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

Contents

[**Chapter One: The National Childhood Development Survey (1958): Youth Transitions in Education and Employment** 5](#_Toc151287511)

[**Introduction to Chapter One** 5](#_Toc151287512)

[**Literature Review: NCDS Timeframe and Context** 7](#_Toc151287513)

[**Story of transitions for NCDS youth** 8](#_Toc151287514)

[**Structural Barriers to successful transitions – the role of Sex and social-class** 17](#_Toc151287515)

[**The role of social theory** 22](#_Toc151287516)

[**Data and Methods** 28](#_Toc151287517)

[**Introduction to the NCDS data** 29](#_Toc151287518)

[**Introduction to measures for subsequent analysis** 32](#_Toc151287519)

[**Descriptive Statistics** 51](#_Toc151287520)

[**Modelling Main Economic Activity** 58](#_Toc151287521)

[**Discussion and Conclusion** 79](#_Toc151287522)

[**Testing Measures of Parental Social Class** 83](#_Toc151287523)

[**Discussion and Conclusions** 92](#_Toc151287524)

[**Missing Data in the NCDS** 93](#_Toc151287525)

[**Missing Data** 93](#_Toc151287526)

[**Multiple Imputation by Chained Equations** 99](#_Toc151287527)

[**Discussion and Conclusions** 114](#_Toc151287528)

[**Discussion and Conclusions for Chapter One** 115](#_Toc151287529)

[**Data Citation** 123](#_Toc151287530)

[**Bibliography** 123](#_Toc151287531)

List of Tables

[Table 1.1 Sweeps Included for Analysis 30](#_Toc150774991)

[Table 1.2 Participation in the NCDS from birth to 23 years 32](#_Toc150774992)

[Table 1.3 Frequency Statistics for Economic Activity 34](#_Toc150774993)

[Table 1.4 Educational Attainment Count Variable by Economic Activity 39](#_Toc150774994)

[Table 1.5 RGSC Class Schema 46](#_Toc150774995)

[Table 1.6 NS-SEC Class Schema 48](#_Toc150774996)

[Table 1.7 Examples of Occupations from Analytical NS-SEC 49](#_Toc150774997)

[Table 1.8 Descriptive Statistics for Economic Activity 53](#_Toc150774998)

[Table 1.9 Descriptive Statistics by Economic Activity 54](#_Toc150774999)

[Table 1.10 Goodness-of-fit summaries for explanatory variables and Economic Activity 60](#_Toc150775000)

[Table 1.11 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity 60](#_Toc150775001)

[Table 1.12 Mlogit of Economic Activity 68](#_Toc150775002)

[Table 1.13 Sensitivity analyses of alternative measures of parental social stratification 86](#_Toc150775003)

[Table 1.14 Missing data patterns for NCDS 98](#_Toc150775004)

[Table 1.15 Comparison of CRA NS-SEC vs Imputed NS-SEC 109](#_Toc150775005)

[Table A.1 Goodness-of-fit summaries for explanatory variables and Economic Activity (CAMSIS) 121](#_Toc150775006)

[Table A.2 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity (CAMSIS) 121](#_Toc150775007)

[Table A.3 Goodness-of-fit summaries for explanatory variables and Economic Activity (RGSC) 122](#_Toc150775008)

[Table A.4 Model building goodness-of-fit summaries for multiple logistic regression model of Economic Activity (RGSC) 122](#_Toc150775009)

List of Figures

[Figure 1.1 Transitional Pathways for NCDS Cohort 7](#_Toc150775010)

[Figure 1.2 Predicted Probabilities of Economic Activity by NS-SEC 73](#_Toc150775011)

[Figure 1.3 Predicted Probabilities of Economic Activity by Sex 74](#_Toc150775012)

[Figure 1.4 Predicted Probabilities of Economic Activity by Educational Attainment 75](#_Toc150775013)

[Figure 1.5 Predicted Probabilities of Economic Activity by Housing Tenure 76](#_Toc150775014)

List of Abbreviations

A’level Advanced Level

BCS British Cohort Survey

CRA Complete Records 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 (1958): Youth Transitions in Education and Employment**

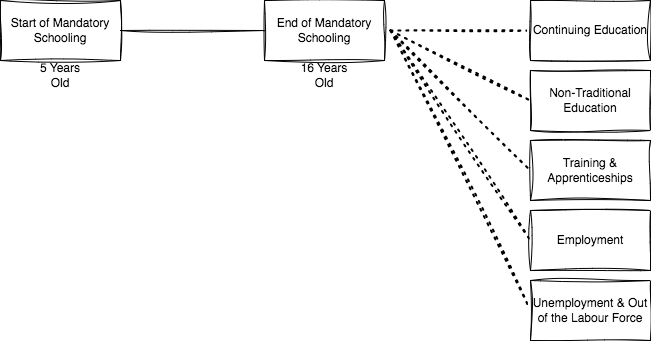
"Let us be frank: most people have never had it so good". – Harold Macmillan (1957)

# **Introduction to Chapter One**

The transition from mandatory education into a form of economic activity is one of the first significant life choices an individual in the UK has to make. This choice forms a bridge between the phase of education and preparation for the world of work and the phase of active employment (Mayer, 2009). The notion of an individual’s ‘choice’ is a complicated affair prominent in youth transition literature (Micklewright, 1989; Schoon, 2010). The role of structural influences, such as social class, sex, and housing tenure, can potentially influence individuals' choices across their life course. These structural influences may provide opportunities or hinder individuals' decision-making when selecting their economic activity post-mandatory schooling. The influence of structure upon choice is dependent and influenced by the socio-historical context in which the choice is made. The cohort of individuals analysed in this chapter comes from the National Childhood Development Survey (NCDS). This cohort of 17,638 individuals was all born in the same week in March 1958 and grew up in an identical socio-historical context. The NCDS cohort left mandatory schooling at age 16 in 1974. During this period, the UK was experiencing a large-scale shift in its labour market. This shift consisted of a collapse of traditional manufacturing occupations with simultaneous growth in the service economy and a collapse in the youth labour market. The NCDS cohort experienced a society in which Harold Macmillan proclaimed so confidently that ‘’You never had it so good’’ (Hamnett et al., 1989). The 1960s saw the ‘white heat’ of Wilson’s technological revolution, transforming British society and its subsequent labour markets into a service-based, consumer-based economy (Hamnett et al., 1989). These large-scale economic impacts had reverberations linked to specific social stratification areas. Most apparent consisted of a relative decline in traditionally working-class occupations and a growth of women in the labour market.

The NCDS provides an ability to study the influences of structural inequalities within a socio-historical context upon an individual’s choice and opportunity post-mandatory schooling. The following chapter identifies five core transitionary pathways: employment, non-traditional education, school, training & apprenticeships, and unemployment & out of the labour force. These five transitionary pathways are analysed through structural inequalities influencing an individual’s decision-making in sorting into one of these five economic activities post-mandatory schooling. Figure 1.1 illustrates the transitionary pathways possible for the individuals within the NCDS cohort.

Figure 1.1 Transitional Pathways for NCDS Cohort



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

This section provides an overview of the literature on youth transitions. This literature review focuses on existing research outlining the school-to-work transition and examining the impact of structural inequalities on that transition. A significant focus is placed on the role of social class and sex. A broader focus also examines the nature of choice and opportunity for the NCDS cohort and how structural inequalities impact this. Major transition trajectories have been explored within this review. 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 expanded upon by reviewing the theoretical literature to provide a holistic overview of the school-to-work transition during the NCDS period.

During the NCDS cohort, young people were in full-time education until they were 16. The NCDS youth were part of a larger cohort of children impacted by the rising school-leaving age (ROSLA) in 1972. At 16, individuals were typically expected to sit some form of examination. This was a mixture of Certificate of Secondary Education (CSE) (Pearson qualifications, 2023a) and Ordinary level (O’level) (Pearson qualifications, 2023b). After this examination period and the end of mandatory schooling, individuals are open to a level of choice in what and where they wish to take their lives. Some choose to stay within education and attend a sixth-form college and go on to a university; others go straight into full-time employment, others go to training, and some experience 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**

The youth of the NCDS cohort all followed the same pathway from age five – the start of mandatory schooling – to age 16 – the end of mandatory schooling. From the age of 16, NCDS youth had multiple pathways. Some would choose to enter the labour market and employment straight away. Others would seek to elongate their educational pathway by staying within school. Traditionally, this would mean joining a sixth-form college and taking Advanced levels (A’levels). During the NCDS timeframe, non-traditional educational pathways were also available. These were typically technical colleges offering a set of non-traditional qualifications. Beyond education, there was also the option of joining a training & apprenticeship program. These were mainly geared towards specialised manufacturing labour. Finally, some would enter a period of unemployment or opt to be out of the labour force.

Individuals do not make choices in a vacuum. Given their biographical life course (Elder, 1994) up until the point of finishing mandatory education, stratifying influences play a role in what choices an individual is likely to select and what opportunities they can also participate in. Education is one such stratifying influence. Education regulates the individual by implementing age-graded barriers and hierarchical and time-related credentials. The type of education an individual experiences and the product of their educational attainment captures some of this structural inequality. As with education, the labour market is another potential stratifying influence. The labour market determines gainfully employed and unemployed through its regulatory function, influencing employment trajectories (Mayer, 2004). Occupations also have a stratifying influence on society. The structure and hierarchy of occupations determine social position via segmentation and segregation, partly determined by previous employment and education systems. Stratification is vital in the underlying choices and opportunities that influence individuals' societal pathways.

The outcome of a child's post-mandatory schooling impacts their life chances across their ‘life course.’ Functionally, the study of youth transitions is the study of the life course, education systems, occupation, and labour markets that 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). 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 must be acknowledged for adequate sociological inquiry. Hence, a desire to provide a thorough context of the NCDS cohort's socio-historical context. The following section provides this socio-historical context through empirical research on existing youth transition literature within the NCDS.

Part of this context forms an understanding of what a ‘transition’ looked like for NCDS youth. Some youth transition theorists argue that in the mid-20th century, transitions were smoother and more simplified (Vickerstaff, 2003; Leuze, 2010). During the time of the NCDS, the range of choice has been argued to have been narrower compared to later cohorts, thus owing to a more homogenised pathway (Goodwin and O’Connor, 2005). It is accurate that the NCDS cohort had large homogenous clusters related to transitioning out of mandatory education. However, the delineation between school and employment is not a strict binary – with many children engaging in the youth labour market whilst still in education (Bynner, 2012). Children's choices on what they wish to do after mandatory education are influenced by labour market restructuring and recession and other factors such as the child's family background, parents, teachers, siblings, and contemporaries. While the pathways the NCDS cohort may have been able to choose from were narrow, this is not synonymous with smooth or straightforward. For example, individuals may have faced a seemingly homogenous experience 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 need to be considered.

The NCDS cohort was a homogenous transitory cohort, even more so comparatively. Evidence states that of the NCDS cohort, 96 per cent of men could be grouped into six of the most significant transition ‘clusters’ compared to 90 per cent for the British Cohort Study (BCS) (Martin et al., 2008). This suggests that 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). This is supported by Anders and Dorsett (2017), where transition patterns among school leavers entering the labour market were examined. They found that under the NCDS cohort, there was a large (91 per cent) number of people entering 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 a school-to-employment transition was dominant. However, training was also a vital transition pathway – above that of continuing full-time education (Schoon *et al.*, 2001). The NCDS cohort was, however, caught in a period of severe diminishing influence of apprenticeships. For example, the number of apprenticeships in British manufacturing declined from 240,400 in 1964 to 155,000 in 1979 (Blanchflower and Lynch, 1992). There was a severe gender bias regarding apprenticeships at this time - when the NCDS cohort was 16 years old, 40 per cent of male employees were apprenticed compared with only 8 per cent of females (Blanchflower and Lynch, 1992).

The relative importance of employment and apprenticeship training over educational pathways suggests that the NCDS cohort experienced a pre-credentialed labour market post-mandatory schooling (Bynner, 2005). The NCDS cohort experienced a labour market that did not place challenging roadblocks to employment based on educational credentials. It was not until the 1980s that failing to get qualifications hindered getting work in Britain (Bynner, 2005).

The labour market in the pre-1980s was able to absorb people into large numbers of unskilled jobs (ibid). Those who did struggle to get jobs in the NCDS cohort (Bynner, 2005) 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 indicates that it was only in 1975 that O’levels moved from a pass/fail system to a graded one (Pearson qualifications, 2023b).

It has been established that the NCDS cohort exhibited a homogenous transitional experience. The NCDS birth cohort did not experience a straightforward, smooth school-to-work transition. Teenagers who 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). It is estimated that children aged 16 worked an average of six to nine hours a week and modal earnings in the range of £1-£2 a week while still in full-time mandatory education (Dustmann *et al.*, 1996). 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 oversimplification for the time. Children were engaging in schooling and employment before choosing 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). Students employed during mandatory education were less likely to choose to continue 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 are not the same as smooth transitions. Structural inequality adversely impacts the relative smoothness of an individual’s transitional experience.

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 significant period of restructuring – some have also argued that the ‘collapse’ of the youth labour market also contributed to a relative amount of instability (Bynner, 2012). The notion that the youth labour market ‘collapsed’, indicating a sudden affair, is not precisely accurate. The youth labour market saw a relative decline post-war as part of broader economic restructuring. Nevertheless, the decline of the youth labour market still impacted the options available to NCDS youth. The labour market during school-to-work transitions for the NCDS cohort was unstable and comparatively heightened uncertainty (Leuze, 2010). The collapse of the youth labour market in the early 1980s was not a sudden affair (Bynner, 2012). 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 was facing during their biographical lifespan (Bynner, 2012). The importance of the collapse of the youth labour market relates to introducing uncertainty at a critical stage of development within a young person’s life (Maclure, 1978). This uncertainty can adversely impact individuals' life domains (Mayer, 2009). These periods of instability are documented in detail with monthly employment histories (Leuze, 2010), suggesting that the collapse of the labour market impacted many individuals. This heightened instability during a time of transition for the youth of the NCDS presents an influencing factor in the role of choice and opportunity. When the labour market was facing severe restructuring, a collapsing youth labour market, and a significant economic recession, the choices and opportunities of young people seeking to transition into the world of work would be constrained and influence their choices. These constraints, for example, created a likelihood for disengagement with school and alienation from education (Farrall, Gray and Mike Jones, 2020).

