# The role of families in the stratification of attainment: Parental occupations, parental education and family structure in the 1990s.

Christopher James Playford

Thesis submitted for the degree of Doctor of Philosophy
School of Applied Social Science
University of Stirling
2011

## Abstract

The closing decades of the 20th century have witnessed a large increase in the numbers of young people remaining in education post-16 rather than entering the labour market. Concurrently, overall educational attainment in General Certificate of Secondary Education (GCSE) qualifications in England and Wales has steadily increased since their introduction in 1988. The 1990s represent a key period of change in these trends. Some sociologists argue that processes of detraditionalisation have occurred whereby previous indicators of social inequality, such as social class, are less relevant to the transitions of young people from school to work. Sociologists from other traditions argue that inequalities persist in the stratification of educational attainment by the family backgrounds of young people but that these factors have changed during this period.

This thesis is an investigation of the influence of family background factors upon GCSE attainment during the 1990s. This includes extensive statistical analysis of measures of parental occupation, parental education and family structure with gender, ethnicity, school type and housing tenure type within the Youth Cohort Study of England and Wales. These analyses include over 100,000 respondents in 6 cohorts of school leavers with the harmonisation of data from cohort 6 (1992) to the Youth Cohort Time Series for England, Wales and Scotland 1984-2002 (Croxford, Ianelli and Shapira 2007). By adding the 1992 data to existing 1990s cohorts, the statistical models fitted apply to the complete set of 1990s cohorts and are therefore able to provide insight for the whole of this period. Strong differentials by parental occupation persist throughout the 1990s and do not diminish despite the overall context of rising attainment. This relationship remains net of the other factors listed, irrespective of the measure of parental occupation or the GCSE

attainment outcome measure used. This builds upon and supports previous work conducted using the Youth Cohort Study and suggests that stratification in educational attainment remains a significant factor. Gender and ethnicity remain further sources of persistent stratification in GCSE attainment.

Following a discussion of the weighting system and features of the Youth Cohort Study as a dataset, a thorough investigation of missing data is included, with the results of multiply imputed datasets used to examine the potential for missing data to bias estimates. This includes a critique of these approaches in the context of survey data analysis. The findings from this investigation suggest the importance of survey data collection methods, the limitations of post-survey bias correction methods and provide a thorough investigation of the data. The analysis then develops and expands previous work by investigating variation in GCSE attainment by subjects studied, through Latent Class Analysis of YCS cohort 6 (1992). Of the four groups identified in the model, a clear division is noted between those middle-attaining groups with respect to attainment in Science and Mathematics. GCSE attainment in combinations of subjects studied is stratified particularly with respect to gender and ethnicity. This research offers new insight into the role of family background factors in GCSE attainment by subject combination.

## Declaration

I declare that none of the work contained within this thesis has been submitted for any other degree at any other university. The contents found herein have been composed by the candidate, Chris Playford.

## Acknowledgements

This work was supported by the Economic and Social Research Council (grant number ES/F023251/1).

Many thanks to Professor Vernon Gayle. You have been consistently encouraging and positive and a great supervisor, I am incredibly grateful for your help and direction. Equally, thanks to Dr Paul Lambert for your enthusiasm, discernment and clarity. You have both been very supportive, patient and helped me to think clearly when I have found it difficult to do so.

Thank you to my Mum and Dad for your love and for listening and encouraging. Thank you also to Ruth and the Tudges, you are a wonderful family. Lettie and Richard, I am very grateful for all the times you have kindly put me up, you have made this so much easier!

Susan, Kevin, and Roxanne, you are excellent people to be friends with and have worked with. Thank you for helping me to keep on keeping on. Finally, thanks to all those who have passed an encouraging word, listened and been supportive. It makes all the difference.

# Contents

# List of tables and figures

## Figures

## Tables

# 1 Introduction

The principal aim of this thesis is to investigate the association between family background factors and the attainment of young people at the end of compulsory schooling. This is the first major point of divergence between young people in terms of whether they remain in school, go to college, enter work or become unemployed (Payne 2003; Croll 2009).Introduced in 1988, the General Certificate of Secondary Education (GCSE) is the principal qualification for those leaving school at age 16 in England and Wales (Leckie and Goldstein 2009). GCSE attainment and the route taken by young people at the end of compulsory schooling are associated, as is noted by Gayle, Lambert and Murray (2009a: 14):

*"Poor GCSE attainment is a considerable obstacle which precludes young people from pursuing more advanced educational courses. Young people with low levels of GCSE attainment are usually more likely to leave education at the minimum school leaving age and their qualification level frequently disadvantages them in the labour market."*

It is for this reason that any patterning in the attainment of GCSE attainment is of importance and provided the initial impetus to study the topic. The main research question to be addressed is: *"What is the role of family background in the stratification of GCSE attainment in the 1990s in terms of parental occupations, parental education and family structure?"* To clarify this question, a number of terms used require definition (these definitions are considered in further detail later within the thesis).

Family background is considered using a broad range of measures but is focused upon parental occupations, parental education and family structure. These were selected as indicative of the inequality in the family circumstances existing between young people. As Bottero (2005: 3) notes: *"Inequality in one generation affects inequality in the next. The resources that are available to us growing up as children affect the success of our schooling, and so our eventual occupational careers, and the lifestyles we adopt as adults. However, this means there is also an impact on the next generation, since our social position influences the resources to which our children have access, and so their life-chances too."* Stratification refers to the patterning of inequality (Scott 1996; Bottero 2005). In the context of this thesis, this refers to inequality between young people's GCSE attainment according to difference in family background measures.

The 1990s are a key period of change in the history of stratification in educational attainment due to the context and policies leading up to the decade. During the mid-part of the 20th century it was common for the majority of school-leavers to leave at the earliest opportunity with relatively low levels of qualifications (by the standards of today) and move straight into work (Sanderson 2007). Between the mid-20th century and the 1990s, a number of factors have been identified as responsible for a reduction in choice for those with lower or no educational attainment at age 16 (Gayle, Lambert and Murray 2009b). These are the collapse of the youth labour market (Gayle, Berridge and Davies 2003), the decline of apprenticeships (Ainley and Rainbird 1999), the failure of vocational educational schemes (Stoney and Lines 1987), the growth of the service sector (Sanderson 2007) and the corresponding growth in further and higher education (White 2007).

In addition to the factors above, UK educational policy moved away from the pure models comprehensive schooling ideals of the 1960s (Benn and Chitty 1997). Gillard (2007) describes how during the 1980s and 1990s successive UK governments promoted agendas of parental choice, diminishing LEA control, and the introduction of the national curriculum and GCSE qualifications. It is argued by some that these policies may have led to increased overall educational attainment but also widened the gap between those able to take advantage of the changes to schools and the education system, principally those with more affluent parents (Smith and Noble 1995; Ball 2003). Chapter 2 contains detail of these contextual changes and educational policy.

Some sociologists have used the changing context (identified in chapter two) as evidence for processes of detraditionalisation whereby traditional influences upon attainment have reduced and that individual young people are less constrained by their social circumstances (Heelas, Lash and Morris 1996). Other sociological traditions suggest the measurement of multiple family background factors and educational attainment is valuable in monitoring how the influence of these factors change over time (Rose 1998; Rose and Pevalin 2003; Rose and Harrison 2009). A discussion of these theoretical traditions is included in chapter three.

In order to evaluate these traditions, a number of specific research questions are developed. These are as follows:

1. Has the association between parental occupation and their children's educational attainment increased, decreased or remained stable through the 1990s?

2. How do different measures of parental occupation and parental education report this association? Are the different measures consistent? (a sensitivity analysis of measures of parental occupation)

3. Have the differences in attainment between pupils from lone-parent and couple families increased or reduced through the 1990s?

4. What influence does missing data potentially have upon the estimates in the models fitted? (particularly with regard to parental occupation)

5. How do the subjects studied and grades attained vary between young people from different family backgrounds?

The first part of evaluating these questions is through considering previous studies. Applied research into educational attainment in the 1990s is included in chapter four. This draws heavily upon the Youth Cohort Study, a large scale dataset covering the decade. This builds upon the theoretical discussion in chapter three, to empirically examine the association between family background factors upon educational attainment. A range of family background factors are considered to

identify the strength of association by each factor throughout the decade, with a particular interest in measures of parental occupation, parental education and family structure. The final part of the chapter evaluates the findings of research into the influence of growing up in a lone parent family upon attainment.

Using data from the 1990s Youth Cohort Study (YCS), this thesis builds upon the applied work detailed in the chapter 4 to examine the whole of the decade (much of the research contains a limited number of cohorts). With reference to the specific research questions, the selection of the Youth Cohort study is justified as it is the most relevant dataset for collecting educational information on successive 1990s cohorts. Further discussion of the advantages and disadvantages of the YCS are provided with respect to the questions posed in chapter 5. It is recognised that whilst the YCS has inevitable missing data issues and challenges in variable design: it remains a valuable and under-utilised data resource with respect to the collection of educational data from a representative sample of school leavers in England and Wales. A detailed account of the data management with the Youth Cohort Study is included, with particular reference to harmonising the measures within cohort 6 of the Youth Cohort Study with the work of Croxford, Iannelli and Shapira (2007).

A full descriptive analysis of the Youth Cohort Study variables used within the models is covered in chapter 6. Included are comparisons with Office for National Statistics (ONS) estimates for the measures used to test representativeness of the YCS data. Furthermore, prior to modelling measures of central tendency and dispersion are described so that it is clear that key assumptions in the modelling process are met.

Following this analysis, statistical models investigating the influence of different family background factors on educational attainment are introduced. A sensitivity analysis of alternative parental occupation schemes is made for methodological purposes and to check the reliability of the estimates. Then attempts to control for rising attainment during the 1990s are introduced to examine the persistence of these differentials. Further models are fitted to understand differing functional forms of educational attainment measures. In particular gaining five A* to C passes at GCSE is a key threshold in continuing in post-school education. It is also important to use these techniques to attempt to overcome changes in the marginal distributions of the underlying variables. For example, the number of young people having parents in manufacturing occupations has declined since the 1950s due to the reduction of people working in the sector (Sanderson 2007). It is therefore important not to over or under estimate the influence of this variable upon attainment due to this reduction in numbers.

Missing data is a problem in secondary data research particularly when there is the potential for respondents from different backgrounds to respond differentially to particular questions. Chapter eight contains an investigation of missing data in the Youth Cohort Study. Amongst other stratifying factors, the measurement of parental occupation in stratification research is of great importance. Methods to correct for problems with non-random missing data are used and explained to assist with the validity of the findings in chapter seven.

As will be demonstrated in chapter four, educational attainment has been investigated with volume measures but little has been done to investigate subject level variation in attainment by family background, and nothing particularly using

the YCS. Further understanding of how stratification is reflected in grades attained in GCSE subjects and models of family background factors are included in chapter nine. This is through latent class analysis of attainment in particular combinations of GCSE subjects. This provides a suitable method for investigating this variation.

In summary, the purpose of this research is to better understand stratification in educational attainment. Whilst there have been widespread social changes to the experiences of young people as they prepare to leave school, there has been enduring social stability. Using a variety of methods, this thesis is a unique attempt to use Youth Cohort Study data to investigate the multivariate nature of social disadvantage. This includes data previously under-analysed due to the complexity of data management required. It will explore the family background of young people from the traditional approach of occupational stratification. Then more advanced modelling techniques in the multiple imputation of missing data and latent class analysis are used to explore how missing data may influence the models fitted and how stratification is reflected in the subject choices of young people. The first issue to be considered is further detail of the social changes of the late 20$^{th}$ century for young people growing up in the UK and why the 1990s are such an important decade. This is the topic considered in the next chapter.

## Outline of the thesis

In summary, this thesis comprises of the following chapters:

- Chapter Two describes the changing educational context of the late 20th century and educational policy in the UK leading to and during the 1990s.
- Chapter Three contains a review of the literature concerning educational sociology and the sociology of the family with reference to inequality.
- Chapter Four provides detail of applied research into the stratifying factors in educational attainment with emphasis upon papers based on Youth Cohort Study data. Comparison is made with Scotland and other countries.
- In Chapter Five, the questions, methods and methodology underpinning the research are discussed. Also included is a detailed account of data management with the Youth Cohort Study.
- Chapter Six is a descriptive analysis of the Youth Cohort Study covering the distributions of educational variables and the correlates with family background measures.
- Chapter Seven introduces initial modelling approaches, focusing on linear and logistic models of educational outcomes with multivariate family background measures and time-varying covariates.
- In Chapter Eight, the issue of missing data in parental occupational measures in the Youth Cohort Study is discussed and approaches to overcome these problems are introduced.
- Chapter Nine builds upon the modelling in chapter eight by introducing Latent Class Analysis of educational subject choice to examine the differences in subject choice by family background.
- Finally in Chapter Ten the findings of the previous modelling work are discussed and conclusions are considered.

# 2 Historic Context and Policies of the 1990s

Young people's transitions from school to work have changed markedly in recent decades. This chapter describes the educational context and policies in the years leading up to and throughout the 1990s in order to understand the importance of the decade as a period of study and the implications for the study of educational attainment and family background. The changes in the youth labour market, the relative failure of vocational training schemes and the rise of post-compulsory participation are some of the main reasons for why this is the case (Gayle, Lambert and Murray 2009a).

The occupational structure of Britain since 1945 has changed substantially with a large growth in professional and managerial occupations, a decline in the numbers of people in manual employment particularly in manufacturing industries, and a significant redistribution of employment patterns across socio-demographic groups (Sanderson 2007). There have also been changes in labour market participation patterns. Fewer young people go directly from school to work at the minimum school leaving age than was the case 60 years ago (Sanderson 2007). Growing levels of general unemployment during the 1970s and 1980s, were particularly pronounced among young people (Casson 1979; Maguire and Maguire 1997). As Jackson (1985: 68) noted: *"Youth unemployment rates are more sensitive to changes in economic conditions than overall employment rates"*

Regions within the UK with occupational structures based upon manual industries which traditionally recruited young people with low skill post-16 experienced high levels of youth unemployment (Gallie, Marsh and Vogler 1994). This led to concern

among policy makers due to the association between unemployment and poverty and the future career routes of these young people (Jackson 1985).

Within the context of declining labour market opportunities in the 1970s and 1980s, school leavers were often the most vulnerable to unemployment due to being the least qualified and experienced among staff and the most likely to change employment as a consequence of dissatisfaction with their early employment (Casson 1979).

## Apprenticeships and Youth Training Schemes

With the decline in manufacturing employment, the number of traditional apprenticeships declined. Increasingly following the Second World War it became individual employers' responsibility to manage their own training of employees in many trades (such as construction) whereas previously this had been more centrally co-ordinated (Ainley and Rainbird 1999). In some industries the rise of subcontracting and self employment substantially undermined the training infrastructure (Austrin 1977). As the industry was not externally regulated this contributed further to declining intake of apprentices as without a central pool of apprentices funded by the industry and individual companies unable to pay for apprenticeship schemes, companies were reliant on individual young people funding their own training. Fuller and Unwin (2009) claim that the change from apprenticeships to vocational education represents a change from sponsored schemes provided by companies within industries to detached schemes of training provision, such as the youth training schemes which were part provided by colleges (see also Raffe 1988; Fuller et al. 1999). These changes are important in the context of the late 1980s because of the reduction in non-educational routes available to young people at the end of post-compulsory education. In the previous

section, it was noted that labour market opportunities for school leavers declined. The further reduction in the number of apprenticeships occurred at a time of rising unemployment among young people (Gospel 1995). Increasing youth unemployment in the 1970s appears to have continued with the ratio of unemployed 16-17 year olds to unemployed 18-24 year olds being much greater in 2008 than 1992 (Bell and Blanchflower 2009; Cregan 2001).

To counter the rising unemployment among the young during the 1980s and to promote improved vocational training for young people without qualifications, the Conservative government introduced the Young Workers Scheme (1982-1986) and the Youth Training Scheme or YTS (1983-1990) (Chapman and Tooze 1987; Deakin 1996). The net employment effects of these schemes appears to have reduced youth unemployment in periods of recession and to provide subsidised training for employers during periods of increased youth employment (Deakin 1996). Of these two effects, Deakin (1996) argues that the YTS was reasonably successful at reducing unemployment among young people during times of economic recession but was not well adjusted to skills shortages during economic upturns where more young skilled workers might be required. He recommended the improvement of vocational education via improved core skills for young people at school (i.e. maths and sciences) and improved courses in the post compulsory education sector. Stoney and Lines (1987) claimed that the YTS changed the orientation of vocational schemes from a previously college focus to targeting what employers directly required. There is contended by Ainley (1990) who argued that there was a lack of effective communication between the National Council for Vocational Qualifications (created in 1986) and leading industrial bodies, claiming that employers have been reluctant to employ on the basis of new vocational qualifications. Ainley (1990)

argues this is because the new vocational qualifications have greater relative emphasis on written examinations and coursework rather than practical competence.

In Britain, the change in the occupational structure to a service[1] based economy led to reduced demand for vocationally qualified young people in traditional (and traditionally male) subject areas (details of which are in the next section, see Gallie 2000; Crompton 2008; Hills, Sefton and Stewart 2009). In the context of the late 1980s, this entailed that the skills provided in many traditional training programmes were less important than the attitude of young people for work in a de-skilled, post-industrial economy (Ainley 1988). Vocational qualifications are a fragmented and specialised form of credentials in a labour market in which it appears non-specific academic education is more suited to the contemporary labour market, largely as a consequence of the growth of jobs in the service sector (see next section; Wolf 2002). With the reduction in the number of jobs available for young people in the 1980s, criticism of the education system increased (Roberts 1984). These criticisms included accusations of falling standards, technological illiteracy and poor careers information provision. Against these criticisms Roberts (1984) argues that education cannot equate to first-hand experience of employment and that the companies and industries which benefit from an educated workforce must accept some responsibility for proper training. Furthermore, government schemes could not completely overcome the economic conditions faced by the declining youth labour market (Roberts 1984).

[1] *"The service sector covers the wholesale and retail trades, hotels and restaurants, transport and communication, financial services, real estate and business activities and government and other services."* (ONS 2000:1)

## *Increasing emphasis on Education*

Corresponding to the decline of the UK manufacturing sector there has been a large growth in the service sector particularly in finance, retail and personal services (Matthews, Feinstein and Odling-Smee 1982; Hills, Sefton and Stewart 2009). The UK service sector increased from 53% to 73% of total employment between 1973 and 1993 (Crompton 2008). Seventy percent of Gross Domestic Product (including government) was generated by service sector industries in 2000 (ONS 2000b).

The number of individuals employed who are part-time, self-employed or on temporary contracts has increased throughout the 20th century (Gallie 2000). The nature of employment has become more flexible and with lower job security (OECD 1997; Wolf 2002; Clasen and Clegg 2003). Those young people who leave school at the end of post-compulsory education are generally those who do not have the grades to continue on to further or later higher education (Gray, Jesson and Tranmer 1994). There is a high cost to training employees (Roberts 1984) and it is suggested that employers appear less willing to employ young people without academic qualifications, and regard academic qualifications attained as a proxy for potential productivity (Jenkins and Wolf 2003; Sanderson 2007). This is in contrast with the arguments of Ainley (1990) regarding vocational qualifications which have not been popular with employers. Fuller and Unwin (2009) argue that this is because vocational training has become fragmented, downgraded, marginalised and poorly integrated into educational systems.

Concurrent with the decline in the youth labour market, the decline of apprenticeships and the attempts to enhance vocational education, participation in education post-16 increased particularly during the 1990s (Further Education

Funding Council 2000; ONS 2006b; ONS 2007b). Raffe (1992) conceptualises these changes as push and pull factors, akin to supply and demand in post-compulsory education. The push factors are the changing economic base of the UK (more service sector jobs) and overall rising attainment levels leading to greater numbers aspiring to continue in education post-16. The pull factors are the increased provision in higher education, as a consequence of demographic decline in the relevant age group and policies to increase participation in higher education, and a declining youth labour market (Gray, Jesson and Tranmer 1994). These factors may appear to be self-reinforcing (i.e. fewer alternatives to remaining in education encourages a young person to remain in education). The purpose of describing this typology is to understand the context at the beginning of the 1990s, the importance of this period and the circumstances facing young people, rather than extended debate about which of the factors is more important.

The large growth in the number of young people in higher education was particularly marked at the end of the 1980s (Tinklin and Raffe 1999; Archer, Hutchings and Ross 2003; Forsyth and Furlong 2003a; Clark, Conlon and Galindo-Rueda 2005). Between 1970/71 and 2003/04 the number of students in UK further education rose from 1.7 million to 4.9 million (ONS 2006b). During the same period the number of students in UK higher education rose from 630,000 to 2.4million (ONS 2006b). Sanderson et al. (2007) note this has largely been due to the growth of professional and managerial occupations and the widening of access to full time education throughout the 20$^{th}$ century (see also Gallie 2000; Machin 2003). Increased provision of post-compulsory education has also been partially driven by rising educational attainment, whereby young people with better grades than their predecessors are eligible (according to the requirements of schools and colleges) to

study for a range of post-compulsory courses (Gayle, Berridge and Davies 2003; Clark, Conlon and Galindo-Rueda 2005).

It is difficult to deduce whether rising attainment has led employers to expect more highly qualified workers or whether these qualifications are genuinely necessary for the employment provided. The possibilities of being over or under educated exist where people are employed in jobs lower than their ability (McIntosh 2003; Sanderson 2007). There has also been a discrepancy between the matching the demand for educated workers in specific industries with the qualifications gained by these young people (Dolton and Vignoles 2000; Sanderson 2007). However, the incomes of those possessing high-level qualifications did not fall during the 1990s implying that there was not an oversupply of qualified workers (Blundell et al. 2000; Elias and Purcell 2004).

Job turnover appears to have increased with fewer longer term permanent jobs than in the 1950s (Gallie 2000). This is perhaps misleading as job security varies by occupation (Wolf 2002; Schroeder et al. 2008). In the context of increased participation in education, Wolf (2002) argues that this places those without qualifications (who do not remain in education) at further disadvantage. The proportion of the population without any qualification has remained relatively stable through the 1990s at 12-15% whilst the qualifications of the remainder of the population have risen (Elias, Hogarth and Pierre 2002). There is a clear negative correlation between unemployment and education. Unemployment is less likely for those young people with degrees and this relationship has strengthened between 1984 and 2008, according to data from the Labour Force Survey (Bell and Blanchflower 2009).

The increasing number of young people going into higher education is particularly important as the association between admission to higher education and the social background of young people is persistent and strong (Kogan 1993; Lynch and O'Riordan 1998; Forsyth and Furlong 2000; Forsyth and Furlong 2003a; Forsyth and Furlong 2003b; Gayle, Berridge and Davies 2003). Whilst there has been a significant growth in the numbers attending higher education institutions, these institutions are stratified by the social background of students attending (Cheung and Egerton 2007).

Compared with previous decades, by the 1990s alternatives to remaining in education for young adults had diminished and these factors encouraged extended periods in post-compulsory education. These changes potentially have placed greater emphasis on the qualifications gained at the end of post-compulsory education. In addition to the contextual factors, there are a number of policy initiatives which have influenced, and interacted with, the relationship between family background and educational attainment among youths, and these will now be discussed.

## *Policy*

Alongside the changes in the educational context and youth labour market of the 1990s were changing policies of the period concerning education and employment. In particular there was a radical shift from the education policies of the 1960s and 1970s (which are now briefly described) to the introduction of market based principles in educational provision of the Conservative government of the 1980s and 1990s (Gewirtz, Ball and Bowe 1995).

During the 1940s, 1950s and 1960s it was widely felt that the schools selection process of the pre-1960s, in particular the 11-plus examination, through selection of the most able pupils into grammar schools discriminated according to social background (Benn and Chitty 1997). The development of comprehensive schools appears to have been driven by the failures of the selective school system, in particular the levels of inequality between students, the large numbers leaving school at the earliest opportunity and the inefficiency of allowing potentially talented students to be consigned to an inferior schooling due to selection on attainment at a relatively young age (Gillard 2007). The 1964 and 1966 general election victories for Labour gave momentum to the growth of comprehensive education but the crucial bill was not passed prior to the Conservative election victory of 1970. Ultimately, the consequence was that comprehensive schools were not implemented uniformly across England and Wales, whereby the majority of local authorities adopted comprehensive systems by the mid-1970s, but some regional authorities retained aspects of the previous system (such as grammar schools) much longer, and some do to the present day[2] (Benn and Chitty 1997; Jones 2003; Gillard 2007).

[2] Examples include Essex, Kent, Lancaster and Buckinghamshire.

Local authorities in the 1970s were informed by the Conservative government that there was no pressure to pursue comprehensive change (Benn and Chitty 1997). In addition to this, the recession of the mid-1970s led to cutbacks in funding for education and a growing division between the political parties over the course of educational reform (Galton, Simon and Croll 1980). In this context, the Labour government did not pursue comprehensive schooling with the vigour expected. During the 1980s, the Conservatives increasingly sought to base educational policy around principles of consumer choice as part of the 'new public management agenda' (Benn and Chitty 1997; Ranson 2008).

## 1988 Education Reform Act

This introduction of market based principles by the Conservative government in the 1980s was firstly introduced in 1985 by the linkage of teacher appraisal and performance related pay and concurrent Education Acts of 1980, 1984 and 1986 which sought to weaken the control of Local Education Authorities (LEAs). A key piece of legislation though was the 1988 Education Reform Act (Gewirtz, Ball and Bowe 1995; Gayle, Lambert and Murray 2009b). This act introduced a large number of measures which increased the power of parents in school admissions and reduced the power of Local Education Authorities. These included the introduction of a National Curriculum and GCSE examinations to replace O-levels, attainment targets at age 7, 11, 14 and 16, the introduction of the Office for Standards in Education (OFSTED) and targets for school level performance based upon testing and league tables and the introduction of governors at schools (often local parents). Further to this, schools could opt out of local authority control and become "grant maintained", being directly funding from government if parents and governors wished it (Benn and Chitty 1997; Glennerster 2007).

The introduction of these measures of school assessment and reduction of LEA control was intended to enable parents to exercise increased choice over the school their child attended. Grant maintenance and the introduction of governors enabled control of funding to be re-routed directly from central government to schools. Businesses could be involved in sponsoring city technical colleges. In addition parents could apply for assisted places funding at independent schools when fees could not otherwise be afforded (David 1995). All of the above can be presented as a strong movement away from comprehensive education which may have had implications for the socially disadvantaged (Gewirtz, Ball and Bowe 1995; Ball 2003; Ball 2006). Perhaps the most important of these measures with respect to studying educational attainment in the 1990s is the growth of accountability through the attainment targets at ages 7, 11, 14 and 16, as these have highlighted where young people and schools are doing well or poorly. This has increased the transparency of attainment differences between young people.

## The introduction of new GCSEs

In 1988, the General Certificate of Secondary Education (GCSE) was introduced replacing O-Levels, CSEs and 16+ examinations (Mobley et al. 1986; Ashford, Gray and Tranmer 1993). The new examinations and curriculum were not without their critics however. The increased focus on skills training (which appears to have been in response to employers concerns) could also be considered by the educational establishment as "anti educational" in principle (Roberts 1984; North 1987). North (1987) further criticised the potential for a reduction in the standard of education received and predicted the decline of less popular academic subjects.

Since the introduction of GCSEs, participation in education post 16 increased greatly (Ashford, Gray and Tranmer 1993). Whilst as individuals some young people sought greater credentials to increase their expectation of increased lifetime returns, others sought to remain in education or training to avoid potential unemployment (Cregan 1999; Cregan 2001; Breen and Goldthorpe 1997).

GCSEs are the principal qualifications for those leaving school at age 16 and a key measure in school league tables (Leckie and Goldstein 2009). This has implications for the comparison of young people and schools as it is used both as a measure of educational attainment for a young person but also in monitoring progress in schools. There is a major incentive for young people to do well for personal reasons but also for the schools concerned as poor results reflect badly in school accountability measures.

## Conservative Reforms 1989-1997: Promoting choice

The Conservative governments led by Margaret Thatcher and John Major in the last decade of Conservative administration continued to promote choice and specialisation in schools (David 1995; Gillard 2007). *"By the beginning of the 1990s, the whole approach to parent-school relations had shifted from one about how to ensure some measures of equity to how to ensure parental rights and responsibilities in order for individual parents to be able to influence each child's educational success in formal examination situations"* (David 1995: 68). Glennerster (2007) emphasises the key features of the Conservative reforms. First, decentralising financial responsibility to schools and removing the powers of local authorities. Second, schools were free to compete with one another and attract as

many pupils as they saw fit. Funding would be linked to schools with the greatest number of pupils.

In 1991, the then Education Secretary, Ken Clarke announced a change in direction from the original plan for the National Curriculum to consist of 10 subjects of study for all pupils. With the exception of English, Maths and Science, students were free to choose which subjects they studied encouraging further differentiation and diversity (Benn and Chitty 1997). Smith and Noble (1995) argue that the early 1990s were a period of decreased funding with evidence for falling per-pupil expenditure in the early 1990s and a net fall in total expenditure once costs were taken into account.

The *1992 Education White Paper, 1993 Education Act* and *1996 Education White Paper* all further promoted the growth of Grant Maintained status for schools on the basis of parental demand and a diverse and increasingly specialised series of schools and colleges (Chitty 2004). The effect of this specialisation appears to be greater inequality between schools in terms of popularity, funding and grades achieved (Walford 1994). Several different vocational qualification schemes were tried and failed with the government ignoring the advice of the *1996 Dearing report* to integrate vocational education within a coherent system of post-compulsory education and choosing to maintain the distinction between the A/AS level route and vocational qualifications (Benn and Chitty 1997; Gillard 2007). Essentially the policy measures of the early nineties emphasised choice and it was apparent that there were correspondingly widespread selection practices according to the popularity of some schools above others. Indeed, the more popular schools were

able to determine their own admissions, potentially compounding social inequalities in selection to them (West, Hind and Pennell 2006).

## New Labour Reforms 1997-2000

The New Labour administration of 1997-2010 initially chose to continue with Conservative policies, particularly emphasising the importance of parental school choice through further measures of teacher accountability and publication of school league tables of examination performance (Chitty 2004; Smithers 2007; Barker 2008). The focus of policy was concentrated on admission to schools, in particular monitoring and controlling catchment areas for different schools, in response to evidence of targeting strategies employed by advantaged parents for gaining entry to 'better' schools, for example moving house to live in the catchment area for a school with higher attainment (Gillard 2007; West, Barham and Hind 2009). The *1997 Education White Paper* continued to place emphasis on performance targets and minimum standards in attainment at age 11, progress in performance for individual pupils, increased inspection by OFSTED and increased emphasis on 'standards' (Chitty 2004). Raffe (2005) argues that, *"this vision was based on specialisation, collaboration, frontline control and strong accountability, compared with the old system based on uniformity, isolation, centralised control and weak accountability"* (Raffe 2005: 53). Evidence suggests however, that there remains significant 'unfairness' in selection procedures by a minority of schools (West, Hind and Pennell 2006).

The proposed changes to the governing structure of schools faced resistance by many in parliament due to concerns about selection (Glennerster 2007). Vocational qualifications and post-compulsory education appears to have had less emphasis

placed upon it than education till age 16 and in particular GCSE attainment (Hodgeson and Spours 2001). Education Action Zones (EAZ) have been introduced in areas of under-performing schools and family disadvantage in an attempt to target the young people with poorest levels of educational attainment (Whitty 2002; Bell and Stevenson 2006). The evidence for the success of EAZs is limited (Bell and Stevenson 2006).

Despite these changes there is strong evidence of continuity in policy rather than wholesale change from the Conservative approach (Smithers 2001). Policy models treat parents and children as rational autonomous individuals who may decide on the optimal choices for their future routes. However, *"critics see this system as likely to take government policy backwards to a pre-comprehensive era, allowing for new tiers of selective schooling to emerge, reintroducing and further intensifying existing social class divisions."* (Olssen, Codd and O'Neill 2004: 201)

## Commentary on educational policies

There has been much criticism of the reforms that have been made, particularly with regard to the 1980s and 1990s promotion of parental choice which appears to have worsened the plight of poorer students. Whilst the majority of state schools are comprehensives, the comprehensivisation policy agenda of the 1960s appears to have been subsequently abandoned in favour of market based initiatives. *"Instead, reform has amounted to little more than an ad hoc series of disjointed, half-hearted, sometimes transient concessions that have marginally mitigated the effects of a fundamentally unequal education system"* (Lodge and Blackstone 1982: 220).

The introduction of market principles into educational provision was an attempt to answer the fundamental problems namely the *"...system's lack of democracy, its*

*low standards and its failure to organize itself around principles of universal access"* (Jones 2003: 142). As Ball (2006: 121) notes, *"markets, of any kind, are complex phenomena. They're multi-faceted, untidy, often unpredictable and both creative and destructive."* The principle of individual choice represents a step away from one of universal access.

The persistence of differential schooling systems throughout this period has led to the clear processes of segregation of intake between those attending higher attaining grammar schools and those who did not. *"One of the most telling findings was the educationally depressing effect that the remaining grammar schools had upon the comprehensive schools in their own vicinities in 1994: distorting their intakes (with fewer high attaining pupils and unrepresentative social mixes), lowering their numbers and staying-on rates and depressing their examination results..."* (Benn and Chitty 1997: 463). Individual stories of meritocratic success by young people should not legitimate or obscure the structural inequalities that exist in education (Whitty 2002).

Jones (2003) argues that the changes to the occupational structure divided those parents according to whether this change had led to increased personal prosperity or not. *"Cracks had widened between different sections of the working class. For some restructuring meant increased prosperity, access to better-paid-jobs in parts of the service sector and – in some respects – choice. Home owning, private health provision and higher education all experienced significant growth in the 1980s. For others, it had an opposite meaning: millions were relegated to an insecure existence on the margins of employment"* (Jones 2003: 108). Glennerster (2007) argues that the legacy of this agenda was one of great inequality which creates a great tension

between social cohesion and *"market-driven economic upheaval"* (Jones 2003: 110). Bradley and Taylor (2002) argue that the quasi-market in education has made the system more efficient but less equitable.

Another contributing factor is that more affluent parents may have persistently viewed state sector education as a greater "investment risk" in terms of their children gaining the requisite attainment to maintain their advantaged position (Ball 2006). By this, Ball is implying that parents in advantaged occupations wish to see their children become highly qualified in order to then have similarly advantaged and remunerated occupations. Attending a state school may decrease the likelihood of their children gaining higher levels of attainment, so many seek schools with higher average levels of attainment. *"Such investments require resources, skills and capitals that are unevenly distributed across the population, but with which the middle classes are particularly well endowed. The education market with all its risks is well accommodated to the dispositions and interests of the middle classes. Their assiduous engagement with choice, their use of their capitals, their particular sense of responsibility, all contribute to their social reproduction and the assurance of their social advantages"* (Ball 2006: 275). Jones (2003) claims that the Conservative (and New Labour) agenda of promoting choice in education was due to pressure from parents to do so, in order to maintain this advantage.

What is clear from the contextual and policy changes for young people growing up in the 1990s is that inequality of opportunity in education and educational attainment remains prevalent (Smith and Noble 1995). Smith and Noble (1995) claim that this inequality in attainment occurs through the effects of experiencing an adverse social background whilst at school (in terms of income, housing, job

opportunities of parents, and local environment), the legacy of early socialisation in the home, and disadvantage in the educational environment in school. More recently, Croll (2009: 7) notes that: *"Whilst education can break such intergenerational cycles of disadvantage, it can also act to reinforce them: for example, if education policy is not designed with egalitarian notions in mind."*

Judge (2006) argues that the market-approach to education which encouraged diversity fails to recognise that some schools and young people will be left behind. "*Any system which deliberately elevates the status and achievements of some schools (be they grammar schools or city academies or specialist schools) without serious weighing the implications for all schools is certain to fail."* (Judge 2006: 49). In particular, in exercising choices not only do some parents have more access to certain advantaged choices but they can limit the opportunities of others as educational provision is a finite resource (Jonathan 1990; Olssen, Codd and O'Neill 2004). There is continuing evidence that pupils from disadvantaged social backgrounds do badly in formal educational systems (Whitty 2002). In particular,

*"Low-achieving children from less advantaged social backgrounds are very much more likely to leave education at 16 than children with similar levels of achievement who come from families in favourable positions with regard to education, income and type of employment. These are the young people whose needs are least likely to be met by current curriculum arrangements and rigid academic/vocational distinctions and are also the young people who most need to be persuaded of the value of educational participation."*

(Machin 2006: 415)

## Summary

The 1990s are a key period in educational research for a number of reasons. In the post war UK there has been a decline of the traditional routes into employment for young people. The reduction in the number of apprenticeships and relative failure of youth training schemes and vocational programmes in the 1980s led to a decreased likelihood of young people entering work immediately post-school. The failure of vocational educational schemes appears to be partly a consequence of the shift from a manufacturing to a service based economy.

