts to 63 per cent complete cases for sweep 4 (age 23). Missing data could potentially skew the substantive findings. In the next section, this model will undergo a sensitivity analysis alongside other social stratification measures to assess the first issue raised. After this, another section seeks to handle missing data within this model in an attempt to understand the potential impact missingness has had upon the interpretation of this model.

## **Sensitivity Analysis of Independent Variables**

There are a variety of socio-economic measures used by social scientists. It is not common practice within social stratification research (Lambert and Barnett, 2021) but a sensitivity analysis of social stratification measures provides the most well-informed assessment about which social stratification measure to use within a given model. NS-SEC, CAMSIS, and RGSC are three of these measures. The analytical distinctions between these three measures has already been discussed. Given the historical nature of the NCDS cohort, a sensitivity analysis would provide an interesting insight to the temporal sensitivity of these socio-economic measures, as well as presenting results that demonstrate the best model fit. The subsequent sensitivity analysis will compare like-for-like models of economic activity each using a different socio-economic measure. The base model – NS-SEC – is then compared to the CAMSIS and RGSC models. Whilst it is not appropriate to compare log odds across regression models, the following sensitivity analysis will instead compare models following substantive conclusions. Goodness-of-fit statistics are provided and are assessed via AIC, BIC, and a range of R2 measures.

### **Testing Measures of Parental Social Class**

There are strong correlations between parental social class measures. Parental NS-SEC and Parental RGSC has a significant Chi-Square statistic at the p<0.001 level. Parental NS-SEC and CAMSIS have a significant Anova at the p<0.001 level.

Three separate multinominal logistic regressions are presented in table 1.12. The first model has been described at length in the previous section and uses NS-SEC. The second model uses CAMSIS and the third uses RGSC. These models all present using log odds. A further table assessing all models using average marginal effects is found within the appendix in table A5.

The substantive conclusions drawn from the analysis of the NS-SEC model do not change across all three models. In fact, it is remarkable how similar the coefficients are across the models. Looking specifically the social stratification measures across the models, all produce substantively identical conclusions. This provides evidence to suggest that none of the socio-economic measures diverge greatly and all are as temporally sensitive (or insensitive) as one another. This is an interesting finding, considering the construction of each of these various socio-economic measures happened at different time-points, the similarity within substantive findings is remarkable. Especially when considering the fact each socio-economic measure operates under slightly different analytical frameworks.

The goodness-of-fit statistics are similar for all three models. Differences in R2 measures exist but the small nature of these differences indicate the amount of variance explained across the three models remains consistent. AIC and BIC differences are also small. The most parsimonious model is the CAMSIS model. Considering BIC penalises models for estimating additional parameters it is not entirely surprising that it considers the CAMSIS to be a better fit than the NS-SEC schema. These differences are however extremely small. Given the identical substantive conclusions alongside the near identical goodness-of-fit statistics there is no overall preferred model statistically speaking. As such, going forward the preferred model of choice for subsequent analysis will be the NS-SEC model. This choice was primarily made due to its relative use in contemporary sociological literature compared to the RGSC and to a lesser degree CAMSIS.

Table 1. 12 Sensitivity analyses of alternative measures of parental social stratification

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | | |  |  |  |  |  |  |
|  | **NS-SEC** | | | **CAMSIS** | | | **RGSC** | | |
| Economic Activity | **Coef.** | **S.E** | **Sig.** | **Coef.** | **S.E** | **Sig.** | **Coef.** | **S.E** | **Sig.** |
| Employment | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
|  |  |  |  |  |  |  |  |  |  |
| Post-Schooling Education |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 1.60 | (0.11) | \*\*\* | 1.60 | (0.11) | \*\*\* | 1.61 | (0.11) | \*\*\* |
| Sex |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Male* | -0.43 | (0.10) | \*\*\* | -0.44 | (0.10) | \*\*\* | -0.43 | (0.10) | \*\*\* |
| Housing Tenure |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Don't Own Home* | -0.71 | (0.10) | \*\*\* | -0.68 | (0.10) | \*\*\* | -0.72 | (0.10) | \*\*\* |
| NS-SEC |  |  |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.47 | (0.16) | \*\* |  |  |  |  |  |  |
| *Intermediate Occupations* | -0.66 | (0.30) | \* |  |  |  |  |  |  |
| *Small Employers* | -1.05 | (0.19) | \*\*\* |  |  |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -0.71 | (0.16) | \*\*\* |  |  |  |  |  |  |
| *Semi-Routine Occupations* | -1.05 | (0.17) | \*\*\* |  |  |  |  |  |  |
| *Routine Occupations* | -1.04 | (0.16) | \*\*\* |  |  |  |  |  |  |
| CAMSIS |  |  |  | 0.28 | (0.04) | \*\*\* |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |
| *Professional* |  |  |  | (.) | (.) |  | (.) | (.) |  |
| *Managerial and Technical* |  |  |  |  |  |  | -0.56 | (0.28) | \* |
| *Skilled Non-Manual* |  |  |  |  |  |  | -0.46 | (0.29) |  |
| *Skilled Manual* |  |  |  |  |  |  | -1.00 | (0.27) | \*\*\* |
| *Partly Skilled* |  |  |  |  |  |  | -1.14 | (0.29) | \*\*\* |
| *Unskilled* |  |  |  |  |  |  | -1.52 | (0.33) | \*\*\* |
| Reading Scores | 0.04 | (0.01) | \*\* | 0.04 | (0.01) | \*\* | 0.04 | (0.01) | \*\*\* |
| Maths Scores | 0.01 | (0.01) | \* | 0.01 | (0.01) |  | 0.01 | (0.01) |  |
| Intercept | -1.35 | (0.20) | \*\*\* | -3.32 | (0.22) | \*\*\* | -1.25 | (0.31) | \*\*\* |
|  |  |  |  |  |  |  |  |  |  |
| School |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 2.94 | (0.09) | \*\*\* | 2.93 | (0.09) | \*\*\* | 2.94 | (0.09) | \*\*\* |
| Sex |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Male* | 0.37 | (0.08) | \*\*\* | 0.35 | (0.08) | \*\*\* | 0.36 | (0.08) | \*\*\* |
| Housing Tenure |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Don't Own Home* | -0.56 | (0.08) | \*\*\* | -0.52 | (0.08) | \*\*\* | -0.58 | (0.08) | \*\*\* |
| NS-SEC |  |  |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.67 | (0.14) | \*\*\* |  |  |  |  |  |  |
| *Intermediate Occupations* | -1.16 | (0.27) | \*\*\* |  |  |  |  |  |  |
| *Small Employers* | -1.21 | (0.16) | \*\*\* |  |  |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -1.06 | (0.14) | \*\*\* |  |  |  |  |  |  |
| *Semi-Routine Occupations* | -1.09 | (0.14) | \*\*\* |  |  |  |  |  |  |
| *Routine Occupations* | -1.48 | (0.14) | \*\*\* |  |  |  |  |  |  |
| CAMSIS |  |  |  | 0.36 | (0.03) | \*\*\* |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |
| *Professional* |  |  |  | (.) | (.) |  | (.) | (.) |  |
| *Managerial and Technical* |  |  |  |  |  |  | -0.97 | (0.24) | \*\*\* |
| *Skilled Non-Manual* |  |  |  |  |  |  | -0.97 | (0.25) | \*\*\* |
| *Skilled Manual* |  |  |  |  |  |  | -1.57 | (0.23) | \*\*\* |
| *Partly Skilled* |  |  |  |  |  |  | -1.63 | (0.25) | \*\*\* |
| *Unskilled* |  |  |  |  |  |  | -1.86 | (0.27) | \*\*\* |
| Reading Scores | 0.06 | (0.01) | \*\*\* | 0.06 | (0.01) | \*\*\* | 0.06 | (0.01) | \*\*\* |
| Maths Scores | 0.04 | (0.01) | \*\*\* | 0.04 | (0.01) | \*\*\* | 0.04 | (0.01) | \*\*\* |
| Intercept | -2.19 | (0.18) | \*\*\* | -4.76 | (0.20) | \*\*\* | -1.84 | (0.26) | \*\*\* |
|  |  |  |  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 0.09 | (0.11) |  | 0.11 | (0.11) |  | 0.11 | (0.11) |  |
| Sex |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Male* | 1.94 | (0.08) | \*\*\* | 1.93 | (0.08) | \*\*\* | 1.94 | (0.08) | \*\*\* |
| Housing Tenure |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Don't Own Home* | -0.25 | (0.07) | \*\*\* | -0.25 | (0.07) | \*\*\* | -0.25 | (0.07) | \*\*\* |
| NS-SEC |  |  |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.50 | (0.15) | \*\* |  |  |  |  |  |  |
| *Intermediate Occupations* | -0.40 | (0.27) |  |  |  |  |  |  |  |
| *Small Employers* | -0.19 | (0.15) |  |  |  |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -0.17 | (0.14) |  |  |  |  |  |  |  |
| *Semi-Routine Occupations* | -0.26 | (0.14) |  |  |  |  |  |  |  |
| *Routine Occupations* | -0.45 | (0.14) | \*\* |  |  |  |  |  |  |
| CAMSIS |  |  |  | 0.04 | (0.03) |  |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |
| *Professional* |  |  |  | (.) | (.) |  | (.) | (.) |  |
| *Managerial and Technical* |  |  |  |  |  |  | -0.51 | (0.27) |  |
| *Skilled Non-Manual* |  |  |  |  |  |  | -0.40 | (0.28) |  |
| *Skilled Manual* |  |  |  |  |  |  | -0.42 | (0.26) |  |
| *Partly Skilled* |  |  |  |  |  |  | -0.43 | (0.27) |  |
| *Unskilled* |  |  |  |  |  |  | -0.81 | (0.28) | \*\* |
| Reading Scores | 0.01 | (0.01) |  | 0.01 | (0.01) |  | 0.01 | (0.01) |  |
| Math Scores | 0.02 | (0.01) | \*\*\* | 0.02 | (0.01) | \*\*\* | 0.02 | (0.01) | \*\*\* |
| Intercept | -1.90 | (0.17) | \*\*\* | -2.35 | (0.18) | \*\*\* | -1.72 | (0.28) | \*\*\* |
|  |  |  |  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 0.12 | (0.28) |  | 0.12 | (0.28) |  | 0.11 | (0.28) |  |
| Sex |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Male* | -0.01 | (0.14) |  | -0.01 | (0.14) |  | -0.02 | (0.14) |  |
| Housing Tenure |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Don't Own Home* | 0.13 | (0.16) |  | 0.10 | (0.15) |  | 0.14 | (0.16) |  |
| NS-SEC |  |  |  |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.46 | (0.34) |  |  |  |  |  |  |  |
| *Intermediate Occupations* | -0.39 | (0.58) |  |  |  |  |  |  |  |
| *Small Employers* | -0.52 | (0.33) |  |  |  |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -0.72 | (0.32) | \* |  |  |  |  |  |  |
| *Semi-Routine Occupations* | -0.26 | (0.29) |  |  |  |  |  |  |  |
| *Routine Occupations* | -0.33 | (0.28) |  |  |  |  |  |  |  |
| CAMSIS |  |  |  | -0.06 | (0.07) |  |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |
| *Professional* |  |  |  | (.) | (.) |  | (.) | (.) |  |
| *Managerial and Technical* |  |  |  |  |  |  | 1.02 | (1.03) |  |
| *Skilled Non-Manual* |  |  |  |  |  |  | 0.35 | (1.06) |  |
| *Skilled Manual* |  |  |  |  |  |  | 0.71 | (1.02) |  |
| *Partly Skilled* |  |  |  |  |  |  | 0.94 | (1.03) |  |
| *Unskilled* |  |  |  |  |  |  | 0.90 | (1.03) |  |
| Reading Scores | -0.06 | (0.02) | \*\* | -0.06 | (0.02) | \*\* | -0.05 | (0.02) | \*\* |
| Math Scores | -0.02 | (0.01) |  | -0.02 | (0.01) |  | -0.02 | (0.01) |  |
| Intercept | -1.34 | (0.33) | \*\*\* | -1.49 | (0.34) | \*\*\* | -2.54 | (1.04) | \* |
|  |  |  |  |  |  |  |  |  |  |
| Number of observations | 7915 | | | 7915 | | | 7915 | | |
| McFadden’s Adjusted Pseudo R2 | 0.26 | | | 0.26 | | | 0.25 | | |
| Cox-Snell Pseudo R2 | 0.51 | | | 0.51 | | | 0.51 | | |
| Nagelkerke Pseudo R2 | 0.54 | | | 0.54 | | | 0.54 | | |
| AIC | 16122.66 | | | 16106.37 | | | 16151.18 | | |
| BIC | 16457.53 | | | 16301.72 | | | 16458.14 | | |
|  | \*\*\* p<.001, \*\* p<.01, \* p<.05 Data Source: NCDS  Note: Complete Case Analysis for NS-SEC, CAMSIS, RGSC | | | | | | | | |

### **Discussion and Conclusions**

This section has sought to provide a sensitivity analysis of socio-economic measures into the model of economic activity to understand if there is any temporal sensitivity amongst the socio-economic measures and also to understand the best fit model amongst these measures. Overall, the findings suggest that there is harmony in socio-economic measures when it comes to the substantive findings of the economic activity model and the near identical goodness-of-fit statistics suggest no one model is better than any other. Overall, this section concludes by going forward with the NS-SEC model for any subsequent statistical analysis.

Perhaps the most interesting finding out of this sensitivity analysis stems from the near identical results from the NS-SEC and RGSC model. The NS-SECs predecessor – the Goldthorpe schema – was claimed by the likes of Marshal to be superior to the likes of the RGSC schema (Rose and Pevalin, 2002). For the sensitivity analysis to present substantive findings of the Goldthorpe schemas successor being identical to the likes of RGSC does present some challenges to the apparent superiority of this analytical construction.

When comparing AIC and BIC statistics there is a slight advantage given to the CAMSIS model over both RGSC and NS-SEC. This would at first suggest the selection of the CAMSIS model for substantive interpretation going forward. This would be an erroneous conclusion, however. Firstly, the differences in AIC and BIC statistics between say CAMSIS and NS-SEC models are marginal at best – this is most likely due to such statistics favouring a continuous measure over a categorical one. Secondly, the intent on model selection is not entirely rested upon model parsimony but also how best it presents evidence toward understanding given social phenomena. Given that one of the core elements of this research is to understand the nature of structural inequalities and how they influence choice and opportunity for youth transitions post mandatory education, a social class measure theoretically is preferable to a continuous measure such as CAMSIS. The combination of these two factors presents a compelling case for selecting NS-SEC as the model going forward. If the AIC and BIC statistics pointed to a large disparity between the NS-SEC and CAMSIS models, model selection may have resulted in a different conclusion.

Overall, this section on sensitivity analysis has used contemporary statistical techniques to update prior literature on youth transitions within the NCDS cohort and has also aided in model selection going forward. The next section seeks to continue this tradition of employing contemporary statistical techniques by attempting to deal with missingness within the NS-SEC preferred model.