The development of this concept of uncertainty and risk within the NCDS starkly contrasts the theory of ‘late modernity’ - entailing notions of risk and uncertainty in a society that provides individuals with more choice, promoting greater risk (Beck, Giddens and Lash, 1994). The literature has demonstrated that the NCDS cohort experienced complicated pathways and transitions. Structural inequalities impacted the choices and opportunities within these different pathways. The notion of ‘Late Modernity’ (Giddens *et al.*, 1991; Beck, Giddens and Lash, 1994) is based upon the idea that in the past, more concrete certainties have given way to 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 that the NCDS cohort did not have explicitly concrete certainties – it exhibited relatively homogenous pathways. However, the details within these pathways were often complex and influenced by uncertainty and risk of the time. There is debate over how fluid certainty and choice have become; Gayle et al. (2009) provide 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. Educational attainment – and staying within education post-mandatory schooling - protects from unemployment (Bynner, Wiggins and Parsons, 1996). Those of the NCDS cohort that stayed within education post-mandatory schooling initially had higher unemployment levels due to exogenous shocks of rising national unemployment – due to the aforementioned labour market restructuring and economic recession. Whilst experiencing short-term levels of unemployment, in the long run, individuals who stayed on within education had a long-term advantage in income over their peers who did not stay on within education (Payne, 1987).

Looking in more detail at educational attainment within the NCDS, individuals in the UK who choose to stay on at school post-16 were a small minority and were low by Organisation for Economic Co-operation and Development (OECD) standards compared to other Western countries (Micklewright, 1989). Compared to their non-manual peers, individuals from manual backgrounds were less likely to stay on post-16 (Micklewright, 1989). Research concurs 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).

Whilst individuals from manual backgrounds were less likely to continue to 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) find 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 who 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 leads to subsequent full-time employment (Schoon *et al.*, 2001). School 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 stable employment. Vocational-based education is generally considered a smoother transition from school to work than academics. While this short-term benefit is worth considering, long-term disadvantages such as lower employment and wages impact those individuals with lower vocational education (Brunello and Rocco, 2017). This phenomenon has broken down post-NCDS with the breakdown of traditional apprenticeship and training programs in the UK (ibid).

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

Throughout the story of NCDS youth, a common theme of structural barriers and inequalities influencing choice and opportunity has been identified. This next section seeks to explore these structural dimensions more closely. The roles of sex, social class, and housing tenure will be explored in greater detail in an attempt to provide clarity to the current empirical consensus on these forms of social stratification about NCDS youth.

#### **Sex**

Choice and opportunity within the school-to-work transition of the NCDS youth are influenced and impacted by structural inequality factors like sex (Dolton, Joshi and Makepeace, 2002; Makepeace, Dolton and Joshi, 2004; Cebulla and Tomaszewski, 2013). Women's roles within the labour market have marked differences from their male peers (Dex and Bukodi, 2012). Whilst women are more likely achieve their educational aspirations than men (Cebulla and Tomaszewski, 2013) 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, part-time work began to grow in the 1950s. Part-time work is pervasive among women in the NCDS, especially when returning from giving birth (Dex *et al.*, 1998).

There was some slight decline in gender segregation within the labour force for the NCDS cohort due to the ‘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 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; Schoon, 2010). However, equalised rates of male labour force participation have 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 the human capital improvements made by women would have ultimately attracted a 17 per cent greater wage premium if held by a man (Makepeace *et al.*, 1999).

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

When limiting attention of social mobility to solely full-time workers, mobility does not significantly vary by gender (Bukodi, Goldthorpe and Kuha, 2017), though some research (Savage and Egerton, 1997; Savage, 2011) does emphasise the impact gender has on social mobility. Part-time female workers have highly varied pathways (Connolly and Gregory, 2010). While 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 the 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 on transition outcomes and later life chances. When looking at educational attainment (Holm and Jæger, 2011), it is essential 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 critical role in the life chances of young people than parental income (Feinstein, Duckworth and Sabates, 2004; Field, 2010). Early success in education confers an advantage in later educational attainment and labour market experience (Dolton, Makepeace and Marcenaro‐Gutierrez, 2005). 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), it persists when translating educational attainment into the most successful occupational outcomes – those from privileged backgrounds are more likely to gain access to the highest-paying occupations, leveraging their educational qualifications.

Educational attainment translates to higher levels of income in later life—individuals with higher educational ability experience faster wage growth than their lower-ability peers (Galindo-Rueda, 2003). Early successful ability is influenced, however, by a structural class effect. Those from working-class backgrounds are less likely to succeed in early life stages than their non-working-class peers (Machin and Vignoles, 2005). Some argue that this is due to poorer families being less likely to invest in education over their more affluent peers (Chevalier and Lanot, 2001). However, the nature of what constitutes ‘investment’ in an individual’s education is left unclear and subject to speculation.

Low levels of qualifications and educational attainment are related to higher propensities toward unemployment (Bynner and Parsons, 2000). The propensity toward experiencing unemployment also has a social class effect, with the growth in unemployment during the 1970s being attributed to the subsequent decline in the manufacturing sector linked to working-class labour (Schoon *et al.*, 2001). Those who are unemployed also appear to hold the lowest levels of employment commitment when they eventually enter employment (ibid). 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 a vital role in establishing adult future earnings (Gregg, 2001). 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). The influencing impact that social class has on attainment and propensity toward unemployment appear to have long-term consequences for later life chances. The impacts of social class on youth transitions from school to work are felt in the short and long term.

Work-related training, or training on the job, has been lauded as a way for those who 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 boost the already well-educated and leave those less educated behind. In a later study, (Arulampalam and Booth, 2001) reaffirm 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 who happen to be well-educated are related to those who come from advantaged social class positions demonstrates that advantage breeds advantage (Machin and Vignoles, 2005). Those from less affluent backgrounds who engage in work-related training 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 backgrounds 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 increases the likelihood of unemployment at the adult stage of the life course (Gregg, 2001).

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

Throughout 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 dispense with static ‘snapshot’ notions of sociology. Instead, it views the individual in a constant web of changing temporal contexts influencing the agent (Elder, 1994). The life course approach is best suited for analysing 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). The term ‘life course’ is a concrete multilevel phenomenon defined via individuals' social trajectories through structured pathways of given institutions that form the developmental experience of a given individual (Elder, 1994). 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 within the life course support an analysis focusing on inequalities concerning race, class, gender, and other structural aspects of social life (Bernardi et al.. 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 specific points in their lives 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, Johnson and Crosnoe, 2003), also known as ‘’bounded agency’’ (Evans, 2007). The bounded agency is a concept that argues that the agency of the individual is situational and bound to the circumstances of place and time (Bernardi et al.,

By focusing on a life course perspective, analysis can extend beyond static moments. This allows research to be expanded about individuals and between individual analyses. A life course approach appreciates the fact that structured pathways are temporally grounded and, as such, acknowledges 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 on the interdependence between life domains (Mayer, 2009). This means that outcomes within one domain (e.g., school) are interrelated with the outcomes and behaviours of other domains (e.g., 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).

The second key theme established through this literature review is the relevance of structural inequalities on individual choice and opportunity. This role of structure is, in one way, expressed through the social theory of structuration (Giddens, 1989). Structuration argues that structural factors like social class, gender, and ethnicity still play an essential role in shaping the lives of individuals and are indeed determinants for the individual pursuing the ‘imperative of living a life of one’s own’ (Beck, 2002). In response to this theory of structuration, the theory of individualisation argues that in place of these ‘collective guides’ (Gayle et al., 2009), individualised identities that have greater scope beyond the mere structures (Murray, 2011) they inhabit can create complex and subjective lifestyles that deviate from the much more rigid structures detailed above (Gayle et al., 2009).

If the individualisation thesis were 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 altered, and even in some cases become more complex, but that does not mean there is support for ‘detraditionalisation’ (ibid). There has always been an element of navigation and choice within youth transitions. However, 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). In other words, individualisation fails to account for structural elements' still apparently strong influence on a person.

A more robust critique towards structuration theory – the declining relevance of social class – argues that structures are essential. However, the specific role of social class 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 notably related to housing tenure – are more influential on outcomes than 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). However, it must be stated that new structuralism as a theoretical perspective grew out of the socio-historical context of the 1980s. Some effects of the new structuralist perspective ought to exist within any analysis of the NCDS – albeit a relatively weak effect compared to a cohort in the 1980s.

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

The central theme of this literature review that has unfolded relates to the notion of choice and opportunity. Concepts of choice and opportunity are constrained through structural inequalities and socio-historical contexts. Within the life course approach, the structure/agency dichotomy is replaced with an agency within the structure (Diewald and Mayer, 2008) theoretical modelling. The concept of agency is impacted based on different temporal foci (Hitlin and Elder, 2007). Understanding the life course requires a multidimensional notion of agency (Hitlin and Kirkpatrick Johnson, 2015). Hitlin and Elder (2007) break down 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); 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).

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 depends on socio-temporal aspects that leave room for individual decision-making or ‘agency’ (ibid). Throughout an individual’s life course, there are times when, due to social position and other times due to socio-historical constraints, the individual experiences differing forms of social pressures and the ability to impress themselves upon social structures in the forms of choice. The empirical literature reviewed thus far has illuminated such cases – concerning class position and the likelihood of entering higher education (Micklewright, 1989), for 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 the opportunity to discover how that impacts later life decisions and outcomes. These decisions, in other words, are youth ‘transitions’ (Hitlin and Elder, 2007). Treating the individual as an active agent in shaping their biographies is crucial 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 NCDS provides relevant data to analyse the choices and opportunities that individuals in the 1970s experienced and the influence that structures had upon the role of choice and opportunity. As illustrated in Figure 1.1, individuals leaving mandatory schooling at age 16 experienced five possible pathways of choice: employment, continuing schooling, non-traditional education, training and apprenticeships, or unemployment and out of the labour force. The NCDS dataset allows the construction of a dependent variable that produces this five-category economic variable that details individuals' choices in the 1970s. Due to the categorical nature of this dependent variable, a multinomial logistic regression is applied to study the influence that structural inequalities have on choice.

This relationship between structure and choice 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, social class, and sex are 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 choices 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 and other structural consumption cleavages, such as housing tenure. The model also accounts for individual merit or ability through prior attainment. After establishing the initial multinomial logistic regression model, a sensitivity analysis will be conducted on various social stratification surrounding class measures to assess the most appropriate measure to include within the model. Finally, an analysis of missing data involving multiple imputations will be conducted to assess the impact, if any, on the substantive findings of the multinomial logistic regression model.

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

This work will use the National Child Development Study (University College London, UCL Institute of Education, and Centre for Longitudinal Studies, 2023)[[1]](#footnote-1). 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 a week in March 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 September 1974, or month 201.

Table 1.1 Sweeps Included in Analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | 1958 | 1965 | 1969 | 1974 | 1981 |
| Sweep Number | 0 | 1 | 2 | 3 | 4 |
| Age | Birth | 7 | 11 | 16 | 23 |

The NCDS cohort originated in 1958 (when participants were born) and continues today. Only data up until age 23 (wave 4) is considered for analysis. Table 1.2 details the sample size of the NCDS. At birth in 1958, the total cohort consisted of a sample 17,638 with 17,415 participants. By 1974, age 16, the total cohort had increased to 18,558. This is because the original sample was 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 actual participants from the 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. After considering 960 dead and 1,196 emigrants, the eligible sample is 16,402. 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). For 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 filled out the survey independently without parental or guardian assistance (like at age 16) or having it done for them by their parents, teachers, and medical professionals. Data were collected by a paper and pencil-based survey. Economic activity as an outcome was reported through a retrospective employment history question that had to be filled out from 16 every month until the current age of 23. This level of detail does make the possibility of refusal a possibility.

According to Hawkes and Plewis (Hawkes and Plewis, 2006), non-response: others (cases where there is no data for this sweep but there is for later sweeps, and ‘temporary emigrants’) make up 10.7 per cent of non-response. Around 6.3 per cent of the remaining missingness is categorised as ‘eligibility unknown’ (ibid). Eligibility unknown relates to those who either died or permanently emigrated. There is, overall, a substantive amount of missingness within the data used for analysis. An issue with sample attrition for the NCDS is that the size of specific ethnic minority populations when the sample was first collected was small, meaning that attrition analysing ethnic minority populations is challenging (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; Silverwood *et al.*, 2021). 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 the 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 |

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 using the UCL CLOSER search platform, which promotes a resource that allows all NCDS cohort variables 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 the analysis. Manual construction of codebooks for some variables and manual inspection of original documentation was called for. From this, variable selection became a much smoother process. Ultimately, economic activity, educational attainment, sex, social class, and housing tenure were selected.

#### **Economic Activity**

The primary outcome variable of interest is the main economic activity of month 201. This is what individuals did after mandatory schooling in September at age 16. For example, the economic activity individuals engaged with after year 11 in the English and Welsh school system context. September was selected to allow time for children to gain their O’level results. This economic activity variable [ec201] was a retrospective work history collected at age 23. Participants were asked to note their current economic activity from age 16-23 each month. This variable comes from sweep 4 (Age 23) of the NCDS. Everyone's economic activity was recorded retrospectively by the participants at age 23 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. These 28 values comprised a combination of main categories: employment, education, training, and unemployment. Individuals could, for example, be coded as being in full-time employment and doing an apprenticeship scheme, etc. These 28 unique values were recorded as follows: five of these collapsed into the 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, non-traditional education, school, training & apprenticeships, and unemployment & out of labour force. Main Economic Activity is determined based on whether 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 the 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 | 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 necessary to get at the nuance of some of the economic activity data. For example, much data was coded as full-time employment – including training schemes, apprenticeships, Technical and Vocational Educational Initiative (TVEI), and TOPs schemes[[4]](#footnote-4).

Re-coding this variable translates into five categories: employment, non-traditional education, school, training & apprenticeship, and unemployment & out of the labour force. Employment 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 who, after completing mandatory schooling at age 16, decides 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 are 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. However, 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 1981. At this time, the primary educational qualifications were either the Certificate of Secondary Education (CSE) (Pearson qualifications, 2023a), introduced in 1965, or the Ordinary level or O’level, introduced in 1951 (Pearson qualifications, 2023b). The O’level was understood to be a higher level than CSEs, and fewer people achieved O’level grades. This is the best and most advanced ability measure for the age of 16 and makes a good measure of educational attainment for those at 16 after mandatory schooling ends.

Researchers have advocated for establishing education measures to facilitate better 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. However, this would be attractive for analysis 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, and as such, a like-for-like measure may not be the most attractive (Murray, 2011). A measure of attainment, GCSEs and O’levels provide considerable barriers to entry for young people pursuing future goals (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 measures 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 [n4655], the second, on condition of the first, then asks how many O’levels that person had passed [n4656] – passing in this context refers to if an individual’s O’level grade was within the grade boundaries A-C. Combining these two variables produces a single count variable that includes the number of zeros. This attainment variable was then 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. The second reason for recoding is one of practicality. Keeping O’levels as a count variable illustrates a truncated position of several O’levels, making a binary dummy more sensible – 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 | Nine or More | Total |
| Economic Activity |  |  |  |  |  |  |  |  |  |  |  |
| Employment | 3,024 | 586 | 368 | 251 | 234 | 181 | 113 | 70 | 35 | 25 | 4,887 |
| Non-Traditional 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 [n622\_4]. Its inclusion for analysis is because, during the timeframe of the NCDS, sexed dynamics played an essential 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.. 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 (Payne, 1995; Lindley, 1996; Gayle et al., 2009). Race as a variable [n2017] for inclusion in this model presents too many statistical issues to be an effective measure.