It is suggested that this has led to a rise in the number of young people remaining in education due to the lack of alternatives but also because of rising attainment enabling young people to do so (particularly on the advantageous A-level route). There has been a growth in both the provision and take-up of further and higher education. Educational attainment at age 16 is important as it enables young people to gain access to these courses and the level of attainment indicates the range of options available to a young person. Young people with the highest GCSE attainment are the most likely to study A-levels and continue into Higher Education and potentially advantaged occupations (Goldstein and Thomas 1996; Yang and Woodhouse 2001).

This period is also important because of the policy reforms of the late 1980s and early 1990s. The policies of governments between 1988 and 2000 have moved away from the comprehensive schooling policies of the 1960s. In particular, the Education Reform Act (ERA) 1988 is seen by commentators as the principle piece of legislation signifying this shift. The ERA led to a reduction in local authority controls, the introduction of the national curriculum, increasing monitoring and inspection of

schools through OFSTED and testing/attainment targets. These were developed and maintained by successive Conservative and Labour governments through the 1990s around market based principles.

Most importantly, the 1988 Education Reform Act heralded the introduction of new GCSE qualifications and the National Curriculum and this appears to mark a watershed in educational policy. The reforms of the Conservative governments of 1989-1997 and the New Labour government of 1997-2000, promoted increased specialisation and accountability in schools as part of market based principles of education. The majority of commentators upon educational policy have suggested that this has led to greater inequality between young people as a consequence of family background. This is also a concern among many sociologists of youth and this is the subject of the next chapter.

# 3 Educational Sociology and Sociology of the Family

In the previous chapter, it was argued that the decline of the traditional route from school to work has led to increased numbers of young people remaining in education post-16. This is the position whereby it is acknowledged that the changes identified have significantly altered youth transitions compared with prior decades (Gayle, Lambert and Murray 2009a: 2): *"Within the 'changing times consensus', authors agree that the transformation was driven by a series of interrelated social and economic changes."*

This thesis is concerned with the stratification of educational attainment by parental occupation, parental education, family structure and other family background factors, which in practice means understanding the multivariate nature of these factors. This is an empirical rather than theoretical investigation. The questions posed in chapter one are focused upon the measurement of these family background factors through the 1990s using nationally representative data. This chapter is included to provide theoretical context about the nature of the debate in differing sociological traditions during this period. There are many theories regarding which factors are relevant to the association between family background and educational attainment (Bates and Riseborough 1993). The purpose of discussing these traditions is to provide insight to the empirical analysis in this thesis, which is useful both to sociological and empirical audiences (for an example of a similar approach, see Bills 2004).

Theories of detraditionalisation are introduced as these have grown in prominence in recent years, as a consequence of the changing times consensus among

sociologists of youth (Gayle, Lambert and Murray 2009b). Those supporting these theories argue that individuals are less constrained by traditional structural divisions, such as gender, ethnicity and social class (with respect to educational attainment and subsequent employment). By contrast, the way in which this has been approached in traditional sociology is through theories of social class, whose supporters argue that these structural divisions persist. The theories of social class referred to are from the Weberian tradition in which: *"Social classes are clumps of occupations with similar life-chances, linked by common mobility patterns"* (Bottero 2005: 38).

This leads into an account of the development of occupational measurement schema and how the measurement of stratifying factors occurs within a family setting and why the attributes of parents may influence young people's attainment. Part of the reason for introducing these traditions is to better understand the role of these factors in intergenerational social mobility, that is the difference between a young person's circumstances growing up and their likely future circumstances (Lambert, Prandy and Bottero 2007: 2). A further reason is to evaluate the arguments for and against detraditionalisation. Finally, two further traditions have also been introduced to understand potential reasons for why family background may influence filial educational attainment. Firstly, human capital which is concerned with the investment in education by parents, and secondly, social capital or the development of human relations and ties. These are not the focus of this thesis but are included as part of the context in theoretical debate around the family and educational attainment. The concluding part of this chapter considers the relevance of these theories to empirical social research.

## Detraditionalisation

Not all theorists consider family background to be an important influence upon the educational attainment of young people. The contextual changes noted in chapter two are cited by some sociologists as evidence of processes of detraditionalisation. Giddens (1990) argues that the pace and scope of change plus the changing nature of institutions are evidence for the discontinuity from traditional to modern social orders. In particular traditional structural ties between individuals, such as social class, have loosened (Heelas, Lash and Morris 1996). Part of the detraditionalisation thesis argues that individuals are freer to live reflexively and make choices irrespective of traditional social orders (Beck 1992). Beck and Beck-Gernsheim (2002) claim that the concepts of social location (such as class, occupational group, social layers) are less relevant due to the ambiguity of social position through the lifecourse of individuals. These authors argue that the stable and predictable elements of these typologies need to be replaced with more dynamic concepts that accommodate the precarious nature of individuals' lives and decision making. From this perspective, it is argued that the role of education has for youth transition changed. *"Formal education in schools and universities, in turn provides individual credentials leading to individualised career opportunities in the labour market"* (Beck and Beck-Gernsheim 2002: 32). Individuals compete to gain the skills that will serve them well in securing employment. Others have criticised this general depiction arguing that social structures, institutions and lifestyles are more stable and predictable than in this depiction (Goldthorpe 2007a; McGovern et al. 2007).

Beck and Beck-Gernsheim (2002) emphasise that contemporary competition isolates individuals within formerly homogenous social groups and individualises their transitions. Essentially young people become more 'self-biographical' in their routes post-schooling and less constrained by traditional structural divisions by gender,

social class and ethnicity. Correspondingly there has been a growth of qualitative and cultural studies alongside studies using social survey data which seek to understand the identities of young people and their relationships with traditional collective identities such as gender and ethnicity (Heinz 2009). For example, Brannen and Nilsen (2002) note that young people appear to be categorised according to their attitudes towards the future and whether they are more adaptable or contingent, compared to emphasising predictability and security. However, this raises the question of whether the social background of a young person would be likely to influence their attitude. White (2007) for instance, emphasises the conservative, family-influenced choices (i.e. opting in to stability) of many young school leavers.

These theoretical assertions fail to account for the persisting association between parental occupation and their children's educational attainment, i.e. that the children of parents in advantaged occupations tend to achieve higher grades in school qualifications (see for example Scott 2004; Shavit, Yaish and Bar-Haim 2007). Further to this, access to a wider choice of post-compulsory education routes is not the universal experience of all but differential according to the educational attainment of young people (Furlong 2009). For example, these agendas of choice, strategy and risk may be more accessible to some groups of young people, such as middle class girls who tend to have higher levels of educational attainment in year 11 (Arnot and Mac an Ghaill 2006). This emphasises a contradiction within the individualisation thesis that there may be different choices available to young people yet such choices appear to still be structurally divided (Atkinson 2007). As Gayle, Lambert and Murray (2009b: 22) note, *"If the process of de-traditionalisation is taking place we would expect that the influence of individual-level factors (e.g.*

*social class, ethnicity and gender) have on young people's transitions (i.e. participation in education, employment and training) would be in decline."*

There is strong evidence to suggest that entirely decentralised or individualised transitions are not empirically supported. There are high levels of stability in the patterns of post-school transitions which are strongly associated with the educational attainment of young people, particularly in terms of gaining access to higher education (Furlong and Cartmel 1997; Forsyth and Furlong 2000; Forsyth and Furlong 2003b; Forsyth and Furlong 2003a; Furlong et al. 2006). There are strong associations between pupils family background and their educational attainment (Shavit, Yaish and Bar-Haim 2007). Indeed, similar patterns of stability have been emphasised in cross-nationally comparative research (e.g. Shavit and Müller 1998; Blossfeld et al. 2008). Reay (2006) argues that the individualisation thesis does not represent the collective experiences of young people. In particular: *"...regardless of what individual working-class males and females are able to negotiate and achieve for themselves within education, the collective patterns of working-class trajectories remain sharply different from those of the middle classes, despite over a hundred years of universal state schooling"* (Reay 2006: 294).

It is apparent that great inequality persists in the opportunities for young people according to their social and family background (Thompson 2009). The evidence for the associations between attainment and post-compulsory participation and attainment and family background are considered in further depth in chapter four. Such evidence suggests that, *"there is, after all, little doubt that marked inequalities exist within modern societies; that such inequalities are structured and have some degree of persistence over time. There is further little doubt that some of this*

*inequality can be described in terms of differential life chances related to the disposal of skills and resources on the market"* (Morgan 1996: 69).

## Social class and the stratification of occupations

Liberal theories of social class (based around the work of Max Weber) have traditionally been concerned with the patterning of life-chances by occupational grouping and the role these play in the perpetuation of social relations between such groups (see Scott 1996; Savage 2000; Bottero 2005). These theories of social class are concerned with the increase of class mobility and reduction over time of class-linked inequalities of opportunity (Erikson and Goldthorpe 1992; Goldthorpe 1996). In contrast with the detraditionalisation thesis, such inequalities of opportunity persist and there remains a strong and relatively stable association over time between educational attainment and being from an advantaged class background (Featherman, Lancaster Jones and Hauser 1975; Mare 1981; Goldthorpe 1996; Erikson et al. 2005).

Class based analyses have been criticised as not capturing the dynamics of contemporary social life particularly with the emergence of gender and ethnicity as alternative sources of inequality in educational attainment and has been portrayed as increasingly fragmented (Roberts 1977; Savage 2000). However, Furlong and Cartmel (1997: 5) note that, *"class still has an impact on people's life chances, but as a result of the fragmentation of social structures, collective identities have weakened."* Savage (2003) argues that class analysis needs to respond to the criticisms made by individualists, in particular that class based analyses need to be subtler in their investigation of how these identities are hidden and used. For

example, the strategies used by middle class families to perpetuate this advantage for their children particularly with regard to education (Ball 2003; Devine 2004).

The criticisms of class based analyses may in part be due to the massive expansion in the last 50 years of those occupations classified[3] as middle class, due to the growth of the service sector (Devine 2004; Breen and Jonsson 2005). In 1911, 75% of workers were in manual occupations. In 1991, this had fallen to 38% (Gallie 2000). Correspondingly during this period, there was a large increase in the numbers employed in professional occupations, as managers and administrators and as clerical workers. It is also important, given the extension of youth transitions (described at the beginning of this chapter) that the timing of the measurement of an individual's occupation may affect the outcome if this is too soon after leaving education (Egerton and Savage 2000). Further to this, there are problems of measuring women's occupations using traditional class schema as these may not accurately reflect their stratification position particularly for married women with children (Erikson 1984).

Partially as a consequence of these criticisms and in order to better represent the UK occupational structure, attention during the 1980s was largely concerned with the development of various occupation based class measurement schemes, from the Registrar General's Social Class scheme (Rose 1995) and Goldthorpe's Class schema (Goldthorpe, Llewellyn and Payne 1980; Goldthorpe and Llewellyn 1983) to the National Statistics Socio-Economic Classification which are based on occupation (Rose 1995; Rose and Pevalin 2003; Savage 2000; Savage 2003). Rose and Pevalin (2003) argue that:

[3] The subject of how occupations are classified and the schemes used to do so is discussed later in this section.

*"This one [tradition] sees individuals in similar socio-economic circumstances as occupying as a consequence common positions in the social structure in terms of social power and this concentrates on the relational aspects of inequality as well as the distributive ones. In other words, individuals possess certain resources as a result of their situations and consequently face a range of possibilities and constraints in terms of their behaviour. Those who share similar resources, and thus similar structural positions, will share similar possibilities and constraints in terms of 'life chances' (e.g. chances for educational attainment, health, material rewards and social mobility."*

(Rose and Pevalin 2003: 29)

As indicated above, inequalities in the occupational structure can be measured using a range of plausible alternative measures (Rose and Harrison 2009). However, social stratification is more than simply occupational inequality (Scott 1996). As was noted in the first chapter, stratification refers to the patterning of inequality across the lifecourse (Scott 1996; Bottero 2005). Morgan (1996: 41) identifies this position: *"I shall treat the term 'stratification' as the more general term dealing with the various ways in which hierarchical social divisions within society might be recognised or signified."* This is also why it is important to analyse parental occupations, as opposed to income, as these provide further information about family circumstances (Blau 1999; Bradley and Taylor 2004).

Measuring female and male occupations within a similar scheme can also be difficult due to the differences in the male and female labour market, the relative lack of dispersion in female occupations and greater likelihood of females being out of work

to care for children. Different measurement schemes have emerged to counter for this but it is also possible to use measures of family circumstances by taking the higher status occupation where two adults are living as a couple (Erikson 1984; Prandy 1990; Lambert, Prandy and Bottero 2007).

There is a need to consider parental occupation alongside factors such as gender, ethnicity and school type which together are broader than the traditional measurement of class. *"Social stratification occurs when structured social inequalities are systematically interrelated in the way that they shape people's life chances and are involved in the formation of large scale collectivities that stand in hierarchical relations to one another"* (Scott 1996: 191). Essentially individual choice is limited by the choices of those around us (Bottero 2005). What is sociologically important is that inequality in one generation potentially affects inequality in the next. *"The resources that are available to us growing up as children affect the success of our schooling, and so our eventual occupational careers, and the lifestyles we adopt as adults."* (Bottero 2005: 3) Therefore, the measurement of a range of family background measures is important in the study of educational attainment and there is a patterning in the educational attainment of young people and consequent decisions made and routes taken over time according to their background. These two points are discussed further in the next two sections on the family in the stratification of educational attainment, and social mobility and educational attainment.

## Family and the stratification of educational attainment

Morgan (1996) describes a 'tentative' model between the state, class, family and gender. Family in this sense includes households (e.g. those routinely sharing a meal under one roof) and those connected by relationships (kin or marriage). This thesis is focused upon Morgan's (1985) interaction between class/family namely:

a) *"The ways in which the family reproduces systems of patterned inequality*
b) *The ways in which systems of class or inequality place limits upon or provide opportunities for various modes of family living."*

(Morgan 1985: 67)

The family is an important institution in the reproduction (and removal) of these inequalities in terms of education, income, wealth and life chances of the individuals involved. In part this is due to dependency of offspring upon their parental resources. Furthermore, *"Certain tendencies within working-class families – earlier marriage, large families etc. – will combine with other economic features such as lower job security, poor housing and educational disadvantages to produce a mutually reinforcing system of disadvantage and poverty. The family cannot be said to 'cause' the condition of poverty but it is a major institution in its reproduction"* (Morgan 1985: 98). It is therefore important to measure a range of family background factors.

Part of the measurement of stratification must take account of factors such as parental education and family structure (such as lone parent compared with dual parent families) in order to understand, *"the processes by which the social structure as a whole reproduces itself over generations..."* (Morgan 1985: 102). The household and family structure in which a young person is living whilst growing up

is likely to affect the life chances they face (Morgan 1996). Goldthorpe (1983) argues that families may form a better unit for stratification analysis than individuals, as this captures more information about the wider circumstances in which an individual lives. By examining parental occupation and parental education, this provides information about the circumstances of these families whilst the individual young person's educational attainment remains the subject of interest. It is likely that the children of lone parents and dual parent families have different opportunities but it is not clear whether remains the case net of parental occupation.

## Social Mobility and educational attainment

Education is influential in studies of social mobility, particularly with regard to the relationship between socio-economic background, educational attainment and socio-economic attainment. Bills (2004) locates the association between family background and educational attainment in a Status Attainment model (see figure 3.1). This is founded upon Blau and Duncan's book *The American Occupational Structure* (Blau and Duncan 1967).

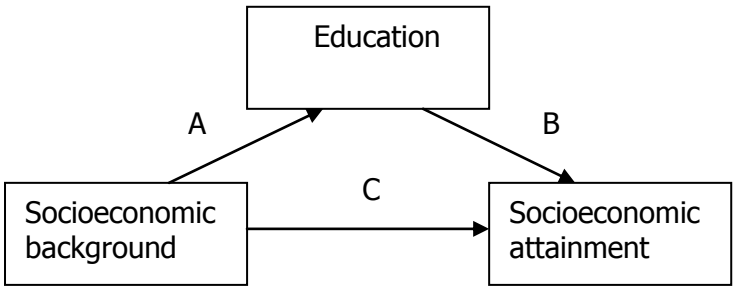



**Figure 3.1 A model of the socio-economic life cycle**

Source: (Bills 2004; adapted from Blau and Duncan 1967)

Studies of social mobility are concerned with the dynamics of the occupational structure, with education being part of this process (Blau and Duncan 1967). *"Intergenerational mobility is defined as the difference between social origins (during childhood) and current social circumstances"* (Lambert, Prandy and Bottero 2007: 2). Intragenerational mobility refers to the changing occupational classification during the lifecourse of an adult (Saunders 1990). Low levels of social mobility imply strong association between the occupation of parents and their children's occupations (using occupational classification schema). Studies of social mobility are often comparative and based on class schema (see for example Erikson and Goldthorpe 1992; Breen 2004). Both these studies use the Erikson-Goldthorpe scheme of social class as the basis for measuring social origins and destinations (Goldthorpe and Llewellyn 1983). However there is debate as to whether these patterns of occupational inequality are robust when measured using different classification schemes (Ganzeboom 2006; Lambert, Prandy and Bottero 2007). Ganzeboom, Luijkx and Treiman (1989) and Erikson and Goldthorpe (1992) contest whether there has been temporal variation in intergenerational mobility (in terms of socioeconomic background and socioeconomic attainment).

Erikson and Goldthorpe advocated a pattern of stability over time which came to be known as the 'constant flux hypothesis.' Breen (2004) argues that the evidence supports the hypothesis put forward by Featherman, Jones and Hauser (1975) that mobility is essentially externally driven by changes to occupational structure of countries and thus are rather rigid, albeit with some evidence of change (towards greater mobility over time) in some European countries. In general research using the Erikson-Goldthorpe scheme suggests that there may be absolute mobility as a consequence of these changes but less relative mobility (Goldthorpe 1980; Erikson

and Goldthorpe 1992; Breen 2004). Lambert, Prandy and Bottero (2007) however observe that when longitudinal data are available, there is a trend of slowly increasing social mobility, though they acknowledge though that intergenerational inequalities persist. It is worth cautioning that 'social mobility' as a term has been criticised as mobility includes upwards and downwards mobility (Goldthorpe and Jackson 2007). Frequently 'social mobility' is used when 'upward mobility' is the concept being referred to (particularly among politicians).

Education (as the precursor to occupational routes as an adult) remains a key mechanism in intergenerational social mobility (Savage and Egerton 1997; Breen and Jonsson 2005). Goldthorpe argues that the problem facing young people from different backgrounds is stability of the difference in the cost-benefit of remaining in education post-16. *"It is, moreover towards the end of the period of children's compulsory education, when crucial educational choices have to be made and when the question of opportunity cost first arises, that the earnings curves of parents in different classes are likely to be at their most divergent"* (Goldthorpe 1996: 493).

While education may be getting more equal in quality for all temporally (see a wealth of research on school quality and effectiveness, e.g. Teddlie and Reynolds 2000; Fitz-Gibbon 1996; Hoy, Bayne-Jardine and Wood 2000; Hopkins 2002; Gray 2004; Kyriakides 2008; Creemers and Kyriakides 2010; Strand 2010) the structural inequalities in resources available to families (prior to children going to school) have widened and it is these persisting differentials that continue to drive the difference in attainment (Goldthorpe 1996). The probability of successfully attempting to gain extra qualifications may also be crucial in the decision making of young people and this is also contingent on their existing qualifications (Breen and Jonsson 2005).

The study of social mobility is complicated by the changes to the occupational structure over time. The measurement of class position and occupational stratification is therefore crucial to establishing the basis upon which studies of social mobility are founded. Halsey, Heath and Ridge (1980: 197) note that, *"as a supplier of children, the service class increased from less than 1 in 10 in the first cohort (born 1913-22) to approximately 1 in 5 in the fourth (born 1943-52),"* as a consequence of the changing occupational distribution identified within chapter 2. Breen and Jonsson (2005) acknowledge this and describe the growth in microlevel (loglinear) models to address the problem of conflating change in the marginal distributions of occupational classes, such as that described by Halsey, Heath and Ridge (1980). As another example of changing the marginal distribution in educational qualifications, 71% of males (and 85% of females) born 1900-09 did not have a qualification. By contrast, only 8% of males (and 9% of females) born 1970-79 did not have a qualification (British Election Surveys 1979-97, cited in Smith 2000). This section highlights the persisting value of careful and systematic measurement of educational attainment and family background measures.

## The role of education in the stratification of occupations

The purpose of introducing stratification and social mobility research is to demonstrate the continued importance of measuring parental occupation and filial educational attainment. The final section of this chapter is to introduce theories regarding the reasons for why this association exists. Bills (2004) describes two general models for describing the relationship between educational attainment and occupations: the meritocracy model and the credentialist model. These respectively have a role in the assumptions behind two theories of the role of education in work: human capital and social capital.

## Human Capital

Human capital is an economic theory of differential rewards which can serve as a bridge between parental occupation, parental education and family structure (Becker 1991; Becker 1993). In essence, education enhances income prospects over the life course of an individual net of ability and family backgrounds (advantages). *"Human capital analysis assumes that the schooling raises earnings and productivity mainly by providing knowledge, skills, and a way of analyzing problems"* (Becker 1993: 19). This supports a meritocratic argument that educational qualifications represent the direct abilities and skills a person possesses and therefore their job performance, and is based on functionalism, that aspects of a society contribute positively to the functioning of that society (Bills 2004; see Merton 1957; Parsons 1967). The features of parental investment in their children are in human or nonhuman capital. Typically human capital comes in the form of contacts, values and skills whereas nonhuman capital is typically financial (Becker 1991). Both contribute towards a child's future income. Having lesser financial resources to give to their children, the less-advantaged parents need to invest further in human capital to compete with regard to their children's income as, *"richer families can pay for the training of their children, including the earnings forgone when children spend time in training rather than at work"* (Becker 1993: 22). The risk of failure to gain further qualifications is greater for young people from poorer families as the financial cost of remaining in education (earnings forgone) is greater relative to overall family income for poorer families (Breen and Goldthorpe 1997). This places importance on the contacts, values and skills that the parents have in being able to pass this on to their children. The contacts that parents have will more likely be with other adults of similar educational and occupational level (homophily) implying that parents from less-advantaged backgrounds (particularly in regard to occupation) are

less likely to have similar contacts, values and skills to those parents from more advantaged backgrounds (Bottero and Prandy 2003).

The processes of family income and wealth inequality tend to be amplified by the different family structures and family formation patterns. The more children a family has, the smaller the proportion of nonhuman capital (i.e. financial resources) that may be spent upon each child (Becker and Tomes 1976). Women from lower grade occupations and with lower educational attainment tend to have children younger and tend to have more children (Ekert-Jaffe et al. 2002). The temporal stability of family structure is also important. Heterogeneous families experiencing transition in forms, such as divorce or multiple episodes of single parenthood interspersed with periods of cohabitation, may tend to face situations of instability in their income, suggesting that it is important to include family structure in the analysis of family background (Morgan 1985; Steele et al. 2005). Statistical models of family background factors will therefore include a measure of family structure.

## Social Capital

As opposed to human capital, which is based around economic relations, social capital is concerned with human relations and ties (Field 2008). Social capital is the strength and number of relationships between actors (Portes 1998). There are several key authors, notably Bourdieu, Coleman and Putnam and each has a slightly different formulation on the concept. Social capital appears to bear similarity to Becker's (1993) non-human capital but differs in the application by each of the above authors. It is based upon a different tradition, credentialism, whereby educational qualifications are used as a means of controlling access to advantaged occupations rather than direct indicators of academic ability (see Bills 2004: 56; Bourdieu and Passeron 1990; Collins 1979).

For Bourdieu, success in educational qualifications is due to the possession of cultural capital (Sullivan 2002). The concept of cultural capital refers to how individuals are situated within a hierarchy of preferences (Bourdieu and Passeron 1990; Field 2008). These internalised tastes and aspirations are referred to as *habitus* by Bourdieu. Young people from advantaged family backgrounds will tend to prefer and aspire to educational attainment and more advantaged occupations as a consequence of exposure to this in their family setting (Bourdieu and Passeron 1990; Bottero 2005). De Graaf, de Graaf and Kraaykamp (2000: 107) suggest that higher educational attainment of young people from advantaged backgrounds may be the consequence of greater encouragement to read when young from parents, prevalence of books in the family home and a *"stimulating learning environment at home." Family habitus* conceptually links the cultural accounts of decision making with stratification in the opportunities and attainment of young people (see Brannen 2006). Educational qualifications are a form of cultural capital as forms of knowledge and competences that are valued by their social strata (i.e. their families) and by the designers of the examination system (Bottero 2005). The cultural capital available to young people from more advantaged backgrounds in post-compulsory education also is utilised via family contacts and networks (Field 2008).

In his studies of educational attainment, Coleman (1990) suggested that the influence of community norms on the expectations of parents, pupils and teachers was responsible for the higher rates of dropping out from school among young people from disadvantaged backgrounds in the USA (Coleman and Hoffer 1987; Field 2008). Putnam (1996) concentrates on the civic ties that bind individuals in

communities and the perceived loosening of these ties in the more individualised context of contemporary America. This is a national level form of social capital from a political science perspective (Seaman and Sweeting 2004). Education is important in this form of social capital as, *"...well-educated people are much more likely to be joiners and trusters, partly because they are better off economically, but mostly because of the skills, resources, and inclinations that were imparted to them at home and in school"* (Putnam 1996: 34). These theories suggest why family background has a continuing role in educational attainment.

## A critique of Human Capital and Social Capital theories

The term 'social capital' has had such diverse applications that it is now difficult to define the original meaning (Portes 1998). Whilst Woolcock (2001) attempts to unify the forms of social capital, there is still ambiguity about what is measured. Both human and social capital theories have been criticised as covering a wide range of resources which are difficult to measure (Block 1990; Sullivan 2002; Blackburn 2003; Bills 2004). Social capital approaches also seem to underemphasise that individuals can competitively use their status and privilege to further their own position, for example through increasing their capital (Field 2008).Human Capital theories appear to be tautological in that, *"...the attainment of high wages is often taken to indicate the presence of high rates of human capital"* (Bills 2004: 45), yet the subject of enquiry is the relationship between human capital (in the form of education) and socio-economic attainment (in the form of wages).

The debate between meritocrats and credentialists (described by Bills 2004) is not critical to the empirical investigation of the association between socio-economic background and educational attainment. Both suggest (for different reasons) that

young people seek educational qualifications in order to improve their chances of gaining an advantaged occupation. What constitutes social capital, in the field of education, appears to be defined from a middle class perspective and prioritises middle class forms of social capital above working class social capital (Ball 2003; Savage 2003; Reay 2006). These social networks appear to strongly reflect and reinforce inequality in resources and it is this inequality which it is more important to measure (Bottero 2005). The influence of social capital upon educational attainment appears to be similar to stratification and class position, where parents in advantaged occupations use their connections and resources to advance their children's education (Allatt 1993).

## Summary

At the beginning of this chapter the justification provided for introducing these theoretical traditions was to provide theoretical context to the 1990s as a period and to discuss the relevance of these theories to empirical social research. The changing times consensus reflects agreement among theorists that these contextual changes are a significant alteration from the processes of the first half of the 20th century (Gayle, Lambert and Murray 2009a). The 1990s as a decade represents a crucial period of study set against the contextual and policy changes identified in chapter two.

The debate between detraditionalisation theorists and those supporting stratification by family background represents an attempt to understand what these changes mean for the sociology of youth. Detraditionalisation theorists argue that the previous class-based way of understanding such changes is no longer relevant. Those observing these changes from the perspective of social stratification argue that measurement of differences between young people must change to reflect this. The social stratification tradition suggests that a broader range of measures needs to be included beyond occupation.

The empirical test of this is that if detraditionalisation theorists are correct then it would be expected that the influence of parental occupation would be stronger at the beginning of the 1990s and weaker at the end of the decade. Furthermore, there would be a lesser effect of gender and ethnicity upon educational attainment.

All of the theoretical traditions discussed suggest that it is important to measure wider family background factors and educational attainment. It would expected from these theoretical positions that these would likely be change over time in the

association. This is particularly relevant to the study of intergenerational social mobility.

Stratification as a term is used to suggest the patterning in inequality by this broader range of measures. Various empirical measurement schemes have been developed with respect to occupational measurement. There is therefore value in comparing these with respect to influence upon educational attainment. The family context relating to stratification suggests there is complexity in the manner in which family background characteristics coexist and that this must be evaluated empirically using a range of measures, including parental education and family structure as indicators of family background. Gender and ethnicity are also factors by which educational attainment is stratified and this will be discussed further in chapter four.

The sections on human and social capital suggest potential means by which parental occupation and parental education may influence educational attainment in terms of physical resources and social networks. Critics suggest that these mechanisms are hard to operationalise in practice and the nearest means of evaluating these with secondary survey data is by using available measures of parental attributes as proxies for advantage. This thesis does not seek to evaluate this but the theoretical context of these traditions is relevant in understanding potential reasons for the underlying inequalities in family background.

# 4 Stratification in Educational Attainment: The Youth Cohort Study

This chapter is a review of the applied research into the principal family background factors in the stratification of educational attainment. The review of empirical research in this chapter is principally a-theoretical in nature. Whilst theoretical traditions in the previous chapter were noted with regard to how these were relevant to empirical research, a direct evaluation of these theories does not locate well with empirical survey data analysis and is not the focus of this thesis.

In the previous chapter the review of theoretical traditions suggested the importance of measuring a range of family background factors over time. Consequently, measures of parental occupation, parental education and family structure but other variables are included (such as gender and ethnicity). This is followed by a section examining research using alternative measures of stratification.

The importance of studying the 1990s has been repeatedly emphasised and the majority of the chapter is a review of research using the Youth Cohort Study, a major dataset covering the educational attainment of young people in the 1990s in England and Wales. Further justification for the selection of the Youth Cohort Study as a dataset is provided in chapter five. Particular emphasis is given to the *Employment Department Research Series Reports* covering this period. These are mostly concerned with educational attainment and participation in post-compulsory education but some other relevant outcomes are included.

Following this is an investigation of the findings of work into the effects of family structure upon educational attainment during the 1990s using data from the *DSS/PSI Programme of Research into Low Income Families and the Families and Children Study* and consideration of potential difference in attainment between married and cohabiting parents. This chapter concludes with a summary of the main themes identified from research in these areas and the gaps in existing research identified.

## *Rising Attainment*

Since the inception of the General Certificate of Secondary Education (GSCE) in 1988 (as identified in chapter two) there has been a consistent rise in the number and proportion of students gaining 5 or more GCSE passes at grades A* to C (ONS 2001; ONS 2008; ONS 2009a). This threshold is important as it is usually a pre-requisite for gaining admission to A/AS level courses (and other more prestigious courses) which is the most common route to admission to higher education and higher grade occupations (Payne, Cheng and Witherspoon 1996; Gayle, Berridge and Davies 2003). GCSEs are the principal qualifications for those leaving school at age 16 and a key measure in school league tables (Leckie and Goldstein 2009). More recently the benchmark of 5 or more GCSE passes at grades A* to C now includes A* to C grades in English and Mathematics. In 1990, 35% of students gained 5 or more GCSE passes at grades A* to C. By 1999 this had risen to between 48% (YCS data) and 51% (Gayle, Lambert and Murray 2009b). By 2005/06, the figure had risen to 59.2% (ONS 2008) and more recently to 69.7% in 2008/09 (ONS 2009a). Rising GCSE attainment was preceded by a steady rise in O-level attainment (see figure 4.1). The 1990s is a key period in this trend, particularly as the decade follows the introduction of the GCSE as a qualification. It is clear however, that

there is considerable variation among the attainment of young people according to a range of social background factors. It is also clear that subject choice has a role in measuring attainment. For earlier data we do not have access to pass rates by subject but it is possible that overall attainment may be driven (in part) by changing curriculum design and changing subject choice.

## *Gender*

Since the introduction of GCSEs, the attainment of girls has improved consistently and relative to boys (Stobart, Elwood and Quinlan 1992; Arnot et al. 1998; Burgess et al. 2004; Arnot and Mac an Ghaill 2006). The gender effect has reversed since the 1970s and 1980s with growing concern about the lower attainment of boys (Younger and Warrington 2005; Gayle, Lambert and Murray 2009b). As can been seen in figure 4.1, this has continued through the 1990s.

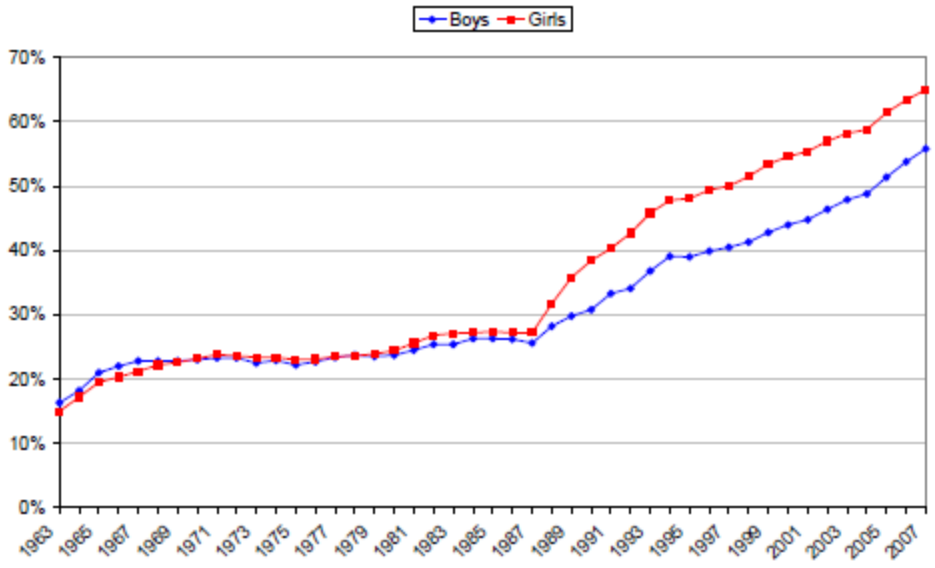



**Figure 4.1** **Percentage of 15 year old pupils in England achieving 5+ A*-C GCSEs (or equivalent), by gender, 1963-2007**

Source: Broecke and Hamed (2008: 4)[4]

[4] See also ONS - http://www.statistics.gov.uk/cci/nugget.asp?id=1892

There has been much concern at this divergence in attainment (Epstein et al. 1998; Raphael Reed 1999; Francis 2000). The point of noted divergence appears to be in 1988, which is when the GCSE was introduced as the standard qualification at the end of compulsory schooling. Stobart, Elwood and Quinlan (1992) note that the growth of the differential in educational attainment between girls and boys may be driven by the increased inclusion of coursework in GCSEs. They also observe that the differential is greater in subjects other than mathematics and some sciences (Stobart, Elwood and Quinlan 1992). However, Arnot et al. (1998) identify that gender gaps exist in GCSE performance as also existing at Key Stage 1 (after two years of compulsory schooling, i.e. age 7). Using data from the Department of Education and Employment from 1995, they note that girls do better than boys in reading, maths and science and make more progress between 7 and 16.