## **Missing Data in the NCDS**

### **Missing Data**

Missing data is an essential component of any longitudinal data analysis – the major concern being that missing data and non-response is bound to affect the inferences made by the analysis of longitudinal studies (Hawkes and Plewis, 2006: 479; Silverwood *et al.*, 2021). The various factors that account for sample attrition in the datasets outlined above has the potential to present real issues as it relates to comprehensive data analysis. For the purposes of analysis those that exit the sample due to death or emigration are considered ‘natural’ exists from the original sample. Those however that either cannot be found, reject continued participation etc are individuals that we hold partial data on – being able to utilise this partial data within my analysis could be beneficial.

When dealing with missing data there are three primary types of classification. The first is missing completely at random (MCAR), meaning that missingness does not depend on observed or unobserved values. The second, being missing at random (MAR), meaning that given observed values missingness does not depend on the unobserved ones. Finally, missing not at random (MNAR) meaning that missingness depends on unobserved values (Silverwood et al 2021: 2). If data is found to be MAR then approaches like multiple imputation (MI), inverse probability weighting are made available – the former being extensively documented with the NCDS in particular in (Hawkes and Plewis 2006).

When dealing with missing data there are multiple methods to tackle the problem. The first is listwise deletion. Listwise deletion removes all observations from the data which have a missing value in one or more of the variables included in analysis. This is also known as Complete Case Analysis (CCA). The CCA approach is unpredictable, there is no way to know the consequences for this loss of information (Carpenter and Kenward, 2012).

A second method that deals with missing data is the use of survey weights. Survey weights take into account missingness, Inverse Probability Weighting (IPW) creates weighted copies of complete cases so as to remove selection bias introduced by missing data. Whilst IPW is a method of dealing with missing data, alternatives such as multiple imputation are regarded as much more efficient (Seaman *et al.*, 2012; Seaman and White, 2013).

A third method involves Multiple Imputation (MI). This method substituted missing data with substituted values[[12]](#footnote-12). MI is an attractive method because it is practical and widely applicable (Carpenter and Kenward, 2012).

Full-information Maximum Likelihood (FIML) is another method for dealing with missingness. For the regression based analysis including interactions with data from at least two stages of the life course, (Silverwood *et al.*, 2021) as the current analysis is, multiple imputation is plausible and more flexible than FIML. This flexibility stems from the ability to include auxiliary variables more easily within the imputation phase as well as being readily able to after imputing data sets obtain point estimates and standard errors at ease (Carpenter and Kenward, 2012). Recently, there has been some debate surrounding FIML vs MI approaches.

Paul Allison in a series of articles (Allison, 2012a, 2012b, 2015) argues that FIML is 1) simpler to implement, 2) FIML has no incompatibility between an imputation model and an analysis model, 3) FIML produces a deterministic result rather than a different result every time, and 4) FIML is asymptomatically efficient. Firstly, MI does have greater variability than FIML but that increased choice in model selection is not necessarily a negative so long as proper procedures are followed – in fact greater variability of choice has the potential to make MI a more attractive candidate for dealing with missingness over FIML. Secondly, MI models only run into an incompatibility problem when the MI model is inconsistent from the CCA model – something that with appropriate testing and open science practices detailing the model construction, shouldn’t happen. Thirdly, MI models are deterministic provided the same seed is used each time you run the imputation. The only time this would not be plausible would be a scenario where open science practices were not followed, and fellow researchers could not access the MI seed[[13]](#footnote-13). Finally, the argument that FIML is asymptotically efficient only holds true to a certain extent. MI models reach asymptotic efficient by running an infinite number of imputations – though you can reach near to full efficiency with a relatively small number of imputations, Allison (Allison, 2015) argues around 10[[14]](#footnote-14). Overall, whilst FIML does offer some advantages, there is nothing so considerable as to desire FIML over MI. So long as open science procedures are upheld, most major critiques of MI are dealt with. As such subsequent analysis uses CCA and MI to compare the substantive conclusions between the two and to understand if missingness impacts interpretation.

When dealing with MI the subsequent question that naturally follows is how many imputations is sufficient? Silverwood et al (2021: 21) suggest that anything around 50 imputations would be sufficient for reliable estimation of point estimate and estimating p-values with little error. Though sometimes with large samples with sizeable missingness more imputations may be required.

There are 12,450 individuals identified in the NCDS who indicated some form of economic activity at age 23. After using the variable related to the outcome of tracing and interview there are a total of 12,536 individuals within this sample[[15]](#footnote-15). There are 4,638 observations with missing data on at least one of the variables included for analysis. Of the missingness amongst variables: 86 were missing in economic activity, 26 in educational attainment, 1893 in housing tenure, 3779 on NS-SEC, 1,747 on reading scores and 1,751 on cognitive ability. Sex has no missing data as it was recorded at wave 0 (so all individuals were included).

Patterns of missing data are presented in table 1.13. Within the NCDS sample, 63 per cent have complete records on all variables, 16 per cent are missing values at socio-economic measures, a further 8 per cent missing on socio-economic measures, housing tenure and cognitive ability, 5 per cent missing at socioeconomic measures and housing tenure, 4 per cent missing at cognitive ability, and 2 per cent missing values for housing tenure. Further missingness in the sample not presented in the table are <1 per cent. In total, 7,898 cases have a complete observation at all variables, 4,638 cases have incomplete observations.

An overview by (Power and Elliott, 2006) suggests that after accounting for death and emigration, sample loss over time is mainly attributed to individuals moving within the UK and not responding to requests to trace them. As mentioned previously refusal at age 23 was 7.1 per cent. With an eligible sample of 16,402, this corresponds to 1,181 people dropping out due to refusal. Taking the eligible sample after death, emigration, and refusal to 15,221. At age 23, the NCDS notes that 12,503 people were successfully traced and conducted a full interview, with a further 33 completing a partial interview (Power and Elliott, 2006). This amounts to 12,536 people successfully interviewed, meaning that 2,686 people who did not die, emigrate, or refuse to take part in the survey are missing from sweep 4 (age 23).

Table 1. 13 Missing data patterns for NCDS

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Percent Complete (%) | Educational Attainment | Economic Activity | Housing Tenure | NS-SEC | Reading | Maths |
| 63 | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** |
| 16 | **✓** | **✓** | **✓** |  | **✓** | **✓** |
| 8 | **✓** | **✓** |  |  |  |  |
| 5 | **✓** | **✓** |  |  | **✓** | **✓** |
| 4 | **✓** | **✓** | **✓** | **✓** |  |  |
| 2 | **✓** | **✓** |  | **✓** | **✓** | **✓** |

Of the missing data, 63 per cent of cases are complete, the largest proportion of missing data comes from the three socio-economic variables: NS-SEC, CAMSIS, and RGSC. Individuals that are not active within the labour force, are difficult to code within socio-economic measures. NS-SEC categorisation is based upon occupation type, authority duties, and the size of organisation (Rose and Pevalin, 2002). CAMSIS occupational scores only relate to the working population within a country. The RGSC also suffers similar problems with reference to categorisation of retired, unemployed, and out of the labour force individuals. Unemployment rates during this time period including people out of the labour force are around 5-7 per cent (ONS, 2023). This combined with refusal rates of 7.1 per cent (Power and Elliott, 2006) presents a reasonable case in explaining potential missingness of the largest missing variables in the model.

A complete case analysis is only valid if data can be considered MCAR. If data is found to be MAR, then steps should be taken to address this potential bias in analysis. The gold standard to deal with MAR related data is MI (Treiman, 2009). The following section seeks to compare a CCA and MI approach to estimate if there are any differences in the substantive conclusions reached. If differences are found, implications are then discussed.

### **Multiple Imputation by Chained Equations**

Multiple Imputation by Chained Equations is a tool developed to address missing data on all variables within a given model at the same time. It does this by filling in missing values in multiple variables iteratively by using chained equations Multiple imputation models are estimated using the mi suite in Stata. This suite is compatible with the svy suite and so can also adjust for complex survey design.

Whilst multiple imputation does help when it comes to missingness, it does have some drawbacks. Goodness-of-fit statistics for example are not able to be used – R2 and BIC most prominently. Therefore, it is not possible to assess the more appropriate or parsimonious model.

Disadvantaged socio-economic background in childhood, worse mental health and lower cognitive ability in early life, and lack of civic and social participation in adulthood have been found to be consistently associated with non-response (Silverwood *et al.*, 2021). These variables are easily translated into auxiliary variables for imputation. Using the NCDS missing data guide (ibid), each predictor of non-response at sweep 4 (age 23) was recoded. These correspond to region, number of persons per room, sex of child, social class of mother’s husband, family moves since child’s birth, dad reads to child, area of world in which mother was born, number of family moves since child’s birth, number of household amenities, number of family moves since child’s birth, sum of favourable learning environments. Some auxiliary variables such as cognitive ability was taken out as they were already included in the main model of analysis.

Whilst there is no clear consensus on the number of optimum imputations required to estimate the model (Silverwood *et al.*, 2021), 50 imputations is advised with checks performed after values have been imputed to see if more imputations are required. Basic tests were performed with 5 imputations before increasing this to 50. After 50 imputations were performed, basic tests were replicated, the results remained consistent and stable across imputations.

The imputation model naturally includes all analytical variables included in the previous chapter. The imputation model also includes several auxiliary variables to add in maximising the plausibility of the MAR assumption in order to reduce bias due to missingness (ibid). These are broken down into two types. The first are variables that are predictive of both the probability of missingness and the underlying missing values themselves. The second are variables that are predictive of the underlying missing values only. Missing data in the NCDS is largely derived by non-response at a given sweep and auxiliary variables are selected from pre-determined sets of variables predictive of non-response at sweep 4 as seen in (Silverwood *et al.*, 2021). The fact that missingness within the NCDS is largely driven by sweep non-response rather than item non-response means it is even more important to include auxiliary variables within the imputation.

There were 18 variables identified for inclusion in the imputation model. However, variables that are predictive of the chance of missing values but are not predictive of the underlying missing values themselves will not add any information to the model (ibid). Thus, the decision was made not to include such variables in the imputation model that were not predictive of economic activity following the advice from the handling missing data NCDS guide (ibid). From this 8 out of 18 variables are associated with economic activity in a substantive way and so are included as auxiliary variables in the imputation model.

Prior to imputation it is best to explore the distribution of variables comparative to complete and non-complete cases. In the presence of a MCAR mechanism all distributions should be the same comparatively. If this is not the case, then this is suggestive of a MAR or MNAR mechanism. Variables such as maths scores present clear imbalances – a mean of 17.60 in complete cases and 16.06 in non-complete cases. These imbalances present themselves in every variable within the model except for sex. This is unsurprising considering that sex as a variable presents zero missingness. The distributions of the variables thus far present some indications of a MAR or MNAR mechanism being present.

With a greater number of variables as our model has, convergence issues are probable. This risk is increased if a model has many categorical variables. Failure to converge was a consistent problem. Without resorting to re-coding analytical variables, the decision was made to drop one of the auxiliary variables in order to produce an imputed model[[16]](#footnote-16).

The means and standard deviations of imputed values from each iteration of the imputation were checked and tabulations of each categorical variable to check the distributions of each variable against the imputations.

The next models presented will be a comparison of a complete case analysis using NS-SEC from the previous chapter and the imputed model in table 1.14. The CCA model has 7,915 observations. Using a variable within the NCDS dataset (add what variable this is) that noted down how many individuals were successfully contacted for sweep 4 (age 23) of the NCDS, there are 12,536 individuals within this sweep. The imputed dataset thus has 12,536 observations compared to the 7,915 observations of the CCA model.

Overall, there is similarity between the complete case analysis and the imputed model. The substantive conclusions between CCA and MI models are near identical. There are some very slight differences in the log odds across educational attainment, sex, and housing tenure but these slight differences are not large enough to impact the substantive conclusions presented in the interpretation of the CCA model. The imputed model confirms the substantive conclusions made from the CCA model.

Table 1. 14 Comparison of CCA NS-SEC vs Imputed NS-SEC

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | CCA NS-SEC | | | Imputed NS-SEC | | |
| Economic Activity | **Coef.** | **S.E** | **Sig.** | **Coef.** | **S.E** | **Sig.** |
| Employment | Ref. | (.) |  | (.) | (.) |  |
|  |  |  |  |  |  |  |
| Post-Schooling Education |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 1.60 | (0.11) | \*\*\* | 1.52 | (0.09) | \*\*\* |
| Sex |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  |
| *Male* | -0.43 | (0.10) | \*\*\* | -0.40 | (0.08) | \*\*\* |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  |
| *Don't Own Home* | -0.71 | (0.10) | \*\*\* | -0.77 | (0.08) | \*\*\* |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.47 | (0.16) | \*\* | -0.45 | (0.15) | \*\* |
| *Intermediate Occupations* | -0.66 | (0.30) | \* | -0.52 | (0.28) | \* |
| *Small Employers* | -1.05 | (0.19) | \*\*\* | -0.94 | (0.18) | \*\*\* |
| *Lower Supervisory & Technical Occupations* | -0.71 | (0.16) | \*\*\* | -0.73 | (0.15) | \*\*\* |
| *Semi-Routine Occupations* | -1.05 | (0.17) | \*\*\* | -0.98 | (0.16) | \*\*\* |
| *Routine Occupations* | -1.04 | (0.16) | \*\*\* | -0.98 | (0.14) | \*\*\* |
| Reading Scores | 0.04 | (0.01) | \*\* | 0.04 | (0.01) | \*\*\* |
| Maths Scores | 0.01 | (0.01) | \* | 0.01 | (0.01) | \*\* |
| Intercept | -1.35 | (0.20) | \*\*\* | -1.42 | (0.18) | \*\*\* |
|  |  |  |  |  |  |  |
| School |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 2.94 | (0.09) | \*\*\* | 2.89 | (0.08) | \*\*\* |
| Sex |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  |
| *Male* | 0.37 | (0.08) | \*\*\* | 0.43 | (0.06) | \*\*\* |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  |
| *Don't Own Home* | -0.56 | (0.08) | \*\*\* | -0.59 | (0.07) | \*\*\* |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.67 | (0.14) | \*\*\* | -0.60 | (0.14) | \*\*\* |
| *Intermediate Occupations* | -1.16 | (0.27) | \*\*\* | -1.03 | (0.26) | \*\*\* |
| *Small Employers* | -1.21 | (0.16) | \*\*\* | -1.04 | (0.15) | \*\*\* |
| *Lower Supervisory & Technical Occupations* | -1.06 | (0.14) | \*\*\* | -0.97 | (0.13) | \*\*\* |
| *Semi-Routine Occupations* | -1.09 | (0.14) | \*\*\* | -1.02 | (0.13) | \*\*\* |
| *Routine Occupations* | -1.48 | (0.14) | \*\*\* | -1.45 | (0.13) | \*\*\* |
| Reading Scores | 0.06 | (0.01) | \*\*\* | 0.06 | (0.01) | \*\*\* |
| Maths Scores | 0.04 | (0.01) | \*\*\* | 0.04 | (0.00) | \*\*\* |
| Intercept | -2.19 | (0.18) | \*\*\* | -2.21 | (0.16) | \*\*\* |
|  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 0.09 | (0.11) |  | 0.04 | (0.09) |  |
| Sex |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  |
| *Male* | 1.94 | (0.08) | \*\*\* | 1.96 | (0.06) | \*\*\* |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  |
| *Don't Own Home* | -0.25 | (0.07) | \*\*\* | -0.31 | (0.06) | \*\*\* |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.50 | (0.15) | \*\* | -0.43 | (0.15) | \*\* |
| *Intermediate Occupations* | -0.40 | (0.27) |  | -0.27 | (0.26) |  |
| *Small Employers* | -0.19 | (0.15) |  | -0.06 | (0.15) |  |
| *Lower Supervisory & Technical Occupations* | -0.17 | (0.14) |  | -0.09 | (0.14) |  |
| *Semi-Routine Occupations* | -0.26 | (0.14) |  | -0.18 | (0.14) |  |
| *Routine Occupations* | -0.45 | (0.14) | \*\* | -0.41 | (0.14) | \*\* |
| Reading Scores | 0.01 | (0.01) |  | 0.01 | (0.01) | \* |
| Maths Scores | 0.02 | (0.01) | \*\*\* | 0.02 | (0.00) |  |
| Intercept | -1.90 | (0.17) | \*\*\* | -2.09 | (0.15) | \*\*\* |
|  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | Ref. | (.) |  | (.) | (.) |  |
| *Five or More O-levels* | 0.12 | (0.28) |  | -0.06 | (0.23) |  |
| Sex |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  |
| *Male* | -0.01 | (0.14) |  | 0.11 | (0.10) |  |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  |
| *Don't Own Home* | 0.13 | (0.16) |  | 0.19 | (0.14) |  |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | Ref. | (.) |  | (.) | (.) |  |
| *Lower Managerial* | -0.46 | (0.34) |  | -0.34 | (0.31) |  |
| *Intermediate Occupations* | -0.39 | (0.58) |  | -0.19 | (0.54) |  |
| *Small Employers* | -0.52 | (0.33) |  | -0.34 | (0.31) |  |
| *Lower Supervisory & Technical Occupations* | -0.72 | (0.32) | \* | -0.59 | (0.29) | \* |
| *Semi-Routine Occupations* | -0.26 | (0.29) |  | -0.14 | (0.27) |  |
| *Routine Occupations* | -0.33 | (0.28) |  | -0.20 | (0.26) |  |
| Reading Scores | -0.06 | (0.02) | \*\* | -0.05 | (0.01) | \*\* |
| Maths Scores | -0.02 | (0.01) |  | -0.02 | (0.01) | \* |
| Intercept | -1.34 | (0.33) | \*\*\* | -1.44 | (0.29) | \*\*\* |
|  |  |  |  |  |  |  |
| Number of observations | 7915 | | | 12536 | | |
| Average RVI |  | | | 0.28 | | |
| Largest FMI |  | | | 0.36 | | |
|  | \*\*\* p<.001, \*\* p<.01, \* p<.05 Data Source: NCDS  Note: Comparison of CCA NS-SEC vs Imputed NS-SEC model | | | | | |