The issue with the NCDS data, however, is that white people make up 96% of all participants. The resulting ethnic minority categories are thus too small to conduct helpful analysis. Initially, the resultant variable was parametrised as ‘white’ and ‘non-white’. Two significant issues resulted in the race variable being dropped from the 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 regarding 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.

#### **Housing Tenure**

Previous analyses have used housing tenure regarding educational attainment and labour market outcomes (Di Salvo and Ermisch, 1997; Duta et al., 2021). For subsequent analysis, tenure measures whether an individual lives in their home or not [n1152][[5]](#footnote-5). 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. Including housing tenure in this model allows a more direct investigation of this sentiment. Enabling the evaluation of the 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). Inequalities can manifest in gender, ethnicity, social class, etc. These social inequalities impact individuals' aspirations, desires, educational outcomes, labour market position, and destinations.

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

Social class as a variable has constant and consistent debate throughout sociological literature (Bottero, 2004) – 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 vital aspect of this chapter is to find the most appropriate measure of social class to fit the given analysis models. 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). Multiple measures of social stratification are reflected upon. The following section seeks to establish the significant measures of social class and weigh their common strengths and weaknesses, which may affect model parsimony.

Longer-term structural transformations of society will alter the underlying distribution of stratification over time (Lambert and Barnett, 2021). Whilst the Treiman constant[[6]](#footnote-6) is often hailed as the single most crucial 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. For example, the decline of the manufacturing industry in the UK may have a qualitative impact on transitions into the type of skilled manual occupations and a quantitative impact on the number of occupations available.

The NCDS provides occupational coding measures for the father’s socio-economic position using a variety of measures (Gregg, 2012). The measures provided are the Registrar General Class Schema (RGSC), National Statistics Socio-Economic Classification (NS-SEC), and the Cambridge Social Interaction and Stratification Scale (CAMSIS). Occupational codes were constructed for parents of NCDS youth in 1969. Unfortunately, no such occupational measures were taken for mothers, making it impossible to employ a semi-dominance approach (Connelly, Gayle and Lambert, 2016). The first occupational measure gives a complete six-class categorisation[[7]](#footnote-7). The last measure is continuous, so no recording was required. The occupational coding conducted by (Gregg, 2012) was subsequently merged with the primary data of sweeps 0-4 (up until age 23). The occupational coding data was then cleaned and re-coded into relevant schemas.

The following variables used for subsequent analysis are all considered socio-economic variables. Two are social class schemas, and one is a stratification scale. By comparing cohort substantive findings, multiple socioeconomic measures are considered to see whether there are different patterns for different dimensions of social stratification both within cohorts and across them.

The rationale for including the RGSC, NS-SEC, and CAMSIS is based on an attempt to analyse the impact of different social stratification measures on a given model. The NCDS and occupational coding data (Gregg, 2012) provide a limited number of social stratification variables to construct. NS-SEC, RGSC, and CAMSIS are the most practical social stratification measures that can be constructed using the NCDS dataset. Even when constructing these specific social stratification measures, issues were encountered. For example, NS-SEC was forced to use a simplified coding scheme due to a lack of information regarding fathers’ occupational status. 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 and CAMSIS (Gregg, 2012).

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

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

The Registrar General’s Social Class 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 (Stevenson, 1913). This measure of social stratification was later re-developed in 1921 and again in 1928 by stating that class was more closely equated with occupation than material factors of income or wealth in explaining certain phenomena[[8]](#footnote-8) (Stevenson, 1928). The measure is built upon the assumption that society is graded based on a hierarchy of occupations (Murray 2011). The schema is broken into six categories, from unskilled manual occupations to higher-level professionals (ibid). The RGSC once formed the basis of all commonly used social classifications within Britain (Szreter 1984). Alternative measures like the National Statistics Socio-Economic Classification have become prominent. The RGSC has been used within social stratification research and is often included as a measure of social class in datasets (Gregg, 2012). The RGSC, first developed in 1911 (Stevenson, 1913), means that as a measure of social stratification, it existed for 47 years before the NCDS existed. Compared to other social stratification measures, such as NS-SEC, 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 essential part of the subsequent sensitivity analyses.

The RGSC rests upon a theoretical assumption that social inequality exists within society and that individuals are socially stratified by unequal rewards (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 comprises an upper-middle, middle, and working class (Prandy, 1999). This assumption is baked into the theoretical implications of the aforementioned unidimensional hierarchy. The RGSC schema also follows an explicit hierarchical ordering split into two halves: a non-manual dimension at the top half of the scheme and a manual dimension at the bottom half of the scheme, as seen in Table 1.5.

The Full RGSC class schema is detailed below:

Table 1.5 RGSC Class Schema

|  |  |  |  |
| --- | --- | --- | --- |
|  | Class | Occupations | Example Occupations |
| Non-Manual | I | Professional Occupations | Accountant |
| II | Intermediate Occupations | Police Officer |
| IIIN | Skilled Non-Manual Occupations | Clerical Worker |
| Manual | 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 and SOC90 (Gregg, 2012). Amongst the social stratification variables that are provided, full-auto, semi-auto, and verification processing variables are provided. Semi-auto processing social stratification variables are used [N2SRGSC] within subsequent analysis (Gregg, 2012) as suggested.

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

Rose and Pevalin developed the NS-SEC schema (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 Erikson-Goldthorpe-Portocareo (EGP) perspective (Rose and Pevalin, 2002). The EGP scheme, like NS-SEC, rested on theoretical assumptions of labour market positions, employment status, and relations and was initially developed by Goldthorpe (Goldthorpe, 1980).

The NS-SEC was developed from the EGP perspective (ibid). Employment relations are central to the NS-SECs' ideas on social class – and the development of social class schemas. 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). This differentiation of employment relations 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 complete NS-SEC classification schema has 14 operational categories related to employment relations but can be broken down into as few as three analytical categories. The occupational codes offered (Gregg, 2012) provide NS-SEC in its operational category form that is then broken down into the NS-SEC nine analytical class variety (Rose and Pevalin, 2010). This ability to break down the social class schema is attractive – particularly when using data with limited sample sizes or complications related to multiple imputation convergence. Whilst this is true, the collapsing of NS-SEC into a reduced schema ultimately risks capturing the theoretical implications of employment relations that NS-SEC as a class schema seeks to capture. For NCDS data, collapsing the NS-SEC analytical schema is not needed, and thus, the integrity of the theoretical orientation surrounding its construction is maintained.

Table 1.6 NS-SEC Class Schema

|  |  |
| --- | --- |
|  | Analytical Variables for NS-SEC |
| Operational Categories |  |
| L1  Employers in large establishments | 1.1 Large Employers and higher managerial occupations |
| L2  Higher managerial occupations |
| L3  Higher professional occupations | 1.2 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 occupations |
| 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 the operationalisation of the complete NS-SEC class schema [N2SNSSEC]. 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 |
| 1.1 Large Employers and higher managerial occupations | Chief Executives, Managers and directors in finance |
| 1.2 Higher professional occupations | Lecturers, Judges, Doctors |
| 2 Lower Managerial and professional occupations | Managers in retail, Nurses, School Teachers |
| 3 Intermediate occupations | Paramedics, Teaching Assistants |
| 4 Small employers and own account workers | Bricklayers, Carpenters, Shopkeepers |
| 5 Lower supervisory and technical occupations | Mechanics, Plumbers, Skilled Construction Supervisors |
| 6 Semi-routine occupations | Sales Assistants, Veterinary Nurses |
| 7 Routine occupations | Cleaners, Welding Trades, Hairdressers |

##### **CAMSIS**

CAMSIS argues that individuals are embedded within socially moderated spaces and networks within which they engage in various social and economic interactions, different from interactions with persons 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 continuous nature of CAMSIS means that numerical values are attached to occupations, meaning the relative value of each occupational value is only meaningful compared to other occupations on the same scale (Connelly et al. 2016). This is meaningful when it comes to the interpretation of the CAMSIS measure within models of analysis, as the value of the coefficient is always going to be concerning the comparison to other occupations along the CAMSIS scale. The most considerable difference between CAMSIS and other social stratification measures discussed is that CAMSIS does not believe 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 a much more dynamic, constantly re-constitutive process (Bergman and Joye, 2001).

While CAMSIS stands in contrast to the other social stratification measures mentioned, they share some similarities. CAMSIS contends - as do the NS-SEC and RGSC – that occupational groups are the primary 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.

The CAMSIS scale typically has a mean of around 50 and a standard deviation of around 15. Occupations such as cleaners would find themselves at the lower end of this scale, with occupations like judges being around 86.

As with RGSC and NS-SEC, CAMSIS codes are provided within the occupational codes provided by (Gregg, 2012). However, CAMSIS had to be re-coded using SOC codes within the Gregg (Gregg, 2012) files. The original files were created using SPSS; the Stata files contained errors within the CAMSIS constructions. The CAMSIS variables had all their decimal places moved to the left. This had two consequences. The first was that the mean and s.t.d were far off the typical 50 and 15 that CAMSIS typically represents. The second was that because the decimal place moved one to the left, it also only displayed data to one decimal place, meaning that data was lost in this error. For example, if someone had a CAMSIS score of 44.49, the value in the Stata files would be 4.4. This meant that simply moving the decimal place one to the right (by multiplying by 10) was not a possibility. CAMSIS was thus recoded using SOC codes [N2SSOC90]. Like NS-SEC, details on the employment status of individuals’ fathers were unavailable, so a ‘simplified CAMSIS’ was constructed. After this recoded, a comparison was made between this recode and multiplying the original CAMSIS values by 10. The former was much closer to the mean of 50, s.t.d of 15, which is expected from CAMSIS, as seen in Table 1.8.

## **Descriptive Statistics**

Table 1.8 shows the frequencies and summary statistics for the NCDS. Overall, 38.17 per cent of the sample is in full-time employment. Whilst 30.42 per cent remain in school, 8.82 per cent moved on to full-time post-school education. Unemployment and being out of the labour force make up 3.10. Finally, 19.48 per cent of the sample are in some training or apprenticeship scheme.

Regarding Educational Attainment, 64.48 per cent of individuals received less than 5 O’levels, with the remaining 35.52 per cent receiving five or more O’levels. Sex presents a relatively equal split between men (48.87 per cent) and women (51.13 per cent). Regarding homeownership, 48.07 per cent of individuals grew up in a home owned by their parents compared to 51.93 per cent that did not.

The NS-SEC categories all see a relatively even distribution between 10-20 per cent except for the most significant category – 7, at 23.86 per cent – and the smallest categories –1.1 and 1.2, at 3.30 per cent and 4.85 per cent. RGSC is much more unevenly distributed than NS-SEC, with skilled manual workers making up 41.44 per cent of individuals and professionals only making up 4.29 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 42.12 and a standard deviation 12.91.

For a breakdown of descriptive statistics by the outcome variable of economic activity, see table 1.9:

Table 1.8 Descriptive Statistics for Economic Activity

|  |  |  |
| --- | --- | --- |
|  | n | % |
| Economic Activity | | |
| *Employment* | 3,225 | 38.17% |
| *Non-Traditional Education* | 745 | 8.82% |
| *School* | 2,570 | 30.42% |
| *Training/Apprenticeships* | 1,646 | 19.48% |
| *Unemployment and OLF* | 262 | 3.10% |
| Educational Attainment O’levels | | |
| *Less than 5 O’Levels* | 5,447 | 64.48% |
| *Five or More O’Levels* | 3,001 | 35.52% |
| Sex of Respondent | | |
| *Female* | 4,235 | 50.13% |
| *Male* | 4,213 | 49.87% |
| Housing Tenure of Respondent When Child | | |
| *Own Home* | 4,061 | 48.07% |
| *Do not Own Home* | 4,387 | 51.93% |
| NS-SEC Social Class of Father when Respondent Child | | |
| *1.1* | 279 | 3.30% |
| *1.2* | 410 | 4.85% |
| *2* | 1,038 | 12.29% |
| *3* | 824 | 9.75% |
| *4* | 1,024 | 12.12% |
| *5* | 1,372 | 16.24% |
| *6* | 1,485 | 17.58% |
| *7* | 2,016 | 23.86% |
| RGSC Social Class of Father when Respondent Child | | |
| *1* | 362 | 4.29% |
| *2* | 1,738 | 20.57% |
| *3* | 924 | 10.94% |
| *4* | 3,501 | 41.44% |
| *5* | 1,205 | 14.26% |
| *6* | 718 | 8.50% |
|  |  |  |
|  | Mean | SD |
| CAMSIS Score of Father when Respondent Child | 42.12 | 12.91 |
|  |  |  |
| n |  | 8448 |
| Data Source: NCDS [Sweeps 0-4] | | |