Corresponding with rising attainment among girls there appears to have been an increase in staying on rates through the 1990s. In England, the percentage of girls staying on till age 18 rose from 37% in 1986 to 59% in 1999 (ONS 2001). The corresponding figures for boys in this period were 46% in 1986 and 60% in 1999. These figures may have increased as a consequence of increasing attainment as those with higher attainment tend to stay on in education (Courtenay and McAleese 1993b). Broecke and Hamed (2008) used Youth Cohort Study (YCS) data from young people aged 16 in 2003/2004 and a matched dataset comprising of information from the National Pupil Database, the Pupil Level Annual Schools Census, and the Higher Education Statistics Agency Student Record to investigate gender gaps in participation in Higher Education. They identified that female participation in Higher Education has overtaken male participation as a consequence of greater attainment at age 16 (in 2007, 65% of girls achieved 5+ A*-C GCSEs or

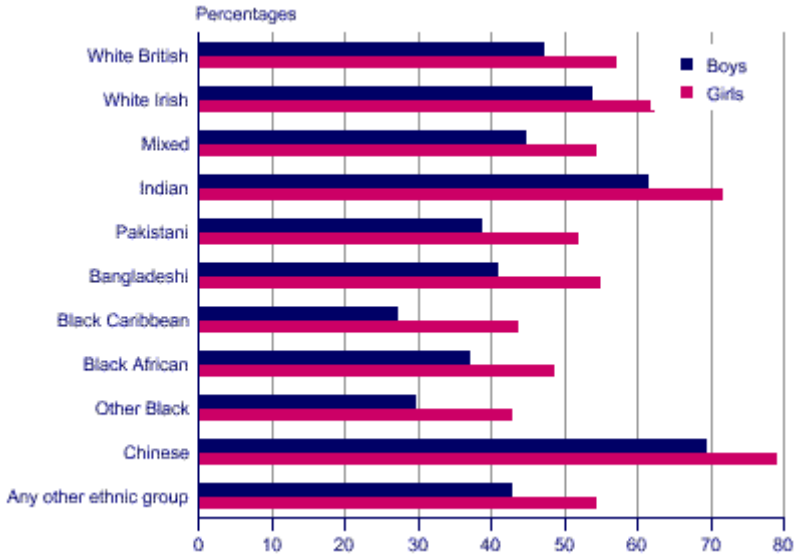
equivalent, compared to 55.8% of boys), greater staying-on rates post-16 (82% of girls and 72% of boys remained in full-time education in 2006) and greater take up and performance in A-levels than boys.

Yang and Woodhouse (2001) studied gender differences in progression from GCSE to A and AS Level using examination data from 722,903 students between 1994 and 1997 collected as part of the *ESRC Applications of Advanced Multilevel Modelling Methods for the Analysis of Examination Data project* (O'Donoghue et al. 1997). Using multi-level models, they demonstrated that female students with grades at GCSE averaging C or above tended to enter for fewer A/AS levels than similarly qualified males, and to obtain lower overall scores. This is surprising given the greater attainment relative to boys. It suggests that there are other factors that influence the decisions faced by girls to boys perhaps indicative of their social background and earlier differentials in educational progress (Strand 1999) .

## *Ethnicity*

There is considerable inter-linkage between the variation in GCSE attainment by ethnicity and parental occupation. As Payne (1995a: 4) identifies, *"holding constant year 11 GCSE results, institution, sex and other courses taken, working class young people and young people from minority ethnic groups did less well (i.e. received lower grades) at GCE A-level..."* By the term 'holding constant', Payne (1995a) is describing that if a young person is from a minority ethnic group or have parents in working class occupations, they are less likely to gain higher grades at A-level regardless of their previous GCSE results, school type, sex or whether they studied other courses.

Plewis (2009) looked at ethnic group differences in educational progress between the ages of 7 to 11 using the National Pupil Database. The progress made in educational attainment between these ages for Indian, Pakistani and Bangladeshi groups is faster than it is for white boys. Attainment at age 14 was lower for Black-Caribbean students independent of the effect of social class (Strand 2011). Rothon (2007) argues that social class operates in a similar manner across ethnic groups and that there is no specific ethnic effect mitigating how disadvantage operates for particular ethnic groups. For all ethnic groups, girls tend to outperform boys (Gillborn and Gipps 1996). This is illustrated below with ONS data fro*m 2004:*




Pupils achieving 5 or more A*-C at GCSE/GNVQ: by sex and ethnic group, 2004, England

**Figure 4.2 Pupils achieving 5 or more A*-C at GCSE/GNVQ: by sex and ethnic group, 2004, England**

Source: ONS (2006a)

The work of Drew, Gray and Sime (1992), Drew (1995), and Demack, Drew and Grimsley (2000) is particularly relevant as they used YCS data to examine the relationship between ethnicity and educational attainment. Drew, Gray and Sime (1992) investigated the labour market and educational outcomes of Afro-Caribbean young people aged between 15-19 using pooled data from YCS cohorts 2 (1985)

and 3 (1986), a sample of over 28,000 young people. They emphasised the importance of differences in material affluence between young people and their corresponding differences in educational attainment at age 16. Fourteen percent of Afro-Caribbean, 26% of Asian and 9% of White young people's fathers were unemployed. In terms of parental occupations (using socio-economic group or SEG[5]) 8% of Afro-Caribbean, 9% of Asian and 21% of White young people were from professional backgrounds. Sixty two percent of Afro-Caribbean, 60% of Asian and 47% of White were from manual backgrounds, suggesting that Afro-Caribbean and Asian young people tended to be from less advantaged backgrounds. Afro-Caribbean young people tended to have lower O-level or CSE attainment than their Asian and White counterparts (see also Drew and Gray 1990). 60% of Afro-Caribbean, 50% of Asian and 47% of White young people gained no higher grade passes (a higher grade pass being defined as an O-level pass or CSE level 1). 11% of Afro-Caribbean, 24% of Asian and 26% of White young people gained four or more higher grade passes (Drew, Gray and Sime 1992).

Using a logistic regression model predicting staying on in full-time education at 16 with the base groups being those with 4+ higher grade passes, white, parents in professional occupations, male, graduate parents and living in two-parent families the relative effects of lower attainment, ethnicity and parental occupation become clearer. Those with lower attainment were significantly less likely to stay-on than those with 4+ higher grade passes. However, controlling for the characteristics previously mentioned, the odds of Afro-Caribbean young people staying-on were three times those of White young people and for Asian young people ten times those of White young people. Within the same model, those young people with

[5] See Croxford (2006a) for an overview of how occupational variables were derived in the YCS.

parents in lower grade occupations, boys, those with non-graduate parents and those living in single parent families were all less likely to stay-on compared to the base category. From this model once attainment is controlled for ethnic origin was the most important characteristic in staying-on. Given that 62% of Afro-Caribbean young people were from manual backgrounds and 60% gained no higher grade passes, Afro-Caribbean young people were particularly likely to stay on compared to White young people (given that their socioeconomic background and attainment would be contributing to decrease their overall likelihood of doing so).

For those young people not staying on in full-time education post-16, participation in youth training schemes appears to have been associated with attainment, ethnicity, parental occupation and the local unemployment rate. Those with 1-3 higher grade passes at age 16, Afro-Caribbean young people and those with parents in manual occupations were more likely to choose the YTS than those with 4+ higher grade passes, white young people or those with parents in professional occupations. This is consistent with the findings of Gray and Sime (1990) that the YTS was often chosen by middle qualified young people. In the second model, living in an area with a higher local unemployment rate was added to the previous model. This tended to increase the likelihood of participating in the YTS and removed the effect of parental occupation (attainment and ethnicity remained significant effects). The effect of ethnicity is mediated by the local unemployment rate. Drew, Gray and Sime (1992) suggested that the greater odds of staying on despite comparatively poor qualifications for Afro-Caribbean young people represents an attempt to insulate themselves from potential labour market discrimination (see also Biggart and Furlong 1996) and that higher local unemployment increased the likelihood of

staying on, particularly on those courses (such as youth training schemes) which are less dependent upon prior attainment.

A logit analysis of unemployment at 19 using cohort 3 data suggests that the route taken, attainment, ethnic origin, parental occupation, gender and parental education significantly influenced whether a young person was unemployed or not. Of particular interest is that the odds of being unemployed for young people with no higher grade passes at age 19 were 3.5 times those of young people with 2 or more A-levels, net of the factors listed. The odds of being unemployed for young people with parents in manual occupations were 1.6 times those of young people with parents in professional occupations, net of the factors listed. Therefore, having parents in manual occupations decreased the initial likelihood of gaining higher grade passes and increased the likelihood of unemployment even once the effect of attainment have been controlled for within the model.

Drew (1995) investigated race, education and work extensively using data from YCS cohorts 2 & 3 (those eligible to leave school in 1985 and 1986). The ethnicity of young people is classified into three categories, Asian, Black and White. This is somewhat simplistic as an improved classification of ethnicity demonstrates that different Asian groups have very different educational attainment, for example Indian and Pakistani young people (Demack, Drew and Grimsley 2000). Drew (1995) identifies a strong gradient in attainment by parental occupation across all ethnic groups with the children of parents in professional occupations consistently outperforming those from intermediate, manual or other backgrounds in attainment at age 15/16 (using SEG). In a logistic regression model of staying-on in education post-16, the background of young people had a major influence on this decision

controlling for prior attainment. With base groupings of those young people with 4+ higher grade passes, white, parents in professional occupations, male, with graduate parents and living in two parent families, it is clear that Afro-Caribbean and Asian young people were much more likely to stay-on than White young people. Young people from intermediate and manual backgrounds were less likely to stay-on than those from professional backgrounds.

The attainment achieved by young people largely dictated the route that they took. Those with the highest attainment tend to stay-on and chose the A/AS level route. Eighty three percent of Asian young people with 4+ higher grades at age 15/16 choose to study for 2+ A-levels compared with 68% of Afro-Caribbean and 61% of White young people with equivalent attainment. Asian and Black young people were much more likely to attempt to improve their O-level performance than White young people if it was below 4 higher grade passes (O-level pass or CSE level 1). Those on youth training schemes (YTS) tended to be male and without any higher grade passes. This contrasts with Courtenay and McAleese's (1994) work which suggests that those of moderate attainment (highest GCSE grade being D-G) tended to participate in YTS. There did not appear to be a higher local adult unemployment rate for the areas in which minority ethnic groups lived which might have explained participation in YTS (as an alternative to unemployment).

In a logit model examining this, Drew (1995) confirmed that the highest odds of participating in the YTS by educational attainment was for those young people with 1-3 higher passes compared to those with 4+ higher passes. Afro-Caribbean and Asian young people were more likely to be on YTS than White young people, net of

the effects of the local unemployment rate, parental occupation, gender and parental education.

Importantly, ethnic minority young people were more likely than white young people to have parents in manual occupations, to have unemployed parents and to live in inner city areas. However, young people from ethnic minorities invested more heavily in post-compulsory education (given their attainment) than White young people suggesting attempts to counter this. Drew (1995) identifies that Afro-Caribbean and Asian young people were still more likely to be unemployed and less likely to have found full-time employment by age 18-19 than White young people suggesting significant discrimination and barriers still existed.

Using data from cohorts 4 (1988), 5 (1990), 6 (1992), 7 (1993) and 8 (1995) of the YCS (80,000 young people eligible to leave school between 1988 and 1995) Demack, Drew and Grimsley (2000) examined educational attainment at age 16 by ethnicity, gender and social class. The proportion of students gaining 5+ A-C grades at GCSE rose from 26% in 1988 to 41% in 1995. This is reflected in a points score measure (A grade at GCSE = 7 points to G grade at GCSE = 1 point) where during the same time period the mean points score attained by young people rose from 24 to 35.

The proportion of young people split by ethnic background or parental occupation classification (classified using Registrar General's Social Class, RGSC) is fairly stable between cohorts. The RGSC measure is constructed from socio-economic group (SEG) based upon the higher status occupation between parents (Demack, Drew and Grimsley 2000). Two groups in attainment by ethnicity were identified using

both the 5 A-C pass measure and the points score measure: a higher attaining group consisting of Chinese, Indian and White pupils; and a lower attaining group consisting of Pakistani, Black and Bangladeshi pupils. Across all of these measures, girls consistently outperformed boys.

The magnitude of the association between educational attainment and parental occupation (RGSC) is strong. To compensate for small ethnic minority group sub-samples, parental occupations were collapsed into non-manual, manual and unskilled/uncategorised. Consistently across ethnic groups, young people from non-manual backgrounds had higher attainment to those from non-manual and unskilled/uncategorised backgrounds. There was more variation by ethnicity among the non-manual classes than the manual classes. The strongest gradient (in 1995) by parental occupation is for Indian and White young people. For Indian young people 64% of those from non-manual backgrounds and 40% from manual backgrounds gained 5+ A-C passes. The corresponding percentages are 55% and 27% for White young people and 30% and 25% for Black young people (although for Black young people the sub-sample sizes are betweeen 30 and 100 in these categories and the non-manual category is quite broad in terms of occupations). Parental occupation appears to have had less of an effect for Black young people than White, as the differential in attainment between manual and non-manual backgrounds is smaller. From these data, there appears to be a widening gap in education attainment by gender, ethnicity and parental occupation during this period: *"When those pupils obtaining five or more high grade passes are considered, or the GCSE scores, the ethnic differences were larger than the gender differences and the social class differences were the largest of all."* (Demack, Drew and Grimsley 2000: 137). It is important to note though that Black, Pakistani and

Bangladeshi young people tend to have parents in manual occupations. Ethnicity and parental occupation are closely linked as: *"a person's social position reflects the cumulative influence of their parents' educational and employment-related experiences, as well as any difficulties or discrimination that are (or have been) associated with these"* (Demack, Drew and Grimsley 2000: 138).

## *Parental Education*

There is a strong association between parental educational attainment and their children's educational attainment, which is frequently studied by psychologists and economists (particularly with regard to human capital, see for example Chevalier and Lanot 2002; Chevalier 2004; Eccles and Davis-Kean 2005). Parental education and parental occupation are closely related (as would be expected) with those parents with greater educational attainment being more likely to be in more advantaged occupations. However, parental education measures have been included in models within this thesis to identify whether there remains a separate and distinct effect.

Lampard (2007) used data from BHPS wave 13 to investigate the relative effects of parental education and parental occupation upon whether or not their children gained service sector jobs (including controls for qualifications attained by the young person). While this is not an educational outcome, it is important to note that, "...*the increase in explanatory power provided by parental education is quite limited where both parents' occupations (when an individual was 14) are known."* More importantly the largest proportion of variance explained in the models is directly from the educational attainment of the young people concerned. As with much social mobility research, parental characteristics having prior influence on the

educational attainment of the young people and the subsequent routes to employment that they take (see figure 3.1; Blau and Duncan 1967).

Korupp, Ganzeboom and Van der Lippe (2002) compared the influence of mothers' and fathers' occupational and educational status on their children's educational attainment using a pooled dataset containing data from the Netherlands, West Germany and the USA. They observed that, *"the most prominent pattern discovered is that the influence of the parental education has a 'modified dominance' pattern... The influence of the higher educated parent is more pronounced than the influence of the lower educated parent. We find no dominance pattern for the influence of parental occupations. The influence of the lower status parent is slightly higher than the influence of the higher status parent"* (Korupp, Ganzeboom and Van der Lippe 2002: 31). It is possible that the parental educational variables are capturing some unmeasured aspect of parental occupation data, perhaps due to lack of resolution within the classification scheme for parental occupation. These effects are also likely to be influenced by homophily, whereby the parents of young people have partners with similar occupational or educational backgrounds to themselves (Ganzeboom, Treiman and Ultee 1991; Blossfeld and Timm 2003; Henz and Mills 2011).

Elliott, Dale and Egerton (2001) explored the importance of mothers qualifications upon their labour market decisions using data from the 1958 National Child Development Survey. Women with higher level qualifications (particularly if they are occupationally specific) were more likely to be in work (or return to work) at age 33 than those with no qualifications or non-occupationally specific qualifications. This demonstrates how parental qualifications may influence both the parental

occupation and whether they return to employment following childbirth. The inclusion of parental education information in addition to parental occupation measures in models of children's educational attainment may be of value as some parents may not be employed at the time of the survey but their education level may indicate their longer term employment prospects.

## *Parental Occupation*

There is a strong association between parental occupation and educational attainment prior to undertaking GCSE qualifications (see figure 4.3).

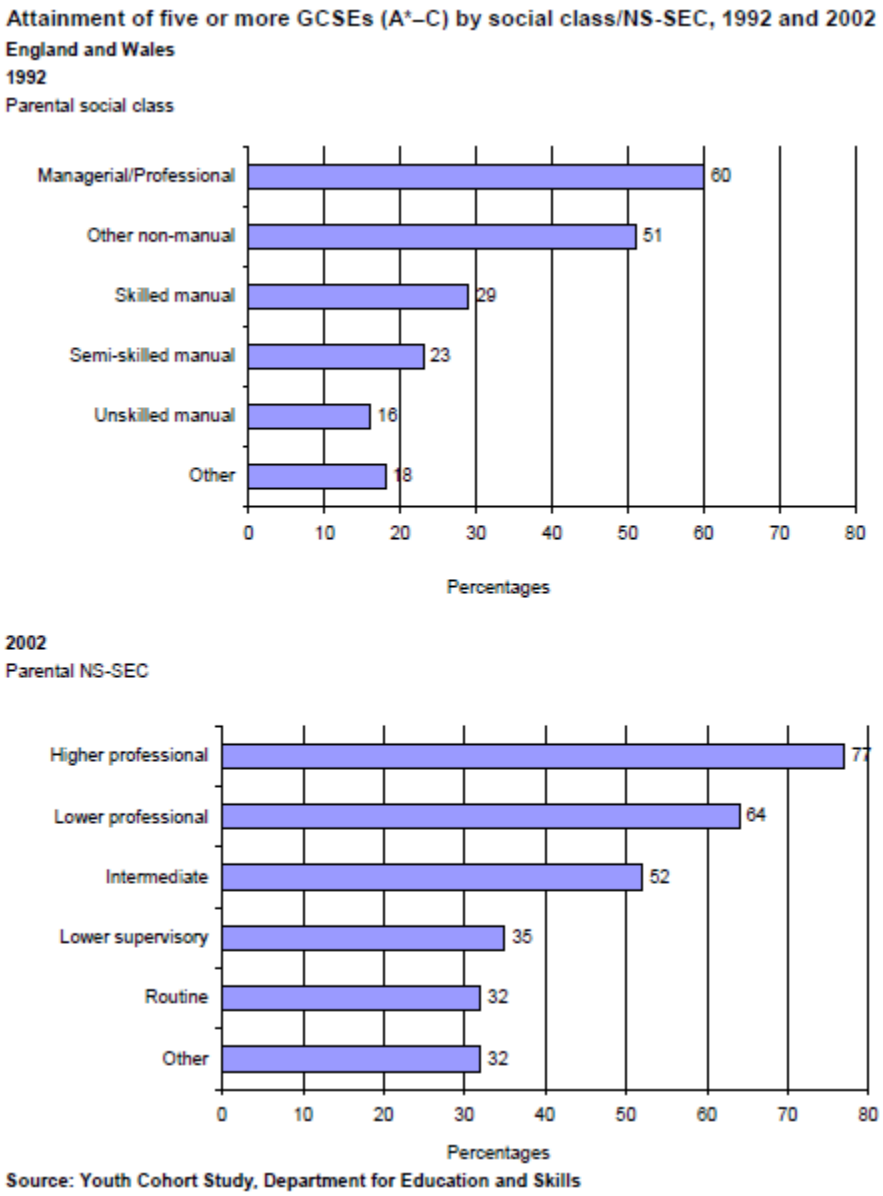



**Figure 4.3 Attainment of five or more GCSEs A*-C by social class/NS-SEC, 1992 & 2002**

Source: Babb (2005: 9)

Seventy seven percent of young people in 2002 with parents in higher managerial and professional occupations gained 5 or more GCSE passes at grades A*-C,

compared with 32% of with parents in routine occupations[6]. Lupton, Heath and Salter (2009) further illustrate the clear differential in GCSE attainment by parental occupation using SEG/NS-SEC from 1989 to 2004 using YCS data.

**Table 4.1** **Attainment of five GCSE grades A*-C in year 11, by social class (1989-1998), England and Wales.**

| Parental Occupation (SEG) | 1989 | 1991 | 1992 | 1994 | 1996 | 1998 |
|---|---|---|---|---|---|---|
| Managerial/professional | 52 | 58 | 60 | 66 | 68 | 69 |
| Other non-manual | 42 | 49 | 51 | 58 | 58 | 60 |
| Skilled manual | 21 | 27 | 29 | 36 | 36 | 40 |
| Semi-skilled manual | 16 | 20 | 23 | 26 | 29 | 32 |
| Unskilled manual | 12 | 15 | 16 | 16 | 24 | 20 |
| Other/not classified | 15 | 18 | 18 | 20 | 22 | 24 |
| Gap between highest and lowest (excluding other) | 40 | 43 | 44 | 50 | 44 | 49 |
| Ratio of highest to lowest categories | 4.3 | 3.9 | 3.8 | 4.1 | 2.8 | 3.5 |

Source: Lupton, Heath and Salter (2009: 79), DCSF YCSE SFR04/2005

This has been studied extensively in the form of social class schema or measures of parental occupation (Erikson et al. 2005). Recently this has been emphasised by Gayle, Lambert and Murray (2009b) in terms of participation in education in the October following Year 11 using data from YCS cohorts 5, 7, 8, 9 and 10 (those young people eligible to leave school in 1990, 1993, 1995, 1997 and 1999). There is a clear negative association between advantaged parental occupations (using RGSC, NS-SEC and CAMSIS) and the decision to leave school at the earliest opportunity.

## Alternative Measures of stratification

Using alternative measures of stratification confirms the strong influence of parental occupation upon educational attainment. Ermisch, Francesconi and Pevalin (2001) investigated children who have experienced poverty (defined as 60% of current

[6] for details of parental SEG and NS-SEC in the YCS, see Croxford (2006a)

median income, both currently and in previous years) using the British Youth Panel element of the British Household Panel Survey Data 1994-1999. Children who had experience poverty were greatly disadvantaged in terms of educational attainment compared to those from more affluent backgrounds being less likely to gain 5 or more A-C passes at GCSE, less likely to stay-on at school post-16 and more likely to become unemployed (Ermisch, Francesconi and Pevalin 2001).

Free School Meals (FSM) is a measure used by both the ONS (ONS 2007c) and social research. Cassen and Kingdon (2007) explore patterns of stratification in attainment. The ONS study looked at the characteristics of high attaining[7] young people at Key Stages 2, 3 & 4. At Key Stage Four (age 15/16), of the 58 900 high attaining young people in the National Pupil Database only 3.1% were registered as receiving FSM (ONS 2007c). This finding reinforces the linkage between poorly paid parental occupations (or unemployment) and low levels of attainment. This also has significant implication for subject choice. *"At Key Stage 4, 91% of mathematics A* grades and 89% of English A* grades are achieved by high attainers[8]. Among Key Stage 4 high attainers, 77% achieved grade A or A* in mathematics and 78% achieved these grades in English, whereas the equivalent rates for the rest of the cohort were 5% and 6%"* (ONS 2007c: 30). Cassen and Kingdon's uses the *Pupil Level Annual School Census (PLASC)*[9], a dataset containing 577 201 pupils in state secondary schools reaching the age of 16 in 2003. They note that of the 14% of pupils who receive FSM, nearly 47% of these get no passes above a D grade at GCSE. There is a strong difference between boys and girls on this measure with 54% of FSM boys and 41% of FSM girls failing to achieve above a GCSE D grade.

[7] High Attainment in this instance was defined as the top 10% of pupils at Key Stage 4 using capped GCSE point score.
[8] See previous footnote
[9] Based on the National Pupil Database

Furthermore, this continues a trend of previous low levels of attainment at a younger age.

Free School Meals is a basic indicator of poverty and more prevalent among the children of parents with lower grade occupations (using SEG). *"Twenty five percent of FSM pupils in 2003 had parents in 'routine' occupations, and a further 33% had both parents unemployed"* (DfES 2006; Cassen and Kingdon 2007: 15). However, Cassen and Kingdon (2007) acknowledge that it is an imperfect measure, particularly as so few young people receive FSM. It is inferior to measures such as NS-SEC which are relatively more stable over time, and are able to provide much greater information about the stratification in parental occupations and detail of the life circumstances of young people's family backgrounds (see for example Lupton, Heath and Salter 2009; Rose and Pevalin 2003).

## *Previous research using the YCS*

The following papers constitute a review of the *Employment Department Research Series Reports* which used YCS data. There are a number of themes that run through this series which are linked with aspects of educational attainment. The purpose of reviewing these papers is first substantive (the nature of the topics investigated and how these related to parental occupation, education and family structure/living arrangements). Second, to provide a thorough contextualisation of the dataset prior to analysis. This review contains the majority of research reports between 1991 and 1995, part of the period most relevant to the thesis. These reports have therefore been grouped thematically within this range, into the following topics, pre GCSE (1988), cohorts 4 & 5 descriptive reports (context 1988-1990), changing post-compulsory participation (1986-1988), transitions from school – early 1990s, further and higher education, and other YCS research.

## *Pre GCSE (1988)*

Roberts and Chadwick (1991) examined the transition of young people into the labour market between 1984 and 1987 using YCS cohort 1, comparing sweeps 1-3. The distributions of young people in education and employment in 1984-87 are very different to the late 1980s and 1990s. The proportion of young people moving from full-time education to full-time employment between age 16 and age 19 are much greater than in later cohorts, indicative of the labour market context of the period. The percentage of young people in full-time education fell from 39% in sweep 1 to 18% by sweep 3. Correspondingly the percentage in full-time employment increased from 29% in sweep 1 to 59% by sweep 3.

Roberts and Chadwick (1991) categorised the transitions made by young people by post-compulsory education year 2 (PC2) as full-time post-compulsory education, via the YTS, through full-time study combined with the YTS and 'traditional' (i.e. involving none of the previously mentioned, a 'catch all' category). The full-time post-compulsory study route was still very much for the minority, nineteen percent of female and 11% of male respondents. A third of both males and females were on youth training schemes as their main activity. Fifty one percent of males and 38% of females followed the traditional route.

For males of the highest level of attainment (4+ higher grade passes) at age 15/16, fifty two percent followed the traditional route compared with 30% choosing full-time study. For females of similar attainment, 35% followed the traditional route compared with 39% in full-time study. Using measures of parental occupation with YCS cohort 1 is difficult due to the small number in the sweep 3 sample who remained in full-time education (less than 1% of young people in full-time education identify as having parents with Professional occupations (RGSC I). The proportion of young people with parents in RGSC classes II & III (non manual) is much greater than for those who followed the traditional or Youth Training Scheme (YTS) routes.

According to Roberts and Chadwick (1991) the Youth Training Scheme (and other vocational initiatives) appeared to "mop up" middle qualified students (those with 1-3 higher grade passes) during recessionary periods and prevented youth unemployment rates being higher. The lowest qualified were more likely to be unemployed. Roberts and Chadwick (1991) suggested the problems facing poorly qualified young people in the 1990s. *"A greater medium-term danger, in our view, is of a growing underclass being left behind, especially if the emphasis on 'quality'*

*education, training, skills and young people becomes overwhelming"* (Roberts and Chadwick 1991: 52).

Jesson, Gray and Sime (1991) described the changing patterns of post-16 participation in full-time education using YCS cohort 3 (those young people who were age 15/16 in 1986). The principle differences among those choosing to remain in post-compulsory education appear to be among the ethnicity and gender of young people staying-on and grades they attained to enable them to do this. There has been a considerable growth in the number and proportion of ethnic minorities remaining in education post 16. Furthermore, as girls have outperformed boys in educational attainment at age 16 they have increasingly decided to continue to further study. Eighty one percent of young people with 4 or more higher passes (defined as those with GCE O-level grades A-C or CSE grade 1) were in full-time education in the year after post-compulsory education. This compares with 43% of those with 1-3 higher passes (GCE O-level grades A-C or CSE grade 1) and 16% of those with no higher passes. The dominant route for those young people with 4+ higher passes was to study for A-levels accounting for 68% of those with this level of attainment. This compares with 4% of those with 1-3 higher grade passes and 0% of those with no higher grade passes. This demonstrates the relationship between higher levels of attainment at age 16 (using the criteria described) and increased likelihood of continuing in education post-16.

Those young people who were most likely to stay-on had parents in professional occupations, graduate parents, were of non-white ethnicity (see also Drew 1995), and attended independent schools. Sixty three percent of those with Professional parental occupations (RGSC I & II) remained in education compared with 41% of

those with Intermediate (RGSC III non manual) and 29% of those with Manual parental occupations (RGSC III manual, IV & V). Sixty eight percent of those with graduate parents and 39% of those with non-graduate parents remained in education. Sixty nine percent of those of non-white and 42% of those with white ethnicity remained in education. Ninety three percent of those attending independent schools and 39% of those attending maintained schools at age 15/16 remained in education.

Staying-on in education post-16 was associated with higher attainment (at GCE O-level or CSE) but even when controlling for attainment, young people from families with parents in professional occupations were more likely to remain in education post-16. Among those young people with 4+ higher grades, 88% of those from professional backgrounds, 79% of those from intermediate backgrounds and 65% of those from manual backgrounds remained in education. Among those with no higher passes, 24% of those from professional backgrounds, 17% of those from intermediate backgrounds and 15% of those from manual backgrounds remained in education.

Jesson, Gray and Sime (1991) modelled whether a young person stayed-on in education post-16. Two levels were specified in the model, type of school attended to 16 and local unemployment rate for the area in which the school was located. The purpose of such a model is to examine whether there are differences between certain school types and local areas but also the differences between young people within those school types and local areas. Gaining higher grade passes (GCE O-level grades A-C or CSE grade 1) at age 15/16 was the strongest influence but there was a significantly increased likelihood in staying-on for girls, those from non-white

ethnic backgrounds, those with graduate parents and a small influence of the local unemployment rate. Those young people from manual backgrounds were significantly less likely to remain in education controlling for the prior factors. Remaining in education between Post Compulsory education year 1 (PC1) and Post Compulsory education year 2 (PC2) is strongly associated with attainment in PC1 (for those whose courses finished at the end of PC1).

Logistic regression modelling of whether young people remain in education post-16 identified a rising likelihood of participation between 1986 and 1988 for all levels of qualification with similar effects to the previous multi-level model. Jesson, Gray and Sime (1991) suggest that the decline of the youth labour market is a potential incentive to remain in education net of prior attainment, due to lack of alternative routes. Those with qualifications at age 16 and those without have different decision making processes: the cost-benefit of staying on differs according to socio-economic circumstances. Those with no qualifications were likely to be from manual backgrounds and do not have either the grades or the financial resources to stay in education.

Gray and Sime (1990) explored the changing context and nature of youth transitions in the UK, in particular the extension of post-16 educational routes and delay of entering the labour market between 1975 and 1986. The proportion of 16 year olds in employment fell from over 60% in 1976 to less than 20% by 1986 (DES Statistical Bulletin). They investigate these routes between the ages of 16 and 19 using data from YCS cohort 1 (those aged 15/16 in 1984) and the *ESRC 16-19 Initiative regional samples data* (from Swindon, Liverpool, Sheffield and Kirkcaldy). Using these data they identified a series of common routes which fit a simplified

typology of routes which ended in full-time education, employment, government schemes or unemployment.

Using this typology, they fitted a series of logistic regression models predicting the differing odds of attaining these various outcomes according to the effects of educational qualifications, gender, social class and ethnic background. As would be expected, the key determinant of staying in education is educational attainment. Those whose parents were in manual occupations were consistently less likely to remain in education than similarly qualified young people with parents in non-manual occupations. However, the effect of parental occupation upon staying-on for Black young people was reduced for those with higher levels of attainment. The differences in staying-on rates by manual/non-manual background were small for Black young people with higher attainment. This demonstrates that Black young people were consistently more likely to stay-on in education for all levels of attainment than White young people. Girls were consistently more likely to stay on than boys across all levels of attainment.

The probability of being unemployed was statistically independent of parental occupation (classified as manual/non-manual with no further details supplied), gender, region or ethnicity and was most strongly related to educational attainment at age-16. There was a further effect of whether a young person lived in the inner-city or not with: those living in the inner-city bring consistently more likely to be unemployed. The inner-city effect was greatest for those with low levels of educational qualifications. Young people from manual parental backgrounds were less likely to continue in education post 16 but this did not increase their chances of being out of work.

Roberts and Chadwick (1991) and Courtenay and McAleese (1994) assert that the growth of youth training schemes in the 1980s led to them being taken up largely by those with moderate qualifications (without O-levels/GCSEs higher grade passes at A-C) seeking to improve their chances of getting a job or those without alternatives (i.e. facing unemployment due to lack of local jobs), suggesting that this is not perhaps an active choice. Gray and Sime (1990) commented on the delayed point of decision making with regard to young people and the increased participation in further/higher education reflecting changing aspirations and avoiding unemployment. The effects of local labour markets were investigated and reported as affect the immediacy of gaining employment and the ability to insulate those with poor qualifications from being unemployed. It is clear that qualifications in the mid-1980s were the key means of remaining in education and avoiding unemployment.

## *Cohorts 4 & 5 descriptive reports (context 1988-1990)*

The following reports provide a contextual understanding of the 2 years prior to analysis of the 1990s data. Courtenay and McAleese (1993b) reported descriptive statistics from cohort 4 sweep 1 who were aged 16-17 in 1989. They note that the proportion of young people in full-time education within this age range rose from 41% in 1987 to 48% in 1989.

**Table 4.2 Activity at 16-17 years old by characteristics and experience (YCS 4 sweep 1)**

| | Full-Time Education | Full-Time Job | YTS | Out of Work | Other | Weighted Base |
|---|---|---|---|---|---|---|
| **Total Sample** | 48% | 25% | 22% | 4% | 2% | 14,116 |
| | | | | | | |
| **Gender** | | | | | | |
| Boys | 43% | 28% | 25% | 4% | 1% | 7,227 |
| Girls | 53% | 22% | 19% | 4% | 3% | 6,889 |
| | | | | | | |
| **Year 11 (5th year) school type** | | | | | | |
| Up to 18 comprehensive | 47% | 25% | 22% | 4% | 2% | 8,151 |
| Up to 16 comprehensive | 39% | 27% | 27% | 5% | 2% | 3,925 |
| Other state maintained | 38% | 35% | 22% | 3% | 2% | 1,005 |
| Independent | 93% | 4% | 2% | * | 2% | 1,036 |
| | | | | | | |
| **Parents' socio-economic group** | | | | | | |
| Non-manual | 65% | 18% | 14% | 2% | 1% | 5,312 |
| Skilled manual | 39% | 29% | 26% | 4% | 2% | 5,041 |
| Semi-skilled manual | 35% | 29% | 29% | 6% | 2% | 1,588 |
| Unskilled manual | 27% | 31% | 32% | 6% | 3% | 465 |
| Unclassifiable/not stated | 36% | 26% | 28% | 7% | 3% | 1,710 |
| | | | | | | |
| **Ethnic Group** | | | | | | |
| White | 47% | 25% | 22% | 4% | 2% | 13,137 |
| Black | 68% | 10% | 15% | 4% | 3% | 198 |
| Asian - subcontinent | 74% | 9% | 12% | 4% | 1% | 334 |
| - other | 84% | 4% | 5% | 4% | 3% | 124 |
| Other ethnic group | 60% | 14% | 14% | 6% | 6% | 88 |
| Not stated | 25% | 32% | 32% | 6% | 6% | 234 |
| | | | | | | |
| **5th year highest qualifications** | | | | | | |
| GCSE grades ABC | 65% | 19% | 13% | 2% | 2% | 8,601 |
| Other GCSE | 23% | 32% | 37% | 5% | 2% | 4,095 |
| No GCSE | 11% | 38% | 34% | 14% | 4% | 1,412 |

Source: Courtenay and McAleese (1993b: 7)

Boys were less likely to continue in full-time education but more likely to have a full-time job. There were strong patterns in staying-on full-time by school type, parental socioeconomic group and ethnicity for the same age group. It is important to note that there is no information on the derivation of parental socioeconomic group (SEG) for YCS cohorts 1-8 (1984-1995) (see Croxford 2006a).

Between 1987 and 1989, the percentage gaining an A-G pass rose from 85% to 90% and those gaining an A-C pass increased from 54% to 61%. This was heavily stratified though by school type, parental SEG and ethnicity.