Additional checks on the imputed model are produced via postestimation statistics through RVI and FMI. A high RVI corresponds to large amounts of missing data and/or that they are weakly correlated with other variables within the imputed model. The closer the RVI is to zero corresponds with the less effect missing data has upon a variable’s variance. The FMI on the other hand relates to the proportion of the total sampling variance that is due to missing data. The higher the FMI is relating to the greater number of imputations required for reliable results. The number of imputations should be roughly equivalent to the highest FMI percentage (which has been followed in this model). The highest FMI in the imputed model is 36 per cent, indicating a requirement of at least 36 imputations – the MI model has 50 imputations. The relative efficiency indicates efficiency. The closer it is towards one indicates that the analysis has the right number of imputations.

The average RVI score was 0.28, meaning that on average there is a small impact that missing data has upon the model’s variance. According to the RVI scores across categories, NS-SEC across all economic activity and tenure within unemployed and out the labour force is the only two analytical variables that have consistently above average (greater than 0.30) RVI scores. For all other analytical variables missing data has little effect upon their variance. Housing tenure and NS-SEC both have moderate (greater than 0.25) FMI scores indicating that a substantive amount of the total sampling variance is due to missing data. The FMI value of Intermediate occupations for unemployed & OLF is the highest FMI value from the model with a value of 0.36. This value indicates that 50 imputations was an appropriate number of imputations. All variables have a relative efficiency close to one – none of the variables have a relative efficiency below 0.99. This lends support to the notion that 50 imputations are an appropriate number.

With regards to parameter-specific degrees of freedom (DF) and percentages of increase in standard errors due to nonresponse. The closer RVI is to zero, the larger the degrees of freedom regardless the number of imputations. The highest degrees of freedom correspond to sex. This suggests that the loss of information due to non-response is the smallest for this analytical variable. This is unsurprising given the level of missingness related to sex.

### **Discussion and Conclusions**

The fact that there are no substantive differences between the complete case analysis and the imputed model suggests that there is little evidence for a MAR assumption. This indicates that missingness in these categories has no substantive impact upon the resulting interpretation and analysis of results. As such, the imputed model is no better than the complete case analysis for interpretation. Whilst this section does present a lot of work that amounts to a preference for a complete case analysis model, the use of multiple imputation and discussion of missing data was important to discuss. There was no a priori evidence to suggest that the missing data present within this model was not MAR. Due to this, and due to the ability to check this assumption, there was a statistical due diligence to present these findings.

The implementation of tools seeking to deal with missingness within this model provides contemporary statistical techniques to the research of youth transitions within the NCDS. In essence whilst substantively finding identical results from complete case analysis to multiple imputed models, this exercise in dealing with missingness has allowed for the updating of youth transitions literature within this field of study. Going forward there is a relative level of confidence that this research of looking into the choices and opportunities of youth transitions from mandatory education has attempted to control for all statistical possibilities.

## **Discussion and Conclusions for Chapter One**

The overall empirical finding from analysis is that structural inequalities matter for NCDS youth in influencing their choice and opportunity as it relates to transitions from mandatory education. Social class, sex, housing tenure, and educational attainment all have substantively significant impacts upon an individual’s choice and opportunity with respect to pathway selection post-mandatory education.

More socio-economically advantaged children tend to maintain a pathway of elongated education, compared to their less advantaged peers who are more likely to enter employment. There is a caveat within this however, pathways expressed in non-traditional educational pathways as well as training & apprenticeship programs offer an opportunity for those at the lower end of socio-economic advantage. Most prominently with the latter of these pathways, individuals whose fathers are from skilled manual backgrounds are most likely to take up this opportunity – though these individuals are most likely to be men rather than women. This finding is consistent with previous studies related to the class divide within training & apprenticeships (Booth and Satchell, 1994) as well as the sex divide within the segregated labour market of the time (Dex and Bukodi, 2012).

Sex based differences within pathways are most evident within this training & apprenticeship category, although differences are also apparent with women more likely to enter employment over men. Whilst training & apprenticeship programs during the timeframe of the NCDS were dominated by traditionally masculine fields, the explanation as to why more women than men entered straight into employment is slightly more nebulous. One possibility relates to rational choice theory (Goldthorpe, 1998), in that given their structural position, women saw entering employment directly after mandatory schooling as the most secure and stable pathway for their given life course.

Other structural factors like housing tenure also matter – but not to the extent of traditional structural inequalities such as social class and sex. These findings of structural effects impacting life chances is something found within previous literature (Saunders, 2021). This empirical evidence suggests that over and above matters of social class, how you live – in the case of housing tenure – presents a clear barrier to entry when it comes to the choices and opportunities individuals make post-mandatory education.

The fact that previous structural inequalities that have manifested during previous life domains (Mayer, 2009) influence life chances in other life domains is indicative of support for promoting a life course perspective within this research. One aspect of these structural effects that has not been reflected on within this research is the recency onto which things like housing tenure and social class position may have upon constricting choice and opportunity. There is a possibility that an individual that was in a household that rented for most of their life and only recently bought their home would have a different magnitude of effect compared to an individual whose household had always owned their own home. Similarly, to this, an individual that has had multi-generational social class stability may have a different magnitude of effect compared to an individual whose family has very recently experienced upward (or downward) social mobility. This does not fit the purview of this research but is certainly something to not for future research within this area.

The findings from this empirical research appear to confirm the relatively influential impact that structures have upon individual life chances. Individual choice is constrained and influenced by structural factors such as social class, sex, and housing tenure. Something that is supported by previous literature (Hutchison, Prosser and Wedge, 1979; Connolly, Micklewright and Nickell, 1992; Booth and Satchell, 1994; Schoon *et al.*, 2001; Dolton, Joshi and Makepeace, 2002; Dolton, Makepeace and Marcenaro‐Gutierrez, 2005). Echoing back to discussions on agency and structure, this empirical evidence is suggestive of an agency within structure understanding of the life course – individuals certainly choose their given pathways and opportunities, but structures continue to influence and constrict these opportunities. Furthermore, whilst these empirical results appear to confirm much of previous empirical literature on this subject, the arguments proposed by new structuralist theorists (Saunders, 2003, 2021) are not as sound. Whilst evidence has been found that concurs with the premise first emphasised in Saunders (2003) that housing tenure offers an independent and substantive impact on life chances. The argument that it is dominant above that of social class has no evidence within these results.

The inclusion of prior educational attainment alongside structural effects such as social class, sex, and housing tenure provide a much more complex relationship with individuals choices and opportunities post-mandatory education. The findings provide substantial empirical evidence that prior educational attainment does matter in relation to pathway selection. Most notably these effects whether an individual seeks to continue along a traditional educational pathway. In saying this, lack of educational success at 16 does not block youth off from entering several pathways. It appears to influence traditional education, but this is not the case for employment, non-traditional educational pathways, and training & apprenticeships. This demonstrates that less academically able youth desire alternative pathways beyond employment, and that given the socio-historical context of the NCDS cohort, the option was there to allow young people to enter these alternative pathways.

This work provides methodological innovation by conducting sensitivity analysis of socio-economic measures of social stratification. Sensitivity analysis of NS-SEC, CAMSIS, and RGSC demonstrate that NS-SEC is a robust and strong measure of social class which is suitable for use within this model using NCDS data . The findings from this sensitivity analysis confirm that all three measures promoted an almost identical substantive interpretation, with almost identical BIC and AIC statistics. As such the robustness of NS-SEC is characterised by the equal levels of robustness that CAMSIS and RGSC demonstrate. Choosing NS-SEC as the dominant model through the analysis was based upon a theoretical desire to understand class-based dynamics. Through its implementation, social class was found to have a resounding impact upon individuals’ choices and opportunities in relation to transitional experiences. The results are also innovative by assessing missingness within the complete case analysis model. Missingness was first descriptively detailed before strategies for handling such missingness was discussed. A multiple imputation model found that missingness has no impact upon the substantive findings of the complete case analysis model. Whilst this means that the substantive findings remain the same as previously detailed, the implementation of dealing with missingness was an important contemporary statistical strategy that previous literature within this field typically overlooked. Both the implementation of sensitivity analysis and multiple imputation techniques thus serve a methodological innovation beyond that of prior literature within the field.

Overall, the literature has been updated, with prior literature being confirmed in some cases (Connolly, Micklewright and Nickell, 1992; Booth and Satchell, 1994; Dolton, Makepeace and Marcenaro‐Gutierrez, 2005) and challenged in others (Saunders, 2003, 2021). Social theories related to youth transitions and the impacts on structural effects and agency within the realm of choice and opportunity have also been contended with. Finally modern statistical techniques have been applied to update prior literature to modern statistical standards by employing sensitivity analyses and providing tools to deal with potential missingness.

Going forward, whilst previous literature has been confirmed and updated, questions remain that are important to reflect on for future inquiry. As has been mentioned, structural influence is dependent upon the given pathway of choice – with different structural influences mattering more for some pathways than for others. A closer inspection of these differences is called for. On top of this, so far this research reflects upon the entry into, or the transitional stage of going from school-to-work. Whilst identification of several pathways has been made apparent, prior literature related to the relative smoothness of these transitions is paramount to focus on next. Structural effects matter for the transition itself, but the question of whether they manifest throughout the life domain (Mayer, 2009) onwards is relatively important in the discussion of structural effects throughout the life course.

The next section will attempt to replicate the analysis conducted in this section using the British Cohort Study (BCS). The BCS is a nationally representative birth cohort survey that was conducted in a week in 1970. Much of the data in the BCS has been harmonised with the NCDS which allows for a detailed comparison of trends between cohorts. The analysis of the next section will reflect on the differences and similarities that have been made within the conclusion of this section focusing once again upon the nature of structural inequalities and their influence upon choice and opportunities for youth transitions.

# **Chapter Two: The British Cohort Study: Youth Transitions in Education and Employment**

"a different country… You have to blink and rub your eyes". – Jacques (1982)

# **Introduction to Chapter Two**

Continuing the theme set in chapter one, chapter two attempts to replicate the previous analysis of entry from school into work. This chapter focuses on the British Cohort Study, the third birth cohort study in the UK that started in 1970. As in chapter one, this chapter will focus on the pathways and choices made by individuals in the BCS after they reached 16 and ended mandatory schooling. Once more focus will be placed upon structural inequalities of social class, sex, and housing tenure to understand young people’s transitional experiences during this period. An attempt is made to replicate the NCDS chapter as precisely as possible to enable a comparison of cohort transitional experiences.

This chapter thus begins with a literature review of present literature on the topic of BCS youth and their transitional experiences. Then this chapter will move on to a replication analysis of the model used in chapter one, with sensitivity analysis and handling missing data sections. Finally, this chapter will conclude with a comparison of NCDS and BCS youth, pointing out any similarities or diversions in their biographical life courses.

The data used comes from sweeps from birth up to age 30 using the BCS. The present chapter continues the tradition of the previous by updating prior literature with modern statistical techniques. This chapter will look at four primary economic activity outcomes: employment, education, training & apprenticeship, and unemployment & out of the labour force. Due to the nature of the BCS, missingness will form a key discussion within this overall chapter.

## **Literature Review: BCS Timeframe and Context**

This section provides an overview of the literature within the field of youth transitions of the BCS cohort. This review focuses on existing research outlining the school-to-work transition and examination of the structural impacts upon that transition within the context of the BCS cohort. Initially literature will focus upon the historical and temporal context of the BCS cohort to ground the empirical research on transitions. As with chapter one, major transition themes are identified as they relate to employment, education, training, and unemployment. Each are influenced to some degree by structural factors that impact individual choice and opportunity. The changing nature of the labour market, and British polity during the BCS timeframe have had a substantive impact on the role of training and apprenticeships within an individuals first major transition from mandatory schooling into the world of economic activity.

At the time of the BCS cohort, young people were in full time education until the age of 16 – like the NCDS cohort. At this age, individuals were typically expected to sit some form of examination – for the BCS cohort they were some of the last individuals to sit the O-level at 16, prior to its replacement with the GCSE. After this period of mandatory schooling there were options of continuing within education, moving on to training under the YTS scheme, entering employment, or becoming unemployed or out of the labour force. The relative diversity of options comparative to the NCDS cohort was restricted. Traditional apprenticeship schemes were gone, as was unemployment benefit for individuals aged 16-18. These effects will be discussed at length in the literature review below.