Table 1.9 Descriptive Statistics by Economic Activity

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Economic Activity of Respondent in September when they are 16 | | | | | |
|  | **Employment** | **Non-Traditional Education** | **School** | **Training & Apprenticeships** | **Unemployment & OLF** | **Total** |
| Educational Attainment O-levels | | | | | | |
| *Less than five O'level passes* | 2,930 (90.9%) | 404 (54.2%) | 438 (17.0%) | 1,429 (86.8%) | 246 (93.9%) | 5,447 (64.5%) |
| *Five or more O'level passes* | 295 (9.1%) | 341 (45.8%) | 2,132 (83.0%) | 217 (13.2%) | 16 (6.1%) | 3,001 (35.5%) |
| Sex of Respondent | | | | | | |
| *Female* | 1,958 (60.7%) | 522 (70.1%) | 1,293 (50.3%) | 304 (18.5%) | 158 (60.3%) | 4,235 (50.1%) |
| *Male* | 1,267 (39.3%) | 223 (29.9%) | 1,277 (49.7%) | 1,342 (81.5%) | 104 (39.7%) | 4,213 (49.9%) |
| Housing Tenure of Respondent When Child | | | | | | |
| *Own Home* | 1,073 (33.3%) | 462 (62.0%) | 1,746 (67.9%) | 706 (42.9%) | 74 (28.2%) | 4,061 (48.1%) |
| *Do not Own Home* | 2,152 (66.7%) | 283 (38.0%) | 824 (32.1%) | 940 (57.1%) | 188 (71.8%) | 4,387 (51.9%) |
| NS-SEC Social Class of Father when Respondent Child | | | | | | |
| *1.1* | 50 (1.6%) | 33 (4.4%) | 159 (6.2%) | 33 (2.0%) | 4 (1.5%) | 279 (3.3%) |
| *1.2* | 43 (1.3%) | 37 (5.0%) | 291 (11.3%) | 38 (2.3%) | 1 (0.4%) | 410 (4.9%) |
| *2* | 213 (6.6%) | 131 (17.6%) | 544 (21.2%) | 129 (7.8%) | 21 (8.0%) | 1,038 (12.3%) |
| *3* | 234 (7.3%) | 94 (12.6%) | 360 (14.0%) | 121 (7.4%) | 15 (5.7%) | 824 (9.8%) |
| *4* | 401 (12.4%) | 84 (11.3%) | 269 (10.5%) | 241 (14.6%) | 29 (11.1%) | 1,024 (12.1%) |
| *5* | 546 (16.9%) | 126 (16.9%) | 354 (13.8%) | 314 (19.1%) | 32 (12.2%) | 1,372 (16.2%) |
| *6* | 680 (21.1%) | 97 (13.0%) | 305 (11.9%) | 340 (20.7%) | 63 (24.0%) | 1,485 (17.6%) |
| *7* | 1,058 (32.8%) | 143 (19.2%) | 288 (11.2%) | 430 (26.1%) | 97 (37.0%) | 2,016 (23.9%) |
| RGSC Social Class of Father when Respondent Child | | | | | | |
| *1* | 35 (1.1%) | 33 (4.4%) | 257 (10.0%) | 36 (2.2%) | 1 (0.4%) | 362 (4.3%) |
| *2* | 416 (12.9%) | 199 (26.7%) | 849 (33.0%) | 238 (14.5%) | 36 (13.7%) | 1,738 (20.6%) |
| *3* | 260 (8.1%) | 112 (15.0%) | 384 (14.9%) | 156 (9.5%) | 12 (4.6%) | 924 (10.9%) |
| *4* | 1,538 (47.7%) | 285 (38.3%) | 759 (29.5%) | 806 (49.0%) | 113 (43.1%) | 3,501 (41.4%) |
| *5* | 567 (17.6%) | 81 (10.9%) | 226 (8.8%) | 272 (16.5%) | 59 (22.5%) | 1,205 (14.3%) |
| *6* | 409 (12.7%) | 35 (4.7%) | 95 (3.7%) | 138 (8.4%) | 41 (15.6%) | 718 (8.5%) |
| CAMSIS Score of Father when Respondent Child | 38.78 (10.35) | 43.94 (13.35) | 47.87 (14.80) | 39.67 (11.04) | 37.09 (10.27) | 42.12 (12.91) |
| N | 3,225 (38.2%) | 745 (8.8%) | 2,570 (30.4%) | 1,646 (19.5%) | 262 (3.1%) | 8,448 (100.0%) |
| Data Source: NCDS [Sweeps 0-4] | | | | | | |

From Table 1.9, some observations can be made. An individual’s educational attainment is widely different when stratified by their economic activity. Those that enter employment have a split of 91 per cent having achieved less than five O’levels compared to 9 per cent of their peers that achieved five or more O'levels. The reverse is true for those who stayed at school, whereby 83 per cent of individuals achieved five or more levels. The split of educational attainment is almost identical for those who entered employment and those who entered a training and apprenticeship scheme, as well as those who entered a period of unemployment or out of the labour force. It is only non-traditional education that has a somewhat even split between individuals who have achieved five or more levels and those who have not.

From observing the descriptive statistics, economic activity is stratified heavily by sex. Whilst there is an even split of women and men staying on in school, those who decide to enter training and apprenticeship schemes are dominated by men (82 per cent). Comparatively, those who enter non-traditional education are primarily women (70 per cent). Women are also a majority in entering employment and being unemployed or out of the labour force. Men are only the majority in one economic activity category – training and apprenticeships.

Those who lived with parents who did not own their own homes make up the majority (72 per cent) of individuals who are unemployed and out of the labour force category. Most individuals who chose to stay within school or go to non-traditional education also had parents who owned their own home (at rates of 68 per cent and 62 per cent, respectively). For those who chose employment or training and apprenticeship schemes, most came from parents who did not own their own home (at 67 per cent and 57 per cent, respectively.

Looking at NS-SEC, the largest concentration of NS-SEC 1.1, 1.2 and 3 is concentrated within the school category at 6.2 per cent, 11.3 per cent, and 14 per cent, respectively. The lowest concentrations of 1.1, 1.2 and 3 are concentrated within the unemployment and out of the labour force category at 1.5 per cent, 0.4 per cent, and 5.7 per cent, respectively. NS-SEC 2 shares the highest concentration within schools at 21.2 per cent, but unlike NS-SEC 1.1, 1.2, and 3, its lowest concentration is within employment at 6.6 per cent. The largest concentration of NS-SEC 4 and 5 are within training and apprenticeship programs at 14.6 per cent and 19.1 per cent, respectively; however, they deviate from each other concerning their lowest concentration. For NS-SEC 4, the lowest concentration is within schools at 10.5 per cent, whereas for NS-SEC 5, it is within unemployment and out of the labour force at 12.2 per cent. NS-SEC 6 and 7 share the highest concentration of individuals within unemployment and out of the labour force at 24 per cent and 37 per cent, respectively, and a shared lowest concentration within school at 11.9 per cent and 11.2 per cent, respectively. Looking at NS-SEC within each economic activity, there is a linear increase in individuals participating in employment as NS-SEC increases (from 1.1-7). When looking at non-traditional education and comparing the per cent of each NS-SEC category to their total, it is evident that NS-SEC 1.1, 2, and 3 are overrepresented in this economic activity, whilst NS-SEC 4-7 are underrepresented. For schools, NS-SEC 1.1-3 are overrepresented in this economic activity outcome, whereas NS-SEC 4-7 are underrepresented in this category. The exact reverse is true regarding the training and apprenticeship category. The only NS-SEC categories overrepresented in the unemployment and out of the labour force category are NS-SEC 6 and 7, with all else being underrepresented.

For the RGSC schema, an explicit manual/non-manual divide becomes apparent when looking at Table 1.9. Delineating RGSC 1-3 as non-manual and RGSC 4-6 as manual, for employment, training and apprenticeships, and unemployment and out of the labour force economic activities, those in manual occupations are overrepresented, and those in non-manual are underrepresented. The reverse for non-traditional education and school is genuine, whereby manual occupations are underrepresented, and non-manual occupations are overrepresented.

For CAMSIS, with a base total mean of 42.12, like RGSC, there is a delineation between categories of economic activity relating to employment, training and apprenticeships, and unemployment and out of the labour force on the one hand, and on the other, non-traditional education and school. Those in the former group have a CAMSIS mean below the total, and those in the latter have a CAMSIS mean above the total.

## **Modelling Main Economic Activity**

The primary 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 estimates a multinomial logistic regression model with NS-SEC as the chosen social stratification measure (RGSC and CAMSIS will be introduced in a sensitivity analysis later). Table 1.10 details the deviance, change in deviance, change in degrees of freedom, and McFadden’s Adjusted Pseudo R2, AIC, and BIC measures to compare the null model with models of one explanatory variable. Table 1.10 details the exact statistics 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 and NS-SEC are simultaneously equal to 0 can be rejected at the 0.01 level.

The model output uses the reference category of the school. The schooling category contrasts with all other economic activity categories because it has the most significant barrier to entry; continuing schooling expects previous educational merit. School as a reference category is sociologically compelling. Contrasting school with other economic activity destinations like employment or apprenticeships is temporally relevant given the possible impact that increasing the mandatory school leaving age, decline in the manufacturing industry, and rise in part-time work may have on the economic destinations of youth. Less than five O’levels is the reference category for educational attainment, Female is the reference category for Sex, Own home is the reference category for housing tenure, and NS-SEC 7 is the reference category for NS-SEC. The rationale behind selecting NS-SEC 7 as the reference category is twofold: firstly, there is no agreed-upon reference category for NS-SEC, and the typically small number of observations at NS-SEC 1.1 discourages some from using that as the reference – NS-SEC 7 has the single largest among of observations. Secondly, the desire behind this model is to attempt to understand if there is any evidence of the impact of social stratification upon economic activity sorting; one of the most intuitive ways to demonstrate this possibility is by making the most deprived or least deprived category the reference. For these reasons, NS-SEC 7 was selected as the reference category[[9]](#footnote-9).

Table 1.10 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 | 23150.59 | - | - | - | 23158.58 | 23186.75 |
| Null Model + Educational Attainment | 18908.96 | 4241.63 | 4 | 0.18 | 18924.96 | 18981.29 |
| Null Model + Sex | 22159.83 | 990.76 | 4 | 0.04 | 22175.83 | 22232.16 |
| Null Model + Tenure | 22328.74 | 821.85 | 4 | 0.03 | 22344.74 | 22401.07 |
| Null Model + NS-SEC | 21876.24 | 1274.35 | 28 | 0.05 | 21940.24 | 22165.57 |

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

Table 1.11 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 | 23150.59 | - | - | - | 23158.58 | 23186.75 |
| Null Model + Educational Attainment | 18908.96 | 4241.63 | 4 | 0.18 | 18924.96 | 18981.29 |
| Null Model + Educational Attainment + Sex | 17921.39 | 987.57 | 4 | 0.22 | 17945.39 | 18029.89 |
| Null Model + Educational Attainment + Sex + Tenure | 17677.49 | 243.9 | 4 | 0.23 | 17709.49 | 17822.16 |
| Null Model + Educational Attainment + Sex + Tenure + NS-SEC | 17417.09 | 260.4 | 28 | 0.24 | 17505.08 | 17814.92 |

The model fit statistics demonstrate that there are typically distributed residuals and that the model is correctly specified. Table 1.11 suggests that deviance is reduced by 5,733.50 from the null for the full proposed model. AIC and BIC statistics also suggest that the full model best fits those entered. Finally, the full model presents an adjusted pseudo-R2 of 0.24. In other words, the full model explains 24 per cent of the variance of economic activity, leaving 76 per cent unexplained. The following analysis with the full model is a complete records analysis with 8,448 observations.

Before discussing this model's results, a discussion on interpretation must be had. When dealing with multinomial logistic regression, results in the form of coefficients are reported in the default Stata output as log odds. Log odds are notoriously tricky 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 considered the effect on the log odds of moving from the reference category to the category of the X variable. Due to this difficulty in interpretation, some (Norton and Dowd, 2018) have advocated for using odds ratios over log odds. However, odds ratios have their issues, which result in an inability to compare across models and datasets, even if they have the exact model specification (ibid). Sometimes, odds ratios cannot be compared and interpreted within a model (ibid). This issue stems from odds ratios changing if variables are added to the model, even if such additional variables are independent of the other variables. Due to these issues, both log odds and odds ratios provide an underwhelming desire to use them to interpret multinominal logistic models beyond establishing primary substantive effects of ‘higher’ and ‘lower’ (Gayle and Lambert, 2009). The popular alternative to using logs odds and odds ratios is the average marginal effect of an explanatory variable on the probability that equals 1 versus 0. In the case of this model, the average change in probability of being in an economic activity category over school, holding all other variables at their observed values. The rationale for interpreting multinomial logistic models using average marginal effects is based on the fact that the marginal effect is less sensitive to changes in model specification than the odds ratio, and the average marginal effect can be either positive or negative. Finally, average marginal effects for subgroups (like social class) can differ, leading to different implications and interpretations (Norton and Dowd, 2018).

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

The results of the multinomial logistic regression model are reported in Table 1.12. It is impossible to ascertain the significance of variables' parameters other than 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 1.12. 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 the multinominal logistic regression models in this chapter – QV estimates are only produced for the first category in the categorical outcome variable, nor does it work with the sub-command ‘’ib().” that is used to identify a specific reference category of a chosen variable such as NS-SEC – this is because the ‘qv’ command predates the implementation of the subcommand ‘’ib().”. The creation of quasi-variance statistics can be completed 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 producing quasi-variance statistics outside of Stata 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. An alternative solution was identified that did not break the workflow and was committed within Stata. The ‘’ib().” The subcommand issue can be overcome by recoding NS-SEC whereby the reference category is first – in this case, recoding NS-SEC 7 as NS-SEC 1 so that Stata is forced to use that category as the reference. Recoding the outcome variable three times so that each outcome category of the outcome variable is coded once as '1' means that the QV statistics can be procured for each category and then combined later. QV statistics and QV graphs for each category of the outcome variable can be produced within the Stata environment. QV statistics are reported alongside log odds and average marginal effects, whilst QV graphs are reported below.

The output for employment demonstrates that individuals who received five or more O’levels have decreased log odds of employment over the school. Using average marginal effects, there is a 39 per cent decreased probability for an individual to be employed over school if they received five or more O’levels. Educational attainment has the most substantial impact on an individual’s choice to be employed over school. Men had a decreased log odds of being in employment over school. Regarding average marginal effects, this translates to a 17 per cent decreased probability that men would be employed over school. Individuals with parents who did not own their own home when a child had increased log odds of employment compared to school. Translated into average marginal effects, this represents an 8 per cent increased probability for an individual to be employed over school if they lived in a household where their parents did not own their own home as a child. Using NS-SEC 7 as a reference category, every other NS-SEC category has a decreased log odds of employment over school compared to NS-SEC 7. The most substantial single impact on an individual’s choice to bring in employment over the school in respect to NS-SEC categories is NS-SEC 1.2, whereby there is a 12 per cent decreased probability of an individual whose social origins in 1.2 would be in employment over the school in comparison to NS-SEC 7. The weakest single impact concerning NS-SEC categories (and the category of employment overall) is NS-SEC 6, with only a 4 per cent decreased probability of being employed over school compared to NS-SEC 7. For a full breakdown of the marginal effects of NS-SEC in the employment category, see Figure 1.2 and its explanation.

The output for non-traditional education demonstrates that individuals who received five or more O’levels had increased log odds of being in non-traditional education over school. Using average marginal effects, there is a 3 per cent increased probability for an individual to be in non-traditional education over employment if they received five or more O’levels. Men had a decreased log odds of being in non-traditional education over school than women, or in terms of average marginal effects, a 7 per cent decreased probability of being in non-traditional education over employment if the individual is a man. Sex is the single most substantial impact on an individual’s choice to enter non-traditional education. Individuals who did not own their own home compared to those who did have decreased odds of being in non-traditional education over school. In other words, there is a 4 per cent decreased probability of being in non-traditional education over school if the individual resided in a home their parents did not own. Results suggest that individuals whose parents have higher NS-SEC social class occupations (from 1.1-3) have an average marginal predicted probability of being more likely to be in non-traditional education compared to school than their NS-SEC 7 peers. For a full breakdown of social class across each economic activity category, see Figure 1.2 for the predicted probabilities. The 95 per cent quasi-variance comparison intervals suggest significant differences in non-traditional education activity of children whose fathers occupy different NS-SEC categories.