**Table 4.3 Year 11 GCSE attainment by grade (YCS 4 sweep 1)**

| | ABC | Other | None | Weighted Base |
|---|---|---|---|---|
| **All** | 61% | 29% | 10% | 14,116 |
| | | | | |
| **Gender** | | | | |
| Boys | 57% | 31% | 12% | 7,227 |
| Girls | 65% | 27% | 8% | 6,889 |
| | | | | |
| **Year 11 (5th year) school type** | | | | |
| Total maintained | 59% | 31% | 10% | 13,080 |
| Up to 16 comprehensive | 56% | 34% | 10% | 3,925 |
| Up to 18 comprehensive | 60% | 29% | 10% | 8,151 |
| Other maintained | 55% | 35% | 9% | 1,005 |
| Independent | 90% | 3% | 8% | 1,036 |
| | | | | |
| **Parents' socio-economic group** | | | | |
| Non-manual | 76% | 19% | 5% | 5,312 |
| Skilled manual | 57% | 33% | 10% | 5,041 |
| Semi-skilled manual | 47% | 39% | 14% | 1,588 |
| Unskilled manual | 40% | 43% | 18% | 465 |
| Miscellaneous | 70% | 29% | 1% | 131 |
| Unclassifiable | 40% | 38% | 22% | 1,710 |
| | | | | |
| **Ethnic Group** | | | | |
| White | 61% | 29% | 10% | 13,137 |
| Black | 53% | 33% | 14% | 198 |
| Asian - subcontinent | 59% | 34% | 8% | 334 |
| - other | 72% | 23% | 5% | 124 |
| Other ethnic group | 56% | 32% | 13% | 88 |
| Not stated | 38% | 36% | 26% | 234 |

Source: Courtenay and McAleese (1993b: 31)

School type, parental occupational classification (SEG) and ethnicity remained strongly correlated with GCSE attainment (see table 4.3). It is clear from these data that the decision to remain in full-time education at age 16-17 is strongly associated with educational attainment at age 15-16. In addition for both measures there is strong stratification by school type, parental socio-economic group and ethnicity. This is important as the most common reasons young people cite for deciding to stay-on full-time were, to improve their chances of getting a job or to apply to college/university (Courtenay and McAleese 1993b).

Courtenay and McAleese (1994) reported on the YCS data from cohort 4 sweep 2 or those young people aged 17-18 in 1990. The focus of this report is consistent with educational concerns of the early 1990s with a greater emphasis on routes into the labour market reflecting the dominant route taken by young people at the time: the minority of young people progressing into full-time education by age 17-18 in the late 1980s. The percentage of those in full-time education between sweep 1 (age 16-17 in 1989) and sweep 2 (age 17-18 in 1990) of YCS cohort 4 fell from 48% to 41%.

**Table 4.4 Activity at 17-18 years old by characteristics (YCS 4 sweep 2)**

| | Full-Time Education | Full-Time Job | YTS | Out of Work | ET | Other | Weighted Base |
|---|---|---|---|---|---|---|---|
| **Total Sample** | 41% | 35% | 15% | 2% | 4% | 2% | 10,464 |
| | | | | | | | |
| **Gender** | | | | | | | 5,357 |
| Boys | 37% | 38% | 17% | 2% | 4% | 1% | 5,107 |
| Girls | 45% | 33% | 12% | 1% | 4% | 5% | |
| | | | | | | | |
| **Parents' socio-economic group** | | | | | | | |
| Non-manual | 60% | 26% | 9% | 1% | 2% | 2% | 4,067 |
| Skilled manual | 32% | 42% | 18% | 2% | 4% | 3% | 3,736 |
| Semi-skilled manual | 26% | 41% | 22% | 2% | 6% | 3% | 1,198 |
| Unskilled manual | 20% | 43% | 19% | 3% | 11% | 4% | 349 |
| Unclassifiable/not stated | 27% | 39% | 19% | 2% | 8% | 5% | 114 |
| | | | | | | | |
| **Ethnic Group** | | | | | | | |
| White | 40% | 36% | 15% | 2% | 4% | 3% | 3,805 |
| Black | 65% | 11% | 12% | 1% | 10% | 2% | 123 |
| Asian - subcontinent | 67% | 16% | 8% | 1% | 5% | 2% | 247 |
| - other | 77% | 16% | 4% | - | - | 3% | 90 |
| Other ethnic group | 54% | 26% | 12% | 1% | 5% | 1% | 60 |
| Not stated | 20% | 45% | 21% | 3% | 4% | 7% | 139 |
| | | | | | | | |
| **5th year highest qualifications** | | | | | | | |
| GCSE grades ABC | 59% | 27% | 8% | 1% | 2% | 2% | 6,437 |
| Other GCSE | 13% | 49% | 26% | 3% | 6% | 3% | 2,886 |
| No GCSE | 12% | 62% | 3% | 3% | 14% | 6% | 1,141 |

Source: Courtenay and McAleese (1994: 7)

The association between GCSE attainment and remaining in education at age 17-18 is very strong. Of those young people with an A-C grade at GCSE, 59% remained in full-time education compared to 13% of those without an A-C grade (see table 4.4). Twenty six percent of those without an A-C grade (but with a GCSE grade) were on the YTS suggesting that the YTS was participated in mostly by young people with moderate attainment. From observing figure 4.4, it appears that the greatest risk of unemployment is for those young people without any qualifications. Girls were more likely to stay on than boys: 45% of girls compared with 37% of boys being in full-time education at age 17-18. This reflects the higher attainment of girls at GCSE: 65% of girls and 57% of boys having gained at least one A-C grade by this age.

The pattern of attainment by parental socioeconomic group (SEG) noted by Courtenay and McAleese (1993b) appears to be amplified when looking at the proportion of young people in full-time education at 17-18: 60% of young people of parents from non-manual occupations and 20% of young people of parents with unskilled manual occupations remaining in education at this age. This suggests that young people with parents in unskilled manual occupations have lower initial attainment (Courtenay and McAleese 1993b) and that there is a greater tendency to leave the educational route sooner than those with higher attainment. Interestingly, this association between GCSE attainment and remaining in post compulsory education appears weaker according to ethnic group. Seventy seven percent of Asian, 40% of White and 65% of African/West Indian students remained in full-time education at age 17-18. This appears to support the argument that young people from ethnic minority backgrounds insulate themselves from labour market discrimination by remaining in post-compulsory education (Drew, Gray and Sime 1992; Drew 1995).

Park (1994) noted the increasing overall GCSE attainment and proportion of young people in post compulsory education by comparing students in state schools[10] at age 18-19 between cohorts 1-4 (1987-1991). The proportion in full-time education at this age increased from 18% to 27%. It is suggested that the higher attainment among young people at 16 since introduction of GCSE is likely to have encouraged more young people to take A-levels (Park 1994: 11; Courtenay and McAleese 1993b; Courtenay and McAleese 1994).

[10] Only state schools were included because cohort 1 didn't contain information for other school types.

Activity status at 18-19 (in 1991) was clearly stratified by parental socio-economic group (SEG). Of those young people whose parents were from professional and managerial backgrounds, 43% were in full-time education, 43% in a full-time job and 6% were unemployed. By contrast, for those young people whose parents were in unskilled manual occupations, 13% were in full-time education, 54% were in a full-time job and 22% were unemployed. There was a similar association for parental education with the majority of those young people where one or both parents had a degree being in full-time education.

Cohort 4 appears to be the starting point for trends that persist through the 1990s, namely increasing participation post-16 and decreasing participation in full-time work by age 18-19. At age 16-17 (in 1989) 48% of the cohort 4 sample were in full-time education (see table 4.2). By age 18-19 (in 1991) this had fallen to 27% (see table 4.5).

**Table 4.5 Activity status at 18-19 years old and characteristics (YCS 4 sweep 3)**

| | Full-Time Education | Full-Time Job | YTS | Out of Work | Other | Weighted Base |
|---|---|---|---|---|---|---|
| **All** | 27% | 54% | 2% | 10% | 6% | 8189 |
| | | | | | | |
| **Gender** | | | | | | |
| Boys | 26% | 57% | 4% | 11% | 3% | 4193 |
| Girls | 27% | 51% | 2% | 10% | 10% | 3996 |
| | | | | | | |
| **Parents' socio-economic group** | | | | | | |
| Prof. and managerial | 43% | 43% | 1% | 6% | 7% | 1827 |
| Other non-manual | 36% | 47% | 2% | 8% | 6% | 1406 |
| Skilled manual | 19% | 63% | 3% | 10% | 5% | 2931 |
| Semi-skilled manual | 17% | 58% | 5% | 14% | 6% | 936 |
| Unskilled manual | 13% | 54% | 3% | 22% | 8% | 273 |
| Misc/unclassifiable | 18% | 56% | 4% | 15% | 5% | 816 |
| | | | | | | |
| **Parents' education** | | | | | | |
| Both parents have degree | 64% | 21% | 2% | 4% | 11% | 296 |
| Father only has degree | 54% | 32% | 1% | 6% | 7% | 444 |
| Mother only has degree | 55% | 34% | - | 4% | 8% | 149 |
| Neither has degree | 26% | 57% | 3% | 9% | 5% | 3957 |
| Not sure | 16% | 55% | 4% | 17% | 7% | 913 |
| Not stated | 22% | 58% | 3% | 11% | 6% | 2094 |
| | | | | | | |
| **Ethnic Group** | | | | | | |
| White | 25% | 55% | 3% | 10% | 6% | 7696 |
| Black, African, Caribbean | 57% | 19% | 7% | 14% | 3% | 96 |
| Asian | 59% | 28% | 3% | 6% | 4% | 265 |
| None of the above | 36% | 38% | - | 15% | 12% | 44 |
| Not stated | 14% | 64% | 3% | 11% | 9% | 89 |

Source: Park (1994: 14)

**Table 4.6 Highest qualification at 18-19 (YCS 4 sweep 3)**

| | All | A or A/S level | GCSE A-C | Other Grades | None |
|---|---|---|---|---|---|
| **Weighted base** | 8189 | 1897 | 3357 | 2122 | 812 |
| | | | | | |
| **Gender** | | | | | |
| Men | 51% | 49% | 48% | 55% | 62% |
| Women | 49% | 51% | 52% | 45% | 38% |
| | | | | | |
| **Parents' socio-economic group** | | | | | |
| Non-manual | 39% | 65% | 40% | 23% | 18% |
| Skilled manual | 36% | 23% | 39% | 42% | 35% |
| Semi-skilled manual | 11% | 6% | 10% | 16% | 19% |
| Unskilled manual | 3% | 1% | 3% | 5% | 6% |
| Misc/unclassifiable/not stated | 10% | 5% | 8% | 14% | 22% |
| | | | | | |
| **Parents' education** | | | | | |
| Father only has degree | 5% | 12% | 5% | 1% | 1% |
| Mother only has degree | 2% | 4% | 2% | 1% | * |
| Both parents have degree | 4% | 11% | 2% | * | 1% |
| Neither has degree | 48% | 49% | 50% | 46% | 43% |
| Not sure | 11% | 4% | 10% | 15% | 19% |
| Not stated | 26% | 18% | 27% | 29% | 30% |

Source: Park (1994: 29)

Of those young people whose highest qualification was an A or AS level or a GCSE A-C pass (by age 18-19), the proportion of males to females was approximately equivalent. However, of those reporting no qualifications, 62% were male compared with 38% female (see table 4.6). Socioeconomic background (using SEG) and ethnicity are strongly associated with attaining A-level or AS level qualifications. Of those who had an A or AS level by age 18-19, 94% were White, 3% were Asian and 1% were Black suggesting that young people from minority ethnic backgrounds were on vocational routes post-16 and that these courses tended to be shorter in duration (given the higher rates of participation among such groups). Parental education was similarly associated with the highest qualification attained, although to a lesser extent.

Finally in this series of cohort specific reports Courtenay and McAleese (1993a) described cohort 5 sweep 1, those young people aged 16-17 years old in 1991. Increasingly young people chose to stay in full-time education post 16, from 41% in 1987 (cohort 3) to 48% in 1989 (cohort 4) and to 58% in 1991 (cohort 5). For this age group, remaining in full-time education was heavily stratified by school type, parental social occupation classification (using SOC90, see Courtenay and McAleese 1993a: Appendix E), ethnicity, and GCSE attainment by this age.

The relationships described in YCS cohort 4 (1988) are similar to those found in YCS cohort 5 (1990) with persisting associations between remaining in education and gender, school type, parental occupation, ethnicity and GCSE attainment (see table 4.7). The analysis by SOC90 major groups provides more detailed parental occupation groupings (Courtenay and McAleese 1993a; Croxford 2006a). There was a modest rise in those gaining one or more A-C grade at GCSE by age 16-17 from 61% in 1989 to 63% in 1991. GCSE attainment remains the strongest predictor of remaining in full-time education at age 16-17. Courtenay and McAleese (1993a) note the rising participation among those with better qualifications. To reflect the rising levels of attainment, the measurement of GCSE performance included volume from an A-C pass to 4+ A-C grades at GCSE.

**Table 4.7 Activity at 16-17 years by characteristics and experience (YCS 5 sweep 1)**

| | Full-Time Education | Full-Time Job | YTS | Out of work | Other | Weighted Base |
|---|---|---|---|---|---|---|
| **Total Sample** | 58% | 18% | 14% | 7% | 3% | 14,511 |
| | | | | | | |
| **Gender** | | | | | | |
| Boys | 53% | 22% | 16% | 7% | 2% | 7,410 |
| Girls | 64% | 14% | 13% | 6% | 4% | 7,101 |
| | | | | | | |
| **Year 11 (5th year) school type** | | | | | | |
| Up to 18 comprehensive | 58% | 19% | 14% | 7% | 2% | 4,550 |
| Up to 16 comprehensive | 50% | 20% | 19% | 8% | 3% | 7,792 |
| Other state maintained | 62% | 17% | 11% | 7% | 2% | 1,080 |
| Independent | 95% | 3% | 1% | 1% | * | 1,088 |
| | | | | | | |
| **Parents' Social Occupation Classification (SOC)** | | | | | | |
| Managers | 74% | 13% | 9% | 3% | 2% | 2,433 |
| Professionals | 86% | 6% | 6% | 2% | 1% | 1,361 |
| Associate prof./tech | 76% | 12% | 7% | 3% | 2% | 904 |
| Clerical and secretarial | 66% | 16% | 11% | 4% | 3% | 780 |
| Craft and related | 51% | 21% | 19% | 7% | 3% | 3,114 |
| Personal and protect services | 56% | 21% | 13% | 9% | 2% | 857 |
| Sales | 57% | 20% | 14% | 7% | 2% | 519 |
| Plant/machine ops | 41% | 25% | 21% | 10% | 3% | 1,641 |
| Other | 37% | 25% | 23% | 12% | 3% | 889 |
| Unclassifiable/not stated | 46% | 19% | 17% | 13% | 5% | 1,915 |
| | | | | | | |
| **Ethnic Group** | | | | | | |
| White | 58% | 18% | 15% | 7% | 2% | 13,264 |
| Black | 71% | 8% | 8% | 12% | 1% | 188 |
| Asian - subcontinent | 73% | 7% | 8% | 8% | 3% | 629 |
| - other | 89% | 5% | 6% | - | - | 121 |
| Other ethnic group | 67% | 10% | 2% | 14% | 8% | 81 |
| Not stated | 34% | 31% | 19% | 11% | 6% | 229 |
| | | | | | | |
| **Highest GCSE** | | | | | | |
| 4+ ABC grades | 88% | 7% | 3% | 1% | 1% | 5,926 |
| 4+ other grades (incl. Less than 4 ABCs) | 46% | 24% | 20% | 7% | 3% | 6,261 |
| 1-3 other grades | 20% | 26% | 29% | 21% | 5% | 1,118 |
| None | 9% | 33% | 25% | 27% | 7% | 818 |
| Not stated | 21% | 27% | 23% | 22% | 8% | 287 |

Source: Courtenay and McAleese (1993a: 8)

**Table 4.8 Year 11 GCSEs and respondent characteristics (YCS 5 sweep 1)**

| | 4+ A-C | 4+ (no A-C) | 1-3 any grade | None | Not stated | Weighted Base |
|---|---|---|---|---|---|---|
| **All** | 41% | 44% | 8% | 6% | 2% | 14,511 |
| | | | | | | |
| **Gender** | | | | | | |
| Boys | 37% | 45% | 9% | 7% | 2% | 7,410 |
| Girls | 45% | 42% | 7% | 5% | 1% | 7,101 |
| | | | | | | |
| **Activity at survey** | | | | | | |
| FT education | 61% | 35% | 3% | 1% | 1% | 8,469 |
| FT job | 16% | 60% | 11% | 10% | 3% | 2,582 |
| Unemployed | 9% | 62% | 15% | 10% | 3% | 2,085 |
| Either | 19% | 46% | 14% | 16% | 6% | 380 |
| | | | | | | |
| **Ethnic Group** | | | | | | |
| White | 42% | 44% | 7% | 6% | 2% | 13,264 |
| Black | 27% | 52% | 10% | 8% | 3% | 188 |
| Asian - subcontinent | 29% | 50% | 13% | 5% | 3% | 629 |
| - other | 56% | 30% | 12% | - | 1% | 121 |
| Other ethnic group | 40% | 45% | 2% | 8% | 5% | 81 |
| Not stated | 21% | 45% | 13% | 11% | 10% | 218 |
| | | | | | | |
| **Year 11 (5th year) school type** | | | | | | |
| Up to 16 comprehensive | 34% | 48% | 9% | 7% | 2% | 4,550 |
| Up to 18 comprehensive | 38% | 47% | 7% | 6% | 2% | 7,792 |
| Other maintained | 47% | 39% | 9% | 4% | 2% | 1,080 |
| Independent | 87% | 8% | 1% | 1% | 2% | 1,088 |

Source: Courtenay and McAleese (1993a: 17)

This is noteworthy as it appears to show the importance of gaining a number of passes above GCSE grade C, perhaps indicating the growing importance of GCSE attainment by 1991. The majority (53%) of those studying for vocational qualifications gained 4+ grades (no A-C passes) by age 16-17 suggesting the diversity of routes taken by young people with this moderate level of attainment. For details of the other associations between the characteristics of young people and their GCSE attainment, see table 4.8.

In terms of subject choice among those young people studying for academic subjects at age 16-17, English was studied by 28%, mathematics by 30% and between 10-30% studied creative arts, biology, social studies, geography, physics

and history. It would be of great interest to know more about the GCSE subject choice, as the majority of young people studying for academic subjects at age 16-17 would be on the A or AS level route. This provides part of the rationale behind the decision to investigate attainment in GCSE subjects using latent class analysis in chapter 9. The evidence presented within these reports suggest that the influence of parental occupation upon GCSE attainment remains strong. The difference in routes according to GCSE attainment is another noted finding of the reports analysing cohorts 4 and 5 of the YCS.

## *Changing Post-Compulsory Participation (1986-1988)*

Ashford, Gray and Tranmer (1993) assessed in further detail the introduction of GCSE exams and changes in post 16 participation. This was achieved via a cross sectional comparison of pre-GCSE cohort 3 (1986) and post-GCSE cohort 4 (1988) data. The percentage of young people staying on for at least one year of post-compulsory education rose from 43% to 51%. The attainment of these young people improved between these cohorts from 16% to 19% gaining 7+ A-C grades at GCSE and from 13% to 18% gaining 4-6 A-C grades at GCSE. Further to this, the proportion of young people staying on for some post-compulsory education (irrespective of the level of qualifications attained in year 11) rose between the pre and post GCSE cohorts. In 1986, 92% of those with 7 or more A-C grade O-levels remained in education post-16 compared with 96% of those with 7 or more A-C grade GSCES in 1988. Similarly, the percentage of who did not have any graded passes who remained in education post-16 rose from 5% in 1986 to 12% in 1988.

A-levels increasingly became the dominant form of post-16 qualifications, the proportion of young people studying for 2+ A-levels rose from 49% in 1986 to 57% in 1988. Correspondingly, among those young people who had completed two years of post-compulsory education, the expectation that they would continue to university or polytechnic was expressed by 50% of the post GCSE cohort compared to 46% of the pre GCSE cohort. This is virtually identical to the proportions that did in fact continue to do so (46% & 49% respectively[11]).

[11] Note: *"Owing to young people 'dropping in' to education in the PC3 year, the proportion of each cohort in full-time education at PC3 is slightly greater than the proportion of young people with some post-compulsory experience in full-time education at PC3."* (Ashford, Gray and Tranmer 1993: 9)

Vocational courses were typically taken by those with lower grades than those who studied for A-levels. Those who were re-sitting their GCSEs did not tend to go on to study for A-levels. Participation rates generally increased after the introduction of the GCSE due to increasing numbers of students with better grades at age 16. Interestingly the correlation between year 11 points score (based on GCSE/O-levels/CSEs) and UCCA (A-level) points score declined from 0.6 to 0.42 between 1986 and 1988 (after the introduction of the GCSE) (Ashford, Gray and Tranmer 1993).

Ashford and Gray (1993) investigated further part-time participation in education and training amongst 16-19 year olds using YCS cohorts 3 (1986) and 4 (1988). It is clear that three principal post-compulsory routes were identified from the YCS data - those with no post-compulsory qualifications, those studying part-time, and those studying full-time. Their work demonstrated the divergence of routes taken by 16 year olds by attainment. Those in full-time education post 16 typically had greater success in their GCSEs at 16 than those who chose to study part-time. In terms of GCSE, O-level, and CSE examination attainment at age 15/16, 86% of the highest achieving quartile, 50% of the 2$^{nd}$ quartile, 27% of the third quartile and 10% of the lowest achieving quartile chose to study full-time.

Part-time courses tended to be vocational courses, particularly as the young people aged. By three years after the end of compulsory education (aged 18-19, excluding degree students), 67% of part-time students were studying for technical/vocational qualifications compared with 44% of full-time students. By contrast, only 11% of part-time students (excluding degree students) were studying for academic

qualifications. 83% of part-time students were combining part-time study with full-time employment.

Parental employment status appears to be associated with the type of study undertaken by young people. Forty three percent of young people whose parents both worked studied full-time compared with 18% of part-time students. Twenty eight percent of young people in households where neither parent worked studied full-time compared with 18% of part-time students. Forty six percent of young people whose parents lived in owner-occupied housing studied full-time compared with 18% of part-time students. Eighteen percent of young people whose parents lived in council housing studied full-time compared with 19% of part-time students. The proportions of young people not studying post-16 is greater among those with unemployed parents and those living in council housing, perhaps reflecting the differential in attainment at age 16 for these groups. Forty three percent of girls compared with 35% of boys continued into full-time post-compulsory education. A greater majority of young people from ethnic minority backgrounds continued in full-time post-compulsory education than their white counterparts. This further emphasises the importance that qualifications at age 16 have in determining routes in early adulthood as these figures are very similar to the proportions gaining the requisite qualifications to enable them to remain in full-time education (on the A-level route). Ashford and Gray (1993) conclude that part-timers were young people with different backgrounds, qualifications and motivations to the full-time group. Part-timers were typically combining vocational qualifications with employment and this further increases their likelihood of being in full-time work by age 18-19.

In two studies of factors that boost post-16 participation in full-time education, Gray, Jesson and Tranmer (1993) compared YCS cohorts 3 (1986) and 4 (1988). In the second study, Gray, Jesson and Tranmer (1994) used data from YCS cohorts 3, 4 & 5 (those reaching the end of compulsory schooling in 1986, 1988 and 1990). These factors are grouped into the characteristics of individual young people, the operation of (local) labour markets, the effects of different forms of educational provision, and policies/efforts to change the influence of the previous factors. The characteristics of individual young people are consistent with previous findings. Females were more likely to stay on than males, particularly in one-year vocational courses. Parental occupational status and parental education remain strong predictors of staying on, independent of attainment at 16. Young people from non-white ethnic backgrounds were more likely to remain in education. The strongest predictor of staying on in post-compulsory education is educational attainment at age 16.

Gray, Jesson and Tranmer (1994) examined further the local labour market variations in post-16 participation in terms of push and pull factors referred to in chapter 2 (first described by Raffe 1992). Gray, Jesson and Tranmer (1994) applied multi-level models of participation post-16 to see whether there were different between young people in different local labour markets after controlling for individual characteristics. Overall participation rates rose across all local labour markets between 1986 and 1990. Using cluster analysis, local labour markets were characterised in terms of the dominant industries and then clustered into patterns of change or advantage according to the relative decline of certain industries (Standard Industrial Classification groups 3 & 4 – Metal Goods and Vehicle Industries & Other Manufacturing) and growth in others (SIC group 8 – Banking,

Finance and Insurance). These clusters were then ordered, with the reported employment rates among YCS respondents being higher in the advantaged clusters. For example, compared with national figures, cluster 5 had average declines in SICs 3 and 4 and substantially above average growth in SIC 8 and formed the most advantaged type of local labour market with the highest rates of youth participation.

In terms of individual characteristics, the common patterns previously identified (gender, ethnicity and social background) explained much of the variation between young people. When these factors were combined with the local labour market variation in a single model there were small staying-on differences across local labour markets controlling for qualifications, gender, ethnicity, family background & parental information. However, it became clear through the modelling process that the individual level characteristics, particularly gaining qualifications at age-16 and social background factors, were far more influential than the local labour market effects. The effects of local labour markets appeared to be inconsistent over time, having controlled for the attributes of the individual students within a local area. Gray, Jesson and Tranmer (1993; 1994) linked the location of the respondents to employment information about the local area in which they lived. Further research was deemed necessary in order to ascertain differences in staying on rates according to local educational authority, as there was a lack of ability to control for the local labour market variation (Gray, Jesson and Tranmer 1993). This serves to further highlight the difficulties in using the YCS to study particular events and non-individual level types of effect on staying on. From this, Gray, Jesson and Tranmer (1994) conclude that the effect of pull factors had weakened, due to the greater explanation provided by individual level characteristics and non-significance of the local labour market effects in the models. This lends further weight to their earlier

analyses that emphasise the role of qualifications at 16 and the social gradient in attaining these.

## *Transitions from school – early 1990s*

Transitions from school-to-work vary greatly according to the educational attainment of young people. It is clear from the evidence presented regarding the stratifying influence that parental occupation has on educational attainment that there is also consequential stratification of routes and transitions made conditional on this attainment. Payne (1995b) explored different routes beyond compulsory schooling using YCS cohorts 4, 5 & 6 (those reaching age 16 in 1987-88, 1989-90 & 1990-91). For each cohort all 3 sweeps of data were used apart from sweep 3 of YCS cohort 6 which was not available at the time of publication. The proportion of young people (aged 16-17) who were in full-time education rose from 48% in 1989 to 66% in 1992. Similarly the proportion of young people (aged 17-18) who were in full-time education rose from 41% in 1990 to 57% in 1992. This was particular the case for girls. By 1991, 71% of girls and 61% of boys were in full-time education at age 16-17. There was a corresponding decline in the proportion of young people aged 16-17 and 17-18 in full-time jobs or on Youth Training schemes. In particular the proportion of young people on vocational courses rose during these periods.

Payne (1995b) argues that the economic recession encouraged young people to remain in education. Rising attainment also enabled more young people to enter the A-level route. The proportion of young people (at age 15/16) gaining 5 or more A-C passes at GCSE rose from 30% in 1988 to 37% in 1991. This increased in overall attainment was reflected across groups of young people. The greatest rate of increase in staying-on rates were among those with low levels of GCSE attainment but was still low compared to young people with higher levels of attainment. There was a twofold contribution to the increase of 18 percentage points in staying-on rates between 1986 and 1992. Of these eighteen percentage points, improvement

in actual GCSE results across all groups of young people that enabled them to stay-on contributed 8 percentage points, and increased staying-on rates within each band of GCSE attainment contributed 10 percentage points, with the latter being a more important factor than the former (for details see Payne 1995b: 8). For those young people who were border-line in having the required GCSE attainment (5 A-C passes) to continue onto the full-time A/AS level route, there was an increased tendency to take vocational courses instead between YCS cohorts 4 (1988) and 6 (1992) (Payne 1995b).

Staying-on rates rose across all parental socioeconomic groups (using OPCS 1980 classification) but remained heavily stratified. Within given attainment bandings, stratification by parental SEG increased between 1989 and 1992 -those with the poorest results found themselves falling further behind. As Payne (1995b: 11) notes: *"For young people with five or more passes but only 1-4 at grades A-C, there was a gap of 11 percentage points between the staying-on rates of those from high and low socio-economic groups in 1989. By 1992 this gap had widened to 19 percentage points."* In the same period, for those with poorer GCSE results (defined as less than the previous quoted measure), the gap between young people from high and low socio-economic groups increased from 9 to 18 percentage points. So there is a twofold impact of being from lower parental socio-economic groups, first in terms of direct association with GCSE attainment at 15-16, and second in the rates of staying-on holding constant prior attainment (as demonstrated above).

GCSE attainment in year 11 is pivotal in influencing the route taken post-16. For YCS cohort 6 (1992), 60% of young people studying full-time at schools in year 12 (aged 16-17) had 8 or more GCSE grades at A-C. This compares to just over 50% in

sixth form colleges and just over 40% of those in FE/tertiary colleges. Typically those studying in school sixth forms will be continuing to study for A/AS levels whereas those in FE/tertiary colleges will more likely be studying for vocational qualifications (Payne 1995b).

It is possible to see in the context of rising attainment (in these data) that the routes taken by young people at the end of post-compulsory education had further differentiated between the highest attaining and lowest attaining at GCSE. Those with middle levels of attainment increasingly took vocational courses and studied at FE/tertiary colleges. This is important as it marks a key stage of divergence in the routes taken between young people in their future studies/careers. Those with the lowest GCSE attainment tended to be from lower socio-economic groups and therefore highly unlikely to remain in education post-16. It must also be noted that there is inconsistency in the measurement of parental occupation, and this hinders direct comparison between research using different schema (Croxford 2006a).

In the companion to the previous report, Payne compared the qualifications gained between the ages of 16 and 18 by young people on different routes beyond compulsory schooling (Payne 1995a). Using data from YCS cohort 5, these young people were aged 16-17 (sweep 1) in 1991 and 18-19 (sweep 3) in 1993.

There are strong patterns of differentiation in the types of course studied for by school type. By age 18-19, 51% of young people attending school sixth forms had gained A/AS level qualifications. By the same age this figure was 44% for those attending sixth form college and 12% for those attending Further Education/Tertiary colleges. Forty one percent of FE/Tertiary college students gained vocational

qualifications only. Those young people studying for A/AS levels attending school sixth forms had gained a mean 7.4 A-C passes at GCSE in year 11 (age 15-16) compared with 6.6 for those studying for A/AS level courses at FE/Tertiary colleges. This differential continues to age 18-19 for these young people studying for A/AS levels. This difference is reflected (school sixth form compared to FE/Tertiary colleges in each figure) in the number of A/AS level passes[12] (2.4 compared with 1.9), the mean grade achieved[13] (4.9 compared with 4.2), and the mean UCCA[14] score (14.9 compared with 11).

Following these descriptive statistics, Payne (1995a) ran a series of linear regression models on the number of passes and UCCA score achieved by young people aged 18-19 who were studying for A/AS levels. The strongest predictors were those consistent with other research using the YCS, GCSE attainment, gender, ethnicity, and parental occupation. Once these controls had been added to the models, the difference between types of educational institution reduced to non-significance apart from for those who had transferred educational establishment type within years 12 & 13. Educational establishment type was significant in terms of the total number of passes gained on A/AS level courses, although this is a cruder measure. This suggests that the difference in attainment at A/AS level between young people at age 18-19 in different educational establishments is similar to the factors contributing towards performance at age 15-16. This helps explain the selection process in terms of better qualified young people at GCSE continuing to the A-level route at school sixth forms and those not gaining 5 or more GCSEs at grades A*-C (but gaining some graded results) going to FE/Tertiary colleges. Correspondingly,

[12] GCE A-levels are counted as 1 and AS courses as 0.5.
[13] Using the pre-2001 UCAS tariff, A-level grade A=10 B=8 C=6 D=4 E=2, AS levels are worth half for each grade.
[14] Based upon the total UCAS grade for all subjects taken.

given that the entrance requirement for attending vocational qualifications is generally lower, the majority of pupils at FE/Tertiary colleges had lower GCSE attainment at age 15-16 than their counterparts at school sixth form colleges. This also reflects the mode of study for young people on each route, the majority of young people studying for A/AS levels study full-time. Those on vocational courses were less likely to be full-time as they tended to combine work with part-time study. The full-time educational route particularly was more likely to be taken by girls, largely due to their higher GCSE attainment than boys.

Compared with the models developed in Drew, Gray and Sime (1992), ethnicity did not appear to be significantly associated with UCCA score controlling for the previous factors mentioned (Payne 1995a). However, the combination of all ethnicity minority groups into a single category may have led to the higher attainment of some ethnic minority groups offsetting the lower attainment of other groups. The small sample size (2054) in the model led to a large standard error around the ethnic minority estimate and correspondingly large confidence intervals and lack of significance. This is the probable reason for the ethnic minority categories being combined, as further categories would lead to larger standard errors (due to small sub-sample sizes).

Payne, Cheng and Witherspoon (1996) examined the educational and labour market routes taken by 16-18 year olds using YCS cohorts 3 (1986), 4 (1988), 5 (1990), 6 (1992) & 7 (1993). The growth in post-16 full-time education is prodigious, from 48% of 16-17 year olds (year 12) in 1989 to 72% in 1994. This appears to be largely driven by the similar rise in GCSE attainment at age 15/16. 30% of young people gained 5 or more A-C grades in 1988. This had risen to 42% by 1993. Girls'

attainment was persistently higher than boys throughout this period. As is subsequently demonstrated by Thompson (2009), young people attending further education colleges were more likely to have lower attainment in their year 11 GCSEs than those attending school sixth form colleges, the majority having fewer than 5 A-C passes. The majority of young people who had gained 5 or more A-C passes at GCSE continued onto the A-level route post-16.

A series of regression models were fitted estimating the total number of passes at A-level, the mean grade achieved at A-level and the UCCA score based on data from YCS cohort 5 for those young people in full-time education. GCSE attainment at 16 remained the strongest predictor of A-level attainment consistently between these models. Holding prior attainment and other factors constant, there was consistently lower attainment at A-level for young people with parents in less advantaged occupations. Females had higher attainment than males (controlling for educational establishment, prior attainment, studying for other courses, region, ethnicity and parental occupation).

Using data from YCS cohort 5 (1990) A-level subject choice was clearly associated with gender. Girls were more likely to study English and Social Sciences & Humanities. Boys were more likely to study Mathematics and Physical Sciences. There were a notably high proportion of both sexes studying Social Sciences & Humanities. A series of multi-level logistic regression models were fitted examining the A-level subject choice of young people and the difference in effect between schools (as young people in different types of school may choose difference courses). It was clear that prior attainment in the subjects selected was the key factor but Asian young people were significantly more likely to select Mathematics

and Physical Sciences than White young people. The attainment in GCSE Mathematics is crucial as the likelihood of going on to study A-level Mathematics or Physical sciences was greatly reduced for those young people who did not gain an A grade in GCSE Mathematics. Girls were significantly less likely to study Mathematics or Physical Sciences A-levels than boys (controlling for overall GCSE attainment) which appears to bear similarity to the concentration of females in a narrower range of occupations than males in adult life.

Young people with parents in junior non-manual and skilled manual occupations (using SEG) tended to study vocational courses post-16 in preference to academic courses. This was based upon a multinomial logistic regression model comparing staying on in academic versus vocational education post-16 (using data from YCS cohort 5 1990). These young people were also more likely to drop-out of education than young people with parents in professional/intermediate occupations (controlling for GCSE attainment, gender, ethnicity, father's education, mother's education, parental employment and number of siblings).

Payne (2000) studied young people Not in Education, Employment or Training (NEET)[15] using data from YCS cohorts 3 and 8. Two measurement issues exist with analysing young people who are NEET. The first is that NEET combines young people who may be unemployed due to failure to find a job with more affluent young people who have chosen to take a gap year. Second, the timing of measurement is also important as many young people may not be employed or in education during the summer months. From the pattern of transitions, four principle types were identified: Long term period in Education, Employment or Training

[15] For Scotland, see http://scotland.gov.uk/Publications/2006/06/13100205/0, accessed 1st August 2011

(EET), long term NEET, Drop-outs from EET and Drop-ins to EET (within the period of observation). Higher Year 11 GCSE points score decreased the likelihood of spending 6 months plus in a non summer NEET spell (net of other controls). It was also more likely for girls than boys (due to higher chance of being economically inactive or 'doing something else'), for Pakistani/Bangladeshi young people than White young people, for those living in rented accommodation than owner-occupied households and for truants in year 11. In a multinomial logistic regression model of being long term NEET rather than dropping out (i.e. entering education, employment or training), similarly the odds of being NEET increasing with truancy and was inversely related to qualifications gained. When compared with YCS cohort 3 (1986), it is apparent that the likelihood of being categorised as NEET has increased and it is suggested that this is due to the growth in popularity of the gap-year (particularly at the end of the period, aged 18-19) (Payne 2000).