### **Story of transitions for BCS youth**

Within Britain, the 1970s and 1980s were periods of large-scale transformation (Bynner, Ferri and Shepherd, 2019). The 1970 BCS cohort can be characterized by a continuing decline in manufacturing and apprenticeships, high levels of unemployment, greater government intervention in young people’s economic activity, and a growing higher education framework. The 1970s cohort experienced a society in which Harold Macmillan proclaimed so confidently that ‘’You never had it so good’’ (Hamnett, McDowell and Sarre, 1989).

The 1960s saw the ‘white heat’ of Wilson’s technological revolution that transformed British society and its subsequent labour markets into a service based, consumer based economy (Hamnett, McDowell and Sarre, 1989). The 1970s onwards saw a continuing trend post-war of a simultaneous growth of automation and technology alongside a decline of manufacturing, though this came more out of the 1970s recession that devastated the heavy industrial markets of the North of England – the recovery and rebuilding of a service economy was located exclusively within the South of England (Hamnett, McDowell and Sarre, 1989) – half of all jobs created between 1983-87 were created in the south east (ibid). These pressures brought about major labour market and societal transformation for society, and increased uncertainty and risk for the worker (Schoon, 2007; Beck, 2014). A result of this transformation of society, Hutton describes this period of British history as the ‘30/30/40’ society, whereby 40 per cent of the population are permanently in casual employment, 30 per cent are doing fine, and another 30 per cent are struggling, leading to the phrase ‘Getting on, getting by, getting nowhere’ (Bynner, Ferri and Shepherd, 2019).

These ‘new’ jobs were defined by their transferable skills across the service sector (Bynner and Ferri, 2003). As a result, the apprenticeship scheme that was linked to traditionally heavy manufacturing highly specialised training went into decline and was eventually replaced by the Youth Training Scheme in 1983 under the management of the Manpower Services Commission. This in turn would be replaced by Youth Training in 1990 (Droy, Goodwin and O’connor, 2019). The YTS arguably advented the first time in Britain that youth had become a category of large-scale policy intervention beyond that of education (Wallace and Cross, 1990). The YTS modus opernadi was based on keeping kids off the streets and filling unemployment gaps – this became especially apparent during the recession of 1986-7 whereby the unemployment rate for men was 2.6 per cent but 12 per cent were in some form of government training, though this eventually fell below unemployment figures in 1988 post-recession (Bynner and Ferri, 2003). The YTS has experienced sustained sociological critique (Droy, Goodwin and O’connor, 2019), its main issue comes from the fact that it was an attempt at direct intervention from a collapsing youth labour market from a government that was anti-interventionist. It started out as a one-year in 1983 (eventually to two-year in 1986) scheme that provided mostly low-level training that was more comparable to an alternative to unemployment than to higher education or employment (Bynner and Ferri, 2003). Whilst the YTS did maintain a stay average of 400,000 people between the years of 1958-89, it was neither an adequate replacement from the highly skilled training of a traditional apprenticeship, nor was it an adequate form of pay and employment. The description by some (Wallace and Cross, 1990) that the YTS represented a ‘dual-carriageway’ between school and work for school-leavers is adequate only if you envisage the car being in both lanes at the same time. The YTS was also internally stratified. There were some attractive highly trained schemes that it did offer, so called ‘Model A’ schemes that worked directly with employers, however these were very hard to acquire and oftentimes went to those that did not need them the most (Wallace and Cross, 1990). The ‘Model B’ schemes were most numerous and typically what people mean when they describe the YTS. Among these unattractive schemes, individuals were typically sorted into the growing service sector, associated with insecurity and risky employment prospects. This liminal zone of the youth labour market was stratified along gender and class grounds (Droy, Goodwin and O’connor, 2019).

It was for many a stopgap – an unattractive one at that. It would not be accurate to compare the YTS – which was a training scheme, to the much more rigorous training and education of a traditional apprenticeship (Bynner *et al.*, 2002). Most young people felt forced into the YTS scheme due to the Thatcher government cutting unemployment benefits for all people between the ages of 16-18 in 1988. This is arguably the start of the punitive approach toward unemployment and welfare in the late 20th century (Droy, Goodwin and O’connor, 2019). Due to the timing of these unemployment benefit cuts, the 1970 cohort were still able to claim if they so desired, though they still suffered as part of the ‘vulnerable core’ of the labour market through Thatcher’s cutes and de-regulations towards employment rights and the minimum wage (Hamnett, McDowell and Sarre, 1989). The proclamation in 1981 under the New Training Initiative of heralding in universal youth training for all was in reality a poorly thought out scheme that some compared to a stopgap, whilst harsher critiques referred to it simply as ‘slave labour’ (Bynner, Ferri and Shepherd, 2019). The YTA offered cheap, subsidised labour to employers with no requirements to continue an individual’s employment after the scheme was completed (Droy, Goodwin and O’connor, 2019). It would be fair to characterise the YTS as a short-term benefit to businesses whilst leaving the individual worker under-trained, under-paid, and often unemployed.

The initial desired purpose of the scheme was to establish a training scheme comparable alongside German lines (at the time argued to be the best apprenticeship program in Europe). The result however, was a scheme that failed to appropriately train youth, and the best form of vocational training was instead found to be employment itself (Bynner *et al.*, 2002). In fact, the YTS has been found to have had negative consequences for men’s employment prospects (Droy, Goodwin and O’connor, 2019; Goodwin *et al.*, 2020), and overall a negative impact on earnings over the life course (ibid).

The relative decline of apprenticeship schemes and increase in education opportunities due to the increasing pressure on young people to accumulate credentials (Bynner and Ferri, 2003) resulted in a much higher proportion of school leavers in the 1970s onwards staying on within education than their earlier cohorts. Those that did not choose to stay on within education and had little to no qualifications faced a harsh reality of ‘patchwork’ career trajectory, characterised by shifting occupations and periods of unemployment (Bynner, 2005). In 1976, the number of individuals that left school without any qualifications was 21 per cent, in 1986 it was only 9 per cent (Wallace and Cross, 1990). The 1970 cohort were the last to ever experience the dual O-level/CSE composition at 16 – the BCS cohort were in the middle of a massive amount of Education reform in that would come in 1988 with the advent of the Education Reform Act. Men in particular saw a large increased probability of being in full-time education over employment in comparison with the 1958 cohort (Bynner and Ferri, 2003), though large members of men were also entering government training schemes like the YTS. For women too, the decreasing numbers of young women being out of the labour force saw a corresponding increase in labour market participation (Bynner and Ferri, 2003) as well as higher education participation. The expansion of the university system in the late 1960s following the Robbins Report (*Robbins Report*, 1963) provided the supply of higher education places that this new service based labour market so often demanded (Bynner and Ferri, 2003). Though comparative to the continent at the time, European rates of participation in education were changing more rapidly than Britain (Bynner, Ferri and Shepherd, 2019). For most, the transition into adulthood is characterised by an initial movement from mandatory education to some for of employment. The fact that the BCS cohort appears to exhibit an elongated stay within education (Bynner *et al.*, 2002) is some indication of the changing nature of the labour market within the UK – and also provides evidence for the development of an ‘Emerging Adulthood’ (Bynner, 2005). This transitional change is indicative of two potential sources, the first would be a major economic shock, in the form of a recession would encourage individuals to stay in education for longer in order to avoid the initial economic shocks and uncertainty that comes with being employed in a labour market experiencing a downturn. The second, relates to a degree of economic restructuring due to technological change, resulting in different skills and credentials, thus encouraging a longer stay within education to garner such skills and credentials. For the BCS cohort, both of these things occurred (Bynner *et al.*, 2002). Leaving school to enter employment for minimum school age leavers was a much more difficult process compared to 10-20 years earlier – even more so for those living in industrial and manufacturing heartlands (Bynner *et al.*, 2002).

The returns to education for the BCS cohort confer a 17 per cent average increase in income for those individuals that stayed on within education post-mandatory schooling compared to their peers (Boero *et al.*, 2020). This is not entirely surprising considering that education is the most important single, individual predictor of adult incomes and earnings (Breen, 2022), though it does emphasise the importance of reflecting on the stratifying influences during education and their subsequent impacts on choice and opportunity post-education. This single most important predictor is a worrying phenomena when combined with a ‘wastage of talent’ (Bukodi, Bourne and Betthäuser, 2017) whereby young people from disadvantaged backgrounds face barriers to fully realise their academic potential within the British educational system.

Labour market restructuring was part of the increase in home ownership from the 1950s to 1990s. During the period of 1951 only 31 per cent of people owned their own home, in 1991 this rose to 67 per cent (Bynner and Ferri, 2003). Whilst home ownership did rise, that too was stratified by parental social class and income (Blanden and Machin, 2017). For the BCS cohort, having parents that were homeowners when they were aged 16 increases the probability of themselves being a homeowner at 42 by 116 per cent (ibid).

All this historic phenomenon has impacted the relative stability of youth transitions that are apparent for the NCDS cohort. The relative decline in individuals moving straight from school into work after mandatory schooling and the growth of difficult transitions and accumulating human capital via higher education suggests an increase in risk and uncertainty (Anders and Dorsett, 2017).

### Structural Barriers to successful transitions – the role of social-class and sex

**Social Class**

The BCS cohort experienced a stratified post-mandatory schooling experience. When it comes to participation in higher education, those from the most advantaged social origins background were more likely to attend higher education institutions comparative to less advantage individuals (Alcott, 2013). Prior academic attainment explains most of the variance in this stratified higher education participation (around 60%) (Alcott, 2013). With the growth of an ‘Emerging Adulthood’ and an elongated stay within education, participation in education for the BCS cohort has seen the gap between disadvantaged and privileged social origins get wider (Bynner, 2005). These apparent returns to education are stratified according to social class origins, with the advantages offered by certain qualifications differing according to class origins (Bukodi and Goldthorpe, 2011; Parsons, Green and Wiggins, 2016).

**Sex**

The experience for women within the 1970 cohort saw a continuing weakening of gender differences in processes of occupational attainment – a similar trend seen within the 1958 cohort (Bukodi, 2009), though the strength of education in this process appears to remain the same across cohorts (Bukodi and Goldthorpe, 2009). The weakening of gender differences is seen both at the educational level and occupational level in the form of take home income (Bynner, 2005). However, whilst the BCS cohort experienced a decline in gender segregated occupational sorting (Lekfuangfu and Lordan, 2022), those occupations that have the highest share of males maintained relatively high levels of segregation. Whilst it has been emphasised that social class origins have had an impact upon the BCS youth, the changing nature of the labour market has also had ramifications for men and women with regards to their biographical agency and their ability to find routes to stability and security (Schoon, Martin and Ross, 2007).

#### **Conclusion**

The BCS cohort is one that can be characterised by choice. Comparative to previous generations, that choice was much more numerous in the options presented to the BCS youth on what to do next after mandatory education had finished. The ‘Emerging Adult’ could theoretically choose any number of these options; however, the reality is that many of these options constrained the individual either immediately or down their life course. If the desired route from education was to find stable employment, the NCDS cohort would find that simply entering employment itself would provide a viable route to success. For the BCS cohort however, this was not strictly the case. Labour market restructuring and technological innovation on top of a major recession provided a much more complex, elongated transition to a stable occupation (Martin, Schoon and Ross, 2008), resulting in a ‘winding road’ school-to-work transition (Leuze, 2010). Entering employment immediately after mandatory education had the potential to lead to periods of unemployment due to a lack of skills in a new economic landscape (Bynner, 2005). Entering a government training program like the YTS would provide for some equally unsatisfactory results (ibid). Unemployment was a route that was even more restrictive than earlier cohorts due to the cutting off young people from benefits. Thus, the BCS cohort can be characterised as one of an educational turn. Staying within education, to both weather the recession storm and to pick up relevant and sometimes required qualifications was most likely the best option to lead to a stable and successful occupational career. Unfortunately, education – particularly post-mandatory education – was highly stratified. This stratified nature impacted the most privileged – by giving them advantages in the labour market, and the least privileged – by incurring further disadvantage. It should be assumed that, with this context, those individuals that entered education as a route post-mandatory schooling would thus be from more privileged backgrounds, perhaps in an even more striking ‘haves and have nots’ fashion than previously seen post 1944 Education Act reform.

## Data and Methods

Chapter two is a replication analysis of the models presented in chapter one using the NCDS. Therefore, similar to chapter one, the relationship between social origins and economic activity after mandatory schooling is examined used the large-scale, nationally representative data collected from the British Cohort Study. Educational attainment, housing tenure, sex, reading, and maths scores are also included in the model, as they were in the NCDS model. This is to assess choice and opportunity into different forms of economic activity: employment, education, training & apprenticeships, and unemployment & out of the labour force. BCS data is available using the UK Data Service.

Whilst this chapter has attempted its best to replicate the analysis in chapter one there are some substantive differences. Firstly, and most substantially, is that the outcome variable of economic activity after mandatory schooling in chapter two only has four categories to chapter ones five. Chapter two is missing a ‘post-education schooling’ category that encapsulates non-traditional forms of education that did not follow the traditional route of university. For the BCS cohort these non-traditional forms of education had decreased in popularity on top of not being appropriately recorded In the BCS survey. The second substantive change relates to the construction of social class measures (NS-SEC and RGSC) within chapter two. Whilst both chapter one and chapter two use occupational coding data from (Gregg, 2012), for the NCDS codes are only available for fathers of participants, for the BCS cohort both fathers and mothers are made available. Due to this both NS-SEC and RGSC are coded by using mothers’ occupational data to fill in any missing data entries from the father’s data. Other than these two differences, the model presented for analysis in chapter two is identical to that of chapter one. The reason for this is to start to build a historical picture of the changes and developments in choice and opportunities for different cohorts across different time periods.

As with chapter one, after an initial exploration of descriptive statistics multinominal logistic regression will be used to understand the choice and opportunities of BCS youth when it comes to economic activity post-mandatory schooling. After this initial model, a sensitivity analysis of social stratification measures will be employed to assess the most appropriate measure. Finally, an analysis of missing data involving multiple imputation will be conducted to assess the impact of missingness on the substantive findings of the model.

### Introduction to the BCS data

This chapter will use data from the British Cohort Study (University College London, 2022). The BCS70 began in 1970 with data originally collected on 17,198 babies born in England, Scotland, Wales, and Northern Ireland in the week of 5-11th April. This chapter will use data from participants up to the age of 30. Full cohort sweeps were gathered when participants were aged 5, 10, 16, 26, and 30 – with a subsample taken at 21 (Bynner, 2017). Originally, the data was collected by a mixture of medical records, and mother’s input. As the cohort member got older however they started to actively participate in answering some survey questions. Age 26 was the first time the cohort member took direct control of answering the survey itself. This was also a period of transition for the BCS, it typically relied on school records to keep in contact with its cohort members through their registered addresses but after the age 16 sweep when most left mandatory education a large amount of respondents were lost when it came time to contact them for the age 26 sweep (Elliott and Shepherd, 2006).