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 school, in terms of average marginal effects, this corresponds to a decreased probability of 17 per cent. Men compared with women had an increased log odds of being in training & apprenticeships over employment, or a 24 per cent increased probability. Sex is the single strongest predictor of whether an individual chooses to enter training and apprenticeship schemes over school. Results suggest that individuals that did not own their own home compared to those that did have a decreased log odds of being in training & apprenticeships over school. 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 their parents do not own over people that do. Results also suggest that individuals whose fathers were in NS-SEC 1.1-3 had a decreased log odds of being in training & apprenticeships over the school.

The output for unemployment & OLF demonstrates that individuals who received five or more O’levels had decreased log odds of being unemployed and out of the labour force compared to those in school. This translates into a 3 per cent decreased probability of being unemployed and out of the labour force compared to being in school. Men are less likely to be unemployed or out of the labour force than women, with decreased log odds translating to a 1 per cent decreased probability. Individuals from households where their parents do not own their own home have increased odds of being unemployed and out of the labour force compared to being in school. Regarding average marginal effects, this corresponds to a 1 per cent increased probability of being unemployed and out of the labour force compared to being in school. NS-SEC 1.1-5 all have decreased log odds compared to NS-SEC 7 of being unemployed and out of the labour force compared to being in school.

Table 1.12 Mlogit of Economic Activity

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | NS-SEC | | | Average Marginal Effects | | Quasi-variance | | |
| Economic Activity: ‘School’ Reference Category | **Coef.** | **S.E.** | **Sig.** | **Prob.** | **S.E.** | **S.E.** | **LCI** | **UCI** |
| Employment |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Five or More O’levels* | -3.59 | (0.08) | \*\*\* | -0.39 | (0.01) | (.) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Male* | -0.40 | (0.07) | \*\*\* | -0.17 | (0.01) | (.) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Do not Own Home* | 0.69 | (0.08) | \*\*\* | 0.08 | (0.01) | (.) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |  |
| *1.1* | -0.01 | (0.22) |  | 0.01 | (0.03) | 0.20 | -0.41 | 0.40 |
| *1.2* | -0.67 | (0.22) | \*\* | -0.05 | (0.03) | 0.20 | -1.06 | -0.27 |
| *2* | Ref. | (.) |  | (.) | (.) | 0.10 | -0.21 | 0.21 |
| *3* | 0.35 | (0.15) | \* | 0.05 | (0.02) | 0.10 | 0.13 | 0.57 |
| *4* | 0.96 | (0.14) | \*\*\* | 0.08 | (0.02) | 0.10 | 0.75 | 1.16 |
| *5* | 0.86 | (0.14) | \*\*\* | 0.07 | (0.02) | 0.09 | 0.69 | 1.04 |
| *6* | 0.91 | (0.14) | \*\*\* | 0.09 | (0.02) | 0.09 | 0.73 | 1.09 |
| *7* | 1.36 | (0.13) | \*\*\* | 0.13 | (0.02) | 0.08 | 1.19 | 1.53 |
| Intercept | 0.89 | (0.12) | \*\*\* | (.) | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |  |
| Non-Traditional Education |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Five or More O’levels* | -1.71 | (0.09) | \*\*\* | 0.03 | (0.01) | (.) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Male* | -0.84 | (0.09) | \*\*\* | -0.07 | (0.01) | (.) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Do not Own Home* | -0.12 | (0.10) |  | -0.04 | (0.01) | (.) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |  |
| *1.1* | -0.07 | (0.22) |  | -0.00 | (0.02) | 0.20 | -0.48 | 0.33 |
| *1.2* | -0.52 | (0.21) | \* | -0.01 | (0.02) | 0.18 | -0.89 | -0.15 |
| *2* | Ref. | (.) |  | (.) | (.) | 0.10 | -0.21 | 0.21 |
| *3* | 0.06 | (0.16) |  | -0.01 | (0.01) | 0.12 | -0.19 | 0.30 |
| *4* | 0.15 | (0.17) |  | -0.03 | (0.01) | 0.13 | -0.11 | 0.41 |
| *5* | 0.24 | (0.15) |  | -0.02 | (0.01) | 0.11 | 0.02 | 0.46 |
| *6* | -0.02 | (0.16) |  | -0.04 | (0.01) | 0.12 | -0.27 | 0.23 |
| *7* | 0.44 | (0.15) | \*\* | -0.03 | (0.01) | 0.11 | 0.22 | 0.66 |
| Intercept | 0.18 | (0.13) |  | (.) | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |  |
| School | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Five or More O’levels* | -3.24 | (0.09) | \*\*\* | -0.17 | (0.01) | (.) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Male* | 1.52 | (0.09) | \*\*\* | 0.24 | (0.01) | (.) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Do not Own Home* | 0.38 | (0.08) | \*\*\* | -0.01 | (0.01) | (.) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |  |
| *1.1* | -0.14 | (0.25) |  | -0.01 | (0.03) | 0.23 | -0.60 | 0.32 |
| *1.2* | -0.45 | (0.23) |  | 0.00 | (0.03) | 0.20 | -0.86 | -0.04 |
| *2* | Ref. | (.) |  | (.) | (.) | 0.12 | -0.24 | 0.24 |
| *3* | 0.17 | (0.17) |  | -0.00 | (0.02) | 0.13 | -0.08 | 0.43 |
| *4* | 0.97 | (0.16) | \*\*\* | 0.05 | (0.02) | 0.11 | 0.75 | 1.20 |
| *5* | 0.90 | (0.15) | \*\*\* | 0.05 | (0.02) | 0.10 | 0.70 | 1.09 |
| *6* | 0.84 | (0.15) | \*\*\* | 0.04 | (0.02) | 0.10 | 0.70 | 1.04 |
| *7* | 1.08 | (0.15) | \*\*\* | 0.02 | (0.02) | 0.09 | 0.90 | 1.27 |
| Intercept | -0.72 | (0.14) | \*\*\* | (.) | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Five or More O’levels* | -3.97 | (0.27) | \*\*\* | -0.03 | (0.00) | (.) | (.) | (.) |
| Sex |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Male* | -0.37 | (0.14) | \*\* | -0.01 | (0.00) | (.) | (.) | (.) |
| Housing Tenure |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) | (.) |
| *Do not Own Home* | 0.87 | (0.16) | \*\*\* | 0.01 | (0.00) | (.) | (.) | (.) |
| NS-SEC |  |  |  |  |  |  |  |  |
| *1.1* | -0.18 | (0.57) |  | -0.00 | (0.02) | 0.53 | -1.24 | 0.89 |
| *1.2* | -2.07 | (1.03) | \* | -0.03 | (0.01) | 1.01 | -4.11 | -0.04 |
| *2* | Ref. | (.) |  | (.) | (.) | 0.24 | -0.48 | 0.48 |
| *3* | -0.11 | (0.36) |  | -0.01 | (0.01) | 0.27 | -0.66 | 0.45 |
| *4* | 0.60 | (0.31) |  | -0.00 | (0.01) | 0.21 | 0.19 | 1.02 |
| *5* | 0.29 | (0.30) |  | -0.01 | (0.01) | 0.19 | -0.10 | 0.68 |
| *6* | 0.77 | (0.28) | \*\* | 0.00 | (0.01) | 0.15 | 0.47 | 1.07 |
| *7* | 1.20 | (0.27) |  | 0.01 | (0.01) | 0.13 | 0.94 | 1.46 |
| Intercept | -1.48 | (0.25) | \*\*\* | (.) | (.) | (.) | (.) | (.) |
|  |  |  |  |  |  |  |  |  |
| Number of observations | 8448 | | | | | | | |
| McFadden’s R2 | 0.25 | | | | | | | |
| McFadden’s Adjusted Pseudo R2 | 0.24 | | | | | | | |
| Cox-Snell Pseudo R2 | 0.49 | | | | | | | |
| Nagelkerke Pseudo R2 | 0.53 | | | | | | | |
| AIC | 17505.08 | | | | | | | |
| BIC | 17814.92 | | | | | | | |
| \*\*\* p<.001, \*\* p<.01, \* p<.05 Data Source: NCDS [Sweeps 0-4]  Note: Complete Records Analysis | | | | | | | | |

To understand this in a more manageable format, each variable is graphically visualised with their predicted probabilities. This allows for a more intuitive understanding of the multinominal logistic regression and provides a different outlook for interpretation. Graphing predicted probabilities by a 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 compared.

Focusing on NS-SEC, figure 1.2 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. With the exception of 1.1-1.2, where there is a slight decline in people going into employment, there is a general upward trend for individuals to be employed from NS-SEC 1.2-7. For post-school education, there is a relative flatline across the NS-SEC schema. For the school category, there is an almost identical reverse picture of what happens with individuals sorting into the employment category. There is an increase from 1.1-1.2, and from that point onwards, there is a decline from NS-SEC 1.2-7. Training & apprenticeships has a unique dichotomous state. NS-SEC 1.1-3 are split into one group, and NS-SEC 4-7 into another. The former group has a decreased likelihood of entering this category compared to the latter. The peak of this group comes from NS-SEC 4, characterised by small employers and own account workers. The results from this model reaffirm prior research on the division of opportunity within skilled and non-skilled labour as it pertains to training & apprenticeships schemes. Skilled workers in NS-SEC 4 have the highest probability of being in a training and apprenticeship scheme compared to all other categories in the NS-SEC. This is not entirely surprising, given the socio-historical context of the NCDS. Apprenticeship & training schemes were heavily influenced by skilled worker occupations (Booth and Satchell, 1994). Unemployment has a relative flatline effect.

Figure 1.2 Predicted Probabilities of Economic Activity by NS-SEC

A graph showing the number of probabilities of economic activity

Description automatically generated

Moving on to Sex, figure 1.3 demonstrates that women are more likely than men to enter employment post-mandatory schooling. This is also true for less traditional non-traditional education and schools (though comparatively, the effect sizes are much smaller for schools than for non-traditional education). There is a relatively identical number of men compared to women entering school post-mandatory education. The most considerable sex-based effect relates to men being more likely to enter training & apprenticeship pathways compared to women. Given that the ‘feminisation’ 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.

Figure 1.3 Predicted Probabilities of Economic Activity by Sex

A graph showing the number of probabilities of economic activity

Description automatically generated

Moving on to educational attainment, figure 1.4 demonstrates that effect sizes are comparatively larger when discussing attainment over other variables. Those who gained less than five O’levels at school were more likely to enter employment than their peers who did gain five or more O’levels. Predictably, those individuals who gained five or more O’levels were more likely to enter school pathways and less likely to enter employment pathways compared to their peers who received less than five O’levels. Those who earned less than five O’levels were more likely than their academically successful peers to enter training and apprenticeship and non-traditional education pathways. This suggests that for the NCDS cohort, even those who did not do well academically, there were various pathways above and beyond entering employment straight after mandatory schooling ended. For those who did want to continue their education in some manner, pathways were available to them that were not rigidly restricted based on academic merit at earlier life stages.

Figure 1.4 Predicted Probabilities of Economic Activity by Educational Attainment

A graph showing the number of probabilities of economic activity

Description automatically generated

Moving on to housing tenure, figure 1.5 demonstrates that whilst substantive findings exist, the effect sizes compared to other variables are the smallest. Those who 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 non-traditional education or school-based pathways.

Figure 1.5 Predicted Probabilities of Economic Activity by Housing Tenure

A graph showing the number of probabilities of economic activity

Description automatically generated

Alongside the graphical presentation of predicted probabilities, the following figures also visualise the log odds of NS-SEC within each outcome category (except the reference category of education) alongside quasi-variance statistics to overcome the reference category problem.

A graph showing the number of numbers

Description automatically generated with medium confidenceA graph showing the number of numbers

Description automatically generated with medium confidenceA graph with red and black lines

Description automatically generatedA graph showing the number of numbers

Description automatically generated with medium confidence

### **Discussion and Conclusion**

The multinomial logistic regression model indicates that structural inequalities do indeed have an impact on an individual’s choice of sorting into economic activity post-mandatory schooling. Educational attainment was the single most significant effect upon individuals entering into employment and unemployment and out of the labour force compared to school. Sex was the single most significant effect upon individuals entering into non-traditional education and training and apprenticeship programs in comparison to school. NS-SEC had a persistent impact on individual activity sorting post-mandatory education. However, this social class impact is less pronounced than educational attainment or sex. Housing tenure plays a small but statistically significant role in all but one of the outcome destinations for post-mandatory schooling youth. The overall conclusion from this model is that structures do matter. However, some structures matter more than others, and this influence changes depending on what type of economic activity is being discussed.

These findings have several implications for 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 apparent 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 enter employment. Those in NS-SEC 1.1-2 have over a 10 per cent decreased likelihood of being in employment over the school in comparison to their NS-SEC 7 peers. Compared with another structural dimension – such as housing tenure, which only makes up an 8 per cent increased probability of being in employment over school if living in a household that does not own its own home it becomes evident that some structural inequalities matter substantively more than others. Another key example of this relates to the training & apprenticeship pathways. There is a much more pronounced sex-based effect here compared to, say, social class – with men much more likely (24 per cent) to choose this pathway compared to women. Compare this rather substantively significant effect with the impact of being in NS-SEC 1.1 over NS-SEC 7 (4 per cent), and it provides evidence that the influence of certain areas of social stratification depends upon the given type of economic activity. These varied structural effects speak to the complex socio-historical context of the NCDS.

Social class and sex were not the only structural inequalities in 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 central tenant of new structuralism –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 an 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 structuralism arguments that social class has become disaggregated do not find empirical evidence amongst the NCDS cohort, the view that housing tenure is essential in influencing pathway choice does find some support. As stated previously, however, the arguments of new structuralism were primarily borne out of the 1980s. The fact that there are any housing tenure effects upon the NCDS cohort is worth noting going forward with comparative analysis in later chapters.

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, mean this is not feasible. Alongside this, combining 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, 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 8448 individuals. This amounts to 67 per cent of 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 to understand the potential impact missingness has had on 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). However, 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 have already been discussed. Given the historical nature of the NCDS cohort, a sensitivity analysis would provide an exciting insight into 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. While it is not appropriate to compare log odds across regression models, the following sensitivity analysis will 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 have 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 multinomial logistic regressions are presented in Table 1.13. 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 are all presented using log odds and average marginal effects to enhance interpretation and comparison.