Connolly (2006) used cohorts 9, 10 and 11 (those eligible to leave school in 1997, 1999 and 2001) to investigate the relative influence of social class, ethnicity and gender upon GCSE attainment. Comparison is somewhat difficult as there is a shift from using Socio-Economic Group (SEG in 1997 & 1999) to National Statistics Socio-Economic Classification (NS-SEC in 2001) (Croxford 2006a). In a series of logistic regression models for each cohort with the outcome being the attainment of 5 or more A*-C passes, Connolly (2006) noted that social class and ethnicity had a greater influence on GCSE attainment than gender. The effects of gender, social class and ethnicity were all significant, as would be expected. The inclusion of interaction terms did not improve these models significantly and suggest that these effects did not differ between groups. However, no further controls were added to these models and this variation in part might be explained by variables present in

the dataset that were not used in the models (most notably school type). Gayle and Lambert (Gayle and Lambert 2009) also note further issues in the interpretation of the logistic regression models presented by Connolly (2006).

## *Further and Higher Education*

Gayle, Berridge and Davies (2003) analysed the demand for higher education using YCS cohort 9 (those aged 15/16 in 1997). Forty seven percent of young people gained 5 or more GCSEs at grades A*-C in year 11 (age 15/16). A further 26% had 1-4 GCSE passes at grades A*-C. This differs markedly from earlier cohorts where overall attainment was lower (even when measured using lower threshold categories). Mean year 11 GCSE attainment (the number of A-G passes) is strongly stratified by family social class (measured using RGSC). Those young people with parental occupations in RGSC I gain on average 6.46 passes compared with 5.73 from RGSC II, 3.95 from RGSC IIIM, 3.02 from RGSC IV and 2.13 from RGSC V. Attainment (mean number of passes) is higher for girls (4.92) than boys (4.08) and for Other Asian (5.81), Indian (4.83) and Chinese (4.62) young people than White (4.55), Other Black (3.88), African (3.86), Pakistani (3.74), and Bangladeshi (3.55) young people. Independent and selective schools continued to have much greater levels of mean attainment than comprehensives (either LEA controlled or grant maintained).

GCSE attainment is usually a pre-requisite for admission to A-level courses, the most common route to gaining admission to study for a degree (Payne, Cheng and Witherspoon 1996; Gayle, Berridge and Davies 2000; Gayle, Berridge and Davies 2002; Gayle, Berridge and Davies 2003). Ninety seven percent of those young people studying for a degree at age 18/19 were studying for A-levels at age 16/17. This was strongly stratified by parental occupation. Thirty eight percent of young people from professional/managerial backgrounds entered higher education compared with 16% of young people from skill-manual and 7% from unskilled backgrounds.

Gayle, Berridge and Davies (2003) estimated a series of models investigating entry to degree level higher education via the A-level route and separately for all pupils, using different educational measures as independent variables. By investigating interaction effects between GCSE and A-level attainment it is clear that, ceteris paribus, higher GCSE attainment increases the likelihood of entry for given A-level grades. For example, of those with three grade C A-levels, the probability of entry to degree level education was 0.56 for those with 6 GCSE passes (at grades A*-C), 0.59 for those with 7 passes, 0.62 for those with 8 passes and 0.65 with 9 passes. The same relationship held for those with three grade A or three grade E A-levels. Given the same A-level qualifications and holding other factors constant, Chinese and Indian young people are more likely than White young people to enter degree level higher education. Under similar conditions, those living in council housing or rented housing are less likely than those living in other arrangements to gain entry. Interestingly there was a significant but small effect of having graduate parents which might be expected to be larger but is partially explained by the influence of parental education upon GCSE attainment (prior to gaining entry to Higher Education).

In a logistic regression model of entry to the A-level route, gaining 5 or more A*-C grade passes at GCSE is the most significant factor (net of parental occupation, ethnicity, housing type, parental education, school type and region). Those young people from less advantaged backgrounds (in terms of parental occupation) were significantly less likely to enter the A-level route (having controlled for the previously mentioned factors, including prior attainment). The model further emphasises importance of parental education, ethnicity, school type and housing

tenure, all of which were significant influences. When comparing this model with a similar model for cohort 3 data, there is broad similarity in the factors influencing entry to the A-level route although it is noted that there are differences in the conditions affecting the models (Gayle, Berridge and Davies 2003).

The effect of parental occupation is a strong predictor of both GCSE attainment (as seen in other work) but also independently in gaining entry to A-level (and thus degree) courses. Finally Gayle, Berridge and Davies (2003) used a bivariate probit model to investigate the decision to study for a degree as a two stage model. This model controls for the role of GCSE attainment in the sample selection into the pool of A-level pupils with a second model for entry to degree level Higher Education from the A-level pupils. GCSE attainment continues to predict entry to Higher Education net of the first step of predicting gaining admission to study for A-levels.

Thompson (2009) looked at social class and participation in further education using YCS cohort 11 (young people reaching age 15/16 in 2000/2001). The composition of Further Education (FE) colleges (measured at sweep 1, young people aged 16-17) is strongly associated with parental occupation (using NS-SEC). Twenty five percent of young people with parents in semi-routine or routine occupations and 15% of young people with parents in occupations classified as large employers and higher professionals attended FE colleges. Young people from less advantaged backgrounds are less likely to remain in education but of those that remain in education the majority go to FE colleges as a consequence of their lower attainment.

Having fewer than 5 A*-C grade GCSE passes increased the likelihood of attending FE college for all young people but the greatest effect was for young people whose parents occupations were classified as large employers and higher professionals. 29.1% of young people with fewer than 5 A*-C passes and whose parents occupations were classed as large employers and higher professional attended FE college. This compares with 10.2% of young people from the same background but with 5 or more A*-C passes. By contrast, 21.1% of young people with fewer than 5 A*-C passes and whose parents occupations were classed as semi-routine or routine attended FE college. 26.2% of young people whose parents occupations were classed as semi-routine or routine but with 5 or more A*-C passes attended FE college. It is apparent that those young people with higher attainment (mostly those from more advantaged backgrounds) choose to continue in school or sixth form rather than into further education.

## *Other YCS Research*

Casey and Smith (1995) examined the relationship between truancy and attainment using data from YCS cohort 5 (those at the end of compulsory schooling in 1990, sweeps 1-3 being between 1991 and 1993). A lower percentage of reported truancy was recorded for young people attending independents and grammar schools than comprehensives or secondary moderns. Young black people reported a higher rate of truancy than young people from other ethnic groups. Of particular interest is that rates of reported truancy were higher among those young people with parental occupation categorised as low skilled (using SEG). It was also higher among young people whose parents were not working or for those who lived with only one or neither of their parents. The relationship with parental education is inconsistent but this is attributed by the authors to the large amount of missing data on this variable (nearly a quarter). It must be noted that the percentage missing for SEG was 14% of young people.

Nevertheless, this is important as truancy is associated with attitude to school and educational attainment. The mean exam score at sweep 1 was highest among those young people who reported never truanting (exam score was measured using an ordinal 20 category scheme, from 0-19 based upon volume of A-C and D-G passes. For details see Casey and Smith 1995). Following this Casey and Smith (1995) fitted a logistic regression model on whether young people remained in education post-16 (sweep 1). Truancy significantly decreased the likelihood of remaining in education controlling for region, exam score, school type, gender, parental occupation, parental living arrangements, parental education, parental employment, tenure and ethnicity.

As would be expected, the strongest predictors of staying-on remained exam score (at 16), gender, parental occupation, parental education and ethnicity. The truancy effect was relatively small by comparison. Interestingly (in this model) the effect of living in a single-parent family increased the likelihood of staying net of the above controls listed but was much lower for those living with neither parent (when compared with those young people living with both parents). In a series of models investigating the role of truancy in educational and labour market outcomes, Casey and Smith (1995) conclude that truancy has a damaging effect on the chances of remaining in education, being in full-time work or full-time training independent of the controls previously mentioned. It is particularly worrying as these models suggest that truancy is more likely among young people whose parents' occupations are junior non-manual & skilled manual or low-skilled manual compared to those from professional, employer or managerial occupations. It is also more likely among young people living with one or neither parent. These effects are stronger when the model outcome is for lengthier periods of truancy.

Witherspoon (1995) investigated the role of careers advice and guidance as influencing young people's decisions post-16. This was based on YCS cohorts 3 (age 16 in 1985-1986), 4, 5 and 6 (age 16 in 1990-1991). Of those who were advised to stay on, 77% of those with 5 or more A-C passes responded that their family was a source of help compared to 52% of those with no exam results. Those with middle (5+ grades and 1-4 A-C grades) and lower levels (5+ grades and 0 A-C grades) of GCSE attainment were more likely to state careers officers and teachers as a source of help than the highest attainment group. In models of the decision to stay-on, GCSE qualifications remain the strongest predictor of staying on. Family remains the most important source of guidance in terms of advice of whether to stay on in

education post-16. Careers guidance is perhaps most relevant for those with borderline grades (Witherspoon 1995).

## *Family Structure*

During the 20th century, the structure of families has changed notably (ONS 2005). Haskey (1998) notes a number of trends with regard to family structure in the UK (using GHS data 1986-1994). The number of births outside marriage has increased, with greater incidence of cohabiting-couple families and lone parenthood (see also Haskey 2002). There has also been a rise in prevalence of cohabitation, decline in first marriage and growth in divorce. The proportion of remarriages (amongst all marriages) has increased and there has been a consequent growth in married-couple stepfamilies. Finally, there has been a growth in cohabitation between lone parents and new partners, often forming cohabiting-couple stepfamilies (Haskey 1998).

In particular, *"A number of demographic trends have influenced families and households in recent times, most notably: the decline and delay of marriage and childbearing; the rise in divorce; the increase in cohabitation and the rise of births outside of marriage"* (ONS 2007c: 12). For all families with a head aged under 60, the percentage of married-couple families fell from 83% to 71%, the percentage of cohabiting-couple families rose from 5% to 11%, and the percentage of lone-parent families rose from 12% to 18% between 1986 and 1994 (Haskey 1998). From the early 1970s to mid-1990s, there has been a reduction in the numbers marrying by 40% (for both first time marriages and remarriages, see ONS 2007a; Murphy and Wang 1999). Similar trends are evident in other nations but are more marked in the UK (Sobotka and Toulemon 2008). It is important to consider the role of different family structures, namely cohabitation and marriage and lone parenthood upon educational outcomes. The next section is concerned with the growth of cohabitation partnerships and decline in marriage.

## Cohabitation & Marriage

The prevalence of cohabitation as an alternative to marriage has increased (Murphy 2000). Cohabiting partnerships are in general are shorter in duration than marriages and often a precursor to marriage (Kiernan 2004). This is a general statement and it is necessary to better understand the social and material context of the people cohabitating and the age at which this partnership occurs is needed to predict the stability of the cohabiting partnership. Smart and Stevens (2000) emphasise caution in interpretation of the stability of cohabitation, some types of cohabitation are less stable than some types of marriage.

There has been debate around whether cohabitation and marriage differ in terms of quality and stability as a family structure (see for example Kiernan 1992; Kiernan and Estaugh 1993; Cherlin, Kiernan and Chase-Lansdale 1995; Kiernan 1998; Kiernan et al. 2002; Kiernan 2004: compared with ; Morgan 1998; Morgan 2000). Kiernan and Estaugh (1993) note the increase in cohabitation from the 1970s onwards among young people in their twenties and thirties living together as a prelude or alternative to marriage (using the General Household Survey 1989). The duration of these unions was shorter than for marriages. The majority of male and female cohabitees were aged 25-34 with approximately two-thirds of all cohabitees being single having never married. The educational qualifications of single, cohabiting and married women aged 25-34 varied according to whether they had children (see table 4.9). Gaining qualifications was associated with delay the age of first birth of child (Steele et al. 2005). This also appears to be related to material disadvantage. Among cohabiting mothers with children, 50% of respondents' tenure was local authority housing. This compares with 5% of married couples without children (90% being owner-occupiers).

## Table 4.9 Qualifications and Family Structure

| **Table 3.19** | | | | | | |
|---|---|---|---|---|---|---|
| **Highest level of qualification: by marital status and presence of dependent children,[1] 2008[2]** | | | | | | |
| **United Kingdom** | | | | | | Percentages |
| | Degree or equivalent and higher | Higher education[3] | GCE A level or equivalent | GCSE grades A* to C or equivalent | Other qualifications | No qualifications |
| Mothers with dependent children | 19 | 11 | 17 | 29 | 12 | 12 |
| Married/cohabiting mothers | 22 | 11 | 17 | 28 | 12 | 10 |
| Lone mothers | 10 | 9 | 17 | 32 | 14 | 18 |
| Women without dependent children | 19 | 10 | 14 | 20 | 14 | 22 |
| | | | | | | |
| Fathers with dependent children | 23 | 9 | 25 | 19 | 15 | 10 |
| Married/cohabiting fathers | 23 | 9 | 25 | 19 | 15 | 10 |
| Lone fathers | 13 | 7 | 25 | 27 | 14 | 14 |
| Men without dependent children | 21 | 8 | 28 | 14 | 14 | 15 |
| | | | | | | |
| All parents with dependent children | 21 | 10 | 20 | 24 | 14 | 11 |
| Married/cohabiting parents | 23 | 10 | 21 | 23 | 14 | 10 |
| Lone parents | 11 | 9 | 18 | 32 | 14 | 17 |
| All people without dependent children | 20 | 9 | 21 | 17 | 14 | 19 |
| 1 Children aged under 16 and those aged 16 to 18 who have never married and are in full-time education. | | | | | | |
| 2 Data are at Q2 (April–June) and are not seasonally adjusted. People of working age (men aged 16 to 64 and women aged 16 to 59) or those in employment with qualifications. See Appendix, Part 4: Labour Force Survey. | | | | | | |
| 3 Below degree level but including National Vocational Qualification (NVQ) level 4. See Appendix, Part 3: Qualifications. | | | | | | |
| *Source: Labour Force Survey, Office for National Statistics* | | | | | | |

Source: Social Trends 39 (ONS 2009b)

Whether a respondent was employed or not appeared to be closely related to whether they had children. Approximately 90% of cohabiting and married childless women aged 25-34 were working. Among those with children 53% of married women worked compared to 43% of cohabiting women.

Typically, the married mothers worked part-time while a greater proportion of cohabiting women were working full-time. This potentially suggests greater financial pressure to work amongst cohabiting women. Among those with children, there was a much greater incidence of unemployment among the male partners in cohabiting couples than for the male partners of married couples (20% and 4% respectively). This has potential implications as the majority of women with an unemployed partner were themselves unemployed. Three main groups of cohabitees were identified: "...*youthful never-married childless couples, post-marital cohabitants and never-married couples with children. The circumstances, requirements and needs of these three groups are likely to differ, as are policy responses"* (Kiernan and Estaugh 1993: 69).

Kiernan and Estaugh (1993) suggest that there is little to distinguish between married couples and longer-term cohabiting couples. By contrast, Morgan (2000) argues that cohabitation and marriage differ substantially in terms of quality and stability as a family structure, cohabitation not being simply a substitute for marriage but often a precursor to it. Cohabitation unions tend to be shorter, the median duration being 19 months (BHPS data 1991-95, see Ermisch and Francesconi 2000; and Ermisch 1995). Using evidence from the BHPS and 1958 NCDS, Morgan (2000) observes that the majority of those who cohabited go on to

marry or split up with 16% living with their partner for more than 5 years and 4% for more than 10 years.

The minority of cohabiting couples who do not marry have different socio-economic circumstances to those who do. There is a greater likelihood of people who experienced the dissolution of their parents' marriage to have children outside of marriage (Kiernan 1992). Cherlin, Kiernan and Chase-Lansdale (1995) commented in their study (using the 1958 BCS) that those who experienced parental divorce during childhood that the great majority of young people did not leave home due to family friction nor did the majority have children outside of marriage. However, the likelihood of both outcomes was greater for this group.

Kiernan and Mueller (1999) modelled the likelihood of partnership-breakdown using BHPS data 1992-1995 (the term 'partnership-breakdown' included both divorces and episodes of cohabitation that had ended). They identify a higher odds of partnership breakdown among the socio-economically disadvantaged, the young and the poorly qualified (as does Lampard et al. 1994). *"Unlike marriages, the stability of these non-marital partnerships does not seem to be linearly related to the age at which they were formed suggesting the cohabitating unions may be intrinsically more fragile than marriage regardless of the age at the beginning of the partnership"* (Kiernan and Mueller 1999: 388). In the second part of their study Kiernan and Mueller (1999) used NCDS data to suggest that those with qualifications below A-level (controlling for prior attainment) are more likely to dissolve their partnership. When a control for the age of first partnership is included within the model, this effect diminishes which suggests that such partnerships tend to be at a younger age.

## Lone Parenthood

Between 1971 and 1991, there was a large increase in the number of lone parents from 500,000 to 1,300,000 largely as a result of divorce. In the period 1966-1970 there were 237,503 divorces in England and Wales. For 1991-1995, this figure was 797,822 and had risen in every successive 5 year period in between the two periods quoted (ONS 2007c). Nine out of ten lone parents are lone mothers; the proportion of lone parents who are lone fathers remained consistently low (McKay and Marsh 1994; Haskey 2002; ONS 2007a). Lone mothers, as a proportion of the UK population, rose from 7% in 1971 to 22% in 1998/1999 (ONS 2000a).

Feinstein, Duckworth and Sabates (2008) argue that the form of family structure is not independently important for educational outcomes but that it is stability that matters most. Feinstein, Duckworth and Sabates (2008), Ginther and Pollak (2004) and Aughinbaugh, Pierret and Rothstein (2005) argue that family structure is a non-significant influence on child attainment once parental income, education and other socio-economic variables are controlled for (Ginther and Pollak (2004) and Aughinbaugh, Pierret and Rothstein (2005) use US data from the NSLY79).

Ermisch and Francesconi (2001b) modelled whether family structure represents a separate influence from parental occupation, using data from waves 1-5 of the BHPS (1991-1995). These data consisted of two samples, 764 young people and 411 siblings). They conclude that life in a single parent family during childhood is associated with negative outcomes for children as young adults, notably lower educational attainment, higher risk of economic inactivity, and higher risk of early birth. Young people who experienced disruption to their family structure between the ages of 0 and 5 appears to have lower subsequent educational attainment

(Ermisch and Francesconi 2001b; Ermisch and Francesconi 2001a). By contrast, Fronstin, Greenberg and Robins (2001) observed that disruption to family structure between the ages 11 and 15 significantly reduced educational attainment for young people, in their comparable study using NCDS data. Further investigation is required to evaluate these findings.

Part of the lower educational attainment of the children of lone parents may be due to the different backgrounds, in terms of the education and occupations (or unemployment) of lone parents. Berthoud, McKay and Rowlingson's (1999) analysis of the ESRC's Social Change and Economic Life Initiative (SCELI) data suggests that women in semi-unskilled occupations were more likely than those in professional occupations to have children while unmarried (measured using SEG, the same relationship holds for partner's occupation). This is particularly more likely for mothers who were living in publicly rented accommodation or were of West Indian origin. Those in full-time education were much less likely to have a child prior to marriage. Berthoud, McKay and Rowlingson (1999: 372) argue that, *"those from poor socio-economic backgrounds may see motherhood as providing better status and identity along with more meaningful work than that offered by low-paid jobs."*

This finding is similar to that of Payne and Range (1998) who investigated lone parenthood using the 1958 National Child Development Study (NCDS). They noted that women who gave birth when aged 16-23 were less likely to be in a partnership. This differed from the experiences of older women (who were lone parents) who were likely to have been in partnership which had subsequently terminated. This is important because it suggests that these two groups may be from very different backgrounds. Ermisch and Pevalin (2005) suggest that a teen-birth for a young

woman greatly increases the likelihood of partnering with poorly-educated and unemployment prone men (using data on pregnancy and partnership histories from the 1970 British Cohort Study). This demonstrates that family structure and parental economic circumstances can influence the educational attainment of young people and the subsequent influence this has on their own future occupation and family structure.

Attewell and Lavin (2007) compared the benefits of gaining a degree (for women) between the City University of New York (CUNY) sample of 1,971 former students (re-interviewed in 2000) and the National Longitudinal Survey of Youth (NLSY79) sample of approximately 6,000 female former students (selected as nationally representative). Under the CUNY open access scheme in 1970, students in New York who had low grade attainment (largely due to having attended schools with low levels of attainment) were guaranteed a place at University leading to increased uptake, particularly among ethnic minorities and young people whose parents had lower educational attainment. Using a range of outcomes and tests of bivariate association, regression and propensity matched scoring they identified the benefits that access to higher education brought for these women. Whilst the benefits of gaining a degree were universal, for the CUNY sample these benefits were proportionally greater in terms of household income and their children's educational attainment (among other measures). With regard to family structure, Attewell and Lavin (2007) also reported an increased likelihood of higher educational attainment (in terms of going to college) for children living in stable two-parent families compared with single-parent families. Interestingly, the women who went to college were more likely to have family relationships that lasted than those who did not go to college. This emphasises the effect of family disruption (in terms of parental

relationships) as potentially impacting upon their children's educational outcomes (Morgan 2000; Attewell and Lavin 2007). This was closely related to parental occupation and parental educational attainment with higher educational attainment for the children of parents with higher earnings & occupational complexity and greater educational attainment. All of these findings illustrate well the inter-linkage of social background and parental attributes have upon childhood outcomes in the context of higher education and gaining the requisite level of attainment in the USA.

This suggests that children from poorer backgrounds are educationally disadvantaged and that family structure may contribute to this (Ermisch, Francesconi and Pevalin 2001; Ermisch and Francesconi 2001b). Feinstein, Duckworth and Sabates (2008) conclusion that family structure is not an important explanatory variable in young people's attainment appears overly simplistic. Ermisch and Francesconi's (2001b) assertion that 'adverse family structure' (in particular single parenthood) remains important. Having considered the evidence for the influence of lone parenthood on their children's educational attainment it is apparent that a more detailed investigation is required using a large scale dataset such as the YCS. Parental education and parental occupation remain of critical importance in understanding these processes and it is to these that the next section is concerned, using analysis of two datasets, the DSS/PSI Programme of Research into Low Income Families (PRILIF) and Families and Children Study (FACS).

## DSS/PSI Programme of Research into Low Income Families

A series of studies in the UK have been conducted in the 1980s & 1990s using the DSS/PSI Programme of Research into Low Income Families (PRILIF). This dataset is a longitudinal study which investigates further the relationship between lone parenthood and poverty. The unit of data collection in this study are individual lone parents (rather than families or households). The following section is a review of reports produced considering these data.

Bradshaw and Millar (1991) sampled 1,472 lone parents of which 951 were drawn from one parent benefit records and 521 from income support records which was subsequently compared with the PRILIF Survey 1991 by McKay and Marsh (1994). The majority of lone parents in the Bradshaw and Millar (1991) sample were divorcees (64%) while the remainder were single. The mean age of single lone parents was 26 compared with 33 for separated and 36 for divorced lone parents. Bradshaw and Millar (1991) identified that marital status (at a particular date) does not indicate whether the respondent had experienced pre-marital cohabitation or identify where multiple spells of lone-parenthood had occurred. Measures of these family structures would be valuable as a means of identifying family history of respondents and evaluating the educational outcomes of young people growing up in these circumstances. In addition, Bradshaw and Millar's (1991) study demonstrated the employment-childcare balance faced by lone parents in terms of the relative costs and benefits of employment. The factors affecting decisions around employment also affect married mothers regarding work and child care arrangements. There was a balance to be struck between whether work is worthwhile (i.e. the reservation wage, the level at which pay is sufficient to induce entry to the labour market) particularly in light of benefit thresholds. As a

consequence of this financial pressure there is a dependency on low cost social housing. This emphasises the important linkage between parental occupation and family structure, as lone parents attempt to negotiate the labour market with different circumstances to dual-parent (and dual earner) households. Fifty five percent of the lone parent sample had no recognised qualifications with a further 36% having qualifications below A-level. Seventy two percent of the lone parent sample were on income support.[16] There was a high incidence of unemployment among lone parents. Re-partnering was the most likely route out of reliance on income support. Clearly the children of lone parents were likely to be from economically disadvantaged families and this has implications for their educational attainment.

The initial PRILIF survey contained interviews with two thousand two hundred low income families including 900 lone parent families (McKay and Marsh 1994). Between 1971 and 1991, the number of lone parents increased from 500,000 to 1,300,000, largely due to divorce. The proportion of single (never married) lone parents rose from 37% in 1989 to 46% in 1991 (for further details see table 4.10). There was a sharp divide between the majority of lone parents who relied on Income Support and a minority who did not. Maintenance payments are financial support from a non-custodial parent to the lone parent and were mostly received by divorced lone parents and had a strong independent influence on the likelihood of working (Bradshaw and Millar 1991). At every level of educational qualification, those receiving maintenance payments were more likely to be in work with the most qualified parents and those receiving maintenance payments also being the most

[16] Income Support is an income-related means-tested benefit in the United Kingdom for people who are on a low income, for conditions of eligibility see: http://www.direct.gov.uk/en/MoneyTaxAndBenefits/BenefitsTaxCreditsAndOtherSupport/On_a_low_income/DG_10018708 accessed 21st April 2011

likely to work. The stability of income may also be affected by whether a lone parent has experienced multiple episodes of lone parenthood interspersed with periods of having partners (Bradshaw and Millar 1991).

Ford, Marsh and McKay (1995) compared Bradshaw and Millar's (1991) 1200 lone-parent families with the PRILIF 1991 data (900 lone-parent families) and PRILIF 1993 data, 840 re-interviewed from PRILIF 1991 plus an additional 900 lone-parent families. The proportion of lone parents from different previous marital status and relationship backgrounds was relatively consistent between 1991 and 1993 (see table 4.10). Three-quarters of lone parents left school at 16 or younger with the majority of lone parents had either no qualifications or qualifications below O-level although this figure fell between 1991 and 1993. There may be "age cohort effects" regarding older cohorts when participation in post-16 education was less common (Sanderson 2007).

**Table 4.10 Comparison of Programme of Research into Low Income Families sample with Bradshaw and Millar's (1989) sample**

| | DSS/Bradshaw and Millar 1989 | DSS/PSI 1991 | DSS/PSI 1993 |
|---|---|---|---|
| Marital Status | | | |
| Divorced | 45 | 36 | 35 |
| Separated from marriage | 18 | 18 | 22 |
| All never married | 37 | 46 | 42 |
| Of whom: | | | |
| Separated from cohabitation | | (48) | (56) |
| Never lived as couple | | (52) | (44) |
| Sex | | | |
| Female | 95 | 95 | 94 |
| Male | 5 | 5 | 6 |
| Highest Qualification | | | |
| Below O-level | 12 | 21 | 16 |
| GCE O, City & Guilds | 23 | 22 | 25 |
| GCE A or similar | 5 | 6 | 9 |
| Above A-level | 9 | 10 | 12 |
| None | 50 | 41 | 38 |
| Base (n) | 1342 | 938 | 849 |

Source: Ford, Marsh and McKay (1995: 16)

Ford, Marsh and McKay (1995) identified two principle groups of lone parent. The first group tended to be younger, poorer, social-renting, poorly-qualified mothers, which I shall refer to as the young mothers group. The second group were characterised as older divorcees who were generally better qualified and more affluent (the divorcees group). These subgroups appears to be similar to the subgroups of cohabitees identified by Kiernan and Estaugh (1993). Lone parents, particularly the young mothers group, tend to face financial difficulties with higher levels of debt compared with the divorcees group (Ford, Marsh and McKay 1995). This suggests that the effects of disadvantage (often driven by poorly paid occupations or unemployment) in lone parent households may influence their children's attainment at school (perhaps by young people seeking work due to financial difficulties).

Finlayson and Marsh (1998) compared the 1991, 1993 and 1994 data from the first three waves of the PRILIF study. The 1994 sample consists of 1302 lone parents obtained from a postal survey (a sub-sample from DSS records). The lower than average educational qualifications of young mothers relative to other women is associated with low hourly wage rates. Low hourly wage rates form a substantial barrier to work in relation to the loss of benefits as a consequence of gaining employment (Finlayson and Marsh 1998). As Finlayson and Marsh (1998: 71) explain, *"it is possible to suggest that many lone parents do not really have a problem getting jobs, they have a problem supplying enough hours at a sufficiently high rate of pay per hour to fit the income package offered by in-work benefits and, if they can get them, maintenance payments too."*

**Table 4.11 Percentage unemployed/inactive and percentage with children aged under 5, 1991-94**

| | % Unemployed /inactive 1991 | % Unemployed /inactive 1993 | % Unemployed /inactive 1994 | % with aged under 5 1991 | % with aged under 5 1993 | % with aged under 5 1994 |
|---|---|---|---|---|---|---|
| Single Mother | 76 | 77 | 75 | 70 | 72 | 71 |
| Separated from Cohabitation | 71 | 65 | 67 | 58 | 61 | 53 |
| Divorced Mother | 52 | 51 | 51 | 28 | 24 | 15 |
| Separated from Marriage | 60 | 50 | 56 | 48 | 32 | 35 |
| Base (n) | 938 | 849 | 773 | 938 | 849 | 773 |

Source: Finlayson and Marsh (1998: 8)

As can be seen in table 4.11, single mothers and those separated from cohabitation were particularly likely to be unemployed. This suggests that there are quite distinct types of lone parenthood in respect to the increased age, educational qualifications and material affluence of the mothers who were divorced or separated from marriage (Ford, Marsh and McKay 1995). It is difficult to see whether the effects of the recession of the early 1990s increased the proportions of lone parents in these samples who were unemployed but there was a notable increase in workless lone fathers (despite representing a similar proportion of total in 1994 to 1991 and

1993). The barriers to work suggested have repercussions for material hardship faced by lone parents and their offspring and suggests why wage-supplementation policies may not have had the impact hoped in reducing worklessness among lone parents (Leeming, Unell and Walker 1994; 1999). Low educational attainment among lone parents also tends to be associated with being in poorly paid employment (Ford, Marsh and McKay 1995).

The socio-demographic characteristics of lone parent families appear to have changed little in the early 1990s although the likelihood of being in work increased (see table 4.11). The proportion of lone parents being out of work fell from 63% in 1991 to 51% in 1995, and of those who worked, the proportion working more than 24 hours a week rose from 25% in 1991 to 31% in 1995 (Ford, Marsh and Finlayson 1998). The proportion working fewer hours was relatively stable. Ford, Marsh and Finlayson (1998) argue that the routes for leaving lone parenthood were relatively few:

*"Lone parents in Britain typically have fewer children than couples; they are more likely to occupy social, rented accommodation (council or housing association). They are much less likely to be found in paid work. Their low levels of full-time paid work are explained in part by having to care for young children, their generally low skills base and lack of recent work experience, the fact that most are female and the lower earning power their gender implies."*

(Ford, Marsh and Finlayson 1998: 11)

The most likely reasons for leaving an episode of lone parenthood by 1995 were repartnering or children leaving home (i.e. becoming non-dependent). The younger

the lone parent, the higher the likelihood of repartnering (although this is censored by the duration of the data collection). Those with youngest children aged 16, 17 or 18 (and in full-time education) in 1991 were significantly more likely to no longer be lone parents in 1995, due to their children being about to leave dependency (Ford, Marsh and Finlayson 1998). Models of whether a lone parent has had any spell of leaving lone parenthood between 1991 and 1995 indicate that earlier receipt of maintenance and Income Support was associated with continuing lone parenthood (controlling for parental age, age of youngest child and health). Leaving lone parenthood is not necessarily the end of financial difficulties especially if the end of an episode of lone parenthood is cohabitation. Ford, Marsh and Finlayson (1998) argue that those lone parents who continue as lone parents were those with a preference for independent living and those who had not met a suitable partner.

Finlayson et al. (2000) continued to investigate the PRILIF lone parent cohort using data from 1991 to 1998 (including extra data from 1996 and 1998 as part of efforts to improve response rates).

**Table 4.12 Percentage in work 1991, 1996, 1998**

| | 1991 | 1996 | 1998 |
|---|---|---|---|
| In work 24+ hours/week | 26% | 33% | 38% |
| In work 24+ hours/week | 2% | 12% | 13% |
| In work 24+ hours/week | 9% | 10% | 8% |
| Out of work | 62% | 45% | 41% |
| | (100%) | (100%) | (100%) |
| Base (n) | 938 | 770 | 730 |

Source: Finlayson et al. (2000: 14)

In addition to the rising incidence of working among lone parents (see table) 276 children of parents in the sample had left home by 1996. Of all non-dependent children 25% were working 16 hours or more a week whilst 17% were in full-time

education and 5% were unemployed. Of the children of lone parents who had left home (335 by 1998), 54% were working more 16 hours or more a week, 12% were in full-time education and 15% were unemployed. The notable feature of those children who had left home was the high percentage employed. The majority of those who had left home also had a qualification (this is not defined further). It is unclear what age these young people are as it includes all those eligible to leave education. It is difficult to make further inference from these data as these percentages refer to the children themselves and not their original lone parent household. The sample is small but it is clear that leaving home is associated with being in employment – it is unclear which came first, whether the pressure to leave home led to young people seeking employment, or young people sought employment in order to leave home. What is clear is that the full-time education route is less likely for the children of lone parents in this sample.

The lower standard of living among lone parent families was investigated by Finlayson et al. (2000) using a 7 point scale measure of hardship (lowest=0 and highest=6). This measure is composite scale of debts and items which people cannot afford (see Marsh and McKay 1993). Relative material hardship appears to have lessened through the 1990s based upon the mean hardship score (Finlayson et al. 2000). In a logistic regression model lone parenthood was a key predictor of whether the respondent was still in hardship (defined as a score of 3 or more on the hardship scale) in 1998 having been so in 1991. The likelihood of not being in hardship in 1998 was significantly reduced for those respondents who were in hardship in 1991 and who remained lone parents (controlling for working hours and health status). The group most likely to leave hardship were living without a partner but no longer had a child of dependent age (Finlayson et al. 2000: 111). Two

features of the lone parents are of note. First, *"those in hardship in 1991 were already very like one another (single, young, social tenants, out of work, unwell, unqualified and so on) and very unlike those free of hardship in 1991 (older, working, owner occupiers and so on). Thus differences between those who left hardship and those who did not were open to a large range of chance variations"* (Finlayson et al. 2000: 111). Second, exiting hardship was associated with moves out of lone parenthood. The effect might also be different if different time intervals had been used in re-interviewing as some lone parents may have had multiple episodes of lone parenthood and cohabitation within the duration. The features of these two groups appear to correspond to the young mothers group and the divorcees group identified.

Marsh and Vegeris (2004) continued the work investigating the PRILIF lone parent cohort using data from 1991 to 2001 (the 7$^{th}$ wave, 531 respondents). The primary focus of the 2001 interviews was upon the children of the 1991 respondents. A logistic regression model of a lone parent being in work in 2001 reports a higher probability of working among those parents who were working in 1991, who had qualifications (basic or higher), were aged under 40 in 1991, had formed new partnerships by 2001 or had improved qualifications. A lower probability of working in 2001 was associated with those parents who were social tenants, had had further births since 1991 or had experienced moderate or severe hardship between 1991 and 1998 (Marsh and Vegeris 2004: 63).

The PRILIF sample had lower levels of qualifications than the UK average (for those aged 16 or above). Twenty six percent of the sample had A-levels or above compared with 38% nationally (based on 2001 census data for 16-29 year olds).

This suggests that parental educational measures and family structure measures may partially explain part of a wider process. Marsh and Vegeris (2004: 137) assert that: *"Most consistently, children from a background where the single parent entered into a short-term relationship and, to a less extent, children whose parent remained single during the study period were found to be at a disadvantage. These findings reflect select measures in five broad areas: physical well-being, social adjustment, education and work and household formation."*

## Families and Children Study

The Families and Children Study (FACS) 2001 is a panel survey of Child Benefit Recipients who were lone parents or low/moderate income couples (based on receipt of Child Benefit, Family Credit and having net household incomes up to 35% above the entitlement threshold). Vegeris and Perry (2003) analysed panel data collected in 1999, 2000 and 2001 (3022 families responded to all 3 waves) to investigate the living standards of families and children in the UK. Their findings provide further information confirming and enhancing the work done with the PRILIF data detailed prior to this section.

Vegeris and Perry (2003) compared lone parents with couple families' aspirations for their children. In modelling parental aspirations for their children continuing in education post-16, several factors were of note. Non-white parents or the parents of girls were much more likely to want their children to continue in education. However, when a lone parent had no qualifications or lived in social housing (even when controlling for hardship), they were less likely to want their children to continue in education. Those working more than 30 hours per week or with larger families were not significantly more likely to want their children to continue in

education than those working fewer hours or with few children. This pattern is similar for couple-family's aspirations for their children to continue, particularly with regard to non-white parents and the children of girls. However, the effects of living in social housing were reduced. Those parents with no qualifications had significantly lower aspirations for their children continuing in education when compared with parents with qualifications (Vegeris and Perry 2003).