Alongside the standard sweeps that are detailed in the table below, two other aspects of the BCS are the 21-year-old sub-sample sweep and the BCS economic activity dataset. Both were considered supplementary to the full sample sweeps. The former consisted of a sub-sample of 10 per cent of the participants in the full sample and covered aspects such as economic activity since age 16. The BCS economic activity datasets sole focus was on creating a monthly economic activity record of participants since they left mandatory schooling up to 2016. The economic activity dataset was constructed using activity related data from sweeps 5-10 (Hancock and Peters, 2021). Barring the 21-subsample sweep, economic activity history on the type of activity individuals did post-mandatory schooling at age 16 was not collected until participants were aged 30 in sweep 6. The content covered in the age 30 sweep and the economic activity dataset for the time period of this analysis is virtually identical. The 21 subsample however does provide additional data that is missing in both the economic activity dataset and the sweep 6 dataset. Data was thus merged with the sweep 6 and the subsample to boost the overall sample size of the outcome variable of interest.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Total cohort | Dead | Permanent Emigrants | Non-Response | Other[[17]](#footnote-17) | Participants | (% of eligible sample)[[18]](#footnote-18) | Data Collected From |
| Birth – 1970 | 17287 | 0 | 0 | 716 | 0 | 16571 | 96 | Mother and Medical Records |
| Age 5 – 1975 | 16720 | 567 | 0 | 2812 | 360 | 12981 | 79 | Parents, medical records, cohort members |
| Age 10 – 1980 | 16700 | 587 | 0 | 1108 | 655 | 14350[[19]](#footnote-19) | 89 | Parents, school, tests, medical exam, cohort member |
| Age 16 – 1986 | 16690 | 597 | 0 | 3293 | 1594 | 11206 | 70 | Parents, school, tests, medical exam, cohort member |
| Age 26 – 1996[[20]](#footnote-20) | 16545 | 697 | 45 | 4765 | 2384 | 8654 | 55 | Cohort member |
| Age 30 – 2000 | 16253 | 747 | 287 | 1833 | 2553 | 10833 | 70 | Cohort member |

One of the first things to mention about the above table is that information on permanent emigrants prior to sweep 4 was not recorded, some have attempted to estimate this number in prior sweeps (Plewis, 2004) but it remains an estimation. Another to mention is that unlike the NCDS cohort where the number of participants has a steady decline as the sweeps go by, for the BCS there is a much more tumultuous story. The BCS went through many states of management and many states of how the data was exactly collected, this combined with the rather large period of 10 years from age 16 to age 26 has meant that attrition has been less that steady for the BCS cohort. This, even prior to analysis, suggests that missing data may present itself as a problem for the models to come.

#### Introduction to measures for subsequent analysis

The following section provides an overview of key variables used for subsequent multivariate analysis. Variable selection was a process that involved a combination of using the CLOSER search platform, digital codebooks, and manual searching of the individual BCS databases. This search was made much easier in chapter two compared to chapter one as the analytical variables in question were already chosen in chapter one – thus the goal of variable selection in chapter two was attempting to find the most appropriate similar measurements used in the BCS.

**Economic Activity**

As the main purpose of this chapter is to replicate the analysis of chapter one with a different cohort from a different period of time, the economic activity outcome variable has been selected in much the same way it was for chapter one. The variable of interest was economic activity data for individuals of month 201 – in other words when participants were 16 in September. As the BCS cohort’s mandatory schooling period ended when participants were 16, this was a natural month selection to measure economic activity. The month of September gives enough time for O-level results to be received and have any potential impact upon an individual’s economic activity circumstances.

Economic activity was recorded retrospectively in different sweeps. The first-time economic activity was recorded was in the 21-year-old sub-sweep that contained 10% of participants in the sample. Following this, economic activity was also recorded at age 30, as well as a special dedicated economic activity dataset that holds monthly economic activity data on participants up to present age – though this dataset does have a large amount of missingness on certain months. The economic activity variable is a combined outcome variable of the 21-sub-sample sweep and the age 30 sweep, to enhance responses. In areas that the two datasets overlap with differing values, priority is given to the 21-sub-sample sweep under the assumption that at age 21, someone would be more likely to accurately recall their economic activity in September when they were 16 compared to when they would be 30 years old.

The original economic activity variable in both the 21 sub-sample sweep and the 30-year-old sweep are very similar but have some slight variation. Below is a table providing descriptive statistics of each. Some recoding was required for both. In total, the economic activity variable has four outcomes: employment, education, training & apprenticeships, and unemployment & out of the labour force. Employment is defined as any individual that after mandatory education, entered employment. Education was coded as any individual that stayed within some form of schooling or education post-mandatory period, this for most would be individuals taking A-level examinations. The training & apprenticeships category is coded as all those training schemes that do not fall under education. Finally, the unemployment and out of the labour force category codes anyone that is either unemployed, out of the labour force, or otherwise economically inactive. The response level for this last category is very small and has the potential to adversely impact the statistical power of the model. Unfortunately, there is no viable alternative to this.

[insert raw economic activity variables]

**Educational Attainment**

The BCS cohort members reached compulsory school leaving age in 1986. The BCS cohort were the last group to experience the O-level/CSE split system (XXXX). By virtue of being the last cohort to experience this, the variable dictated as educational attainment is directly replicated from chapter one. The variable itself is a binary variable of number of O-level passes. Comparative to the NCDS cohort, a majority of the BCS cohort gained five or more O-level passes. The table below illustrates that proportionally, more of the BCS cohort were able to gain five or more O-level passes compared to the NCDS cohort.

[insert comparison of O-level passes for NCDS and BCS]

**Sex**

Sex is a variable taken from the birth sweep to ensure the largest possible number of responses. Sex measures the sex of the respondent in a binary male/female format. As seen with chapter one, sex as a variable played an important analytical role in understanding the structural inequalities and barriers that play an important role in choice and opportunity for youth. It’s inclusion in the BCS model is important not just for replication, but as seen in the literature review, the role of women and men in the labour market was still undergoing systematic changes.

**Housing Tenure**

Housing tenure was taken from when respondents were aged 10. There was information on tenure when respondents were aged 16 although the responses were scattered across a number of binary variables that overall had a low level of responses. The age 10 variable on housing tenure was a multiple category variable that has few overall missing cases. Housing tenure as a measure for inclusion in this model is the most important measure to focus upon. The arguments of Saunders (XXXX) and other new-structuralists were born from the time the BCS cohort were economically activity. Arguments related to the ‘death of class’ (XXXX) and the rise of tenure as the most substantive structural explanation for inequality is central to this chapter.

[insert comparison of Tenure]

#### **Social Stratification and Socio-Economic Background: NS-SEC, CAMSIS, RGSC**

A core component of this chapter, as in chapter one, is conducting a sensitivity analysis of social stratification measures. The continuation of a sensitivity analysis in chapter two provides a basis for comparison with chapter one. As seen in chapter one, each of the three models that used a different social stratification measure were relatively and substantively identical to one another. Chapter two seeks to understand if that is a unique phenomenon for the NCDS cohort or is it a pattern that replicates itself in different time periods also.

As mentioned prior, the RGSC and NS-SEC measures for chapter two are slightly different compared to their chapter one counterparts. Whilst the basis of each measure is upon the father’s social class position when the respondent was 10 years old, missing responses are filled in with the mother’s social class position when the respondent was 10 years old. This accomplishes three things. The first is that mother’s social class position is used to fill potential item missingness. The second is that it offers those respondents that do not come from a traditional nuclear family the ability to enter the model itself through taking the mother’s social class position where a father’s is not present. Finally, through both of these accomplishments, the level of missingness and overall responses is increased within the model, enhancing the statistical power of the model overall.

The overall patterns of social class position between the NCDS and BCS have not changed substantively. This lack of change is worth noting considering the relatively large-scale changes that the British economy and society underwent during the years 1958-1980 (XXXX). Below is a table comparing each social stratification measure between the NCDS and BCS cohort.

**Reading and Maths Scores**

Similar to the chapter one, reading and maths scores are included as a measure of prior educational attainment to capture the potential influence on prior educational attainment upon later life chances. The BCS provides these reading and maths scores at age 10.

[need to standardise the reading and maths scores so they are comparable with the NCDS measures]

## Descriptive Statistics

Table xxxx shows the frequencies and summary statistics for the BCS. Overall, 32.12 per cent of the sample is in employment. Whilst 44.75 per cent are in education – making up the plurality of the sample. When it comes to training & apprenticeships, 20.44 per cent of respondents are in this form of economic activity. Only 2.68 per cent of respondents are unemployed & out of the labour force – this is potentially the influence of the YTS.

When it comes to educational attainment, 31.56 per cent of individuals received less than five O-levels, with the reaming 68.44 per cent receiving five or more O-levels. Sex illustrates a slight overrepresentation of men (52.94 per cent) compared to women (47.06 per cent). When it comes to home ownership, 32.31 per cent of individuals grew up in a home that wasn’t owned by their parents compared to 67.69 per cent that did. The NS-SEC categories all see a relatively even distribution of respondents between 10-20 per cent except for ‘Intermediate occupations’ that has a very small 36 respondents – or 1.70 per cent. This will cause issues related to statistical power within the model and due to this category being split over four outcome categories the standard errors for this category are most likely going to be very high. The level of interpretation that will be able to be gained from this specific category within NS-SEC will be low. RGSC is much more unevenly distributed in comparison to NS-SEC. Skilled manual occupations make up 44.70 per cent of respondents, with professional and unskilled occupations making up 4.38 and 5.70 per cent respectively. Similar to comments made about NS-SECs intermediate occupations, the same can be equally said about RGSCs professional occupations. CAMSIS as a metric measure does not have these issues, it has a mean of 4.54. Given its metric nature, alongside the potentially problematic statistical power issues related to some of the categories within both NS-SEC and RGSC, if the sensitivity analysis concludes, similar to chapter one, that all three models are substantively identical it would thus be best to select CAMSIS for further study over NS-SEC and RGSC.

*Table 2 Descriptive Statistics for Economic Activity*

|  |  |  |
| --- | --- | --- |
|  | n | % |
| Economic Activity |  |  |
| *Employment* | 682 | 32.12% |
| *Education* | 950 | 44.75% |
| *Training/Apprenticeships* | 434 | 20.44% |
| *Unemployment and OLF* | 57 | 2.68% |
| Educational Attainment |  |  |
| *<5 O-Levels* | 670 | 31.56% |
| *5 or more O-Levels* | 1453 | 68.44% |
| Sex |  |  |
| *Female* | 999 | 47.06% |
| *Male* | 1124 | 52.94% |
| Housing Tenure |  |  |
| *Own Home* | 1437 | 67.69% |
| *Don't Own Home* | 686 | 32.31% |
| NS-SEC |  |  |
| *Higher managerial, administrative and professional occupations* | 411 | 19.36% |
| *Lower managerial, administrative and professional occupations* | 311 | 14.65% |
| *Intermediate occupations* | 36 | 1.70% |
| *Small employers and own account workers* | 222 | 10.46% |
| *Lower supervisory and technical occupations* | 376 | 17.71% |
| *Semi-routine occupations* | 325 | 15.31% |
| *Routine occupations* | 442 | 20.82% |
| RGSC |  |  |
| *Professional* | 93 | 4.38% |
| *Managerial and Technical* | 507 | 23.88% |
| *Skilled non-manual* | 184 | 8.67% |
| *Skilled manual* | 949 | 44.70% |
| *Partly skilled* | 269 | 12.67% |
| *Unskilled* | 121 | 5.70% |
|  |  |  |
|  | Mean | SD |
| CAMSIS | 4.54 | 1.34 |
| Reading Score | 41.04 | 11.77 |
| Maths Score | 44.35 | 11.34 |
|  |  |  |
| n | 2123 | |
| Data Source: BCS | | |

**Appendix**

Table A. 1 Goodness-of-fit summaries for explanatory variables and Economic Activity (CAMSIS)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome Variable: Economic Activity | Deviance | Deviance (from Null) | d.f. (from Null) | McFadden’s Adjusted Pseudo R2 | AIC | BIC |
| Null Model | 21656.47 | - | - | - | 21664.47 | 21692.38 |
| Null Model + Educational Attainment | 17732.61 | 3923.86 | 4 | 0.18 | 17748.61 | 17804.43 |
| Null Model + Sex | 20709.43 | 947.05 | 4 | 0.04 | 20725.43 | 20781.24 |
| Null Model + Tenure | 20920.88 | 735.59 | 4 | 0.03 | 20936.88 | 20992.70 |
| Null Model + CAMSIS | 20456.28 | 1200.19 | 4 | 0.06 | 20472.28 | 20528.10 |
| Null Model + Reading | 19584.87 | 2071.60 | 4 | 0.10 | 19600.87 | 19656.69 |
| Null Model + Maths | 19373.19 | 2283.28 | 4 | 0.11 | 19389.19 | 19445.00 |

Table A. 2 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity (CAMSIS)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome Variable: Economic Activity | Deviance | Deviance (from Previous) | d.f. (from Previous) | McFadden’s Adjusted Pseudo R2 | AIC | BIC |
| Null Model | 21656.47 | - | - | - | 21664.47 | 21692.38 |
| Null Model + Educational Attainment | 17732.61 | 3923.86 | 4 | 0.18 | 17748.61 | 17804.43 |
| Null Model + Educational Attainment + Sex | 16787.60 | 945.01 | 4 | 0.22 | 16811.60 | 16895.32 |
| Null Model + Educational Attainment + Sex + Tenure | 16522.23 | 215.32 | 4 | 0.23 | 16554.23 | 16665.82 |
| Null Model + Educational Attainment + Sex + Tenure + CAMSIS | 16341.72 | 180.51 | 4 | 0.24 | 16381.72 | 16521.25 |
| Null Model + Educational Attainment + Sex + Tenure + CAMSIS + Reading | 16118.78 | 222.94 | 4 | 0.25 | 16166.78 | 16334.22 |
| Null Model + Educational Attainment + Sex + Tenure + CAMSIS + Reading + Maths | 16050.372 | 68.408 | 4 | 0.26 | 16106.37 | 16301.72 |

Table A. 3 Goodness-of-fit summaries for explanatory variables and Economic Activity (RGSC)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome Variable: Economic Activity | Deviance | Deviance (from Null) | d.f. (from Null) | McFadden’s Adjusted Pseudo R2 | AIC | BIC |
| Null Model | 21656.47 | - | - | - | 21664.47 | 21692.38 |
| Null Model + Educational Attainment | 17732.61 | 3923.86 | 4 | 0.18 | 17748.61 | 17804.43 |
| Null Model + Sex | 20709.43 | 947.05 | 4 | 0.04 | 20725.43 | 20781.24 |
| Null Model + Tenure | 20920.88 | 735.59 | 4 | 0.03 | 20936.88 | 20992.70 |
| Null Model + RGSC | 20642.18 | 1014.29 | 20 | 0.05 | 20690.18 | 20857.62 |
| Null Model + Reading | 19584.87 | 2071.60 | 4 | 0.10 | 19600.87 | 19656.69 |
| Null Model + Maths | 19373.19 | 2283.28 | 4 | 0.11 | 19389.19 | 19445.00 |