Focusing first upon the outcome category of employment, NS-SEC and RGSC exhibit identical substantive findings concerning educational attainment, sex, and housing tenure. The CAMSIS model provides identical findings compared to the NS-SEC model for sex but has differences in the substantive findings for both educational attainment and housing tenure – though these differences are minimal (1 per cent and 2 per cent, respectively). The most significant substantive disparity comes from the interpretation of the differing social stratification variables themselves. Whilst NS-SEC, CAMSIS, and RGSC all show a general pattern that characterises a decreased log odds of being in employment over education as NS-SEC, RGSC, or CAMSIS increases (either in categorical scale or numerical scale), the substantive interpretation from these models is different. The most evident example is a comparison between NS-SEC and CAMSIS models. While CAMSIS concurs with the NS-SEC model that there is a decreased log odds of being in employment over school, this log odds effect is near zero. Translated into average marginal effects, this presents a 0 per cent increased probability of being employed over the school for each unit increase in CAMSIS. This is a substantively different finding to the NS-SEC model that states there is anywhere from a 4 per cent to 18 per cent decreased probability of being in education over school compared to NS-SEC 7 peers. The NS-SEC and RGSC models, on the other hand, whilst differing on exact numbers, provide the same substantive conclusions to one another. There is a general trend that as you go up the class schema, there is a lower probability of being in employment over school compared to their reference category peers.

Moving on to the non-traditional education category and with one exception (with a change of 1 per cent in the CAMSIS model) there are identical findings for educational attainment, sex, and housing tenure. There are, however, substantive differences in the interpretations of each social stratification measure. Whilst NS-SEC sees a small but statistically significant increased probability for those individuals that occupy the top portion of the class schema being in non-traditional education over school compared to their reference category peers, the CAMSIS model presents a 1 per cent decreased probability in individuals being in non-traditional education over school as it increases. This is generally the opposite substantive finding of the NS-SEC model. The NS-SEC model and RGSC model also differ – albeit slightly. The RGSC model only finds an increased probability of individuals being in non-traditional education over school for those that occupy the very top of the RGSC class schema – NS-SEC presents an increased likelihood of being in non-traditional education over school for a much broader range than RGSC does (1.1-3).

For the training & apprenticeship and the unemployment & out of the labour force categories all three models present identical findings for educational attainment, sex, and housing tenure. NS-SEC, CAMSIS, and RGSC models also share similar substantive findings using their social stratification measures.

The goodness-of-fit statistics are similar for all three models. Differences in R2 measures exist, but the minor nature of these differences indicates that the amount of variance explained across the three models remains consistent. AIC and BIC differences are also minor. The most parsimonious model is the NS-SEC model when using AIC and the CAMSIS model when using BIC. Considering that BIC penalises models for estimating additional parameters, it is not entirely surprising that it considers the CAMSIS a better fit than the NS-SEC schema. These differences are, however, minimal. Given the minor substantive differences found within the two models (mainly located within the employment category) alongside the slight preference for NS-SEC using AIC, there is a general preference towards the NS-SEC model. As such, going forward, the preferred model of choice for subsequent analysis will be the NS-SEC model.

Table 1.13 Sensitivity analyses of alternative measures of parental social stratification

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | NS-SEC | | | | | CAMSIS | | | | | RGSC | | | | |
| Economic Activity | **Coef.** | **S.E** | **Sig.** | **Prob.** | **S.E** | **Coef.** | **S.E** | **Sig.** | **Prob.** | **S.E** | **Coef.** | **S.E** | **Sig.** | **Prob.** | **S.E** |
| Employment |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -3.59 | (0.08) | \*\*\* | -0.39 | (0.01) | -3.63 | (0.08) | \*\*\* | -0.40 | (0.01) | -3.60 | (0.08) | \*\*\* | -0.39 | (0.01) |
| Sex |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) |  |  |  | (.) | (.) |  |  |  | (.) | (.) |
| *Male* | -0.40 | (0.07) | \*\*\* | -0.17 | (0.01) | -0.40 | (0.07) | \*\*\* | -0.17 | (0.01) | -0.40 | (0.07) | \*\*\* | -0.17 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) |  |  |  | (.) | (.) |  |  |  | (.) | (.) |
| *Do not Own Home* | 0.69 | (0.08) | \*\*\* | 0.08 | (0.01) | 0.79 | (0.07) | \*\*\* | 0.10 | (0.01) | 0.71 | (0.08) | \*\*\* | 0.08 | (0.01) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.01 | (0.22) |  | 0.01 | (0.03) |  |  |  |  |  |  |  |  |  |  |
| *1.2* | -0.67 | (0.22) | \*\* | -0.05 | (0.03) |  |  |  |  |  |  |  |  |  |  |
| *2* | Ref. | (.) |  | (.) | (.) |  |  |  |  |  |  |  |  |  |  |
| *3* | 0.35 | (0.15) | \* | 0.05 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *4* | 0.96 | (0.14) | \*\*\* | 0.08 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *5* | 0.86 | (0.14) | \*\*\* | 0.07 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *6* | 0.91 | (0.14) | \*\*\* | 0.09 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *7* | 1.36 | (0.13) | \*\*\* | 0.13 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| CAMSIS |  |  |  |  |  | -0.03 | (0.00) | \*\*\* | -0.00 | (0.00) |  |  |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1* |  |  |  |  |  |  |  |  |  |  | -0.96 | (0.23) | \*\*\* | -0.09 | (0.03) |
| *2* |  |  |  |  |  |  |  |  |  |  | Ref. | (.) |  | (.) | (.) |
| *3NM* |  |  |  |  |  |  |  |  |  |  | -0.00 | (0.13) |  | -0.01 | (0.02) |
| *3M* |  |  |  |  |  |  |  |  |  |  | 0.71 | (0.10) | \*\*\* | 0.05 | (0.01) |
| *4* |  |  |  |  |  |  |  |  |  |  | 0.80 | (0.13) | \*\*\* | 0.06 | (0.02) |
| *5* |  |  |  |  |  |  |  |  |  |  | 1.15 | (0.16) | \*\*\* | 0.13 | (0.02) |
| Intercept | 0.89 | (0.12) | \*\*\* |  |  | 2.90 | (0.15) | \*\*\* |  |  | 1.15 | (0.10) | \*\*\* |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Non-Traditional Education |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -1.71 | (0.09) | \*\*\* | 0.03 | (0.01) | -1.71 | (0.09) | \*\*\* | 0.03 | (0.01) | -1.71 | (0.09) | \*\*\* | 0.02 | (0.01) |
| Sex |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | -0.84 | (0.09) | \*\*\* | -0.07 | (0.01) | -0.84 | (0.09) | \*\*\* | -0.07 | -0.07 | -0.84 | (0.09) | \*\*\* | -0.07 | -0.07 |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | -0.12 | (0.10) |  | -0.04 | (0.01) | -0.09 | (0.09) |  | -0.04 | (0.01) | -0.09 | (0.10) |  | -0.04 | (0.01) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.07 | (0.22) |  | -0.00 | (0.02) | (.) | (.) |  |  |  | (.) | (.) |  |  |  |
| *1.2* | -0.52 | (0.21) | \* | -0.01 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *2* | Ref. | (.) |  | (.) | (.) |  |  |  |  |  |  |  |  |  |  |
| *3* | 0.06 | (0.16) |  | -0.01 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *4* | 0.15 | (0.17) |  | -0.03 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *5* | 0.24 | (0.15) |  | -0.02 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *6* | -0.02 | (0.16) |  | -0.04 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *7* | 0.44 | (0.15) | \*\* | -0.03 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| CAMSIS |  |  |  |  |  | -0.01 | (0.00) | \*\*\* | 0.00 | (0.00) |  |  |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1* |  |  |  |  |  |  |  |  |  |  | -0.47 | (0.21) | \* | 0.00 | (0.02) |
| *2* |  |  |  |  |  |  |  |  |  |  | Ref. | (.) |  | (.) | (.) |
| *3NM* |  |  |  |  |  |  |  |  |  |  | 0.12 | (0.14) |  | 0.01 | (0.01) |
| *3M* |  |  |  |  |  |  |  |  |  |  | 0.24 | (0.11) | \* | -0.01 | (0.01) |
| *4* |  |  |  |  |  |  |  |  |  |  | 0.13 | (0.16) |  | -0.03 | (0.01) |
| *5* |  |  |  |  |  |  |  |  |  |  | 0.12 | (0.23) |  | -0.04 | (0.01) |
| Intercept | 0.18 | (0.13) |  |  |  | 0.79 | (0.18) | \*\*\* | (.) | (.) | 0.17 | (0.11) |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| School | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -3.24 | (0.09) | \*\*\* | -0.17 | (0.01) | -3.27 | (0.09) | \*\*\* | -0.17 | (0.01) | -3.26 | (0.09) | \*\*\* | -0.17 | (0.01) |
| Sex |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | 1.52 | (0.09) | \*\*\* | 0.24 | (0.01) | 1.51 | (0.09) | \*\*\* | 0.24 | (0.01) | 1.52 | (0.09) | \*\*\* | 0.24 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | 0.38 | (0.08) | \*\*\* | -0.01 | (0.01) | 0.46 | (0.08) | \*\*\* | -0.01 | (0.01) | 0.40 | (0.08) | \*\*\* | -0.01 | (0.01) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.14 | (0.25) |  | -0.01 | (0.03) | (.) | (.) |  |  |  | (.) | (.) |  |  |  |
| *1.2* | -0.45 | (0.23) |  | 0.00 | (0.03) |  |  |  |  |  |  |  |  |  |  |
| *2* | Ref. | (.) |  | (.) | (.) |  |  |  |  |  |  |  |  |  |  |
| *3* | 0.17 | (0.17) |  | -0.00 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *4* | 0.97 | (0.16) | \*\*\* | 0.05 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *5* | 0.90 | (0.15) | \*\*\* | 0.05 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *6* | 0.84 | (0.15) | \*\*\* | 0.04 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| *7* | 1.08 | (0.15) | \*\*\* | 0.02 | (0.02) |  |  |  |  |  |  |  |  |  |  |
| CAMSIS |  |  |  |  |  | -0.03 | (0.00) | \*\*\* | -0.00 | (0.00) |  |  |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1* |  |  |  |  |  | (.) | (.) |  |  |  | -0.52 | (0.23) | \* | 0.01 | (0.03) |
| *2* |  |  |  |  |  |  |  |  |  |  | Ref. | (.) |  | (.) | (.) |
| *3NM* |  |  |  |  |  |  |  |  |  |  | 0.13 | (0.15) |  | 0.02 | (0.02) |
| *3M* |  |  |  |  |  |  |  |  |  |  | 0.78 | (0.11) | \*\*\* | 0.04 | (0.01) |
| *4* |  |  |  |  |  |  |  |  |  |  | 0.84 | (0.14) | \*\*\* | 0.04 | (0.01) |
| *5* |  |  |  |  |  |  |  |  |  |  | 0.76 | (0.19) | \*\*\* | -0.00 | (0.02) |
| Intercept | -0.72 | (0.14) | \*\*\* |  |  | 1.08 | (0.17) | \*\*\* |  |  | -0.57 | (0.12) | \*\*\* |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -3.97 | (0.27) | \*\*\* | -0.03 | (0.00) | -4.01 | (0.26) | \*\*\* | -0.03 | (0.00) | -4.00 | (0.27) | \*\*\* | -0.03 | (0.00) |
| Sex |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | -0.37 | (0.14) | \*\* | -0.01 | (0.00) | -0.38 | (0.14) | \*\* | -0.01 | (0.00) | -0.38 | (0.14) | \*\* | -0.01 | (0.00) |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | 0.87 | (0.16) | \*\*\* | 0.01 | (0.00) | 0.95 | (0.15) | \*\*\* | 0.01 | (0.00) | 0.90 | (0.16) | \*\*\* | 0.01 | (0.00) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.18 | (0.57) |  | -0.00 | (0.02) | (.) | (.) |  |  |  | (.) | (.) |  |  |  |
| *1.2* | -2.07 | (1.03) | \* | -0.03 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *2* | Ref. | (.) |  | (.) | (.) |  |  |  |  |  |  |  |  |  |  |
| *3* | -0.11 | (0.36) |  | -0.01 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *4* | 0.60 | (0.31) |  | -0.00 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *5* | 0.29 | (0.30) |  | -0.01 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *6* | 0.77 | (0.28) | \*\* | 0.00 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| *7* | 1.20 | (0.27) |  | 0.01 | (0.01) |  |  |  |  |  |  |  |  |  |  |
| CAMSIS |  |  |  |  |  | -0.04 | (0.01) | \*\*\* | -0.00 | (0.00) |  |  |  |  |  |
| RGSC |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1* |  |  |  |  |  | (.) | (.) |  |  |  | -2.03 | (1.03) | \* | -0.02 | (0.01) |
| *2* |  |  |  |  |  |  |  |  |  |  | Ref. | (.) |  | (.) | (.) |
| *3NM* |  |  |  |  |  |  |  |  |  |  | -0.67 | (0.35) |  | -0.02 | (0.01) |
| *3M* |  |  |  |  |  |  |  |  |  |  | 0.47 | (0.21) | \* | -0.00 | (0.01) |
| *4* |  |  |  |  |  |  |  |  |  |  | 0.90 | (0.24) | \*\*\* | 0.01 | (0.01) |
| *5* |  |  |  |  |  |  |  |  |  |  | 1.18 | (0.28) | \*\*\* | 0.01 | (0.01) |
| Intercept | -1.48 | (0.25) | \*\*\* |  |  | 0.82 | (0.31) | \*\* |  |  | -1.33 | (0.21) | \*\*\* |  |  |
| Number of observations | 8448 | | | | | 8448 | | | | | 8448 | | | | |
| McFadden’s Adjusted Pseudo R2 | 0.24 | | | | | 0.24 | | | | | 0.25 | | | | |
| Cox-Snell Pseudo R2 | 0.49 | | | | | 0.49 | | | | | 0.49 | | | | |
| Nagelkerke Pseudo R2 | 0.53 | | | | | 0.52 | | | | | 0.52 | | | | |
| AIC | 17505.08 | | | | | 17583.74 | | | | | 17531.12 | | | | |
| BIC | 17814.92 | | | | | 17724.57 | | | | | 17784.62 | | | | |
| \*\*\* p<.001, \*\* p<.01, \* p<.05 Data Source: NCDS [Sweeps 0-4]  Note: Complete Records 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 are some differences between social stratification measures within each model – particularly concerning the employment category of economic activity. Beyond this, however, there are relatively consistent, stable substantive findings between the models – though CAMSIS, out of the three models, is most likely to have small diverging findings.

Perhaps the most exciting finding of this sensitivity analysis stems from the near-identical results from the NS-SEC and RGSC models. The NS-SEC 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 RGSC does present some challenges to the apparent superiority of this analytical construction.