Parental aspirations for their children attending university do appear to be more related to financial hardship. The previously mentioned factors remain significant with regard to these aspirations. However, financial hardship is more relevant in reducing this aspiration for couple families than for lone parents. Interestingly lone parents working more than 30 hours per week are significantly more likely than those working less than 30 hours per week to want their children to attend university perhaps suggesting the desire for their children to gain better qualifications or employment. This may be related to the benefits entitlement of lone parents and subsequent influence on potential funding support at University for their children. These findings suggest that while financial hardship and resources remain an issue with regard to parental aspirations for their children continuing in education post-16 or to university, there are differences according to ethnicity of the parents, gender of the child and the qualifications that the parents themselves gained.

Vegeris and Perry (2003) suggest there was a reduction in the level of hardship experienced by low/moderate income families between 1999 and 2001. Correspondingly there was an increase in economic activity and decrease in unemployment for lone mothers between 1992 and 2002. There remained a strong

positive association between economic activity (including hours worked) and age of the youngest child of a lone mother (ONS 2002a). Changes in family circumstances remain crucial to understanding the processes of hardship, in particular, "*family break-up, loss of employment, large families, ethnic minority families, and poor family health were all found to limit how far and how fast hardship can be expected to fall*" (Vegeris and Perry 2003: 144).

Barnes et al. (2004) analysed the FACS 2002 data (7352 families). Twenty five percent of the families interviewed were lone parent families. Lone parents were more likely to live in social housing or rented accommodation than couple families and more likely to have a total family income in the lowest quintile. Parents in lone-parent families were generally less well qualified than those in couple families. Forty five percent of lone parents had either no academic qualifications or a highest GCSE grade of D-G compared with 28% of mothers in couple families. There was a strong association between the highest qualification achieved by a parent and the number of hours they worked. Furthermore, the qualifications obtained by those working the least hours (or not at all) tended to be at a lower level (both academic and vocational) for both lone-parents and couple families, more so for lone-parents. This pattern extended to perceptions of how their children were doing at school with those parents who worked less than 16 hours per week, or not at all, being less likely to perceive their children as doing well at school. Lone parents influence on their children's attainment is multivariate in terms of their increased likelihood of being unemployed, increased likelihood of being in social or rented housing and lower level of qualifications attained.

## Summary

The evidence within this literature review, from research using the Youth Cohort Study, suggests that there are strong associations between educational attainment and gender, ethnicity, parental occupation, parental education and family structure. It was identified in the previous chapter that if detraditionalisation theorists are correct then it would be expected that the influence of parental occupation would be stronger at the beginning of the 1990s and weaker at the end of the decade and that there would be a lesser effect of gender and ethnicity upon educational attainment.

A number of themes emerged within this review. GCSE attainment has risen throughout the 1990s and this appears to be stratified by gender, ethnicity, parental education and parental occupation. Attainment in GCSE qualifications at age 16 is important because it increases the likelihood of participation in post-compulsory education and the opportunity to access the A-level route (and potentially higher education). Failure to achieve 5 or more GCSE passes at grades A*-C is associated with alternative routes post-16, typically vocational education, employment or unemployment. This is particularly important as the number and proportion of young people gaining employment in the immediate years after completing compulsory education has declined and a greater proportion of young people remain in education at age 18 in later YCS cohorts.

These findings suggest that the assertions of detraditionalisation theorists are not backed by empirical evidence from the 1990s. Instead, there is persisting stability in the influence of parental occupation on GCSE attainment and on participation in post-compulsory education. Alternative measures of stratification are also considered, particular being in receipt of Free School Meals (FSM) as an indicator of

family circumstances, but it is argued that this is inferior to parental occupational classifications (Rose and Pevalin 2003). It is noted that there has been a change from boys outperforming girls in the 1960s to girls outperforming boys in the 1990s and this gap appears to be widening. Young people from ethnic minorities are more likely than white young people to remain in education post-16, net of educational attainment, but the average pattern is less clear.

The majority of literature available for the YCS concentrates upon the early 1990s. The late 1990s remains relatively less well researched using YCS data. Notable exceptions include Lupton, Heath and Salter (2009) and Gayle, Lambert and Murray (2009b) but few studies include data from YCS6 (1992). The majority of studies include bivariate associations but fewer contain models controlling for multiple sources of variation simultaneously. For example, in a number of studies there was a strong influence of parental occupation and parental education upon attainment and staying-on. In both instances the majority of research observed these as separate bivariate associations, by which I mean the relationship between one independent variable and one dependent variable, which suggests the need to develop multivariate models to control for other known sources of variation (i.e. between multiple independent variables and a dependent variable).

Few studies have analysed attainment in particular GCSE subjects. This is important because many A-level subjects have GCSE pre-requisites and this influences the potential routes taken by a young person at the end of compulsory education. Few papers consider the impact of missing data which may potentially be a source of bias in the data.

In the second part of this chapter, studies of lone parent families were introduced to examine the influence of family structure in comparison to parental occupation and parental education. Methodologically the difficulty in modelling using family structure is that it is heavily associated with many of the other parental attributes. This suggests a further benefit to fitting multivariate models. From the research using Programme of Research into Low Income Families (PRILIF) data, two principle groups among lone parents were identified. The first group tended to be younger, poorer, social-renting, poorly-qualified mothers. The second group were characterised as middle-aged divorcees who were generally better qualified and more affluent. From these groupings, two factors appear particularly important with regard to young people's educational attainment, the material disadvantage of their parents (captured in part by occupation), and whether their parents had a stable relationship status. It would be worthwhile to include these factors in models of educational attainment (consistent with the third research question of this thesis).

Having identified the Youth Cohort Study as a valuable dataset when investigating GCSE attainment and family background in the 1990s, the next chapter considers this dataset in greater detail and the methods to be employed.

# 5 Methodology

In this chapter, the research questions to be addressed and the methods employed are presented. Returning to the questions posed in chapter one, there is one principal question and five specific research questions:

**What is the role of family background in the stratification of GCSE attainment in the 1990s in terms of parental occupations, parental education and family structure?**

1. Has the association between parental occupation and their children's educational attainment increased, decreased or remained stable through the 1990s?

2. How do different measures of parental occupation and parental education report this association? Are the different measures consistent? (a sensitivity analysis of measures of parental occupation)

3. Have the differences in attainment between pupils from lone-parent and couple families increased or reduced through the 1990s?

4. What influence does missing data potentially have upon the estimates in models fitted? (particularly with regard to parental occupation)

5. How do the subjects studied and grades attained vary between young people from different family backgrounds?

The underlying epistemology of this research is to answer middle range questions, such as those listed, and better identify the empirical regularities to which the questions are directed. Principally I wish to draw upon the empiricism of Robert Merton namely developing: *"...theories intermediate to the minor working hypotheses evolved in abundance during the day to day routines of research"* Merton (1957: 5). I intend to develop middle range theories and hypotheses from the data collected which is preparatory work towards the eventual advancement of theory. This seems appropriately cautious in light of Merton's comments that in sociology there are many concepts and few confirmed theories. Regarding the strategies and motivations taken by actors within strata, I concur with Merton on the point that, *"the descriptive account should, so far as possible, include an account of these motivations, but these motives must not be confused, as we have seen with: a) The objective pattern of behaviour; or b) with the social functions of that pattern"* (Merton 1957: 67). By this, it is argued that whilst family strategies for conferring advantage in educational settings are important, the careful measurement of educational attainment according to family background is the primary aim of this research.

Further to this, some researchers may argue that the 1990s is relatively distant from the current period. The importance of the 1990s as a period of study has been emphasised in chapter two, particularly in the context of the "changing times consensus" amongst sociologists of youth (Gayle, Lambert and Murray 2009a) and because of the policy reforms of the late 1980s and early 1990s. There are clear contextual, empirical and theoretical grounds for investigating periods of social change with respect to educational attainment. Goldthorpe (1991: 216) argues that: *"Sociologists have a legitimate, and necessary, concern with such historically*

*defined social change because, as I have earlier suggested, they wish to know how widely over time and space their theories and hypotheses might apply."* This is consistent with the objective of identifying middle range empirical regularities using recent (i.e. 1990s) data.

## *Analysis Methods*

There are a number of reasons for using quantitative survey data analysis in this thesis and for why these methods were selected in preference to others. The questions posed within this thesis are motivated by the desire to describe and understand the relationship between educational attainment and family background factors throughout the 1990s. Given the recent availability of large scale secondary datasets in the field of educational sociology, there is great value in using these resources. Secondary data analysis overcomes the barriers of cost and difficulty of data collection associated with primary analysis. The large volume of data, and range of available measures, enables complex multivariate analytical techniques to be employed, whilst secondary resources also have the attractive methodological feature of enabling the verification and cross-examination of the work of other researchers in the field with directly articulated and comparable methods (e.g. Kiecolt and Nathan 1985). Despite these attractions, it is argued that large scale secondary survey resources relevant to educational sociology are often under-analysed (e.g. UK Data Forum 2007; Gayle, Playford and Lambert 2008). This thesis was funded by an ESRC Quantitative Methods stipend to do survey data analysis.

Qualitative methods may enable the investigation of the family setting in which a young person grows up but it would be not be practicable to employ these methods

at a nationally representative level. The following passage presented in Gayle and Lambert (2006: 1) represents a similar approach to that taken in this thesis:

*"Evaluations of variable analyses in sociology date back at least fifty years (see Blumer 1956). Over the decades a virtual industry producing critiques of variable analyses from various standpoints has developed. We suggest that arguments for and against variable analysis, and in particular the analysis of data from social surveys, have at times resembled a caricature not dissimilar to the Shakespearean feud between the Montagues and the Capulets. In this paper we do not wish to either visit or reopen these debates. However, we would like to note a comment made by Goldthorpe that critics of survey based sociological research ritually characterise it as static and this is simply to ignore the rapid development of survey related work (Goldthorpe 2000: 17)."*

The justification for the use of the Youth Cohort Study as the main data resource is set out later in this chapter (including further advantages and disadvantages of secondary data analysis). The analytical techniques employed in this thesis have been selected because of their appropriateness for answering the research questions. Descriptive techniques have been included in chapter six to describe the properties of the data, extending upon the analyses reviewed in chapter four by the inclusion of all 1990s cohorts of the YCS. While some researchers have analysed 1990s YCS data excluding YCS6 from 1992 (Lupton, Heath and Salter 2009; Gayle, Lambert and Murray 2009b), this work provides complete coverage of YCS data throughout the 1990s, enabling a more comprehensive investigation of educational attainment during this period.

Following the descriptive analysis of the relationships in chapter six, chapter seven is an investigation of the role of the family background measures identified in chapter four. General linear models (in particular linear and logistic regression models) are selected as these permit the multivariate analysis of large scale data resources such as the Youth Cohort Study. These multivariate statistical models permit the investigation of the first three research questions identified in this chapter and allows the secondary data analyst to control for multiple sources of variation in GCSE attainment. This is important in assessing the relative influence of multiple sources of variation and the claims made by detraditionalisation theorists about the decline of traditional structural divisions. If there are no clear patterns in GCSE attainment by these factors or a reduction in their influence, then these theorists would have empirical evidence to support their claims.

Chapter seven contains linear and logistic regression models to examine the effects of parental occupation and parental education on GCSE attainment at age 16 (research questions 1 and 2). Simple cross tabulations and measures of association do not capture the multivariate relationship. Dale and Davies (1994: 5) note that: *"This is precisely the issue that statistical modelling addresses by permitting many interrelationships to be considered simultaneously within a single analysis together with an error structure to represent unmeasured effects."* The purpose of this is to move from descriptive analyses of the YCS data to explanatory and inferential statistics about the magnitude of relationships between variables and whether these can be inferred to the wider population of school leavers in England and Wales (Blaikie 2003).

The benefit of linear regression models above simple correlations (bivariate associations) is that multiple correlation between variables can be included in the models by representing the joint relative influence of several independent variables on a dependent variable. These models estimate the effect of an independent variable *net of the other independent variables* (ceteris paribus) (Dale and Davies 1994; Treiman 2009). There are many reasons for the appeal of this sort of modelling to this thesis. First to distinguish the difference between the children of parents from advantaged and disadvantaged occupations whilst controlling for other relevant variables. Second, to include parental education and family structure in such models in order to assess their relative contribution to outcomes. Third, to investigate interaction effects between these variables to see if the effects are consistent or whether they have influences which are mediated by other factors under consideration. Finally, monitoring change over time and seeking to improve upon current modelling procedures.

Whilst linear and logistic regression models seek to partition the variance between the influence of distinct independent variables upon a dependent variable, there is potential for multicollinearity and autocorrelation between variables (Treiman 2009). This is where explanatory variables within the model of interest predict other variables, for instance age and number of years worked. The influence of multiple variables may be indirect or pass through a mediating variable (Asher 1976) and result in well-recognised problems of interpretation. As Davis (1985) notes, *"(1) Causal analysis in social research depends on assumptions about causal direction; (2) these assumptions depend on empirical beliefs about how the world works. No statistical routine can give you the answers; (3) the assumptions are not arbitrary or whimsical."* To help overcome these potential issues, it is important to explicitly

state the *a priori* assumptions within these models and build upon previous work in the field (such as the studies in the literature review).

Chapter eight contains analysis of missing data in the YCS. Multiple imputation is selected as to explore patterns of missingness within the data and as a form of sensitivity analysis to evaluate reliability of the estimates in chapter seven. A series of datasets are produced using models to estimate the values for missing data (under the condition Missing at Random using Rubin's rules – Rubin 1987). This builds upon the sensitivity analysis of parental occupation measures in chapter seven by considering the influence of missing data upon the models fitted. The estimates produced are the average effect generated from models fitted to the multiply imputed datasets. A fuller discussion of non response and weighting techniques are considered later in this chapter and in chapter eight.

The regression models fitted in chapter seven have been extended in chapter nine with the assortative technique of latent class analysis, part of the family of generalised linear and mixed models. This technique seeks to measure an underlying latent variable, educational attainment. This is directly measured through a number of manifest variables which is attainment in a number of GCSE subjects. Using maximum likelihood estimation, the most likely number of classes within the latent variable are identified. Latent class analysis permits the identification of groups within the range of GCSE attainment by combination of subjects and the likelihood of young people from different backgrounds being in a particular latent class. This technique was selected as the best means to examine categorical data, particularly the importance of gaining A-C passes in GCSE subjects. Five or more GCSE passes at grades A*/A to C is a common benchmark in school league tables

(Payne, Cheng and Witherspoon 1996; Gayle, Berridge and Davies 2003; Leckie and Goldstein 2009).

## *The Youth Cohort Study*

The Youth Cohort Study (YCS) of England and Wales is a nationally representative dataset consisting of a sample of school-leavers since 1984. It is a postal survey with questions on educational qualifications, attitudes to school, family background and details of post-school routes. The Department for Education and Employment (DfEE), Department for Education and Skills (DES), the Department for Children, Schools and Families (DCSF) and now the Department for Education have managed the YCS (Department for Education 2011). Each cohort is re-interviewed at three (and sometimes four) time points or sweeps. This is demonstrated in figure 5.1.

| | Year of Survey | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **85** | **86** | **87** | **88** | **89** | **90** | **91** | **92** | **93** | **94** | **95** |
| **COHORT** | | | | | | | | | | | |
| **1** | 1 | 2 | 3 | | | | | | | | |
| **2** | | 1 | 2 | 3 | | | | | | | |
| **3** | | | 1 | 2 | 3 | | | | | 4 | |
| **4** | | | | | 1 | 2 | 3 | | | | |
| **5** | | | | | | | 1 | 2 | 3 | | |
| **6** | | | | | | | | 1 | 2 | 3/4(A) | |

| | **94** | **95** | **96** | **97** | **98** | **99** | **00** | **01** | **02** | **03** | **04** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **COHORT** | | | | | | | | | | | |
| **7** | 1 | | 2 | | | | | | | | |
| **8** | | | 1 | | 2 | | 3(A) | | | | |
| **9** | | | | | 1 | 2 | 3/4(A) | | | | |
| **10** | | | | | | | 1/2(A) | | 3 | | |
| **11** | | | | | | | | | 1 | 2 | 3 |

| | **04** | **05** | **06** | **07** | **08** | ***09*** | ***10*** | ***11*** | ***12*** | ***13*** | ***14*** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **COHORT** | | | | | | | | | | | |
| **12** | 1 | 2 | 3 | | | | | | | | |
| **13** | | | | 1 | 2 | *3* | *4* | | | | |
| **14** | | | | | | *1* | *2* | *3* | *4* | | |
| | | | | | | | | | | | |
| **LSYPE** | 1 | 2 | 3 | 4 | 5 | *6* | *7* | *8* | *9* | *10* | *11* |

(A) Autumn sweeps
Black = Harmonised Cohort
Pink = Not Harmonised (i.e. YCS Cohort not included in the dataset SN5765)
Blue = Recent Cohorts

**Figure 5.1 The Youth Cohort Study design**

Source: Diagram originally prepared by Ian Noble (supplied on request, Vernon Gayle)

The following points apply to figure 6.1:

*1) YCS Samples are drawn from school registers in the spring of year 11*

*2) YCS Sweep 1 usually takes place in the spring of the year following sampling and subsequent sweeps normally in subsequent springs. An Autumn sweep is shown as (A)*

*3) Occasionally there are two sweeps of the same cohort in both spring and autumn of the same calendar year, these are shown e.g. 3/4(A)*

The YCS is a longitudinal design with a short duration panel. For the majority of cohorts, the first sweep is collected in the spring of the 1st post-compulsory year (PC1), the second is collected in the spring of the 2nd post-compulsory year (PC2) and the third collected in the spring of the 3rd post-compulsory year (PC3). Sweep 1 contains questions which asked about experiences and decisions made until the spring of the 1st post-compulsory year (aged 16/17). These variables are denoted by Croxford, Iannelli and Shapira (2007) as t0 or t1 according to whether they are referring to: a) t0: information relevant to the point at which the young people sat their examinations (the summer before PC1); or b) t1: between this point and the spring of PC1. Sweep 2 contains data collected between sweep 1 and the spring of the 2nd post-compulsory year (aged 17/18). Sweep 3 contains data collected between sweep 2 and the spring of the 3rd post-compulsory year (aged 18/19).

## Sample Design

The sample is designed to be representative of all Year 11 pupils in England & Wales. The Department for Education contacts a sample of young people from an academic year group in the spring following completion of compulsory education and these form the cohorts of the YCS. For cohorts up to and including cohort 5, the sampling frame by which these young people were selected was as follows, *"the*

*DES listed all eligible schools in England (the Welsh Office did the same for Wales) having first ordered schools by school type within Local Education Authority (LEA) and region. The data for this stratification came from the DES form 7 exercise carried out in January each year"* (Drew 1995: 56). The sample was then stratified within school by school register with a 10% sample were selected from this based on pupils aged 16 and birth date being the 5th, 15th and 25th of the month. The design changed for the remaining cohorts of the 1990s (YCS cohort 6 onwards) from a multi-stage stratified random sample to a single-stage simple random sample (Croxford 2006b). The format is as follows. *"In Spring of the sampling year all schools in England and Wales (excluding special schools), both state maintained and private sector, are sent a return form for sampling. This gives a number of dates, e.g. the 5th, 15th and 25th, and all pupils on the Year 11 roll whose birthdates coincide are sampled. Usually three dates are specified giving a simple random sample just under 10%. There are some difficulties with school-level non-response at the sampling stage and to compensate for this there is a further stage of sampling before Sweep 1. Here the initial sample is sub-sampled to give a Sweep 1 final sample that is representative of a population matrix of pupil numbers by school type by sex by region."* (Croxford, Ianelli and Shapira 2007: 5)

The young people selected via the sampling process are then contacted by postal questionnaire, the details of which are covered later in this chapter. The dataset used for this research consists of YCS cohorts 5, 6, 7, 8, 9 and 10 (sweep 1 data). These six cohorts cover all young people surveyed during the 1990s. All bar YCS cohort 6 were available as part of the harmonised dataset SN5765 (available from the data archive) (Croxford, Iannelli and Shapira 2007). Cohort 6 was harmonised

with the other cohorts and further details of this process are included in the section on data management.

The sample size is quite large[17] by the standards of available UK datasets for the study of young people in this age range. This is particularly relevant for those wishing to study young people growing up in the 1990s where there are no alternative birth cohort data to the YCS (Gayle 2005; Gayle, Lambert and Murray 2009b).

**Table 5.1 Frequencies within YCS cohorts**

| Source of data | Year completed compulsory schooling | Freq. | Percent |
|---|---|---|---|
| YCS Cohort 5 | 1990 | 14,511 | 14.3 |
| YCS Cohort 6 | 1992 | 24,922 | 24.5 |
| YCS Cohort 7 | 1993 | 18,021 | 17.7 |
| YCS Cohort 8 | 1995 | 15,899 | 15.6 |
| YCS Cohort 9 | 1997 | 14,662 | 14.4 |
| YCS Cohort 10 | 1999 | 13,698 | 13.5 |
| | | | |
| | Total | 101,713 | 100.00 |

Many modelling procedures assume simple random samples and analyses using stratified samples should account for this or there is a tendency to inflate standard errors due to the failure to recognise the strata (Goldstein 2011). However, as there are no identifiers provided within the data for which school a young person went to and only YCS 5 was not a single-stage simple random sample (within the analyses

[17] Although it is less well suited to regional analysis as the subsample for Wales is much smaller. This thesis has concentrated on England and Wales due to the similarity of educational system and qualifications.

presented in this thesis) it is assumed that the change in sampling design should not have had major effects on the analyses conducted. Table 5.2 suggests that the YCS samples are broadly consistent if slightly higher than national statistics for educational attainment.

**Table 5.2** **National and YCS Surveys, percentage of pupils gaining 5 or more GCSEs at grades A*-C**

| Year | National Figure | All YCS Pupils | | | National Boys | Boys YCS | | | National Girls | Girls YCS | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Lower CI | Upper CI | | Estimate | Lower CI | Upper CI | | Estimate | Lower CI | Upper CI |
| 1990 | 34.5 | 35.1 | 34.3 | 35.9 | 30.8 | 31.6 | 30.5 | 32.7 | 38.4 | 38.7 | 37.6 | 39.9 |
| 1992 | 38.3 | 43.9 | 43.1 | 44.6 | 34.1 | 43.6 | 42.5 | 44.7 | 42.7 | 44.1 | 43.1 | 45.2 |
| 1993 | 41.2 | 42.3 | 41.5 | 43.1 | 36.8 | 38.1 | 36.9 | 39.3 | 45.8 | 46.7 | 45.5 | 47.8 |
| 1995 | 43.5 | 44.9 | 44.1 | 45.7 | 39.0 | 40.6 | 39.5 | 41.8 | 48.1 | 49.3 | 48.2 | 50.4 |
| 1997 | 45.1 | 46.2 | 45.3 | 47.0 | 40.5 | 41.9 | 40.7 | 43.1 | 50.0 | 50.6 | 49.4 | 51.7 |
| 1999 | 47.9 | 51.0 | 50.1 | 52.0 | 42.8 | 46.4 | 45.0 | 47.8 | 53.4 | 55.7 | 54.4 | 57.0 |

*YCS Data Source: Dataset SN5765 & SN3532 (weighted data)*

*Source: (1990-2000) Department for Education & Employment, London*

## The Questionnaire

The sweep one questionnaire consists of 11 parts: the following is based on the cohort 6 questionnaire (Courtenay 1996b; Croxford, Iannelli and Shapira 2007).

1) You and Your School (2 pages)
2) Year 10 and 11 Subjects and Exams
3) Decisions and Advice (2 pages)
4) Since Last Year
5) If you are out of work
6) Work and Training (2 pages)
7) Training for a Job
8) If you are in Full Time Education
9) Part-Time Courses and Part-Time Work

10) Courses and Qualifications

11) You, your family and your plans (2 pages)

Of particular relevance to this thesis are the sections 2, 8 and 11. Section 10 is on qualifications that a young person is currently studying for at the time of the survey. Section two contains the qualifications attained in year 11, the primary dependent variables in this research. Information on family background is mostly in section 11. Section 8 provides information about school type.

The self-reporting of educational attainment is subject to error (as opposed to the LSYPE with results imported from the National Pupil Database). Drew (1995) argues that there was no systematic difference in the reporting of results. Any discrepancies with national figures for educational attainment are more likely to be a product of differential non-response among socio-economic groups of young people.

## Non-response bias

Non-response bias may be defined in terms of general and item non-response. General non-response bias can be considered in terms of failure to complete the survey due to non-contact with the respondent or refusal by the respondent (Lynn and Clarke 2002). Item-non response to specific questions can also be considered in terms of the respondent not knowing the answer or refusing to answer (Shoemaker, Eichholz and Skewes 2002). Both are applicable in the context of postal surveys, in this section I will consider general non-response. There are significant concerns about potential sources of non-response bias in postal surveys, in particular low overall response rates (exacerbated with a repeated contact survey) and differential

non-response according to socio-economic characteristics of respondents (Larson and Chow 2003; Leung et al. 2004). The issue of differential response in terms of socio-economic characteristics of respondents is important if the responses of those not being measured are systematically different from respondents (Groves 2006). For example, if young people with parents in routine occupations are less likely to respond to particular questions than those with parents in higher professional occupations, this may lead to bias in the estimates without weighting or other corrective techniques (for a fuller discussion of this see chapter nine).

Attrition may be problematic due to reduced sample size or if it is different by socio-economic characteristics of young people. Table 5.2 reports the response rates to sweep 3 for each cohort of the YCS.

**Table 5.3 Response rates by YCS cohort**

| Cohort | Year | Time 3 survey in spring of | Response rate at Time 3 | | Un-weighted sample size at Time 3 | |
|---|---|---|---|---|---|---|
| | | | % of Time 1 | % of initial target | England | Wales |
| 5 | 1990 | 1993 | 58 | 42 | 7700 | 696 |
| 6 | 1992 | 1994 | 38 | - | 8658 | 839 |
| 7 | 1993 | 1996 | 45 | 30 | 7630 | 569 |
| 8 | 1995 | 1998 | 64 | 41 | 9530 | 600 |
| 9 | 1997 | 2000 | 43 | 28 | 5948 | 356 |
| 10 | 1999 | 2002 | 53 | 34 | 6835 | 403 |

Source: Croxford, Ianelli and Shapira (2007: 6)

There was considerable attrition between sweep 1 (age 16-17) to sweep 3 (aged 18-19) for YCS cohorts 1 to 4 and only those who responded to the previous sweep were included in the issued sample (complete case analysis). This is problematic for using the YCS as a longitudinal dataset. Efforts were made to improve response

among certain sections of the sample who were of substantive interest to policy researchers. Cohorts 5 and 6 involved sending sweep 2 and 3 questionnaires to all sweep 1 sample regardless of whether they had been respondents to sweep 1 (Lynn et al. 1994; Courtenay and Mekkelholt 1996). There were additional boosts to the sample in cohorts 3, 6 and 7 in areas where government initiatives were occurring, these being the Technical and Vocational Education Initiative (TVEI) in cohort 3, Training Credit Area (TCA) and Training and Enterprise Councils (TEC) in cohorts 6 and 7.

## Survey Weighting

Weighting is used routinely to correct for differential probabilities of an individual being selected according to characteristics of that respondent and also for non response. Young people who are under or over represented according to their presence in the population in question can be weighted to counter the influence that their information contributes to summary statistics.

Lynn et al. (1994) describe in detail some of the issues associated with inability or refusal to answer specific questions in the YCS. Weighting can also be used to correct for under and over representation of certain groups of young people in response to particular questions. Some correction can be made by comparing response by particular items (variables/questions) at sweep 1 to national figures (e.g. Office for National Statistics) by region, school type and gender and weighting appropriately where certain types of respondent are over or under represented. Analysing the characteristics of respondents to sweep 1 can provide insight into those who do not chose to respond to subsequent sweeps of the questionnaire. Lynn et al. (1994) use the example of truancy to demonstrate how new weights can be constructed to correct for differential non-response by truancy categories

compared with the proportion within each category at sweep 1. They conclude that the tradition YCS system of correcting for non-response by weighting matrix works reasonably well and changes to the YCS design that would be likely to lead to greater non-response should be avoided. Further to this, it is advised that it would be more sensible to obtain the individual level examination results directly from the Department for Education instead of asking the respondents directly (as is now the case for the PLASC).

From YCS 6 the weights provided are combined attrition and sampling weights (nationally representative in terms of sex, region, school type and GCSE results ) constructed by the survey contractors for each survey sweep to compensate for differential non-response and probability of selection into the sample (National Centre for Social Research 1999; Payne 2001).

Failure to provide accurate information on educational attainment according to parental occupation is a potentially likely bias relevant to this research. However, in a comparison of the reporting of parental occupation by young people with parental self-reporting of their occupation, over 60% of young people correctly reported their parents' occupation at the 4 digit occupational unit group whilst 74% were able to report it at the 1 digit level (Noble and Moon 2008). It is also feasible that the complexity of certain questions (particularly those concerning qualifications) may deter young people with poor linguistic or reading skills (Shoemaker, Eichholz and Skewes 2002). Croxford (2006b: 7) notes that: "*Young people from lower social classes, ethnic minorities and those with lower attainment are least likely to respond to surveys, and thus are under-represented in the data. In an attempt to compensate for these biases in the results, 'weighting' variables are constructed*

*from known characteristics of the sample and population, but are not wholly satisfactory, and documentation of the weighting procedures used in the surveys is inadequate."*

## General limitations of the YCS

Despite the clear strengths of the Youth Cohort Study as a secondary dataset, there are also some of limitations. Croxford (2006b) identifies these with respect to the survey design of the YCS and the contribution to evidence based policy-making. There are a number of challenges when seeking to monitor changes over time. First, the change from a two-stage (school-pupil) design to a simple random sample in 1992 led to a fluctuation in certain measures (with the example provided being social segregation in schools). Second, the impact of competitive tendering has led to inconsistencies in data curation. Thirdly, the wording of questions and coding of responses has varied between cohorts. Fourth, the documentation of procedures used when the data has been curated, managed and released is sub-optimal. All of these factors present an obstacle to using the YCS as time-series data. These limitations are understandable however, given that the YCS was collected on behalf of the Department of Education for the purposes of monitoring trends in attainment and was not primarily designed for academic research into social stratification.

Despite these limitations it remains an extensive data resource covering education in the 1990s. With further investment in coding and documentation, the YCS could be improved for usage by secondary data analysts. It is clear that the work of Croxford, Iannelli and Shapira (2007) is a great improvement on the original documentation provided by the YCS. The recording of educational qualifications in the YCS could be improved and the recording of vocational qualifications is poor. It is confusing for respondents to have multiple pages within the questionnaire

regarding educational questions which have been achieved at different time points. The layout of the questionnaire is similarly confusing where it is easy to confuse academic and vocational qualifications. More space needs to be allowed within the questionnaire for the recording of qualifications and better guidance to respondents that all grades and qualifications are relevant.

With regard to vocational qualifications, there was a large amount of missing data. This is partly due to the confusing nature of vocational qualifications in England and Wales but was not aided by the design of the questionnaire. Respondents frequently are unaware of the NQF/NVQ level of the qualifications studied. This might have been avoided has the NQF/NVQ level been assigned to the qualifications (imputed) from educational records. As Howieson and Croxford (1996: 28) note:

*"The data on courses and qualifications from the YCS are extremely complex and difficult to navigate – this may perhaps reflect real world courses and qualifications. We would like to warn any subsequent analysts of the YCS not to underestimate the amount of time required to understand the data. We have found from experience that some questions which appear to give appropriate information, such as NVQ level of course, or mode of study, may have inadequate data. On the other hand, it is difficult to ascertain the extent of early leaving from courses, especially part-time courses, because relevant questions are not asked at each sweep."*

Many of these problems could have been overcome by better survey design and earlier following of the procedure for the LSYPE where qualifications information is retrieved from the NPD (National Pupil Database), although it is acknowledged that this would have been much more difficult in the early 1990s compared with today.

Further recommendations for the YCS have been made by other secondary analysts. Taylor and Lynn (1997) make a number of suggestions. Increasing the gap between contacts leads to greater sample attrition particularly among subgroups whose response rates tend to be low. Intermediate contact on the attrition sample reduces the proportion of sample members remaining in the study after sweep 3. Shorter questionnaire tends to lead to quicker responses but not for the attrition sample where there is a marginal impact. Sending a personalised cover letter or advance letters marginally increases response rates.

## The challenges of harmonising YCS6 (1992)

The main reason for harmonising the 1992 data was to gain complete coverage of the 1990s data available within the Youth Cohort Study. This harmonisation was challenging due to the inconsistencies in variable recording between different cohorts of the Youth Cohort Study. Much of the data management work for this thesis has been conducted using the harmonised dataset SN5765 (available from the data archive) (Croxford, Iannelli and Shapira 2007; Croxford, Ianelli and Shapira 2007). Cohort 6 of the YCS was downloaded from the data archive (dataset SN3532, Courtenay 1996b) and harmonised with SN5765 (Croxford, Iannelli and Shapira 2007). Some of the issues facing data management have been commented upon within the limitations section, particularly those addressed by Croxford (2006c). The software used in this analysis has been Stata (StataCorp 2009) and lEM (Vermunt 1997).

It is critical when seeking to compare the different cohorts, that the construction and interpretation of the core variables are consistent. This is relevant for several reasons, first because of the different naming conventions of the variables from

sweep to sweep (for example in YCS 5 this might be w1v176 whereas in YCS 6 this might be sw1v042), second because the variable position in each sweep is inconsistent (i.e. different ordering of recording the survey data), and third in order to understand the categorical construction when harmonising future waves.

The outcomes of interest within this research are educational attainment (GCSE results) at age 16 and the GCSE subject choice of young people. Within the Youth Cohort Study, these are recorded as categorical variables as the GCSE grade (ordinal) for each subject that the young person studies for (nominal). From this these can be transformed into a range of outputs: count variables on the number of passes at a certain grade; continuous (ratio) variables on the number of GCSE passes or passes at a certain grade; or nominal categorical variables of those above and below a threshold. The challenge in recoding the educational attainment information was the lack of summary variables provided. All summary measures related to volume and quality of GCSE passes had to be constructed from the individual GCSE grades attained. For certain complex educational variables cohort 5 of the YCS was downloaded from the data archive (dataset SN3531, Courtenay 1996a) to replicate these variables to ensure consistency with the harmonised dataset SN5765 (Croxford, Iannelli and Shapira 2007).

The principle independent variables of interest within the YCS are parental occupation, parental education and parental living arrangements. Parental occupation was recorded in the following manner. In YCS 6, respondents were asked the following question. *'Please tell us about your parents' (or step parents') jobs. If they are not working at the moment, please tell us about their most recent jobs. A) What is the name of the job? B) What kind of work do they do? C) What*

*sort of a place or organisation to they work for? D) Are they self employed?'* The data curators coded these responses into SOC90 codes. The SOC90 classification and self-employment indicator were then transformed via GEODE[18] into a variety of standardised occupational classification schema.[19] GEODE is an invaluable resource in translating between different occupational classification schemes. These can be expressed as ordinal variable (e.g. RGSC), a categorical variable (e.g. NS-SEC) or as an interval continuous variable (e.g. CAMSIS). There were no indicators of employment status for parental occupations but whether the respondent's father or mother was self-employed was included within the YCS. All respondents were coded as "other employees" apart from the self employed who were coded as "employers – small organisations."

Parental education and parental living arrangements are recorded as nominal categorical variables. Parental education was recorded by the question, *'Do either of your parents have a degree?'* The answer categories were, *'Yes, No, Not sure, I prefer not to say.'* These were coded by the data curators into, *'Yes, No, Don't Know, Not Applicable',* and maintained in this format by Croxford, Iannelli and Shapira (2007).