Table A. 4 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity (RGSC)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome Variable: Economic Activity | Deviance | Deviance (from Previous) | d.f. (from Previous) | McFadden’s Adjusted Pseudo R2 | AIC | BIC |
| Null Model | 21656.47 | - | - | - | 21664.47 | 21692.38 |
| Null Model + Educational Attainment | 17732.61 | 3923.86 | 4 | 0.18 | 17748.61 | 17804.43 |
| Null Model + Educational Attainment + Sex | 16787.60 | 945.01 | 4 | 0.22 | 16811.60 | 16895.32 |
| Null Model + Educational Attainment + Sex + Tenure | 16522.23 | 215.32 | 4 | 0.23 | 16554.23 | 16665.82 |
| Null Model + Educational Attainment + Sex + Tenure + RGSC | 16366.81 | 155.42 | 16 | 0.24 | 16438.81 | 16689.96 |
| Null Model + Educational Attainment + Sex + Tenure + RGSC + Reading | 16131.84 | 234.97 | 4 | 0.25 | 16211.84 | 16490.90 |
| Null Model + Educational Attainment + Sex + Tenure + RGSC + Reading + Maths | 16063.18 | 68.66 | 4 | 0.25 | 16151.18 | 16458.14 |

Table A. 5 Average marginal effects on the probability of Economic Activity

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | NS-SEC | | CAMSIS | | RGSC | |
| Economic Activity | **Prob.** | **S.E** | **Coef.** | **S.E** | **Coef.** | **S.E** |
| Employment | (.) | (.) | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |
| Post-Schooling Education |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | 0.03 | (0.01) | 0.03 | (0.01) | 0.03 | (0.01) |
| Sex |  |  |  |  |  |  |
| *Female* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Male* | -0.07 | (0.01) | -0.07 | (0.01) | -0.07 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Don't Own Home* | -0.03 | (0.01) | -0.03 | (0.01) | -0.03 | (0.01) |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Lower Managerial* | -0.01 | (0.01) |  |  |  |  |
| *Intermediate Occupations* | -0.01 | (0.02) |  |  |  |  |
| *Small Employers* | -0.04 | (0.01) |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -0.02 | (0.01) |  |  |  |  |
| *Semi-Routine Occupations* | -0.04 | (0.01) |  |  |  |  |
| *Routine Occupations* | -0.03 | (0.01) |  |  |  |  |
| CAMSIS |  |  | 0.01 | (0.00) |  |  |
| RGSC |  |  |  |  |  |  |
| *Professional* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Managerial and Technical* |  |  |  |  | -0.01 | (0.02) |
| *Skilled Non-Manual* |  |  |  |  | 0.00 | (0.02) |
| *Skilled Manual* |  |  |  |  | -0.02 | (0.02) |
| *Partly Skilled* |  |  |  |  | -0.03 | (0.02) |
| *Unskilled* |  |  |  |  | -0.05 | (0.02) |
| Reading Scores | 0.00 | (0.00) | 0.00 | (0.00) | -0.00 | (0.00) |
| Maths Scores | -0.00 | (0.00) | -0.00 | (0.00) | -0.00 | (0.00) |
|  |  |  |  |  |  |  |
| School |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | 0.44 | (0.01) | 0.44 | (0.01) | 0.45 | (0.01) |
| Sex |  |  |  |  |  |  |
| *Female* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Male* | -0.00 | (0.01) | 0.00 | (0.01) | -0.00 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Don't Own Home* | -0.03 | (0.01) | -0.03 | (0.01) | -0.04 | (0.01) |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Lower Managerial* | -0.05 | (0.01) |  |  |  |  |
| *Intermediate Occupations* | -0.11 | (0.03) |  |  |  |  |
| *Small Employers* | -0.11 | (0.02) |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -0.10 | (0.01) |  |  |  |  |
| *Semi-Routine Occupations* | -0.09 | (0.02) |  |  |  |  |
| *Routine Occupations* | -0.13 | (0.01) |  |  |  |  |
| CAMSIS |  |  | 0.03 | (0.00) |  |  |
| RGSC |  |  |  |  |  |  |
| *Professional* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Managerial and Technical* |  |  |  |  | -0.09 | (0.02) |
| *Skilled Non-Manual* |  |  |  |  | -0.10 | (0.03) |
| *Skilled Manual* |  |  |  |  | -0.15 | (0.02) |
| *Partly Skilled* |  |  |  |  | -0.16 | (0.03) |
| *Unskilled* |  |  |  |  | -0.16 | (0.03) |
| Reading Scores | 0.01 | (0.00) | 0.01 | (0.00) | 0.01 | (0.00) |
| Maths Scores | 0.00 | (0.00) | 0.00 | (0.00) | 0.00 | (0.00) |
|  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | -0.16 | (0.01) | -0.16 | (0.01) | -0.16 | (0.01) |
| Sex |  |  |  |  |  |  |
| *Female* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Male* | 0.25 | (0.01) | 0.25 | (0.01) | 0.25 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Don't Own Home* | -0.01 | (0.01) | -0.01 | (0.01) | -0.01 | (0.01) |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Lower Managerial* | -0.03 | (0.02) |  |  |  |  |
| *Intermediate Occupations* | 0.00 | (0.03) |  |  |  |  |
| *Small Employers* | 0.04 | (0.02) |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | 0.03 | (0.02) |  |  |  |  |
| *Semi-Routine Occupations* | 0.02 | (0.02) |  |  |  |  |
| *Routine Occupations* | 0.01 | (0.02) |  |  |  |  |
| CAMSIS |  |  | -0.01 | (0.00) |  |  |
| RGSC |  |  |  |  |  |  |
| *Professional* |  |  |  |  |  |  |
| *Managerial and Technical* |  |  |  |  | -0.02 | (0.03) |
| *Skilled Non-Manual* |  |  |  |  | -0.00 | (0.03) |
| *Skilled Manual* |  |  |  |  | 0.02 | (0.03) |
| *Partly Skilled* |  |  |  |  | 0.02 | (0.03) |
| *Unskilled* |  |  |  |  | -0.02 | (0.03) |
| Reading Scores | -0.00 | (0.00) | -0.00 | (0.00) | -0.00 | (0.00) |
| Maths Scores | 0.00 | (0.00) | 0.00 | (0.00) | 0.00 | (0.00) |
|  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |
| *Less than five O-levels* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Five or More O-levels* | -0.02 | (0.01) | -0.02 | (0.01) | -0.02 | (0.01) |
| Sex |  |  |  |  |  |  |
| *Female* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Male* | -0.01 | (0.00) | -0.01 | (0.00) | -0.01 | (0.00) |
| Housing Tenure |  |  |  |  |  |  |
| *Own Home* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Don't Own Home* | 0.01 | (0.00) | 0.01 | (0.00) | 0.01 | (0.00) |
| NS-SEC |  |  |  |  |  |  |
| *Higher Managerial* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Lower Managerial* | -0.01 | (0.01) |  |  |  |  |
| *Intermediate Occupations* | -0.00 | (0.02) |  |  |  |  |
| *Small Employers* | -0.01 | (0.01) |  |  |  |  |
| *Lower Supervisory & Technical Occupations* | -0.01 | (0.01) |  |  |  |  |
| *Semi-Routine Occupations* | 0.00 | (0.01) |  |  |  |  |
| *Routine Occupations* | 0.00 | (0.01) |  |  |  |  |
| CAMSIS |  |  | -0.00 | (0.00) |  |  |
| RGSC |  |  |  |  |  |  |
| *Professional* | (.) | (.) | (.) | (.) | (.) | (.) |
| *Managerial and Technical* |  |  |  |  | 0.03 | (0.01) |
| *Skilled Non-Manual* |  |  |  |  | 0.01 | (0.01) |
| *Skilled Manual* |  |  |  |  | 0.02 | (0.01) |
| *Partly Skilled* |  |  |  |  | 0.03 | (0.01) |
| *Unskilled* |  |  |  |  | 0.03 | (0.01) |
| Reading Scores | -0.00 | (0.00) | -0.00 | (0.00) | -0.00 | (0.00) |
| Maths Scores | -0.00 | (0.00) | -0.00 | (0.00) | -0.00 | (0.00) |
|  |  |  |  |  |  |  |
| Number of observations | 7915 | | | | | |

# **Data Citation**

## **Bibliography**

Alcott, B. (2013) ‘Predicting departure from British education: Identifying those most at risk through discrete time hazard modelling’, *Widening Participation and Lifelong Learning*, 15(4), pp. 46–64. Available at: https://doi.org/10.5456/WPLL.15.4.46.

Allison, P. (2012a) ‘312-2012: Handling Missing Data by Maximum Likelihood’.

Allison, P. (2012b) ‘Why Maximum Likelihood is Better Than Multiple Imputation’, *Statistical Horizons*, 9 July. Available at: https://statisticalhorizons.com/ml-better-than-mi/ (Accessed: 15 May 2023).

Allison, P. (2015) ‘Maximum Likelihood is Better than Multiple Imputation: Part II’, *Statistical Horizons*, 5 May. Available at: https://statisticalhorizons.com/ml-is-better-than-mi/ (Accessed: 15 May 2023).

Anders, J. and Dorsett, R. (2017) ‘What young English people do once they reach school-leaving age: A cross-cohort comparison for the last 30 years’, *Longitudinal and Life Course Studies*, 8(1). Available at: https://doi.org/10.14301/llcs.v8i1.399.

‘Annual report for the 1911 of the registrar-general’ (1913), 182(4708), pp. 1491–1492. Available at: https://doi.org/10.1016/S0140-6736(01)78008-7.

Arulampalam, W. and Booth, A.L. (1997) ‘Who gets over the training hurdle? A study of the training experiences of young men and women in Britain’, *Journal of Population Economics*, 10(2), pp. 197–217. Available at: https://doi.org/10.1007/s001480050038.

Arulampalam, W.N. and Booth, A.L. (2001) ‘Learning and Earning: Do Multiple Training Events Pay? A Decade of Evidence from a Cohort of Young British Men’, *Economica*, 68(271), pp. 379–400. Available at: https://doi.org/10.1111/1468-0335.00252.

Beck, U. (2002) *Individualisation: Institutionalized Individualism and its Social and Political Consequences*. SAGE Publications.

Beck, U. (2014) *The brave new world of work*. John Wiley & Sons.

Beck, U., Giddens, A. and Lash, S. (1994) *Reflexive modernization: Politics, tradition and aesthetics in the modern social order*. Stanford University Press.

Bergman, M.M. and Joye, D. (2001) ‘Comparing Social Stratification Schemas: CAMSIS, CSP-CH, Goldthorpe, ISCO-88, Treiman, and Wright’, *Cambridge studies in Social research*, p. 53.

Bernardi, L., Huinink, J. and Settersten, R.A. (2019) ‘The life course cube: A tool for studying lives’, *Advances in Life Course Research*, 41, p. 100258. Available at: https://doi.org/10.1016/j.alcr.2018.11.004.

Blanchflower, D. and Lynch, L. (1992) *Training at Work: A Comparison of U.S. and British Youths*. w4037. Cambridge, MA: National Bureau of Economic Research, p. w4037. Available at: https://doi.org/10.3386/w4037.

Blanden, J. and Machin, S. (2017) ‘Home Ownership and Social Mobility’, *CEP Discussion Paper* [Preprint].

Blanden, J. and Macmillan, L. (2014) ‘Education and Intergenerational Mobility: Help or Hindrance?’, *Centre for Analysis of Social Exclusion* [Preprint].

Blundell, R. *et al.* (2000) ‘The Returns to Higher Education in Britain: Evidence From a British Cohort’, *The Economic Journal*, 110(461), pp. F82–F99. Available at: https://doi.org/10.1111/1468-0297.00508.

Blundell, R., Dearden, L. and Sianesi, B. (2001) ‘Estimating the Returns to Education: Models, Methods and Results’, *Centre for the Economics of Education* [Preprint].

Boero, G. *et al.* (2020) ‘HOW DOES THE RETURN TO A DEGREE VARY BY CLASS OF AWARD?’, *Higher Education Statistics Agency* [Preprint].

Booth, A.L. and Satchell, S.E. (1994) ‘APPRENTICESHIPS AND JOB TENURE’, *Oxford Economic Papers*, 46(4), pp. 676–695. Available at: https://doi.org/10.1093/oxfordjournals.oep.a042153.

Breen, R. (2022) ‘The stubborn persistence of educational inequality’, *IFS Deaton Review* [Preprint].

Brunello, G. and Rocco, L. (2017) ‘The Labor Market Effects of Academic and Vocational Education over the Life Cycle: Evidence Based on a British Cohort’, *Journal of Human Capital*, 11(1), pp. 106–166. Available at: https://doi.org/10.1086/690234.

Bukodi, E. (2009) ‘Education, First Occupation and Later Occupational Attainment: Cross-cohort Changes among Men and Women in Britain’, *CLS Cohort Studies*, 4.

Bukodi, E. *et al.* (2017) ‘Why have relative rates of class mobility become more equal among women in Britain?: Rates of class mobility’, *The British Journal of Sociology*, 68(3), pp. 512–532. Available at: https://doi.org/10.1111/1468-4446.12274.

Bukodi, E., Bourne, M. and Betthäuser, B. (2017) ‘Wastage of talent?’, *Advances in Life Course Research*, 34, pp. 34–42. Available at: https://doi.org/10.1016/j.alcr.2017.09.003.

Bukodi, E. and Goldthorpe, J.H. (2009) ‘Class Origins, Education and Occupational Attainment: Cross-cohort Changes among Men in Britain’, *CLS Cohort Studies*, 3.

Bukodi, E. and Goldthorpe, J.H. (2011) ‘Social class returns to higher education: chances of access to the professional and managerial salariat for men in three British birth cohorts’, *Longitudinal and Life Course Studies*, 2(2). Available at: https://doi.org/10.14301/llcs.v2i2.122.

Bukodi, E., Goldthorpe, J.H. and Kuha, J. (2017) ‘The pattern of social fluidity within the British class structure: a topological model’, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(3), pp. 841–862. Available at: https://doi.org/10.1111/rssa.12234.

Bynner, J. (1998a) ‘Britain’s birth cohort studies: their use in the study of children’, *Children & Society*, 12(5), pp. 390–395. Available at: https://doi.org/10.1111/j.1099-0860.1998.tb00095.x.

Bynner, J. (1998b) ‘Education and Family Components of Identity in the Transition from School to Work’, *International Journal of Behavioral Development*, 22(1), pp. 29–53. Available at: https://doi.org/10.1080/016502598384504.

Bynner, J. *et al.* (2002) ‘Young people’s changing routes to independence’, *Joseph Rowntree Foundation* [Preprint].

Bynner, J. (2005) ‘Rethinking the Youth Phase of the Life-course: The Case for Emerging Adulthood?’, *Journal of Youth Studies*, 8(4), pp. 367–384. Available at: https://doi.org/10.1080/13676260500431628.

Bynner, J. (2012) ‘Policy Reflections Guided by Longitudinal Study, Youth Training, Social Exclusion, and More Recently Neet’, *British Journal of Educational Studies*, 60(1), pp. 39–52. Available at: https://doi.org/10.1080/00071005.2011.650943.

Bynner, J. (2017) ‘1970 British Cohort Study (BCS70) Twenty one-year Sample Survey’, *CLS Cohort Studies* [Preprint].