When comparing AIC statistics, there is a slight advantage for the NS-SEC model. When comparing BIC statistics, there is a slight advantage for the CAMSIS model. This would, at first, suggest that the selection of the NS-SEC or CAMSIS model for substantive interpretation going forward would be equally weighted. This would be an erroneous conclusion, however. Firstly, the differences in BIC statistics between 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 of model selection is not entirely rested upon model parsimony but also on 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 is theoretically preferable to a continuous measure such as CAMSIS. Combining 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 significant 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 following 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 primary concern being that missing data and non-response are 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 have the potential to present real issues as they relate to comprehensive data analysis. For analysis, those who exit the sample due to death or emigration are considered ‘natural’ from the original sample. Those, however, that either cannot be found, reject continued participation, etc., are individuals that we hold partial data on – utilising 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 missingness does not depend on observed or unobserved values. The second is being missed at random (MAR), meaning that given observed values, missingness does not depend on the unobserved ones. Finally, missing not at random (MNAR) means missingness depends on unobserved values (Silverwood et al. 2021). If data is found to be MAR, then approaches like multiple imputation (MI) and inverse probability weighting are made available – the former being extensively documented with the NCDS in particular (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 with a missing value in one or more of the variables included in the analysis. This is also known as Complete Records Analysis (CRA). The CRA approach is unpredictable; there is no way to know the consequences of 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 records 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[[11]](#footnote-11). 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) more straightforward 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 with the CRA model – something that, with appropriate testing and open science practices detailing the model construction, should not 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 when open science practices were not followed and fellow researchers could not access the MI seed[[12]](#footnote-12). Finally, the argument that FIML is asymptotically efficient only holds to a certain extent. MI models reach asymptotic efficiency by running an infinite number of imputations – though you can reach near full efficiency with a relatively small number of imputations, Allison ( 2015) argues, around 10[[13]](#footnote-13). 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, the subsequent analysis uses CRA 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 are sufficient? Silverwood et al. (2021) suggest that anything around 50 imputations would be sufficient for reliable estimation of the point estimate and estimating p-values with little error. 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 with the sweep at age 23. After using the variable related to the outcome of tracing and interviews, there are 12,536 individuals within this sample[[14]](#footnote-14). 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 and 3779 on NS-SEC. 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 1.14. Within the NCDS sample, 67 per cent have complete records on all variables, 17 per cent are missing values at socio-economic measures, a further 13 per cent are missing on socio-economic measures and housing tenure, and 2 per cent are missing at housing tenure. Further missingness in the sample not presented in the table is <1 per cent. In total, 8,448 cases have a complete observation of all variables.

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 participate in the survey are missing from Sweep 4 (age 23).

Table 1.14 Missing data patterns for NCDS

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N | Percent Complete (%) | Educational Attainment | Economic Activity | Housing Tenure | NS-SEC |
| 8448 | 67 | **✓** | **✓** | **✓** | **✓** |
| 2201 | 17 | **✓** | **✓** | **✓** |  |
| 1636 | 13 | **✓** | **✓** |  |  |
| 251 | 2 | **✓** | **✓** |  | **✓** |
| Total = 12536 |

Of the missing data, 63 per cent of cases are complete, and the largest proportion of missing data comes from the three socio-economic variables: NS-SEC, CAMSIS, and RGSC. Individuals who are inactive in the labour force find it challenging to code within socio-economic measures. NS-SEC categorisation is based on occupation type, authority duties, and the organisation's size (Rose and Pevalin, 2002). CAMSIS occupational scores only relate to the working population within a country. The RGSC also suffers similar problems concerning categorisation of retired, unemployed, and out of the labour force individuals. Unemployment rates during this 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 the potential missingness of the largest missing variables in the model.

A complete records 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 for dealing with MAR-related data is MI (Treiman, 2009). The following section seeks to compare a CRA 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 simultaneously. 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 can be adjusted for complex survey design.

While multiple imputation does help with missingness, it has some drawbacks. Goodness-of-fit statistics, for example, cannot be used – R2 and BIC are the most prominent. Therefore, it is not possible to assess the more appropriate or parsimonious model – it is simply possible to compare the substantive effects between a complete records analysis and a multiple imputation model. For multiple imputation models to be compared to a complete records analysis, the former needs to be ‘‘congenial’’ (White, Royston and Wood, 2011) with the latter. Congeniality or consistency in this respect means that the same variables in the complete record analysis are identical to those included in multiple imputation. Suppose the variables between complete records analysis and multiple imputation models differ. In that case, the correct variance/covariance matrix will not be estimated, and a substantive comparison between the two will become impossible and impracticable due to a loss of statistical power (Von Hippel, 2009; Lynch and Von Hippel, 2013).

Multivariate imputation by chained equations (MICE) is a form of multiple imputation that fills in or imputes missing data within a given dataset through iterative predictive models or *k* imputations. This specification is required when imputing a variable that must only take on specific values, such as the categorical nature of the economic activity response variable within the current analytical model. Using MICE, each imputation *k* is drawn from the posterior distribution of the parameters in the given imputation model, and then the model itself is imputed (Carpenter and Kenward, 2012). To create the *k*th imputation, new parameters are drawn from the posterior distribution. Multiple Imputation following MICE draws from Bayesian influences on the distribution of missing data upon observed data. An essential advantage of Multiple Imputation is that it can be applied for data missing at the response variable or its covariates (Carpenter and Kenward, 2012).

Choosing the number of imputations is difficult. Previous literature suggests that anywhere between 3-5 imputations is sufficient to obtain acceptable properties (Carpenter and Kenward, 2012). However, some modern literature suggests closer to 50 imputations (Silverwood *et al.*, 2021). However, if there is a desire to estimate small p-values or have an MI estimator of the fraction of missing information, greater imputations are required. Carpenter and Kenward (2012) suggest two routes. If an analysis after imputation is clear-cut after a small number of imputations, there is no need to perform more. If, however, after imputation, the inference is less clear-cut, take K = 100, or 100 imputations. Others promote a slightly different interpretation. White et al. (2010) and Bodner (2008) suggest using the Fraction of Missing Information (FMI) as a baseline for the minimum required imputations. If the maximum FMI in a given model is 44 per cent, then 44 imputations are suggested at minimum. When following this assumption, White et al. (2010) found that standard errors and p-values were considerably reduced and stabilised.

After Multiple Imputation is performed, four key statistics are relevant to focus upon: variance total, Relative Variance Increase (RVI), Fraction of Missing Information (FMI), and Relative Efficiency (RE).

The primary usefulness of multiple imputation relies upon its variance estimation. The total variance in multiple imputation is the sum of multiple sources of variance: within imputation variance, between imputation variance and additional sampling variance. The latter is calculated by the within-imputation variance divided by the number of imputations. The variance total is directly related to how standard errors are calculated. Unlike simple imputation methods, multiple imputation estimates SEs so that the SEs for each parameter estimate are the square root of their variance totals.

The RVI or Relative Variance Increase is the proportional increase in total sampling variance due to missing information. Any variable that has a large amount of missingness or is weakly correlated with other variables in the imputation model tends to have larger than average RVIs. Weakly correlated auxiliary variables will always trend towards large RVIs.

The FMI is related to the RVI (which, in turn, is related to the variance total). The FMI is the proportion of the total sampling variance due to missing data. It is estimated based on the percentage of missingness for a particular variable and how correlated this variable is with other variables in the imputation model. When a variable has a high FMI, this can indicate a problematic variable, which may cause convergence issues.

Finally, the relative efficiency or RE relates to how well the actual population parameters are estimated. It is related to both the amount of missingness as well as the number of imputations within an imputation model. The RE is a comparative measure. It compares the relative efficiency of the current model variable to performing an infinite number of imputations. It is relatively easy to achieve a high RE on a given imputation model with few imputations; however, this does not mean that the standard errors within the given imputation model will be calculated accurately.

Auxiliary variables are variables in the data set that are either correlated with a missing variable or variables but are not a part of the primary analytical model of interest. They are included within the imputation model to increase accuracy and statistical power to make the MAR assumption more plausible. Making the MAR assumption more plausible is done by including auxiliary variables – variables that can predict missingness on a given variable. Auxiliary variables are essential when there are high levels of missingness upon a given variable (Johnson and Young, 2011; Young and Johnson, 2011). There is no strict threshold for what an auxiliary variable needs to be included within the imputation; however, some have recommended an r > 0.4 on at least one of the analytical variables within the model (Allison, 2012a). However, this is disputed (Enders, 2010). Others, such as Silverwood et al. (2021), argue that if an auxiliary variable is predictive of the outcome variable, it makes them suitable for inclusion within the imputation model. An auxiliary variable does not have the requirement that the given variable has to have complete information to be valuable – auxiliary variables can still be influential when they have missingness (Enders, 2010).

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 are 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 the child, social class of mother’s husband, family moves since child’s birth, dad reads to a 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.

The imputation model naturally includes all analytical variables included in the previous chapter. The imputation model also includes several auxiliary variables to add to 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 derived mainly 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 primarily driven by sweep non-response rather than item non-response means it is even more critical 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 NCDS guide on handling missing data (ibid). From this, 8 out of 18 variables are substantively associated with economic activity and are included as auxiliary variables in the imputation model.

Prior to imputation, it is best to explore the distribution of variables compared to complete and non-complete cases. In the presence of an 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. 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.

With all the variables in the model being categorical, convergence issues are a possibility. 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 to produce an imputed model[[15]](#footnote-15).

After performing the imputation, it is often helpful to graph the means and standard deviations saved through the tracing subcommand when using MICE – autocorrelation plots would be helpful but are only available for non-MICE related imputations. By graphing variables means and standard deviations through trace plots, for example, over each imputation, any discrepancy or deviation can easily be found. If this were to be the case, this would be problematic for the imputation model and suggest that further imputations would be required (White, Royston and Wood, 2011). The means and standard deviations of imputed values from each iteration[[16]](#footnote-16) were checked to see the distributions of each variable against the imputations. These graphs are seen below. Note that due to the sex variable having zero missingness, no graph was produced, as no imputations on that variable were required. As illustrated, all analytical variables that were imputed have a relatively stable mean and standard deviation across the iteration numbers.

A graph showing different colored lines

Description automatically generated with medium confidenceA graph showing different colored lines

Description automatically generatedA graph showing different colored lines

Description automatically generated with medium confidenceA graph showing the number of numbers

Description automatically generated with medium confidence

The following models presented will compare a complete records analysis using NS-SEC from the previous chapter and the imputed model in Table 1.15. The CRA model has 7,915 observations. Using a variable within the NCDS dataset (add what variable this is) that noted 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 8,448 observations of the CRA model.

Overall, there is a similarity between the complete records analysis and the imputed model. The substantive conclusions between CRA and MI models are nearly identical. There are some very slight differences in the log odds and average marginal effects across the variables. However, these slight differences are not large enough to impact the substantive conclusions presented in the interpretation of the CRA model. The largest single difference in average marginal effects between the CRA and the imputed model amounts to 3 per cent – a difference that ultimately does not change the substantive interpretation of the overall model. The imputed model confirms the substantive conclusions made from the CRA model with some minor variation in log odds and average marginal effects and a reduction in standard errors.