Family structure was a complex question which then had to be disaggregated into a number of variables. The question was phrased: *'Who lives in the same household as you?'* with response categories: *'Father (or stepfather); Mother (or Stepmother); Brothers (How many?); Sisters (How many?); Other relative(s) (How many?); Others – Not relatives (How many?)'* From this question variables were constructed by the curators indicating whether the respondent had brothers/sisters and number

[18] See http://www.geode.stir.ac.uk/ accessed 15th April 2011
[19] See http://www.dames.org.uk/ accessed 15th April 2011

of brother/sisters. Croxford, Iannelli and Shapira (2007) then used these to produce the variable number of siblings. The family structure variable of greatest interest used two variables constructed by the data curators 'Father lives in same household' and 'Mother lives in same household' to which the response categories were: 'Father lives in household, 2 (no label), boarding school, and lives alone' and 'Mother lives in household'. Careful cross tabulation and observation of missing data patterns was then used to replicate the variable of Croxford, Iannelli and Shapira (2007), 'Household at t1' with the response categories: 'Not answered, Father and Mother, Father only, Mother only, and Other response.' Further detail about the household may be desirable but only this data is provided within the YCS.[20]

The harmonisation exercise in this thesis with Croxford, Iannelli and Shapira (2007) also proved valuable in increasing familiarity with the YCS and understanding better the complexity of data recording. It is hoped that future data curators will learn from the findings of this and other secondary analysts work using the YCS. Finally, the absence of key variables (such as more information about family structure) is a shortcoming but much can be analysed with the information available. To conclude, I will emphasise a point made by Croxford (2006c: 8):

*"The youth cohort surveys are an important investment of public money, and with more attention to continuity and comparability, and investment in quality of coding and documentation, they would provide a reliable basis for evaluation of education and training and provision."*

[20] Further controls (additional independent variables) of interest are gender, ethnicity, school type and housing type. All are nominal categorical variables. For further details of these procedures, please see the accompanying STATA syntax files and guidance notes provided in the appendices.

## Summary

At the beginning of this chapter, it was identified that quantitative methods were the best suited to answering the five research questions. This research is positioned within the midrange development of empirical work. In previous chapters it was identified that a complete analysis of 1990s data to investigate this period was desirable. The Youth Cohort Study is the strongest available data to do so. Within this chapter, the structure, sample design, weighting, limitations, questionnaire, data management and coding of the YCS was described and evaluated.

There are clear limitations to the YCS as a postal survey with issues around missing data. The fourth research question seeks to investigate how this missing data may influence the findings, and this is the topic of chapter eight. Whilst it is acknowledged that the YCS can be a challenging dataset to work with, it is the only nationally representative large scale source of school leavers in England and Wales during the 1990s. As Drew (2007) notes: *"...the Youth Cohort Study provides a set of data which is the strongest and most reliable of its kind in England and Wales and enables detailed analysis to be carried out on the progress of young people, either through post-compulsory education, or into the labour market."*

The challenges of working with the YCS have been described in the final section of this chapter, particularly in relation to harmonising YCS6 (1992) with the remaining cohorts. This work is important though as it enables the full series of 1990s cohorts to be analysed.

# 6 Descriptive Statistics

This chapter consists of the distributions and summary statistics for measures of educational attainment, family background (parental occupation, parental education and family structure) and other stratification variables (ethnicity, school type and housing tenure type). It is widely agreed that prior to undertaking more complex analysis, it is important to understand the empirical features of measures, such as statistics of dispersion and central tendency, and patterns of association and the functional form of associations (Gayle 2000; Treiman 2009). This chapter provides further details of the distributions of the variables and descriptive statistics for the complete set of 1990s YCS cohorts.

## *Educational Variables*

The principle educational outcomes of interest are based upon GCSE attainment in year 11 (collected at sweep 1 of the cohorts). These have been expressed in two variables. The first measure is a percentage of those young people gaining five or more GCSE passes at grades A*/A to C, a common benchmark in school league tables (Payne, Cheng and Witherspoon 1996; Gayle, Berridge and Davies 2003; Leckie and Goldstein 2009). The second is a cumulative points score measure based upon grade attained, where a GCSE at grade G is equivalent to 1 point and grade A*/A is equivalent to 7 points (grade A* was introduced in 1994). The points score measure was capped at 84 points (equivalent to 12 GCSE passes at A*/A)[21]. Educational attainment has risen steadily through the 1990s (ONS 2001; ONS 2008;

[21] There are 32 individuals gaining more than 84 points but these are likely to be data errors so have been recoded to 84 points. The weighted mean and standard error for the non-capped score measure are 35.06 and 0.07 and for the capped score measure these are also 35.06 and 0.07

ONS 2009a; Lupton, Heath and Salter 2009; Gayle, Lambert and Murray 2009b). The percentage gaining 5+ A*-C rose from 35% in 1990 to 51% in 1999 as can be seen in figure 6.1. Nationally the figures reported within this period were 35% in 1990 and 48% in 1990[22].

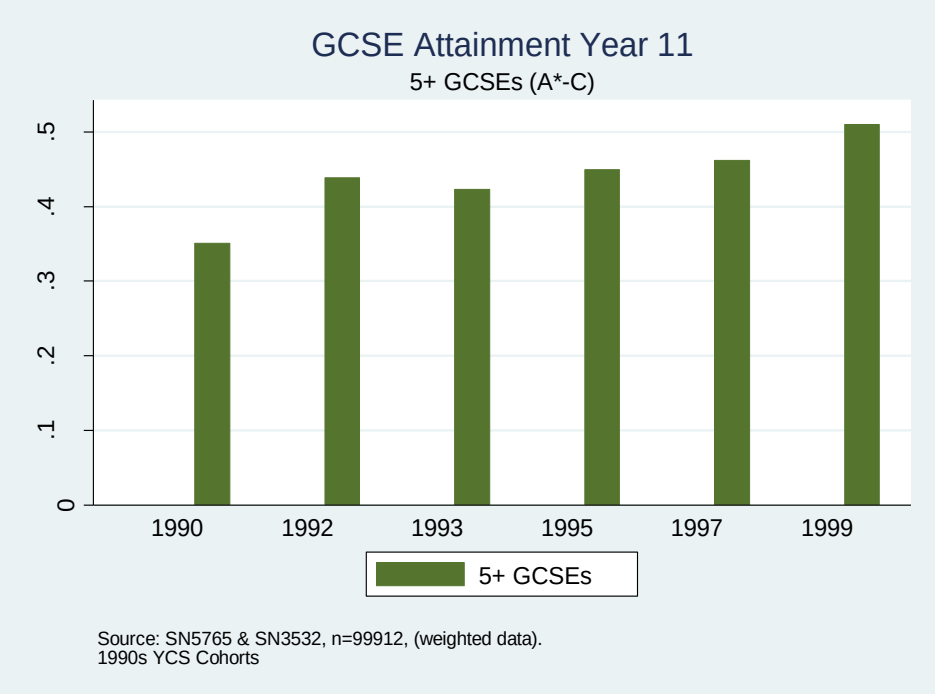



**Figure 6.1** **Percentage of pupils gaining 5 or more A*-C Grade GCSEs, YCS cohorts 5-10 (weighted)**

The median number of A*-C passes at GCSE rose from 2 in 1990 to 5 in 1999 whilst the median number of GCSE courses studies remained constant at 8. The inter-quartile range for number of A*-C GCSE passes was between 0 and 6 1990 compared with between 1 and 9 in 1999 suggesting greater dispersion of attainment among young people in later cohorts (see fig 6.2).

[22] See http://www.bstubbs.co.uk/5a-c.htm#table1 accessed 15th April 2011 - *Source: (1990-2000) Department for Education & Employment, London*

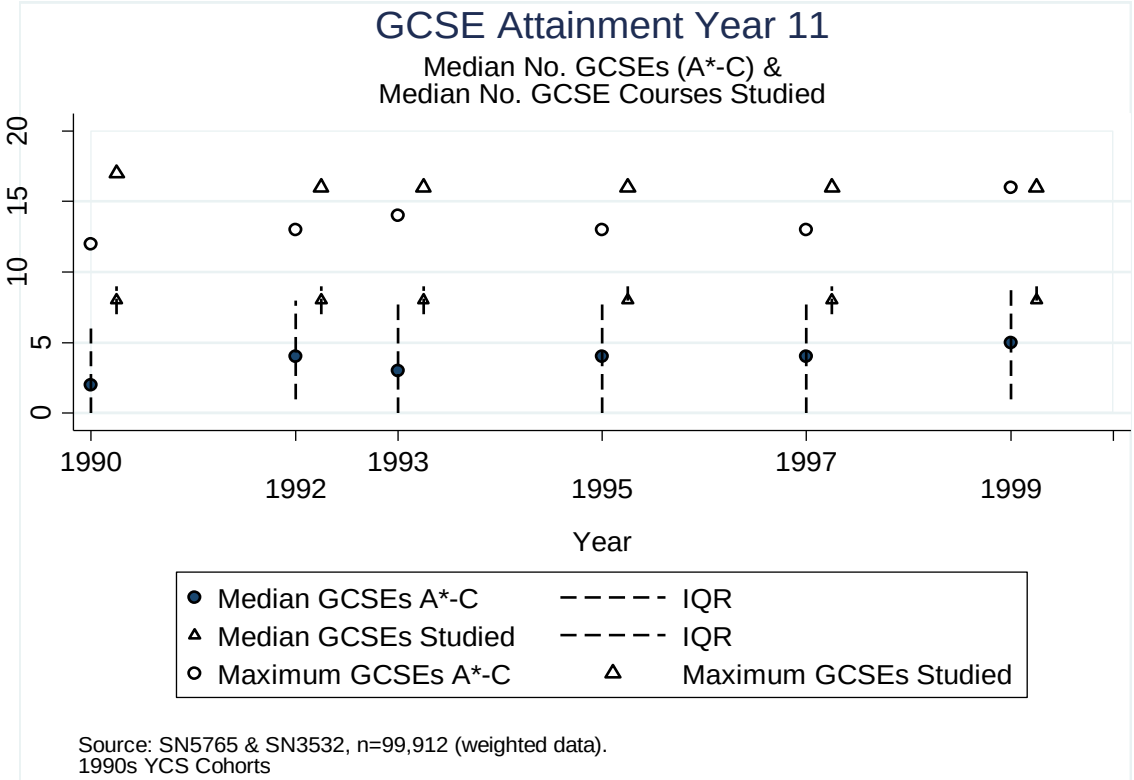



**Figure 6.2** **Median number of A*-C GCSE grades and median number of GCSE courses studied, YCS cohorts 5-10 (weighted)**

In 1990, 31% gained no A*-C GCSE pass, 33% had 1-4 passes and 35% gained 5 or more at these grades. By 1999, these figures were 22%, 27% and 51% respectively (see fig 6.3). Given that the median number of A*-C passes rose between 1990 and 1999 (see figure 6.2) but the categorical measure in figure 6.3 indicates a persistent group of young people with no GCSE, there is greater dispersion in the grades achieved, which is also demonstrated by the inter-quartile range in figure 6.2. This seems to suggest a widening of the gap throughout the 1990s between those young people gaining 5 or more A*-C passes and those who do not (see fig 6.3).

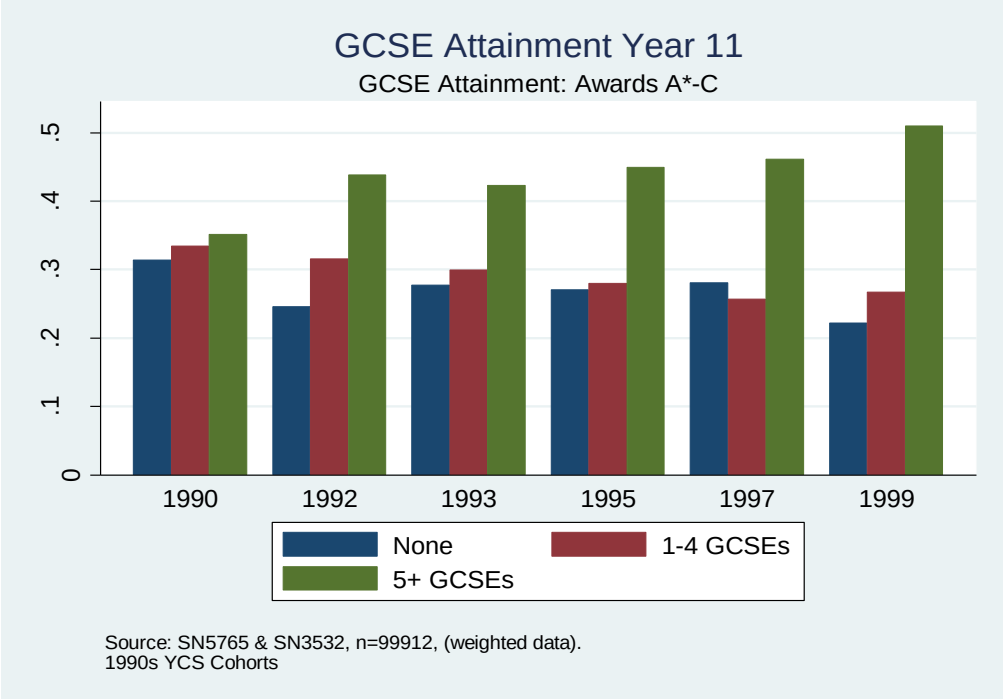



**Figure 6.3** **Proportion of year 11 pupils gaining zero, 1-4 or 5 or more GCSEs at grades A*-C, YCS cohorts 5-10 (weighted)**

Using the points score measure, the median score rose from 29 points in 1990 to 41 points in 1999 (see fig 6.4). The median points score is useful to control for outliers, as there are a small number of young people studying a very large number of GCSEs. As previously noted, the median number of GCSEs being studied for across cohorts is 8. These median points scores are equivalent to 7 GCSE passes at grade D and 1 at G in 1991 versus 1 GCSE pass at grade B and 7 at C in 1999. This is important as this example suggests how the increase in the proportion of young people gaining 5 or more A*-C passes has risen so greatly in figure 6.3.

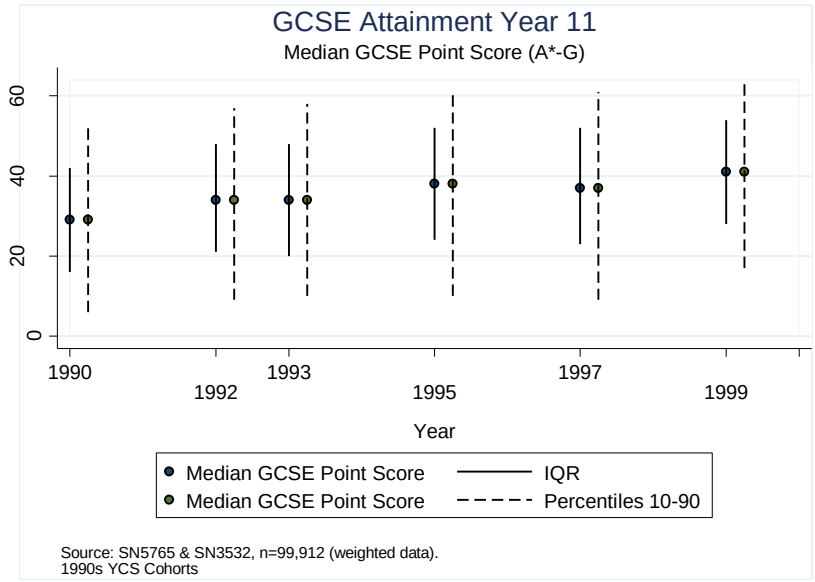



**Figure 6.4 Median GCSE Points Score, YCS cohorts 5-10 (weighted)**

The distribution of GCSE points score is reported in figure 6.5. Five percent of young people recorded as entering exams do not gain any GCSE passes with the modal score being 45-50 points.

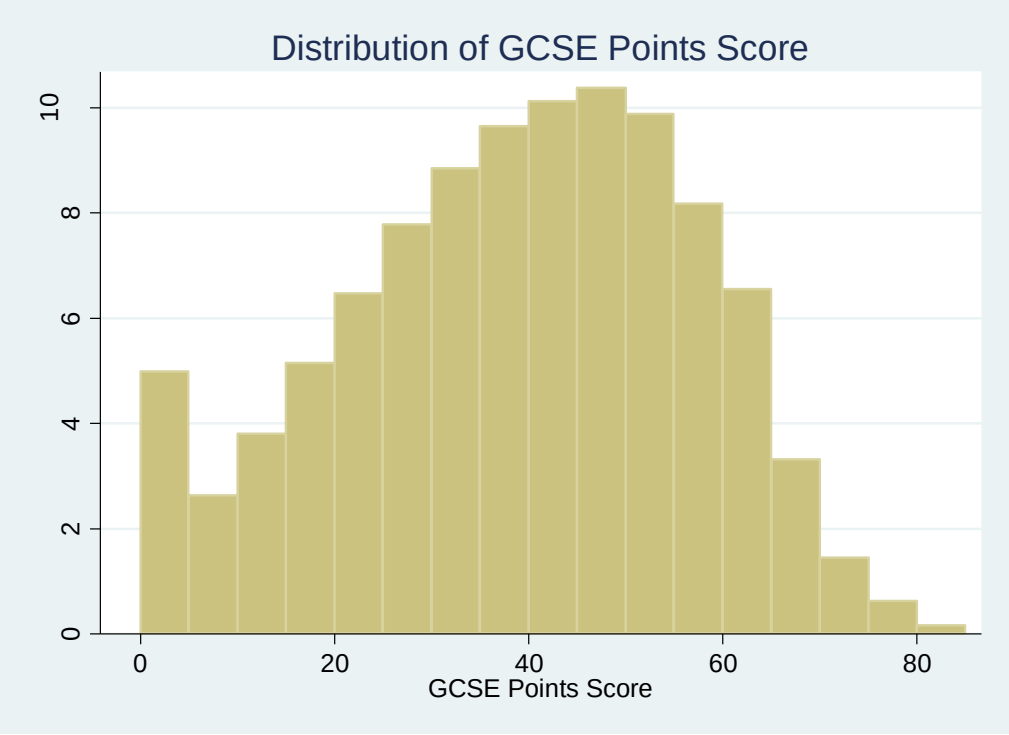



**Figure 6.5 Distribution of GCSE Points Score, YCS cohorts 5-10 (weighted)**

This figure has increased in later cohorts. The mean number of points has risen throughout the 1990s for both male and female young people with females consistently outperforming males (see table 6.1).

**Table 6.1 Mean GCSE Points Score by sex, YCS cohorts 5-10 (weighted)**

| | Male | | Female | |
|---|---|---|---|---|
| | **Mean** | **Standard Error** | **Mean** | **Standard Error** |
| 1990 | 29.26 | 0.15 | 30.81 | 0.20 |
| 1992 | 33.60 | 0.21 | 34.06 | 0.19 |
| 1993 | 32.32 | 0.24 | 36.03 | 0.21 |
| 1995 | 35.06 | 0.24 | 39.20 | 0.22 |
| 1997 | 34.66 | 0.25 | 38.51 | 0.24 |
| 1999 | 38.86 | 0.26 | 42.72 | 0.24 |

Different measures of GCSE performance report rising attainment through the 1990s and increased dispersion around the mean (demonstrated by measures of spread, for example inter-quartile range and standard deviation, see Blaikie 2003). The next subject of interest is changes to the family background factors.

## *Family Background*

The marginal distributions of family background variables over time are considered first followed by distributions of educational attainment according to these factors. Parental occupational classifications are constructed using the semi-dominance approach (i.e. the highest ranking occupation is selected from fathers' and mothers' occupations, see Erikson 1984).

### Parental Occupation

Three principle categorisations of parental occupation are used: NS-SEC (National Statistics Socio-Economic Classification, see Rose and Pevalin 2003), RGSC (Registrar General's Social Class, see Rose 1995) and CAMSIS (Cambridge Social Interaction and Stratification Scale, see Prandy 1999; Prandy and Lambert 2003). It is useful and instructive to compare the findings from the scales to examine consistency in the measurement (see also Lambert and Bihagen 2007). CAMSIS has the advantage of being standardised according to a series of principles which should control for changes in the marginal distribution of the categories. The scale is constructed ranging between 0 and 100 and is based upon occupational advantage and likelihood of a respondent being friends or family of people in similar occupations (Prandy 1990). Limiting the range to between 0 and 100 reduces the influence that changing frequencies within a categorical measure would display as the units are finer. That said the distribution of young people by parental occupation is relatively consistent throughout the 1990s as can be seen in table 6.2 with a slight increase in managerial and professional occupations.

**Table 6.2 Parental NS-SEC by year completed compulsory schooling, YCS cohorts 5-10 (weighted)**

| | Year completed compulsory schooling | | | | | | |
|---|---|---|---|---|---|---|---|
| **Parents NS-SEC 9 class** | **1990** | **1992** | **1993** | **1995** | **1997** | **1999** | **Total** |
| 1.1 Large employers and higher managerial | 5.8% | 6.2% | 5.4% | 5.3% | 5.6% | 7.8% | 6076 |
| 1.2 Higher professional | 9.0% | 7.7% | 9.6% | 10.5% | 9.7% | 12.2% | 9723 |
| 2. Lower managerial and professional | 21.4% | 24.8% | 21.6% | 23.8% | 23.1% | 22.0% | 23341 |
| 3. Intermediate | 14.7% | 14.6% | 14.9% | 15.8% | 15.8% | 13.6% | 15160 |
| 4. Small employers and own account | 15.3% | 7.7% | 14.9% | 13.6% | 15.1% | 15.0% | 13265 |
| 5. Lower supervisory and technical | 6.1% | 8.0% | 5.5% | 5.1% | 5.2% | 4.4% | 6053 |
| 6. Semi-routine | 9.2% | 11.9% | 9.5% | 9.6% | 9.9% | 9.6% | 10289 |
| 7. Routine | 6.7% | 7.7% | 6.0% | 5.3% | 5.5% | 5.2% | 6321 |
| Unclassified | 11.8% | 11.5% | 12.6% | 11.0% | 10.1% | 10.2% | 11485 |
| | | | | | | | |
| **Total** | **14511** | **24922** | **18021** | **15899** | **14662** | **13698** | **101713** |
| | 100% | 100% | 100% | 100% | 100% | 100% | |

The educational attainment of young people in the 1990s differs significantly by family background. There is a moderate correlation between CAMSIS and GCSE points score (r=0.37). In a linear regression between these variables GCSE points score increases by 0.5 for every point on the CAMSIS scale with an intercept of 12.1. The $r^2$ is 0.14 suggesting a modest but significant proportion of variance in educational attainment is explained by parental occupation.

The association between parental occupation and educational attainment is similar when analysed using the mean GCSE points score attained by parental NS-SEC and RGSC measures (see tables 6.3 and 6.4). Females consistently score higher than males, net of the association between NS-SEC/RGSC and mean GCSE points score. Throughout this chapter, the order of the first two NS-SEC categories has been reversed for clarity and consistency with the ordinal nature of other schemes, for example RGSC, and to a lesser extent CAMSIS.

**Table 6.3** **Mean GCSE Points Score and standard deviation by NS-SEC and sex, YCS cohorts 5-10 (weighted)**

| | Mean score | | SD | | |
|---|---|---|---|---|---|
| **NS-SEC** | **Male** | **Female** | **Male** | **Female** | **Number of cases** |
| 1.2 Higher professional | 46.6 | 48.9 | 17.3 | 16.2 | 9663 |
| 1.1 Large employers and higher managerial | 42.2 | 44.8 | 17.6 | 16.7 | 6026 |
| 2. Lower managerial and professional | 39.2 | 42.5 | 17.6 | 16.9 | 23128 |
| 3. Intermediate | 34.8 | 38.8 | 17.2 | 16.4 | 15006 |
| 4. Small employers and own account | 30.8 | 34.3 | 17.3 | 17.4 | 13027 |
| 5. Lower supervisory and technical | 29.6 | 31.8 | 16.2 | 16.4 | 5946 |
| 6. Semi-routine | 27.2 | 30.2 | 16.0 | 16.7 | 10028 |
| 7. Routine | 24.3 | 27.2 | 16.1 | 16.6 | 6150 |
| Missing | 24.6 | 26.9 | 17.0 | 17.3 | 10938 |
| | | | | | 99912 |

**Table 6.4** **Mean GCSE Points Score and standard deviation by RGSC and sex, YCS cohorts 5-10 (weighted)**

| | Mean score | | SD | | |
|---|---|---|---|---|---|
| **RGSC** | **Male** | **Female** | **Male** | **Female** | **Number of cases** |
| I Professional etc | 46.5 | 48.7 | 17.6 | 16.4 | 9037 |
| II Managerial and Technical | 39.5 | 42.8 | 17.7 | 16.8 | 34463 |
| III (N) Skilled non-manual | 33.7 | 36.9 | 16.8 | 16.6 | 20043 |
| III (M) Skilled manual | 27.7 | 30.8 | 16.2 | 16.6 | 15403 |
| IV Partly-skilled | 26.6 | 29.1 | 16.5 | 17.0 | 8092 |
| V Unskilled | 23.1 | 25.7 | 16.2 | 17.2 | 2077 |
| Missing | 24.7 | 26.8 | 17.2 | 17.3 | 10797 |
| | | | | | 99912 |

These findings are consistent when the educational outcome measure is changed to the percentage gaining five or more GCSEs at grades A*-C (see tables 6.5 and 6.6).

**Table 6.5** **Percentage gaining 5+ A*-C passes by NS-SEC, YCS cohorts 5-10 (weighted)**

| Parents NS-SEC 9 class | % 5 A*-C GCSE | Number of Cases |
|---|---|---|
| 1.2 Higher professional | **73.1** | 9663 |
| 1.1 Large employers and higher managerial | **62.8** | 6026 |
| 2. Lower managerial and professional | **57.2** | 23128 |
| 3. Intermediate | **47.0** | 15006 |
| 4. Small employers and own account | **36.1** | 13027 |
| 5. Lower supervisory and technical | **31.8** | 5946 |
| 6. Semi-routine | **26.7** | 10028 |
| 7. Routine | **22.1** | 6150 |
| Missing | **22.6** | 10938 |
| | | |
| Total | **43** | **99912** |

**Table 6.6** Percentage gaining 5+ A*-C passes by RGSC, YCS cohorts 5-10 (weighted)

| RGSC | % 5 A*-C GCSE | Number of Cases |
|---|---|---|
| I Professional etc | **72.5** | 9037 |
| II Managerial and Technical | **57.5** | 34463 |
| III (N) Skilled non-manual | **42.8** | 20043 |
| III (M) Skilled manual | **28.4** | 15403 |
| IV Partly-skilled | **26.3** | 8092 |
| V Unskilled | **20.3** | 2077 |
| Missing | **22.8** | 10797 |
| | | |
| Total | **42.9** | **99912** |

## Parental Education

The variables available within YCS cohorts 5-10 (excluding cohort 6) for studying parental education are, whether father/mother had post-compulsory education, whether father/mother obtained one or more A-levels, and whether father/mother obtained a degree. The parental post-compulsory education variable was constructed to enable comparison with Scottish data and is equivalent to whether father/mother obtain one or more A-levels[23] (see Croxford, Ianelli and Shapira 2007). For YCS cohort 6 the question whether parents had 1 or more A-levels was not asked, only whether they had a degree or not. It was decided to construct a variable based on whether or not the parents had A-levels supplemented by those who had a degree. Therefore a parent was recorded as having A-levels or above if they a) had a degree or b) had A-levels. This recoding therefore ensures some continuity in the measure.

**Table 6.7 Parental Education, YCS cohorts 5-10 (weighted)**

| | Freq. | Percent | | Freq. | Percent |
|---|---|---|---|---|---|
| Father has Degree | 16524 | 16.2% | Mother has Degree | 11000 | 10.8% |
| Father has A-levels | 8184 | 8.1% | Mother has A-levels | 10286 | 10.1% |
| Father Neither | 77005 | 75.7% | Mother has Neither | 80427 | 79.1% |
| | | | | | |
| Total | 101,713 | 100.00 | Total | 101,713 | 100.00 |

It is clear from table 6.7 that higher levels of parental education are associated with improved filial GCSE attainment. The measure constructed is hierarchical (highest qualification) and perhaps under-reports the number of parents with A-levels or above as their highest qualification due to the respondent not being aware of their parents' qualification. There is a gradual increase in parental education through the

[23] There are a few exceptions which cannot be explained.

1990s. In 1990 13.6% of fathers and 7.6% of mothers had a degree. By 1999, these figures had risen to 19.3% and 15.2%. A similar increase in the proportion of parents gaining A-levels was also noted. Unfortunately, the survey in cohort six does not report on A-levels for parents' (only degrees) which may influence the estimates slightly as cohort six is a large sample and it would therefore be expected that a large proportion of the parents of young people within the sample would have A-levels. From table 6.7 it is clear that there were fewer mothers with degrees.

**Table 6.8** **Percentage gaining 5+ A*-C GCSE by parental education, YCS cohorts 5-10 (weighted)**

| | % 5 A*-C GCSE | Number of Cases |
|---|---|---|
| Father has Degree | 68.6% | 16383 |
| Father has A-levels | 56.4% | 8100 |
| Father Neither | 37.8% | 75429 |
| | | |
| Mother has Degree | 68.8% | 10921 |
| Mother has A-levels | 60.2% | 10200 |
| Mother has Neither | 38.7% | 78791 |

There does not appear to be a substantial difference between the influences of different parents (i.e. fathers versus mothers) education on their children's educational attainment from table 6.8. This is similar to table 6.9 where the difference in attainment by sex is consistent across all parental educational levels, potentially suggesting parental homophily with respect to educational attainment (Bottero and Prandy 2003; Bottero 2005).

**Table 6.9 Percentage gaining 5+ A*-C GCSE by parental education and sex, YCS cohorts 5-10 (weighted)**

| | Male | | Female | |
|---|---|---|---|---|
| | **% 5 A*-C GCSE** | **Number of Cases** | **% 5 A*-C GCSE** | **Number of Cases** |
| Father has Degree | **65.7%** | 7988 | **71.9%** | 8395 |
| Father has A-levels | **52.5%** | 3768 | **60.3%** | 4332 |
| Father Neither | **34.4%** | 34506 | **41.2%** | 40923 |
| | | | | |
| Mother has Degree | **66.4%** | 5180 | **71.2%** | 5741 |
| Mother has A-levels | **56.1%** | 4763 | **64.4%** | 5437 |
| Mother has Neither | **35.5%** | 36319 | **42.1%** | 42472 |

The association between GCSE attainment and parental occupation (using RGSC) and parental education is shown in table 6.10 using RGSC. Occasionally there are higher levels of attainment for young people with parents in lower grade occupations but with higher levels of education than the overall association would suggest. It is likely that this is largely due to small sample size corresponding to the lower relative likelihood of parents with high levels of education being in these lower grade occupations. The distribution of female occupations is such that in certain RGSC classes there are very few mothers in these occupations (for example skilled manual occupations, class IIIm). In these unusual cases, the general pattern is sometimes disturbed.

Increased parental education tends to improve the GCSE performance of their offspring for the majority of parental occupations. Parental education has more of an influence when parents (especially mothers) are in more advantaged occupations (often with greater education requirements). Whether this is due to greater inherited ability, greater encouragement of young people to do well at school or greater financial resources to support schooling cannot be deduced from these data.

## Table 6.10 Percentage gaining 5+ A*-C GCSE by RGSC and parental education, YCS cohorts 5-10 (weighted)

| RGSC | Fathers Education | % 5 A*-C GCSE | Number of Cases | Mothers Education | % 5 A*-C GCSE | Number of Cases |
|---|---|---|---|---|---|---|
| I Professional etc | Father has Degree | **79.7%** | 5578 | Mother has Degree | **84.8%** | 3088 |
| I Professional etc | Father has A-levels | **71.5%** | 918 | Mother has A-levels | **78.2%** | 1472 |
| I Professional etc | Father Neither | **60.0%** | 2541 | Mother has Neither | **63.9%** | 4477 |
| | | | **9037** | | | **9037** |
| | | | | | | |
| II Managerial and Technical | Father has Degree | **73.3%** | 7869 | Mother has Degree | **71.6%** | 5992 |
| II Managerial and Technical | Father has A-levels | **65.5%** | 4255 | Mother has A-levels | **68.3%** | 5270 |
| II Managerial and Technical | Father Neither | **51.8%** | 22339 | Mother has Neither | **52.6%** | 23201 |
| | | | **34463** | | | **34463** |
| | | | | | | |
| III (N) Skilled non-manual | Father has Degree | **49.3%** | 1185 | Mother has Degree | **46.7%** | 749 |
| III (N) Skilled non-manual | Father has A-levels | **50.8%** | 1521 | Mother has A-levels | **50.2%** | 1886 |
| III (N) Skilled non-manual | Father Neither | **42.4%** | 17337 | Mother has Neither | **42.6%** | 17408 |
| | | | **20043** | | | **20043** |
| | | | | | | |
| III (M) Skilled manual | Father has Degree | **28.0%** | 797 | Mother has Degree | **28.4%** | 449 |
| III (M) Skilled manual | Father has A-levels | **30.1%** | 682 | Mother has A-levels | **34.0%** | 705 |
| III (M) Skilled manual | Father Neither | **29%** | 13924 | Mother has Neither | **28.8%** | 14249 |
| | | | **15403** | | | **15403** |
| | | | | | | |
| IV Partly-skilled | Father has Degree | **37.4%** | 320 | Mother has Degree | **31.6%** | 211 |
| IV Partly-skilled | Father has A-levels | **26.9%** | 302 | Mother has A-levels | **35%** | 400 |
| IV Partly-skilled | Father Neither | **26.7%** | 7470 | Mother has Neither | **26.5%** | 7481 |
| | | | **8092** | | | **8092** |
| | | | | | | |
| V Unskilled | Father has Degree | **22%** | 45 | Mother has Degree | **28.0%** | 38 |
| V Unskilled | Father has A-levels | **19.2%** | 58 | Mother has A-levels | **23.2%** | 50 |
| V Unskilled | Father Neither | **20.8%** | 1974 | Mother has Neither | **20.6%** | 1989 |
| | | | **2077** | | | **2077** |

## Family Structure

The family structure variable (termed household type in the dataset) is based on the question, "*Who lives in the same household as you?"* with the responses: father and mother, father only, mother only or other response. The majority of young people live with both their father and their mother with the next largest group being those living with only their mother (see table 6.11).

**Table 6.11 Family Structure, YCS cohorts 5-10 (weighted)**

| Household at t1 | Freq. | Percent |
|---|---|---|
| 1. Father and mother | 82,638 | 81.25 |
| 2. Mother only | 12,787 | 12.57 |
| 3. Father only | 2,706 | 2.66 |
| 4. Other response | 1,809 | 1.78 |
| Missing | 1,773 | 1.74 |
| | | |
| Total | 101,713 | 100.00 |

**Table 6.12 Family Structure by year completed compulsory schooling, YCS cohorts 5-10 (weighted)**

| | Year completed compulsory schooling | | | | | | |
|---|---|---|---|---|---|---|---|
| **Household at t1** | **1990** | **1992** | **1993** | **1995** | **1997** | **1999** | **Total** |
| 1. Father and mother | 83.1% | 82.7% | 81.5% | 80.7% | 79.8% | 78.6% | 82638 |
| 2. Mother only | 11.0% | 11.5% | 12.8% | 12.9% | 13.6% | 14.3% | 12787 |
| 3. Father only | 2.2% | 2.5% | 2.5% | 2.9% | 3.1% | 3.0% | 2706 |
| 4. Other response | 2.6% | 0.1% | 2.2% | 2.4% | 2.6% | 1.8% | 1809 |
| Missing | 1.2% | 3.2% | 1.0% | 1.2% | 0.9% | 2.3% | 1773 |
| | | | | | | | |
| **Total** | **14511** | **24922** | **18021** | **15899** | **14662** | **13698** | **101713** |
| | 100% | 100% | 100% | 100% | 100% | 100% | |

In table 6.12, the proportion of young people living with both parents declined slightly through the 1990s with slight rises in the proportion living in mother only or father only households. The proportion giving another response was relatively stable (there may be a coding issue[24] in 1992 where the proportion missing notably rises).

The mean GCSE points score is greatest for young people living with both parents (see figure 6.6 and table 6.13). Those living with a single parent score lower and those who answer "other response" fare notably worse. Unfortunately, there is no further clarification concerning the detail of "other response", the family structure variable is recorded once per cohort (no change can be noted) and there is no family history recorded.