Bynner, J. and Ferri, E. (2003) *Changing Britain, Changing Lives*. Institute of Education Press.

Bynner, J., Ferri, E. and Shepherd, P. (2019) *Twenty-something in the 1990s: Getting on, getting by, getting nowhere*. Routledge.

Bynner, J. and Joshi, H. (2002) ‘Equality and Opportunity in Education: Evidence from the 1958 and 1970 birth cohort studies’, *Oxford Review of Education*, 28(4), pp. 405–425. Available at: https://doi.org/10.1080/0305498022000013599.

Bynner, J. and Joshi, H. (2007) ‘BUILDING THE EVIDENCE BASE FROM LONGITUDINAL DATA: The aims, content and achievements of the British birth cohort studies’, *Innovation: The European Journal of Social Science Research*, 20(2), pp. 159–179. Available at: https://doi.org/10.1080/13511610701502255.

Bynner, J. and Parsons, S. (2000) ‘Marginalization and Value Shifts under the Changing Economic Circumstances Surrounding the Transition to Work: A Comparison of Cohorts Born in 1958 and 1970’, *Journal of Youth Studies*, 3(3), pp. 237–249. Available at: https://doi.org/10.1080/713684379.

Bynner, J., Wiggins, R. and Parsons, S. (1996) ‘AN EXPLORATORY COMPARATIVE ANALYSIS OF DATA COLLECTED IN THE 1958 AND 1970 BRITISH BIRTH COHORT STUDIES’:, *Conference of the International Sociological Association* [Preprint].

Carpenter, J.R. and Kenward, M. (2012) *Multiple imputation and its application*. John Wiley & Sons.

Cebulla, A. and Tomaszewski, W. (2013) ‘The demise of certainty: shifts in aspirations and achievement at the turn of the century’, *International Journal of Adolescence and Youth*, 18(3), pp. 141–157. Available at: https://doi.org/10.1080/02673843.2013.767743.

Chevalier, A. and Lanot, G. (2001) ‘The Relative Effect of Family and Financial Characteristics on Educational Achievement’, p. 32.

Conlon, G. (2001) *The differential in earnings premia between academically and vocationally trained males in the United Kingdom*. London: Centre for the Economics of Education.

Connelly, R., Gayle, V. and Lambert, P.S. (2016) ‘A review of educational attainment measures for social survey research’, *Methodological Innovations*, 9. Available at: https://doi.org/10.1177/2059799116638001.

Connolly, S. and Gregory, M. (2010) ‘Dual tracks: part-time work in life-cycle employment for British women’, *Journal of Population Economics*, 23(3), pp. 907–931. Available at: https://doi.org/10.1007/s00148-009-0249-4.

Connolly, S., Micklewright, J. and Nickell, S. (1992) ‘THE OCCUPATIONAL SUCCESS OF YOUNG MEN WHO LEFT SCHOOL AT SIXTEEN \*’, *Oxford Economic Papers*, 44(3), pp. 460–479. Available at: https://doi.org/10.1093/oxfordjournals.oep.a042058.

Devine, F. (2017) ‘The “new structuralism”: class politics and class analysis’, in *Social Class and Marxism*. Taylor & Francis.

Dex, S. *et al.* (1998) ‘Women’s Employment Transitions Around Childbearing’, *Oxford Bulletin of Economics and Statistics*, 60(1), pp. 79–98. Available at: https://doi.org/10.1111/1468-0084.00087.

Dex, S. and Bukodi, E. (2012) ‘The Effects of Part-Time Work on Women’s Occupational Mobility in Britain: Evidence from the 1958 Birth Cohort Study’, *National Institute Economic Review*, 222, pp. R20–R37. Available at: https://doi.org/10.1177/002795011222200103.

Dex, S., Ward, K. and Joshi, H. (2008) ‘Gender differences in occupational wage mobility in the 1958 cohort’, *Work, Employment and Society*, 22(2), pp. 263–280. Available at: https://doi.org/10.1177/0950017008089104.

Di Salvo, P. and Ermisch, J. (1997) ‘Analysis of the Dynamics of Housing Tenure Choice in Britain’, *Journal of Urban Economics*, 42(1), pp. 1–17. Available at: https://doi.org/10.1006/juec.1996.2009.

Diewald, M. and Mayer, K.U. (2008) ‘The sociology of the life course and life span psychology: integrated paradigm or complementing pathways?’, p. 24.

Dolton, P., Joshi, H. and Makepeace, G. (2002) ‘Unpacking Unequal Pay Between Men and Women Across Cohort and Lifecycle’, *CLS Cohort Studies* [Preprint].

Dolton, P., Makepeace, G. and Marcenaro‐Gutierrez, O.D. (2005) ‘Career progression: Getting‐on, getting‐by and going nowhere’, *Education Economics*, 13(2), pp. 237–255. Available at: https://doi.org/10.1080/09645290500031447.

Dorsett, R. and Lucchino, P. (2013) ‘Visualising the school-to-work transition: an analysis using optimal matching’, *Journal of Social Policy* [Preprint].

Droy, L., Goodwin, J. and O’connor, H. (2019) ‘Liminality, Marginalisation and Low-Skilled Work: Mapping long-term labour market difficulty following participation in the 1980s government-sponsored youth training schemes (YTS)’, *Occasional Papers* [Preprint]. Available at: https://doi.org/10.13140/RG.2.2.28494.92486.

Duckworth, K. and Schoon, I. (2012) ‘Beating the Odds: Exploring the Impact of Social Risk on Young People’s School-to-Work Transitions during Recession in the UK’, *National Institute Economic Review*, 222, pp. R38–R51. Available at: https://doi.org/10.1177/002795011222200104.

Dustmann, C. *et al.* (1996) ‘Earning and Learning: Educational Policy and the Growth of Part-Time Wurk by Full-Time Pupils’, *The journal of applied public economics* [Preprint].

Dustmann, C., Rajah, N. and Smith, S. (1997) ‘Teenage truancy, part-time working and wages’, *Journal of Population Economics* [Preprint].

Duta, A., Iannelli, C. and Breen, R. (2021) ‘Social inequalities in attaining higher education in Scotland: New evidence from sibling data’, *British Educational Research Journal*, 47(5), pp. 1281–1302. Available at: https://doi.org/10.1002/berj.3725.

Elder, G., Johnson, M. and Crosnoe, R. (2003) ‘The emergence and development of life course theory’, in *Handbook of the Lifecourse*. Springer.

Elder, G.H. (1994) ‘Time, Human Agency, and Social Change: Perspectives on the Life Course’, *Social Psychology Quarterly*, 57(1), p. 4. Available at: https://doi.org/10.2307/2786971.

Elliott, J. and Shepherd, P. (2006) ‘Cohort Profile: 1970 British Birth Cohort (BCS70)’, *International Journal of Epidemiology*, 35(4), pp. 836–843. Available at: https://doi.org/10.1093/ije/dyl174.

Evans, K. (2007) ‘Concepts of bounded agency in education, work, and the personal lives of young adults’, *International Journal of Psychology*, 42(2), pp. 85–93. Available at: https://doi.org/10.1080/00207590600991237.

Farrall, S., Gray, E. and Mike Jones, P. (2020) ‘The Role of Radical Economic Restructuring in Truancy from School and Engagement in Crime’, *The British Journal of Criminology*, 60(1), pp. 118–140. Available at: https://doi.org/10.1093/bjc/azz040.

Feinstein, L., Duckworth, K. and Sabates, R. (2004) *A model of the inter-generational transmission of educational success*. London: Centre for Research on the Wider Benefits of Learning, Institute of Education.

Field, F. (2010) *The Foundation Years: preventing poor children becoming poor adults, The report of the Independent Review on Poverty and Life Chances*. The report of the Independent Review on Poverty and Life Chances.

Field, J.A. (2011) ‘Researching the benefits of learning: the persuasive power of longitudinal studies’, *London Review of Education* [Preprint]. Available at: https://doi.org/10.1080/14748460.2011.616320.

Firth, D. (2000) ‘QV Calculator : Quasi-variances in Xlisp-Stat and on the Web’, *Journal of Statistical Software* [Preprint].

Firth, D. (2003) ‘Overcoming the Reference Category Problem in the Presentation of Statistical Models’, *Sociological Methodology*, 33(1), pp. 1–18. Available at: https://doi.org/10.1111/j.0081-1750.2003.t01-1-00125.x.

Furnham, A. and Cheng, H. (2013) ‘Factors influencing adult earnings: Findings from a nationally representative sample’, *The Journal of Socio-Economics*, 44, pp. 120–125. Available at: https://doi.org/10.1016/j.socec.2013.02.008.

Furnham, A. and Cheng, H. (2019) ‘Factors influencing adult savings and investment: Findings from a nationally representative sample’, *Personality and Individual Differences*, 151, p. 109510. Available at: https://doi.org/10.1016/j.paid.2019.109510.

Galindo-Rueda, F. (2003) ‘Employer Learning and Schooling-Related Statistical Discrimination in Britain’, *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.412483.

Gayle, V., Lambert, P. and Murray, S. (2009) ‘School-to-Work in the 1990s: Modelling Transitions with Large-Scale Datasets’, in R. Brooks (ed.) *Transitions from Education to Work*. London: Palgrave Macmillan UK, pp. 17–41. Available at: https://doi.org/10.1057/9780230235403\_2.

Gayle, V. and Lambert, P.S. (2009) ‘Logistic Regression Models in Sociological Research’, *DAMES Node, Technical Paper* [Preprint].

Giddens, A. *et al.* (1991) *Introduction to Sociology*. New York: Norton.

Goldthorpe, J.H. (1998) ‘Rational Action Theory for Sociology’, *The British Journal of Sociology*, 49(2), p. 167. Available at: https://doi.org/10.2307/591308.

Goldthorpe, J.H. and Marshall, G. (1992) ‘The promising future of class analysis: a response to recent critiques’.

Goodwin, J. *et al.* (2020) ‘Returning to YTS: the long-term impact of youth training scheme participation’, *Journal of Youth Studies*, 23(1), pp. 28–43. Available at: https://doi.org/10.1080/13676261.2019.1710484.

Goodwin, J. and O’Connor, H. (2005) ‘Exploring Complex Transitions: Looking Back at the “Golden Age” of From School to Work’, *Sociology*, 39(2), pp. 201–220. Available at: https://doi.org/10.1177/0038038505050535.

Gregg, P. (2001) ‘The Impact of Youth Unemployment on Adult Unemployment in the NCDS’, *The Economic Journal*, 111(475), pp. F626–F653. Available at: https://doi.org/10.1111/1468-0297.00666.

Gregg, P. (2012) ‘Occupational Coding for the National Child Development Study (1969, 1991-2008) and the 1970 British Cohort Study (1980, 2000-2008).’, *CLS Cohort Studies* [Preprint]. Available at: https://doi.org/10.5255/UKDA-SN-7023-1.

Gregg, P. and Tominey, E. (2005) ‘The wage scar from male youth unemployment’, *Labour Economics*, 12(4), pp. 487–509. Available at: https://doi.org/10.1016/j.labeco.2005.05.004.

Guinea-Martin, D. and Elliott, J. (2008) ‘Economic position and occupational segregation in the 1990s: A comparison of the ONS Longitudinal Study and the 1958 National Child Development Study’, *CLS Cohort Studies* [Preprint].

Hamnett, C., McDowell, L. and Sarre, P. (1989) *Restructuring Britain: The changing social structure*. SAGE.

Hancock, M. and Peters, A. (2021) ‘1970 British Cohort Study, Activity Histories (1986 - 2016)’, *UCL Centre for Longitudinal Studies* [Preprint].

Hawkes, D. and Plewis, I. (2006) ‘Modelling non-response in the National Child Development Study’, *Journal of the Royal Statistical Society*, 169(3), pp. 479–491. Available at: https://doi.org/10.1111/j.1467-985X.2006.00401.x.

Hitlin, S. and Elder, G.H. (2007) ‘Time, Self, and the Curiously Abstract Concept of Agency\*’, *Sociological Theory*, 25(2), pp. 170–191. Available at: https://doi.org/10.1111/j.1467-9558.2007.00303.x.

Hitlin, S. and Johnson, M.K. (2015) ‘Reconceptualizing Agency within the Life Course: The Power of Looking Ahead.’, *American Journal of Sociology*, 120(5), pp. 1429–1472. Available at: https://doi.org/10.1086/681216.

Holm, A. and Jæger, M.M. (2011) ‘Dealing with selection bias in educational transition models: The bivariate probit selection model’, *Research in Social Stratification and Mobility*, 29(3), pp. 311–322. Available at: https://doi.org/10.1016/j.rssm.2011.02.002.

Hutchison, D., Prosser, H. and Wedge, P. (1979) ‘The Prediction of Educational Failure’, *Educational Studies*, 5(1), pp. 73–82. Available at: https://doi.org/10.1080/0305569790050109.

Johnson, D. (2002) ‘Lifetime Earnings, Discount Rate, Ability and the Demand for Post-compulsory Education in Men in England and Wales’, *Bulletin of Economic Research*, 54(3), pp. 233–247. Available at: https://doi.org/10.1111/1467-8586.00151.

Jones, G.E. (1986) ‘Youth in the social structure: transitions to adulthood and their stratification by class and gender.’, *University of Surrey* [Preprint].

Joshi, H. (2002) ‘Production, Reproduction, and Education: Women, Children, and Work in a British Perspective’, *Population and Development Review*, 28(3).

Joshi, H. and Davies, H. (1996) ‘Financial dependency on men: Have women born in 1958 broken free?’, *Policy Studies*, 17(1), pp. 35–54. Available at: https://doi.org/10.1080/01442879608423692.

Joshi, H., Makepeace, G. and Dolton, P. (2007) ‘More or Less Unequal? Evidence on the Pay of Men and Women from the British Birth Cohort Studies’, *Gender, Work & Organization*, 14(1), pp. 37–55. Available at: https://doi.org/10.1111/j.1468-0432.2007.00331.x.

Lambert, P. and Barnett, C. (2021) ‘Optimising the use of measures of social stratification in research with intersectional and longitudinal analytical priorities’, in Nico, M. and Pollock, G., *The Routledge Handbook of Contemporary Inequalities and the Life Course*. 1st edn. London: Routledge, pp. 188–198. Available at: https://doi.org/10.4324/9780429470059-18.

Lambert, P.S. *et al.* (2008) ‘The importance of specificity in occupation‐based social classifications’, *International Journal of Sociology and Social Policy*. Edited by R.M. Blackburn, 28(5/6), pp. 179–192. Available at: https://doi.org/10.1108/01443330810881231.

Lekfuangfu, W.N. and Lordan, G. (2022) ‘Documenting occupational sorting by gender in the UK across three cohorts: does a grand convergence rely on societal movements?’, *Empirical Economics* [Preprint]. Available at: https://doi.org/10.1007/s00181-022-02314-5.

Leuze, K. (2010) *Smooth Path or Long and Winding Road? How Institutions Shape the Transition from Higher Education to Work*. Budrich UniPress. Available at: https://doi.org/10.3224/94075542.

Lindley, R.M. (1996) ‘The school-to-work transition in the United Kingdom’, *International Labour Review*, p. 23.