Table 1.15 Comparison of CRA NS-SEC vs Imputed NS-SEC

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CRA NS-SEC | | | Average Marginal Effects | | Imputed NS-SEC | | | Average Marginal Effects | |
| Economic Activity | **Coef.** | **S.E** | **Sig.** | **Prob.** | **S.E** | **Coef.** | **S.E** | **Sig.** | **Prob.** | **S.E** |
| Employment |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -3.59 | (0.08) | \*\*\* | -0.39 | (0.01) | -3.52 | (0.07) | \*\*\* | -0.39 | (0.01) |
| Sex |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | -0.40 | (0.07) | \*\*\* | -0.17 | (0.01) | -0.44 | (0.06) | \*\*\* | -0.17 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | 0.69 | (0.08) | \*\*\* | 0.08 | (0.01) | 0.69 | (0.07) | \*\*\* | 0.09 | (0.01) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.01 | (0.22) |  | 0.01 | (0.03) | -0.07 | (0.22) |  | 0.00 | (0.04) |
| *1.2* | -0.67 | (0.22) | \*\* | -0.05 | (0.03) | -0.73 | (0.21) | \*\*\* | -0.06 | (0.03) |
| *2* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *3* | 0.35 | (0.15) | \* | 0.05 | (0.02) | 0.35 | (0.15) | \* | 0.05 | (0.02) |
| *4* | 0.96 | (0.14) | \*\*\* | 0.08 | (0.02) | 0.90 | (0.14) | \*\*\* | 0.08 | (0.02) |
| *5* | 0.86 | (0.14) | \*\*\* | 0.07 | (0.02) | 0.84 | (0.14) | \*\*\* | 0.08 | (0.02) |
| *6* | 0.91 | (0.14) | \*\*\* | 0.09 | (0.02) | 0.88 | (0.14) | \*\*\* | 0.09 | (0.02) |
| *7* | 1.36 | (0.13) | \*\*\* | 0.13 | (0.02) | 1.39 | (0.14) | \*\*\* | 0.14 | (0.02) |
| Intercept | 0.89 | (0.12) | \*\*\* | (.) | (.) | 0.86 | (0.11) | \*\*\* | (.) | (.) |
|  |  |  |  |  |  |  |  |  |  |  |
| Non-Traditional Education |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -1.71 | (0.09) | \*\*\* | 0.03 | (0.01) | -1.68 | (0.08) | \*\*\* | 0.02 | (0.01) |
| Sex |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | -0.84 | (0.09) | \*\*\* | -0.07 | (0.01) | -0.86 | (0.08) | \*\*\* | -0.07 | (0.01) |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | -0.12 | (0.10) |  | -0.04 | (0.01) | -0.17 | (0.09) |  | -0.04 | (0.01) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.07 | (0.22) |  | -0.00 | (0.02) | -0.07 | (0.23) |  | -0.00 | (0.02) |
| *1.2* | -0.52 | (0.21) | \* | -0.01 | (0.02) | -0.47 | (0.19) | \* | -0.01 | (0.02) |
| *2* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *3* | 0.06 | (0.16) |  | -0.01 | (0.01) | 0.06 | (0.16) |  | -0.01 | (0.01) |
| *4* | 0.15 | (0.17) |  | -0.03 | (0.01) | 0.15 | (0.16) |  | -0.03 | (0.01) |
| *5* | 0.24 | (0.15) |  | -0.02 | (0.01) | 0.22 | (0.15) |  | -0.02 | (0.01) |
| *6* | -0.02 | (0.16) |  | -0.04 | (0.01) | 0.00 | (0.16) |  | -0.04 | (0.01) |
| *7* | 0.44 | (0.15) | \*\* | -0.03 | (0.01) | 0.48 | (0.15) | \*\* | -0.03 | (0.01) |
| Intercept | 0.18 | (0.13) |  | (.) | (.) | 0.13 | (0.12) |  | (.) | (.) |
|  |  |  |  |  |  |  |  |  |  |  |
| School | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
|  |  |  |  |  |  |  |  |  |  |  |
| Training & Apprenticeships |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -3.24 | (0.09) | \*\*\* | -0.17 | (0.01) | -3.20 | (0.08) | \*\*\* | -0.16 | 0.01 |
| Sex |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | 1.52 | (0.09) | \*\*\* | 0.24 | (0.01) | 1.52 | (0.08) | \*\*\* | 0.24 | 0.01 |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | 0.38 | (0.08) | \*\*\* | -0.01 | (0.01) | 0.35 | (0.08) | \*\*\* | -0.02 | 0.01 |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.14 | (0.25) |  | -0.01 | (0.03) | -0.13 | (0.25) |  | -0.01 | (0.03) |
| *1.2* | -0.45 | (0.23) |  | 0.00 | (0.03) | -0.50 | (0.23) | \* | -0.00 | (0.02) |
| *2* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *3* | 0.17 | (0.17) |  | -0.00 | (0.02) | 0.13 | (0.17) |  | -0.01 | (0.02) |
| *4* | 0.97 | (0.16) | \*\*\* | 0.05 | (0.02) | 0.91 | (0.17) | \*\*\* | 0.04 | (0.02) |
| *5* | 0.90 | (0.15) | \*\*\* | 0.05 | (0.02) | 0.86 | (0.15) | \*\*\* | 0.04 | (0.02) |
| *6* | 0.84 | (0.15) | \*\*\* | 0.04 | (0.02) | 0.81 | (0.15) | \*\*\* | 0.03 | (0.02) |
| *7* | 1.08 | (0.15) | \*\*\* | 0.02 | (0.02) | 1.07 | (0.15) | \*\*\* | 0.02 | (0.02) |
| Intercept | -0.72 | (0.14) | \*\*\* | (.) | (.) | -0.80 | (0.13) | \*\*\* | (.) | (.) |
|  |  |  |  |  |  |  |  |  |  |  |
| Unemployment & Out of Labour Force |  |  |  |  |  |  |  |  |  |  |
| Educational Attainment |  |  |  |  |  |  |  |  |  |  |
| *Less than five O’levels* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Five or More O’levels* | -3.97 | (0.27) | \*\*\* | -0.03 | (0.00) | -4.04 | (0.22) | \*\*\* | -0.04 | (0.00) |
| Sex |  |  |  |  |  |  |  |  |  |  |
| *Female* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Male* | -0.37 | (0.14) | \*\* | -0.01 | (0.00) | -0.29 | (0.11) | \*\* | -0.01 | (0.00) |
| Housing Tenure |  |  |  |  |  |  |  |  |  |  |
| *Own Home* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *Do not Own Home* | 0.87 | (0.16) | \*\*\* | 0.01 | (0.00) | 0.94 | (0.14) | \*\*\* | 0.02 | (0.00) |
| NS-SEC |  |  |  |  |  |  |  |  |  |  |
| *1.1* | -0.18 | (0.57) |  | -0.00 | (0.02) | -0.33 | (0.57) |  | -0.01 | (0.02) |
| *1.2* | -2.07 | (1.03) | \* | -0.03 | (0.01) | -2.14 | (1.05) | \* | -0.03 | (0.01) |
| *2* | Ref. | (.) |  | (.) | (.) | (.) | (.) |  | (.) | (.) |
| *3* | -0.11 | (0.36) |  | -0.01 | (0.01) | -0.09 | (0.36) |  | -0.01 | (0.01) |
| *4* | 0.60 | (0.31) |  | -0.00 | (0.01) | 0.61 | (0.29) | \* | -0.00 | (0.01) |
| *5* | 0.29 | (0.30) |  | -0.01 | (0.01) | 0.24 | (0.28) |  | -0.01 | (0.01) |
| *6* | 0.77 | (0.28) | \*\* | 0.00 | (0.01) | 0.73 | (0.28) | \* | 0.00 | (0.01) |
| *7* | 1.20 | (0.27) |  | 0.01 | (0.01) | 1.27 | (0.27) | \*\*\* | 0.01 | (0.01) |
| Intercept | -1.48 | (0.25) | \*\*\* | (.) | (.) | -1.47 | (0.24) | \*\*\* | (.) | (.) |
|  |  |  |  |  |  |  |  |  |  |  |
| Number of observations | 8448 | | | | | 12536 | | | | |
| Average RVI |  | | | | | 0.31 | | | | |
| Largest FMI |  | | | | | 0.40 | | | | |
| \*\*\* p<.001, \*\* p<.01, \* p<.05 Data Source: NCDS [Sweeps 0-4]  Note: Comparison of CRA NS-SEC vs Imputed NS-SEC model | | | | | | | | | | |

Additional checks on the imputed model are produced via post-estimation 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, the less effect missing data has on a variable’s variance. The FMI, on the other hand, relates to the proportion of the total sampling variance due to missing data. The higher the FMI is, the greater the 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 44 per cent, indicating a requirement of at least 44 imputations – the MI model has 50 imputations. Relative efficiency indicates efficiency. The closer it is towards one indicates that the analysis has the correct number of imputations.

The average RVI score was 0.31, meaning that, on average, the missing data has a negligible impact on the model’s variance. According to the RVI scores across categories, NS-SEC across all economic activity and tenure within unemployed and out of the labour force are 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 on their variance. Housing tenure and NS-SEC 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 have a relative efficiency below 0.99. This lends support to the notion that 50 imputations are an appropriate number.

Regarding 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 of 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 records analysis and the imputed model suggests little evidence for a MAR assumption. This indicates that missingness in these categories has no substantive impact on the resulting interpretation and analysis of results. As such, the imputed model is no better than analysing the complete record for interpretation. Whilst this section does present much work that amounts to a preference for a complete records analysis model, the use of multiple imputation and discussion of missing data was essential 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 statistical due diligence to present these findings.

Implementing 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 records 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 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 the 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 on an individual’s choice and opportunity concerning 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) and 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 traditionally masculine fields dominated training & apprenticeship programs during the timeframe of the NCDS, 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); 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 are found in 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 transparent barrier to entry regarding 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 indicates support for promoting a life course perspective within this research. One aspect of these structural effects that have not been reflected 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 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, an individual who 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 not to be considered for future research within this area.

The findings from this empirical research appear to confirm the relatively influential impact that structures have on 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. However, structures continue to influence and constrict these opportunities. Furthermore, whilst these empirical results confirm much of previous empirical literature on this subject, the arguments proposed by new structuralist theorists (Saunders, 2003, 2021) are not as sound. Evidence has been found that concurs with the premise first emphasised by Saunders (2003) that housing tenure offers an independent and substantive impact on life chances. The argument that it is dominant above social class has no evidence within these results.

Including prior educational attainment alongside structural effects such as social class, sex, and housing tenure provides 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 concerning pathway selection. Most notably, these affect whether an individual seeks to continue along a traditional educational pathway. In saying this, lack of educational success at 16 does not block youth from entering several pathways. It appears to influence traditional education, but this is not true 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 a sensitivity analysis of socioeconomic 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 of this sensitivity analysis provided slightly divergent substantive findings. Choosing NS-SEC as the dominant model through the analysis was based upon a theoretical desire to understand class-based dynamics and a slight preference concerning AIC statistics. Through its implementation, social class was found to have a resounding impact on individuals’ choices and opportunities concerning transitional experiences. The results are also innovative by assessing missingness within the complete records analysis model. Missingness was first descriptively detailed before strategies for handling such missingness were discussed. A multiple imputation model found that missingness has no impact on the substantive findings of the complete records analysis model. While this means that the substantive findings remain the same as previously detailed, the implementation of dealing with missing data was an essential contemporary statistical strategy that previous literature within this field typically overlooked. Both the implementation of sensitivity analysis and multiple imputation techniques thus serve as methodological innovations beyond 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. 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 essential to reflect on for future inquiry. As has been mentioned, structural influence is dependent upon the given pathway of choice, with different structural influences matter 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. While the 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 whether they manifest throughout the life domain (Mayer, 2009) is relatively significant in the discussion of structural effects throughout the life course.

The following 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 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 on the nature of structural inequalities and their influence on choice and opportunities for youth transitions.

**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 | 23150.59 | - | - | - | 23158.58 | 23186.75 |
| Null Model + Educational Attainment | 18908.96 | 4241.63 | 8 | 0.18 | 18924.96 | 18981.29 |
| Null Model + Sex | 22159.83 | 990.76 | 8 | 0.04 | 22175.83 | 22232.16 |
| Null Model + Tenure | 22328.74 | 821.85 | 8 | 0.03 | 22344.74 | 22401.07 |
| Null Model + CAMSIS | 22299.86 | 850.73 | 8 | 0.04 | 22315.86 | 22372.19 |

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 | 23150.59 | - | - | - | 23158.58 | 23186.75 |
| Null Model + Educational Attainment | 18908.96 | 4241.63 | 8 | 0.18 | 18924.96 | 18981.29 |
| Null Model + Educational Attainment + Sex | 17921.39 | 987.57 | 12 | 0.22 | 17945.39 | 18029.89 |
| Null Model + Educational Attainment + Sex + Tenure | 17677.49 | 243.9 | 16 | 0.23 | 17709.49 | 17822.16 |
| Null Model + Educational Attainment + Sex + Tenure + CAMSIS | 17543.74 | 133.75 | 20 | 0.24 | 17583.74 | 17724.57 |

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 | 23150.59 | - | - | - | 23158.58 | 23186.75 |
| Null Model + Educational Attainment | 18908.96 | 4241.63 | 8 | 0.18 | 18924.96 | 18981.29 |
| Null Model + Sex | 22159.83 | 990.76 | 8 | 0.04 | 22175.83 | 22232.16 |
| Null Model + Tenure | 22328.74 | 821.85 | 8 | 0.03 | 22344.74 | 22401.07 |
| Null Model + RGSC | 22020.21 | 1130.38 | 24 | 0.05 | 22068.21 | 22237.21 |

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 | 23150.59 | - | - | - | 23158.58 | 23186.75 |
| Null Model + Educational Attainment | 18908.96 | 4241.63 | 8 | 0.18 | 18924.96 | 18981.29 |
| Null Model + Educational Attainment + Sex | 17921.39 | 987.57 | 12 | 0.22 | 17945.39 | 18029.89 |
| Null Model + Educational Attainment + Sex + Tenure | 17677.49 | 243.9 | 16 | 0.23 | 17709.49 | 17822.16 |
| Null Model + Educational Attainment + Sex + Tenure + RGSC | 17459.12 | 218.37 | 36 | 0.24 | 17531.12 | 17784.62 |

# **Data Citation**

## **Bibliography**

Allison, P. (2012a) ‘Handling Missing Data by Maximum Likelihood’, *SAS Global Forum* [Preprint].

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.

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 modernisation: 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 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].

Bodner, T.E. (2008) ‘What Improves with Increased Missing Data Imputations?’, *Structural Equation Modeling: A Multidisciplinary Journal*, 15(4), pp. 651–675. Available at: https://doi.org/10.1080/10705510802339072.

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.

Bottero, W. (2004) ‘Class Identities and the Identity of Class’, *Sociology*, 38(5), pp. 985–1003. Available at: https://doi.org/10.1177/0038038504047182.

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., 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. (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. 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.

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].

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.

Enders, C.K. (2010) *Applied missing data analysis*. New York: Guilford Press (Methodology in the social sciences).

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:, 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.

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. (1989) *Constitution of Society: Outline of the Theory of Structuration*. Polity Press.

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

Goldthorpe, J.H. (1980) *Social Mobility and Class Structure in Modern Britain*. Clarendon.

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. 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.

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.

Johnson, D.R. and Young, R. (2011) ‘Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations’, *Journal of Marriage and Family*, 73(5), pp. 926–945. Available at: https://doi.org/10.1111/j.1741-3737.2011.00861.x.

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.

Lynch, J. and Von Hippel, P.T. (2013) ‘Efficiency Gains from Using Auxiliary Variables in Imputation’, *Cornell University Library* [Preprint].

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].

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).

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.

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].

Rose, P. and Pevalin, D. (2010) *Standard occupational classification 2010*. Basingstoke, Hampshire: Palgrave Macmillan.

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., 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.

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

Stevenson, T.H.C. (1928) ‘The Vital Statistics of Wealth and Poverty’, *Journal of the Royal Statistical Society*, 91(2), p. 207. Available at: https://doi.org/10.2307/2341530.

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, UCL Institute of Education, and Centre for Longitudinal Studies (2023) ‘National Child Development Study.’, *UK Data Service* [Preprint], (13th Release). Available at: https://doi.org/DOI: http://doi.org/10.5255/UKDA-Series-2000032.

Vickerstaff, S.A. (2003) ‘Apprenticeship in the `Golden Age’: Were Youth Transitions Really Smooth and Unproblematic Back Then?’, *Work, Employment and Society*, 17(2), pp. 269–287. Available at: https://doi.org/10.1177/0950017003017002003.

Von Hippel, P.T. (2009) ‘How to Impute Interactions, Squares, and Other Transformed Variables’, *Sociological Methodology*, 39(1), pp. 265–291. Available at: https://doi.org/10.1111/j.1467-9531.2009.01215.x.

White, I.R., Royston, P. and Wood, A.M. (2011) ‘Multiple imputation using chained equations: Issues and guidance for practice’, *Statistics in Medicine*, 30(4), pp. 377–399. Available at: https://doi.org/10.1002/sim.4067.

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.

Young, R. and Johnson, D.R. (2011) ‘Imputing the Missing Y’s: Implications for Survey Producers and Survey Users’, *Proceedings of the AAPOR Conference Abstracts* [Preprint].

1. 13th Release [↑](#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, ‘’DBR’’ means Day Block Release, ‘’LGSS’’ means Local Government Support Scheme, and ‘’Rule6’’ means N/A. [↑](#footnote-ref-3)
4. [↑](#footnote-ref-4)
5. Either in Social Housing or privately rented accommodation. [↑](#footnote-ref-5)
6. The ’constant’ is a concept that argues that occupational positions have the same meaning over time and across different countries. [↑](#footnote-ref-6)
7. Discussed further in the ‘Registrar General Class Schema’ section below [↑](#footnote-ref-7)
8. In the case of his 1928 paper this was in relation to lower mortality rates [↑](#footnote-ref-8)
9. For these reasons when it comes to the sensitivity analysis, RGSC 5 will be used as the reference category of choice [↑](#footnote-ref-9)
10. https://warwick.ac.uk/fac/sci/statistics/staff/academic-research/firth/software/qvcalc/kuvee/ [↑](#footnote-ref-10)
11. Discussed at length in section below on Multiple Imputation [↑](#footnote-ref-11)
12. The seed for MI model is 12346, it can also be found in the .do file within the Github page. [↑](#footnote-ref-12)
13. 50 imputations were used in MI models. [↑](#footnote-ref-13)
14. Variable n4118 used [↑](#footnote-ref-14)
15. The variable in question was acatnn236, a categorical variable. [↑](#footnote-ref-15)
16. Burn-in was 20 during imputation. [↑](#footnote-ref-16)