Machin, S. and Vignoles, A. (2005) ‘Educational inequality: the widening socio-economic gap’, *Fiscal Studies*, 25(2), pp. 107–128. Available at: https://doi.org/10.1111/j.1475-5890.2004.tb00099.x.

Maclure, S. (1978) *Education and Youth Employment in Great Britain*. ERIC.

Makepeace, G. *et al.* (1999) ‘How Unequally Has Equal Pay Progressed since the 1970s? A Study of Two British Cohorts’, *The Journal of Human Resources*, 34(3), p. 534. Available at: https://doi.org/10.2307/146379.

Makepeace, G., Dolton, P. and Joshi, H. (2004) ‘Gender earnings differentials across individuals over time in British cohort studies’, *International Journal of Manpower*, 25(3/4), pp. 251–263. Available at: https://doi.org/10.1108/01437720410541380.

Martin, P., Schoon, I. and Ross, A. (2008) ‘Beyond Transitions: Applying Optimal Matching Analysis to Life Course Research’, *International Journal of Social Research Methodology*, 11(3), pp. 179–199. Available at: https://doi.org/10.1080/13645570701622025.

Mayer, K.U. (2004) ‘Whose Lives? How History, Societies, and Institutions Define and Shape Life Courses’, *Research in Human Development*, 1(3), pp. 161–187. Available at: https://doi.org/10.1207/s15427617rhd0103\_3.

Mayer, K.U. (2009) ‘New Directions in Life Course Research’, *Annual Review of Sociology*, 35(1), pp. 413–433. Available at: https://doi.org/10.1146/annurev.soc.34.040507.134619.

Mayer, K.U. and Schoepflin, U. (2022) ‘The State and the Life Course’, p. 24.

Micklewright, J. (1989) ‘Choice at Sixteen’, *Economica*, 56(221), pp. 25–39. Available at: https://doi.org/10.2307/2554492.

Murray, S.J. (2011) ‘Growing up in the 1990s: Tracks and trajectories of the “Rising 16’s”: A longitudinal analysis using the British Household Panel Survey.’, p. 354.

Neuburger, J. (2010) ‘Trends in the Unequal Pay of Women and Men Across Three British Generations’, *PhD thesis* [Preprint].

Neuburger, J., Kuh, D. and Joshi, H. (2009) ‘Trends in the relative wage opportunities of women and men across three British generations’, *CLS Cohort Studies* [Preprint].

Neyt, B. *et al.* (2018) ‘Does Student Work Really Affect Educational Outcomes? A Review of the Literature’, *Journal of Economic Surveys* [Preprint].

Norton, E.C. and Dowd, B.E. (2018) ‘Log Odds and the Interpretation of Logit Models’, *Health Services Research*, 53(2), pp. 859–878. Available at: https://doi.org/10.1111/1475-6773.12712.

ONS (2023) *Unemployment rate*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/mgsx/lms (Accessed: 1 May 2023).

Paci, P. *et al.* (1995) ‘Is pay discrimination against young women a thing of the past? Atale of two cohorts’, *International Journal of Manpower*, 16(2), pp. 60–65. Available at: https://doi.org/10.1108/01437729510085765.

Parsons, S. (2013) ‘Mass Observation Archive: How to combine information with the British birth cohort studies’, *Discovery UCL* [Preprint].

Parsons, S., Green, F. and Wiggins, D. (2016) ‘Higher Education and Occupational Returns: do returns vary according to students’ social origins?’, *Centre for Longitudinal Studies* [Preprint].

Patacchini, E. and Zenou, Y. (2006) ‘The Dynamics of the Racial Test Score Gap During the School Years in Britain’, *Research Institute of Industrial Economics* [Preprint].

Payne, J. (1987) ‘Unemployment, Apprenticeships and Training: does it pay to stay on at school?’, *British Journal of Sociology of Education*, 8(4), pp. 425–445. Available at: https://doi.org/10.1080/0142569870080405.

Payne, J. (1995) ‘Routes beyond compulsory schooling’, *Policy Studies Institute*, p. 98.

Pearson qualifications (2023a) *About CSEs*. Available at: https://qualifications.pearson.com/en/support/support-topics/understanding-our-qualifications/our-qualifications-explained/about-cses.html (Accessed: 9 May 2023).

Pearson qualifications (2023b) *About O levels*. Available at: https://qualifications.pearson.com/en/support/support-topics/understanding-our-qualifications/our-qualifications-explained/about-o-levels.html (Accessed: 8 May 2023).

Plewis, I. (2004) *National Child Development Study and 1970 British Cohort Study technical report: changes in the NCDS and BCS70 populations and samples over time*. London: Centre for Longitudinal Studies, Bedford Group for Lifecourse and Statistical Studies, Institute of Education, University of London.

Power, C. and Elliott, J. (2006) ‘Cohort profile: 1958 British birth cohort (National Child Development Study)’, *International Journal of Epidemiology*, 35(1), pp. 34–41. Available at: https://doi.org/10.1093/ije/dyi183.

Prandy, K. (1999) ‘Class, stratification and inequalities in health: a comparison of the Registrar‐General’s Social Classes and the Cambridge Scale’, *Sociology of Health & Illness*, 21(4), pp. 466–484. Available at: https://doi.org/10.1111/1467-9566.00167.

Prandy, K. and Lambert, P. (2003) ‘Marriage, Social Distance and the Social Space:: An Alternative Derivation and Validation of the Cambridge Scale’, *Sociology*, 37(3), pp. 397–411. Available at: https://doi.org/10.1177/00380385030373001.

Roantree, B. and Vira, K. (2018) *The rise and rise of women’s employment in the UK*. The IFS. Available at: https://doi.org/10.1920/BN.IFS.2019.BN0234.

*Robbins Report* (1963). Available at: http://www.educationengland.org.uk/documents/robbins/robbins1963.html (Accessed: 28 November 2022).

Roberts, K. (2003) ‘Change and Continuity in Youth Transitions in Eastern Europe: Lessons for Western Sociology’, *The Sociological Review*, 51(4), pp. 484–505. Available at: https://doi.org/10.1111/j.1467-954X.2003.00432.x.

Rose, D. and Pevalin, D.J. (2002) ‘The National Statistics Socio-economic Classification: Unifying Official and Sociological Approaches to the Conceptualisation and Measurement of Social Class’, *Sociétés contemporaines* [Preprint].

Saunders, P. (2003) *Social Theory and the Urban Question*. Routledge.

Saunders, P. (2021) *A Nation of Home Owners*. Routledge.

Savage, L. (2011) ‘Snakes and Ladders: who climbs the rungs of the earnings ladder’, *Resolution Foundation* [Preprint].

Savage, M. and Egerton, M. (1997) ‘Social Mobility, Individual Ability and the Inheritance of Class Inequality’, *Sociology*, 31(4), pp. 645–672. Available at: https://doi.org/10.1177/0038038597031004002.

Schmitt, C. (2021) ‘The impact of economic uncertainty, precarious employment, and risk attitudes on the transition to parenthood’, *Advances in Life Course Research*, 47, p. 100402. Available at: https://doi.org/10.1016/j.alcr.2021.100402.

Schoon, I. *et al.* (2001) ‘Transitions from school to work in a changing social context’, *YOUNG*, 9(1), pp. 4–22. Available at: https://doi.org/10.1177/110330880100900102.

Schoon, I. (2007) ‘Adaptations to changing times: Agency in context’, *International Journal of Psychology*, 42(2), pp. 94–101. Available at: https://doi.org/10.1080/00207590600991252.

Schoon, I. (2010) ‘Becoming Adult: The Persisting Importance of Class and Gender’, in Scott, J., Crompton, R., and Lyonette, C., *Gender Inequalities in the 21st Century*. Edward Elgar Publishing, p. 13500. Available at: https://doi.org/10.4337/9781849805568.00008.

Schoon, I. (2020) ‘Navigating an Uncertain Labor Market in the UK: The Role of Structure and Agency in the Transition from School to Work’, *The ANNALS of the American Academy of Political and Social Science*, 688(1), pp. 77–92. Available at: https://doi.org/10.1177/0002716220905569.

Schoon, I. (2022) ‘Planning for the Future: Changing Education Expectations in Three British Cohorts’, p. 22.

Schoon, I. and Bynner, J. (2003) ‘Risk and Resilience in the Life Course: Implications for Interventions and Social Policies’, *Journal of Youth Studies*, 6(1), pp. 21–31. Available at: https://doi.org/10.1080/1367626032000068145.

Schoon, I., Martin, P. and Ross, A. (2007) ‘Career transitions in times of social change. His and her story’, *Journal of Vocational Behavior*, 70(1), pp. 78–96. Available at: https://doi.org/10.1016/j.jvb.2006.04.009.

Seaman, S.R. *et al.* (2012) ‘Combining Multiple Imputation and Inverse‐Probability Weighting’, *Biometrics*, 68(1), pp. 129–137. Available at: https://doi.org/10.1111/j.1541-0420.2011.01666.x.

Seaman, S.R. and White, I.R. (2013) ‘Review of inverse probability weighting for dealing with missing data’, *Statistical Methods in Medical Research*, 22(3), pp. 278–295. Available at: https://doi.org/10.1177/0962280210395740.

Shanahan, M.J. (2000) ‘Pathways to Adulthood in Changing Societies: Variability and Mechanisms in Life Course Perspective’, *Review of Sociology*, 26(1), pp. 667–692. Available at: https://doi.org/10.1146/annurev.soc.26.1.667.

Sianesi, B., Dearden, L. and Blundell, R. (2003) *Evaluating the impact of education on earnings in the UK: Models, methods and results from the NCDS*. Working Paper Series. IFS. Available at: https://doi.org/10.1920/wp.ifs.2003.0320.

Silverwood, R. *et al.* (2021) ‘Handling missing data in the National Child Development Study: User guide (Version 2).’

Steiner, R., Hirschi, A. and Akkermans, J. (2021) ‘Many Roads Lead to Rome: Researching Antecedents and Outcomes of Contemporary School-To-Work Transitions’, *Journal of Career Development*, p. 089484532110635. Available at: https://doi.org/10.1177/08948453211063580.

Stewart, A., Prandy, K. and Blackburn, R.M. (1973) ‘Measuring the Class Structure’, *Nature*, 245(5426), pp. 415–417. Available at: https://doi.org/10.1038/245415a0.

Stewart, A., Prandy, K. and Blackburn, R.M. (1980) *Social Stratification and Occupations*. Springer.

Stopforth, S. (2020) ‘Parental Socio-Economic Background and Children’s School-Level GCSE Attainment’, *PhD thesis*, p. 306.

Szreter, S.R.S. (1984) ‘The Genesis of the Registrar-General’s Social Classification of Occupations’, *The British Journal of Sociology*, 35(4), p. 522. Available at: https://doi.org/10.2307/590433.

Treiman, D.J. (2009) *Quantitative data analysis doing social research to test ideas*. Jossey-Bass.

University College London (2022) ‘1970 British Cohort Study: Activity Histories 1986-2016’, *Centre for Longitudinal Studies* [Preprint]. Available at: https://doi.org/10.5255/UKDA-SN-6943-4.

University College London, UCL Institute of Education, and Centre for Longitudinal Studies (2023) ‘National Child Development Study.’, *UK Data Service* [Preprint].

Wallace, C. and Cross, M. (1990) *Youth in Transition: the sociology of youth and youth policy*. Psychology Press.

Williams, M. (2017) ‘An old model of social class? Job characteristics and the NS-SEC schema’, *Work, Employment and Society*, 31(1), pp. 153–165. Available at: https://doi.org/10.1177/0950017016653087.

Wright, E.O. (2005) ‘Foundations of a neo-Marxist class analysis’, in E.O. Wright (ed.) *Approaches to Class Analysis*. 1st edn. Cambridge University Press, pp. 4–30. Available at: https://doi.org/10.1017/CBO9780511488900.002.

1. https://github.com/Scott0atley/YouthTransitions [↑](#footnote-ref-1)
2. This latter category can be considered an ‘Other’ category. [↑](#footnote-ref-2)
3. This table contains a number of shortened words. For clarity: ‘’FT’’ means Full Time, ‘’PT’’ means Part Time, ‘’ED’’ means education, ‘’APP’’ means Apprenticeship, ‘’TC’’ means Training Course, ‘’OTH’’ means other, ‘’FTTC’’ means Full Time Training Course, ‘’TOPSTC’’ means Training Opportunities for Young Parents Training Course, ‘’UNEMP’’ means unemployed, and ‘’Rule6’’ means N/A. [↑](#footnote-ref-3)
4. Note: Both ‘’LGSS’’ and ‘’DBR’’ have zero documentation with the NCDS or online to confirm what each abbreviation means. Anytime DBR is mentioned it is alongside a more dominant partner – Full Time Work. As for LGSS, it constitutes a single data point, looking at how it is coded within the Stata files suggests it is grouped alongside TOPs courses – due to this LGSS is paired with TOPs courses throughout analysis. [↑](#footnote-ref-4)
5. [↑](#footnote-ref-5)
6. The issue with the NCDS data however is that white people make up 98.3% of all participants. The resulting ethnic minority categories are thus too small to conduct useful analysis. Originally, the resultant variable was parametrised as ‘white’ and ‘non-white’. There were two major issues with this that resulted in the race variable being dropped from analysis. The first returns to the overall low sample of non-white participants when spread over five different economic activity sub-categories. This low number of observations results in low statistical power and thus would impact the entire model. The second is that missing data is a particular problem when it comes to race. The race variable accounted for 16 per cent of missingness in subsequent models. On top of these two primary concerns, a combined race category into white/non-white presents assumptions surrounding homogeneity within the non-white category that is not theoretically justifiable. [↑](#footnote-ref-6)
7. Either in Social Housing or privately rented accommodation. [↑](#footnote-ref-7)
8. The ’constant’ is a concept that argues that occupational positions have the same meaning over time and across different countries. [↑](#footnote-ref-8)
9. Discussed further in the ‘Registrar General Class Schema’ section below [↑](#footnote-ref-9)
10. https://warwick.ac.uk/fac/sci/statistics/staff/academic-research/firth/software/qvcalc/kuvee/ [↑](#footnote-ref-10)
11. Except reading and maths due to the consistent non-significance across the model [↑](#footnote-ref-11)
12. Discussed at length in section below on Multiple Imputation [↑](#footnote-ref-12)
13. The seed for MI model is 12346, it can also be found in the .do file within the Github page. [↑](#footnote-ref-13)
14. 50 imputations were used in MI models. [↑](#footnote-ref-14)
15. Variable n4118 used [↑](#footnote-ref-15)
16. The variable in question was acatnn236, a categorical variable. [↑](#footnote-ref-16)
17. Other includes those respondents that cannot be accurately traced through any of the aforementioned categories. [↑](#footnote-ref-17)
18. Percentages are based on the participants divided by total cohort. [↑](#footnote-ref-18)
19. The reason sweep 3 has higher participant numbers than sweep 2 etc is due to the way tracking and sampling was handled. Across the BCS, difference organisations took control over this aspect of the survey. [↑](#footnote-ref-19)
20. Age 26 was the first time the cohort member themselves were in complete control of answering the survey itself [↑](#footnote-ref-20)