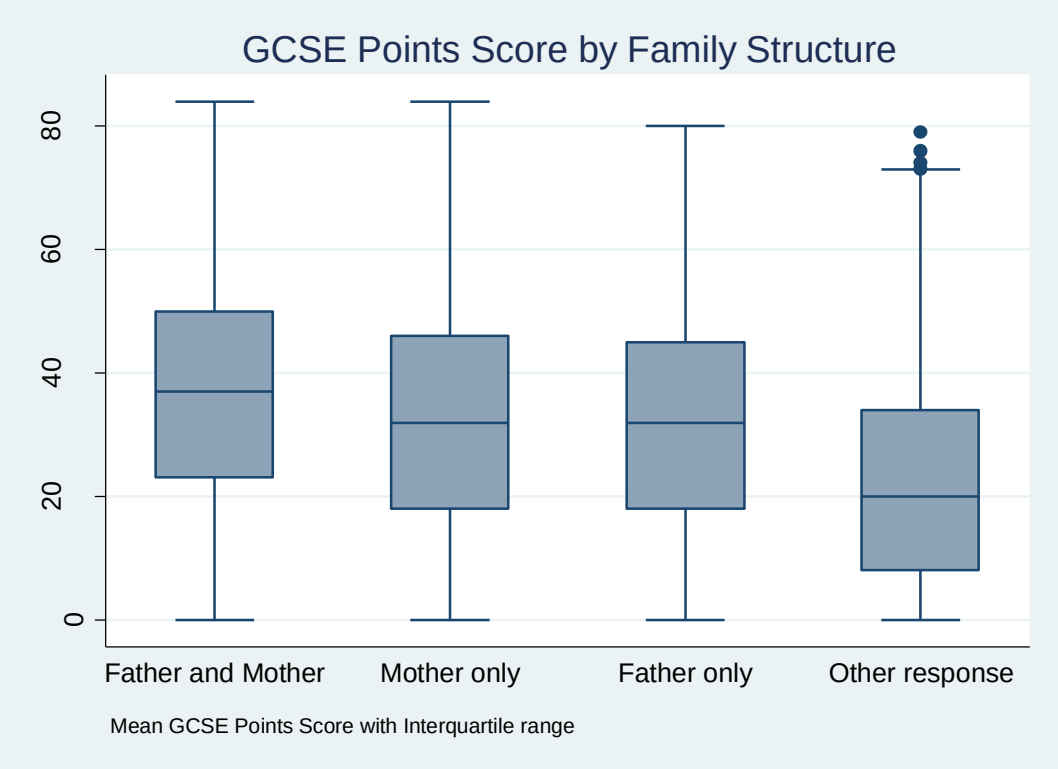



**Figure 6.6 GCSE Points Score by Family Structure, YCS cohorts 5-10 (weighted)**

[24] The family structure variable for YCS6 was constructed from two variables (father lives in same household and mother lives in same household). Due to the way the question was asked, there is therefore a high number of respondents who may have been confused by this format.

**Table 6.13 GCSE Points Score by Family Structure, YCS cohorts 5-10 (weighted)**

| Household at t1 | Mean score | Standard Error | Number of cases |
|---|---|---|---|
| 1. Father and mother | 36.2 | 0.08 | 81464 |
| 2. Mother only | 31.9 | 0.18 | 12488 |
| 3. Father only | 31.5 | 0.39 | 2627 |
| 4. Other response | 22.1 | 0.42 | 1728 |
| | | | 98307 |

The points score measure is supported by GCSE attainment measured by the percentage gaining 5 or more A*-C passes (see figure 6.8). Those young people living with mother only and father only seem to have similar patterns of attainment.

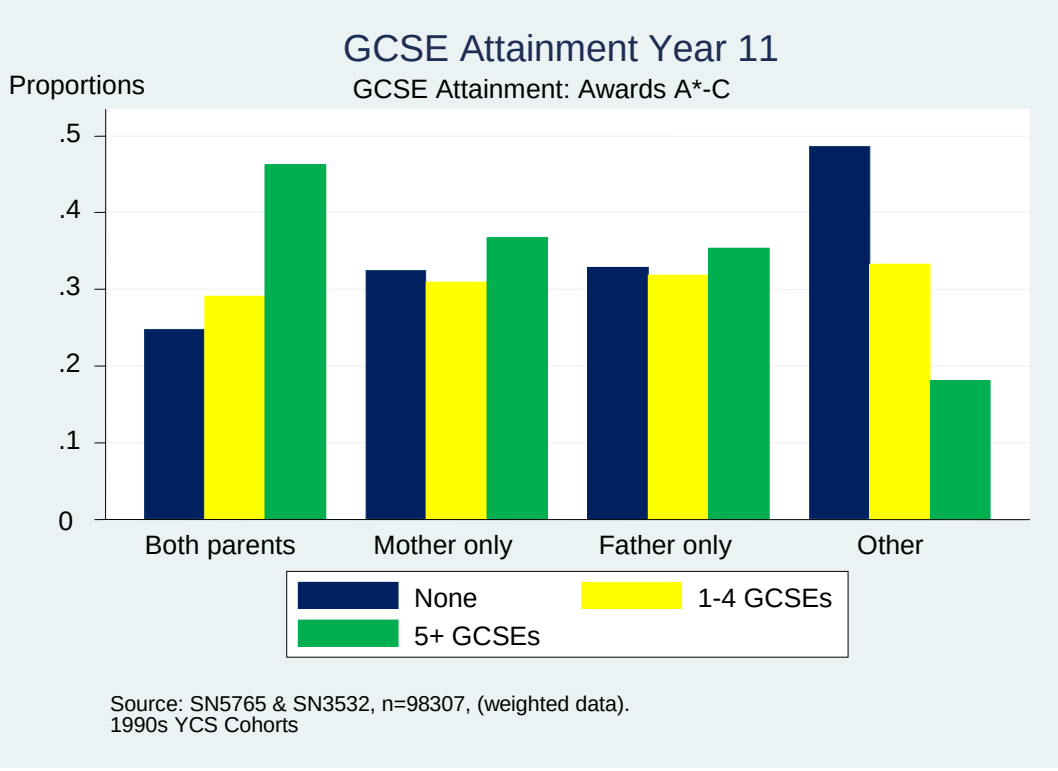



**Figure 6.7 Proportion of year 11 pupils gaining zero, 1-4 or 5 or more GCSEs at grades A*-C by family structure, YCS cohorts 5-10 (weighted)**

The effects of parental occupation appear relatively consistent across family structures (see table 6.14). Young people with parents in more advantaged

occupations do better, irrespective of family structure. Females consistently outperform males. However, among young males the effect of living with a professional mother only appears to boost their GCSE performance (although the sub-sample size is small, 293 cases). Living in a single-parent household appears to be associated with lower attainment (irrespective of whether it is the mother or the father). There does not appear to be a consistent pattern of differences according to family structure within parental occupational classifications.

If a young person gives "other response" to the family structure question, it is assumed that the parental occupation and parental education measures are based on non-resident parents but without further information this cannot be verified. With regard to this group, several points emerge from table 6.14. There are very few respondents within each category when partitioned by RGSC and gender. Their overall attainment is particularly low but markedly so when a parent has occupations categorised as skilled manual, partly skilled or unskilled. As is the case for measures of parental education, it would be best to move to regression models to properly investigate the differences for other family structures by parental occupation with controls for other potential sources of variation.

## Table 6.14 Percentage gaining 5+ A*-C GCSE by Family Structure, RGSC and sex, YCS cohorts 5-10 (weighted)

| | | Male | | Female | |
|---|---|---|---|---|---|
| **Family Structure** | **RGSC** | **% 5 A*-C GCSE** | **Number of Cases** | **% 5 A*-C GCSE** | **Number of Cases** |
| Both Parents | 1 | **70.5%** | 3924 | **76.5%** | 4148 |
| | 2 | **55.5%** | 13887 | **63.4%** | 15767 |
| | 3 | **40.9%** | 7741 | **48.3%** | 8938 |
| | 4 | **26.2%** | 6151 | **33.8%** | 7037 |
| | 5 | **25.3%** | 2785 | **30.6%** | 3439 |
| | 6 | **19.9%** | 643 | **25.1%** | 694 |
| | Missing | **24.1%** | 3064 | **28.6%** | 3246 |
| | | | | | |
| Mother Only | 1 | **80.4%** | 293 | **72.6%** | 362 |
| | 2 | **48.6%** | 1536 | **57.7%** | 1914 |
| | 3 | **35.5%** | 1075 | **42.1%** | 1512 |
| | 4 | **24.0%** | 520 | **33.0%** | 649 |
| | 5 | **24.4%** | 534 | **28.1%** | 782 |
| | 6 | **14.2%** | 246 | **24.2%** | 320 |
| | Missing | **20.2%** | 1192 | **22.7%** | 1553 |
| | | | | | |
| Father Only | 1 | **57.7%** | 105 | **73.6%** | 91 |
| | 2 | **45.9%** | 398 | **56.8%** | 368 |
| | 3 | **40.0%** | 212 | **43.9%** | 204 |
| | 4 | **21.4%** | 277 | **25.6%** | 291 |
| | 5 | **21.0%** | 130 | **27.2%** | 132 |
| | 6 | **16.6%** | 35 | **33.4%** | 43 |
| | Missing | **20.6%** | 176 | **17.5%** | 165 |
| | | | | | |
| Other Response | 1 | **38.0%** | 23 | **43.8%** | 39 |
| | 2 | **26.7%** | 119 | **39.5%** | 242 |
| | 3 | **21.6%** | 78 | **27.9%** | 135 |
| | 4 | **8.8%** | 78 | **12.1%** | 241 |
| | 5 | **10.9%** | 60 | **13.7%** | 123 |
| | 6 | **3.8%** | 25 | **3.0%** | 39 |
| | Missing | **7.7%** | 193 | **13.8%** | 333 |

## *Other Variables of Interest*

The following variables have been identified from the literature review as further stratifying variables, namely ethnicity, school type and housing tenure type.

**Table 6.15 Ethnic classification, YCS cohorts 5-10 (weighted)**

| Ethnic classification | Frequency | Percent | National %[25] |
|---|---|---|---|
| White | 91,447 | 89.91 | 87.0 |
| Black | 1,711 | 1.68 | 2.2 |
| Indian | 2,529 | 2.49 | 2.0 |
| Pakistani | 1,733 | 1.70 | |
| Bangladeshi | 756 | 0.74 | 0.5 |
| Other Asian | 1,140 | 1.12 | |
| Other response | 898 | 0.88 | |
| Missing | 1,499 | 1.47 | |
| | | | |
| Total | 101,713 | 100.00 | |

Ninety percent of young people in the YCS study are white (table 6.15), which compares with 87% nationally (for all age groups). A higher proportion of Indian and Other Asian young people gain five or more A*-C GCSE passes than White young people. By contrast black, Pakistani and Bangladeshi young people are less likely to (see table 6.16).

[25] See http://www.statistics.gov.uk/census2001/profiles/commentaries/ethnicity.asp , accessed 15th April 2011 white percentage for England only.

**Table 6.16 Percentage of pupils gaining 5 or more A*-C Grade GCSEs by ethnicity, YCS cohorts 5-10 (weighted)**

| | Male | | Female | |
|---|---|---|---|---|
| **Ethnicity** | **% 5 A*-C GCSE** | **Number of Cases** | **% 5 A*-C GCSE** | **Number of Cases** |
| White | **41.3%** | 41509 | **47.9%** | 48425 |
| Black | **27.4%** | 680 | **31.7%** | 993 |
| Indian | **45.4%** | 1191 | **49.9%** | 1315 |
| Pakistani | **25.3%** | 822 | **27.5%** | 868 |
| Bangladeshi | **22.2%** | 348 | **27.4%** | 392 |
| Other Asian | **54.1%** | 563 | **60.5%** | 566 |
| Other Response | **38.0%** | 375 | **51.5%** | 511 |

It is likely though that the association between GCSE attainment and ethnic background may also be reflected by parental occupation with different ethnic groups having different parental occupational structures.

Year 11 school type is a strong predictor of educational attainment as a consequence of stratification in the schooling system - i.e. independent schools implement selection based on prior educational attainment and admit fee paying students (Goldstein and Thomas 1996). Therefore, school type attended is closely related to parental occupation as to whether entry to these schools can be afforded and the influence of parental occupation and education upon attainment at an earlier age. Unfortunately there are no measures of educational attainment (e.g. at Key Stage 3) prior to age 15/16 in the YCS.

**Table 6.17 School Type, YCS cohorts 5-10 (weighted)**

| School Type | Frequency | Percent |
|---|---|---|
| Comprehensive to 16 | 31,608 | 31.08 |
| Comprehensive to 18 | 54,792 | 53.87 |
| Grammar | 3,903 | 3.84 |
| Secondary Modern | 3,111 | 3.06 |
| Other state funded | 184 | 0.18 |
| Independent | 7,991 | 7.86 |
| CTC | 100 | 0.10 |
| Missing | 24 | 0.02 |
| | | |
| Total | 101,713 | 100.00 |

Whilst the majority of young people in the YCS went to comprehensive schools (table 6.17), a minority went to grammar and independent schools. Young people at these schools do very much better than their counterparts at comprehensive schools, as can be seen in table 6.18.

**Table 6.18 Percentage of pupils gaining 5 or more A*-C Grade GCSEs by school type, YCS cohorts 5-10 (weighted)**

| | Male | | Female | |
|---|---|---|---|---|
| **School Type** | **% 5 A*-C GCSE** | **Number of Cases** | **% 5 A*-C GCSE** | **Number of Cases** |
| Comprehensive to 16 | **33.0%** | 14198 | **40.0%** | 16749 |
| Comprehensive to 18 | **37.4%** | 24821 | **44.7%** | 28949 |
| Grammar | **88.2%** | 1742 | **92.8%** | 2155 |
| Secondary Modern | **21.0%** | 1382 | **26.5%** | 1676 |
| Other State Funded | **22.7%** | 101 | **36.2%** | 75 |
| Independent | **83.9%** | 3957 | **86.3%** | 3984 |
| CTC | **25.4%** | 46 | **60.0%** | 54 |

Finally, for the minority of young people living in accommodation that is either rented or "other" (compared to owner occupier), there appears to be lower GCSE performance (see table 6.20). This can be empirically evaluated in the modelling

process. As for school type, housing (tenure) type is likely to be closely correlated with parental occupation with those parents in more advantaged occupations being more likely to own their homes.

**Table 6.19 Parental housing type, YCS cohorts 5-10 (weighted)**

| Parental housing type | Frequency | Percent |
|---|---|---|
| Owned | 80,675 | 79.32 |
| Rented | 17,568 | 17.27 |
| Other | 1,760 | 1.73 |
| Missing | 1,710 | 1.68 |
| | | |
| Total | 101,713 | 100.00 |

**Table 6.20 Percentage of pupils gaining 5 or more A*-C Grade GCSEs by parental housing type, YCS cohorts 5-10 (weighted)**

| | Male | | Female | |
|---|---|---|---|---|
| **Housing Type** | **% 5 A*-C GCSE** | **Number of Cases** | **% 5 A*-C GCSE** | **Number of Cases** |
| Owned | **45.7%** | 37635 | **54.0%** | 42079 |
| Rented | **20.0%** | 7110 | **23.8%** | 9790 |
| Other | **36.1%** | 793 | **39.1%** | 931 |

It is therefore proposed that whilst there is important variation according to ethnicity, school type and housing tenure type, the best approach is to include these within models of parental occupation, parental education and family structure. There may be practical reasons for why GCSE attainment differs by these variables but this variation is better understood by more comprehensive models. These models will then allow for adequate control of multiple sources of variation rather than further tabulation of the educational attainment measures used so far.

## Summary

It is apparent from the descriptive analysis in this chapter, and from the literature review in chapter four, that there are many sources of variation to educational attainment. The findings are consistent with those in chapter four but provide additional detail for YCS6 (1992) data and also provide complete analysis of all 1990s YCS cohorts. To properly answer the questions posed in chapter five, multivariate models are required. These enable multiple sources of variation to be included when identifying associations with GCSE attainment. The next chapter includes linear and logistic regression models to achieve this.

There is also a need for greater understanding and improved modelling of the processes that influence young people as they grow up as it can be seen that the factors analysed within this chapter are closely inter-related within wider concepts of advantage within the educational system. This is particularly the case for parental occupation, parental education and family structure but also should include gender, ethnicity and those other variables identified within this chapter.

# 7 Modelling GCSE attainment in the 1990s

As identified in chapter two, during the 1990s participation in post-compulsory education increased and the traditional alternatives available to young people declined. These transformations therefore placed greater emphasis on the importance of qualifications at age 16. In a period when the alternatives for those without qualifications were increasingly limited, grades achieved at 16 could influence whether a young person could secure advantageous post-compulsory education; whether they may in general be able to remain in education, or whether they may achieve a relatively advantageous position outside of the education system (e.g. Drew, Gray and Sime 1992). Concurrent with these changes, the educational policies of the 1990s encouraged greater choice of parents and their influence in the selection of school attended by their children (David 1995; Gillard 2007). In addition, across the period, overall educational attainment rose, but there remained a distribution of differences between higher and lower levels of attainment, and accordingly apparent winners and losers in the new educational context.

Detraditionalisation theorists have suggest that young people are increasingly facing more individualised trajectories from school-to-work (Beck and Beck-Gernsheim 2002). However, there remain important family background factors which consistently predict the life chances of young people on the basis of their educational attainment at aged 16. The studies presented in chapter 4 suggest that there are persisting inequalities in educational attainment for young people. This chapter seeks to investigate these relationships further using data from the 1990s. The Youth Cohort Study has been selected as a representative sample of school leavers during this period. Whilst having limitations, it is the strongest data available

to study this period (see chapter five). Descriptive analysis of the YCS data (see chapter six) reported bivariate associations between family background measures and educational outcomes. In this chapter, statistical models are used to explore the joint relative effects of parental and other background factors upon educational outcomes including all 1990s cohorts.

A statistical modelling approach is desirable because previous research on background influences using the YCS have been largely descriptive in nature (see chapter four). Many studies look at simplistic patterns of association without conducting multivariate analyses which control for other relevant differences between individuals (e.g. Roberts and Chadwick 1991; Courtenay and McAleese 1993b; Courtenay and McAleese 1994; Park 1994). The limitation of these previous approaches is that they cannot conclusively disentangle the distinctive influence of different background factors. Other analyses have employed statistical modelling in a manner to those used below (e.g. Gray and Sime 1990; Jesson, Gray and Sime 1991; Drew 1995; Payne 1995a; Payne, Cheng and Witherspoon 1996), for which cases the forthcoming analysis seeks to replicate, and update their findings using large-scale data and a range of measures in order to check for reliability.

The analysis in this chapter uses statistical modelling to address the following questions:

1. Has the association between parental occupation and their children's educational attainment increased, decreased or remained stable through the 1990s?

2. How do different measures of parental occupation and parental education report this association? Are the different measures consistent? (a sensitivity analysis of measures of parental occupation)

3. Have the differences in attainment between pupils from lone-parent and couple families increased or reduced through the 1990s?

This builds upon the descriptive statistics produced in chapter seven examining family background factors, now using linear and logistic regression models of educational attainment. The analysis examines the influence of a range of family background variables upon GCSE attainment, net of the effect of controls for alternative potential sources of variation in attainment. By including all of the variables of interest, it is possible to see the average net effect of each variable independent of the effects of the others. In addition, different measures of parental occupation are used in separate models in order to compare the reliability of the estimates produced by each scheme.

Different measures of GCSE attainment are included to check for the reliability of findings (GCSE points score, number of A*-C grades achieved, whether a young person gains 0, 1-4 or 5 or more GCSE passes at grades A*-C). Measures that vary over time are built into the models to examine trends in the relative influence of parental occupation during the 1990s. Finally, a number of interaction effects are considered.

## Structure of enquiry

The discussion in chapter six emphasised the need to consider a number of variables simultaneously. Accordingly, the tables below contain models including parental occupation (measured using NS-SEC, RGSC and CAMSIS); the influence of parental occupation in different years; family structure; parental education; sex; ethnicity; school type; and housing tenure type. There are a number of different measures of educational attainment. In summary the models are as follows:

**Table 7.1 List of tables in chapter 7**

The tables have a consistent format but differ according to the measure of parental occupation included: model 1 includes NS-SEC, model 2 RGSC and model 3 CAMSIS. Otherwise the three models in each table are identical. This is to enable the consistent comparison of the three parental occupation measures and to identify whether the estimates for the other covariates are consistent when different parental occupational schemes are used. Each of the three models are run separately but the results are collated on table to enable comparison.

Tables 7.2 – 7.6 contain models examining different GCSE attainment outcomes and functional forms of the dependent variable to demonstrate consistency in the relationships between the independent (family background) variables and the dependent (GCSE attainment) variables. Tables 7.7 and 7.8 examine whether this relationship differs within each cohort (year) of the YCS and to examine trends in GCSE attainment over time. Table 7.10 contains a standardised GCSE points score measure to control for increasing GCSE attainment over time, as demonstrated in tables 7.7 and 7.8 (which will discussed further later in the chapter). The remaining models examine changes according to the independent (family background variables). There are two outcomes modelled for each set of independent variables: part a is a logistic regression model for gaining five or more GCSE passes at grades A*-C in year 11, and part b, an ordinary least squares regression on GCSE points score (truncated) in year 11. Table 7.11 and 7.12 model the relationship using only comprehensive school pupils, consistent with the literature and as the majority of pupils attend mixed ability schools (Goldstein and Thomas 1996). Table 7.13 and 7.14 are complete case analyses. Tables 7.15 – 7.18 examine the interactions between parental occupation and parental education (with details of further interactions contained within the appendix).

Weighting is used within the analyses to correct for attrition in the sample and to improve the estimates for GCSE attainment when compared with national data from the ONS (see table 5.2 in chapter 5). Models were fitted with unweighted data but have not been reported as these are consistent with the findings of the models in this chapter. The parental occupation schemes used are constructed using the semi-dominance approach whereby the parent with the highest ranked occupation on the scale used is selected as being representative of the family parental occupation. This overcomes potential issues of mis-measurement of stratification position by under-representing the background of a young person by measuring a parent who may have a lower grade occupation due to issues such working in a part-time job to accommodate caring for a child. Parental education measures are retained as separate for each parent. Whilst this may be seen to be inconsistent, the large amount of missing occupation data for mothers makes use of the modified dominance approach (using the occupational status of the parent with the higher grade occupation) more practical. This issue of missing data is returned to in chapter eight.

## Modelling using different GCSE Attainment Measures

In binary logistic regression models, the dependent variable is an outcome of either 1 (e.g. the respondent has gained a graded pass in a GCSE subject) or 0 (e.g. the respondent has not gained a graded pass in a GCSE subject). Logistic regression models offer an attractive alternative modelling framework for outcome measures which cannot be represented in a metric functional form as is used in linear regression (see for example Menard 2001). The first outcome is whether a young person gains any GCSE points (i.e. a graded result in a GCSE). Ninety four percent of young people have GCSE points, 4% have none and 2% are missing. This is fairly insensitive outcome measure as the majority of respondents have GCSE points. As a

consequence though, a significantly reduced likelihood of having any GCSE points demonstrates a particularly strong influence (see table 7.2).

In each of the models in this chapter, the first column contains the coefficient (or magnitude of association expressed in log odds), the second column the standard error and the third an indicator of the level of significance at the 10%, 5% and 1% thresholds. The reference category within these models is a female whose parental occupation[26] is measured as NS-SEC 1.2 in model 1 (or RGSC 1 in model 2), whose ethnicity is white, who attended a comprehensive school (till 18), whose parents own their house, who was living with both parents (at age 15/16), and whose parents do not have A-levels. The 'cohort number' term is a continuous time variable which indicates the average change per year or cohort of the YCS. The positive coefficients in models 1 and 2 indicate the rising attainment through the 1990s using the 'any GCSE points' outcome measure.

The coefficients listed with each of the dummy variables indicate the difference between the variable and the reference category. For example, with girls as the reference category, the coefficient for boys for the three models in table 7.2 ranges between -0.32 and -0.29 (depending on the parental occupational scheme used). These coefficients are relative to standard errors of 0.04, which suggests that boys are less likely to gain GCSE points than girls and this association is significant at the 0.01 or 1% level.

Gaining GCSE points is associated with measures of parental occupation, meaning that the higher the occupational status, the higher the average grade attained,

[26] Based on the modified dominance approach using the occupation of the parent with the highest occupational position (see Erikson 1984)

irrespective of whether the occupational measurement scheme is NS-SEC (see model 1), RGSC (see model 2) or CAMSIS (see model 3). Correspondingly, Indians and Other Asians (having positive coefficients) are more likely to gain points than Whites, as are those attending grammar schools than those attending comprehensive schools (till age 18). Those in rented accommodation and other housing are less likely to gain GCSE points than those living in a house which is owned by their parent(s). Those living with their mother only or in another household type are less likely to gain GCSE points than those living with both parents. Parental education (in terms of whether their father or mother has A-levels or above) is non-significant.

**Table 7.2 Logistic regression models for gaining any GCSE points score in year 11**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 0.12 | 0.06 | ** | 0.13 | 0.05 | ** | 0.07 | 0.05 | |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.28 | 0.24 | | | | | | | |
| NS-SEC = 2 | -0.47 | 0.18 | *** | | | | | | |
| NS-SEC = 3 | -0.59 | 0.19 | *** | | | | | | |
| NS-SEC = 4 | -1.01 | 0.18 | *** | | | | | | |
| NS-SEC = 5 | -0.76 | 0.20 | *** | | | | | | |
| NS-SEC = 6 | -1.08 | 0.18 | *** | | | | | | |
| NS-SEC = 7 | -1.36 | 0.19 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.08 | 0.09 | | | | | | | |
| NS-SEC 2 * Cohort | 0.09 | 0.07 | | | | | | | |
| NS-SEC 3 * Cohort | 0.06 | 0.07 | | | | | | | |
| NS-SEC 4 * Cohort | 0.03 | 0.06 | | | | | | | |
| NS-SEC 5 * Cohort | -0.04 | 0.07 | | | | | | | |
| NS-SEC 6 * Cohort | 0.02 | 0.06 | | | | | | | |
| NS-SEC 7 * Cohort | 0.07 | 0.07 | | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.34 | 0.17 | ** | | | |
| RGSC = 3.1 | | | | -0.40 | 0.18 | ** | | | |
| RGSC = 3.2 | | | | -0.85 | 0.18 | *** | | | |
| RGSC = 4 | | | | -0.97 | 0.18 | *** | | | |
| RGSC = 5 | | | | -1.18 | 0.22 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | 0.08 | 0.06 | | | | |
| RGSC = 3.1 * Cohort | | | | 0.01 | 0.06 | | | | |
| RGSC = 3.2 * Cohort | | | | 0.03 | 0.06 | | | | |
| RGSC = 4 * Cohort | | | | 0.01 | 0.06 | | | | |
| RGSC = 5 * Cohort | | | | -0.01 | 0.07 | | | | |
| Parents CAMSIS | | | | | | | 0.02 | 0.00 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.00 | 0.00 | * |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.32 | 0.04 | *** | -0.29 | 0.04 | *** | -0.29 | 0.04 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | 0.20 | 0.20 | | 0.25 | 0.20 | | 0.24 | 0.20 | |
| Indian | 0.43 | 0.19 | ** | 0.34 | 0.19 | * | 0.38 | 0.19 | ** |
| Pakistani | 0.23 | 0.22 | | 0.25 | 0.22 | | 0.30 | 0.22 | |
| Bangladeshi | 0.45 | 0.35 | | 0.40 | 0.33 | | 0.41 | 0.33 | |
| Other Asian | 0.85 | 0.37 | ** | 0.86 | 0.38 | ** | 0.89 | 0.38 | ** |
| Other Ethnicity | -0.33 | 0.25 | | -0.31 | 0.26 | | -0.32 | 0.26 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.07 | 0.05 | | -0.05 | 0.05 | | -0.04 | 0.05 | |
| Grammar | 1.32 | 0.29 | *** | 1.22 | 0.29 | *** | 1.18 | 0.29 | *** |
| Secondary Modern | -0.16 | 0.11 | | -0.13 | 0.11 | | -0.14 | 0.11 | |
| Independent | 0.23 | 0.14 | | 0.26 | 0.14 | * | 0.20 | 0.14 | |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -0.83 | 0.05 | *** | -0.79 | 0.05 | *** | -0.78 | 0.05 | *** |
| Other Housing | -0.42 | 0.14 | *** | -0.41 | 0.14 | *** | -0.37 | 0.14 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.19 | 0.06 | *** | -0.22 | 0.06 | *** | -0.24 | 0.06 | *** |
| Only lives with Father | -0.25 | 0.12 | ** | -0.25 | 0.12 | ** | -0.25 | 0.12 | ** |
| Other Household | -1.07 | 0.11 | *** | -1.10 | 0.11 | *** | -1.12 | 0.11 | *** |
| Father does not have A-level | | | | | | | | | |
| Father has A-level (mod) | 0.13 | 0.07 | * | 0.14 | 0.07 | ** | 0.09 | 0.07 | |
| Mother does not have A-level | | | | | | | | | |
| Mother has A-level (mod) | 0.09 | 0.07 | | 0.10 | 0.07 | | 0.08 | 0.07 | |
| Constant | 4.04 | 0.17 | *** | 3.80 | 0.17 | *** | 1.97 | 0.14 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| **Pseudo $R^2$** | 0.07 | | | 0.07 | | | 0.07 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

The models in table 7.3 report the likelihood of a young person gaining 5 or more GCSE passes at grades A*-C. As was identified in chapter 6, this is a common benchmark in school league tables (Payne, Cheng and Witherspoon 1996; Gayle, Berridge and Davies 2003; Leckie and Goldstein 2009). The models present family background effects upon GCSE performance independently but it is useful to consider these in combination when considering the family circumstances in which a young person lives.

The coefficients for 'Cohort number' are positive for models 1 and 2 indicating the average rise in attainment throughout the 1990s net of other characteristics. There is a strong gradient by parental occupational scheme in all three models in table 7.3. In models 1 and 2 (NS-SEC and RGSC) this is demonstrated by the negative coefficients for lower occupational status groups. In model 3 (CAMSIS) the coefficient is positive, which gives the same substantive conclusion (i.e. higher attainment for higher parental occupational status). The occupation by cohort interactions (an example being 'NS-SEC 5 * Cohort' in model 1) are also negative in models 1 and 2 which suggests that the likelihood of gaining 5 or more A*-C GCSE passes is lower for the children of those in lower occupational groups. This is also supported in model 3 (CAMSIS) but the coefficient is small because coefficient is the increase in log odds per CAMSIS point. It is nevertheless significant at the 1% level. CAMSIS is measured on a 100 point scale compared to 8 categories for NS-SEC and 6 for RGSC. These are important findings which suggest that the attainment of young people having parents in lower occupational groups has worsened during the 1990s.

This example can also be expressed using the logistic regression models in table 7.3 (by converting the coefficients into probabilities) and this may be compared with the

predicted GCSE points score measure in table 7.6. A 16 year old female (in 1990) of white ethnicity, whose parent (with the highest ranked occupation) is employed in a professional or managerial occupation (RGSC 1), who attended a comprehensive (to 16) school living with both parents who own their own home but do not have A-levels themselves would have a predicted probability of 57.9% of gaining 5 or more GCSE passes at grades A*-C and would on average gain 38.8 GCSE points (equivalent to 7 passes at grade C and 1 grade E). If, as may be more likely for a female with a parent in a professional or managerial occupation, a second female of the same background attended a grammar school and both her parents gained A-levels, she would have a predicted probability of 96.6% of gaining 5 or more GCSE passes at grades A*-C and on average have gained 58.9 GCSE points (equivalent to 8 passes at grade A with 2.9 points to spare) demonstrating the importance of school type and parental education.

Returning to the example of the first female, were this female to have instead have a parent in a lower supervisory or technical occupation and attended a comprehensive till 16, then she would have a predicted probability of 33.2% of gaining 5 or more A*-C GCSE passes and on average would have gained 29.3 GCSE points (equivalent to 5 passes at grade C and one at grade D). The points score measure may suggest borderline likelihood of gaining the 5 GCSE passes at grades A*/A-C (often necessary to remain in education post-16 on an A-level course) and the probability measure suggests that this occurs in 1 of every 3 females in such circumstances. This suggests that there may be greater dispersion within the points score measure than in the example provided (i.e. if a young person does better in a few subjects, they would have to do correspondingly worse in others for the points score measure to remain constant). The clustering in attainment in particular GCSE

subjects that is analysed in chapter 9 and this assists in decomposing the average attainment reported in the models within this chapter.

It is useful to use the models to compare how females of similar circumstances would have performed in 1999 when compared with 1990. The probability of gaining 5+ A*-C GCSE passes would have increased from 57.9% to 74.5% for the first female. This includes the overall increase in attainment for all groups plus the occupationally specific effect and the change in the contribution of this effect over time. The third female (with parents in lower supervisory or technical occupations) would have a 40.8% chance of gaining 5+ A*-C GCSE of passes in 1999 compared with 33.2% in 1990. There is a difference in the school type between these two females (comprehensive till 16 compared with comprehensive till 18) but the effect is modest and time invariant. The increase in the probability of gaining 5+ A*-C passes over the decade has risen by 7.6% for the female with parents in lower supervisory or technical occupations compared with 16.6% for the female with parents in professional or managerial occupations. There is a strong effect of parental occupation and this is clearly demonstrated using the logistic regression models in table 7.3. The examples provided also suggest the widening of this difference in GCSE attainment by parental occupational inequalities during the 1990s. Gaining 5 or more GCSE passes at grades A* to C is a clear indicator of future route and this boundary sharpens the distinction by family background measures, parental occupation being a clear indicator of likely 'success' in GCSE performance.

**Table 7.3** **Logistic regression models for gaining five or more GCSE passes at grades A*-C in year 11**

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 0.15 | 0.02 | *** | 0.16 | 0.02 | *** | -0.10 | 0.02 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.26 | 0.08 | *** | | | | | | |
| NS-SEC = 2 | -0.34 | 0.06 | *** | | | | | | |
| NS-SEC = 3 | -0.59 | 0.06 | *** | | | | | | |
| NS-SEC = 4 | -1.00 | 0.06 | *** | | | | | | |
| NS-SEC = 5 | -0.92 | 0.07 | *** | | | | | | |
| NS-SEC = 6 | -1.12 | 0.07 | *** | | | | | | |
| NS-SEC = 7 | -1.20 | 0.08 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | -0.00 | 0.03 | | | | | | | |
| NS-SEC 2 * Cohort | -0.03 | 0.02 | | | | | | | |
| NS-SEC 3 * Cohort | -0.02 | 0.02 | | | | | | | |
| NS-SEC 4 * Cohort | -0.04 | 0.02 | * | | | | | | |
| NS-SEC 5 * Cohort | -0.08 | 0.03 | *** | | | | | | |
| NS-SEC 6 * Cohort | -0.08 | 0.02 | *** | | | | | | |
| NS-SEC 7 * Cohort | -0.11 | 0.03 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.23 | 0.06 | *** | | | |
| RGSC = 3.1 | | | | -0.51 | 0.06 | *** | | | |
| RGSC = 3.2 | | | | -0.95 | 0.06 | *** | | | |
| RGSC = 4 | | | | -0.86 | 0.07 | *** | | | |
| RGSC = 5 | | | | -1.01 | 0.11 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.04 | 0.02 | ** | | | |
| RGSC = 3.1 * Cohort | | | | -0.06 | 0.02 | *** | | | |
| RGSC = 3.2 * Cohort | | | | -0.08 | 0.02 | *** | | | |
| RGSC = 4 * Cohort | | | | -0.12 | 0.02 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.14 | 0.04 | *** | | | |
| Parents CAMSIS Male Scale, Dom. | | | | | | | 0.03 | 0.00 | *** |
| Parents CAMSIS * Cohort | | | | | | | 0.00 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.38 | 0.02 | *** | -0.37 | 0.02 | *** | -0.37 | 0.02 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.32 | 0.08 | *** | -0.34 | 0.08 | *** | -0.32 | 0.08 | *** |
| Indian | 0.27 | 0.06 | *** | 0.20 | 0.06 | *** | 0.24 | 0.06 | *** |
| Pakistani | -0.28 | 0.08 | *** | -0.41 | 0.08 | *** | -0.38 | 0.08 | *** |
| Bangladeshi | 0.09 | 0.14 | | 0.08 | 0.14 | | 0.06 | 0.15 | |
| Other Asian | 0.65 | 0.10 | *** | 0.61 | 0.10 | *** | 0.63 | 0.10 | *** |
| Other Ethnicity | 0.04 | 0.10 | | -0.01 | 0.10 | | -0.03 | 0.10 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.09 | 0.02 | *** | -0.10 | 0.02 | *** | -0.09 | 0.02 | *** |
| Grammar | 2.37 | 0.08 | *** | 2.38 | 0.08 | *** | 2.35 | 0.08 | *** |
| Secondary Modern | -0.78 | 0.05 | *** | -0.80 | 0.05 | *** | -0.81 | 0.05 | *** |
| Independent | 1.61 | 0.05 | *** | 1.60 | 0.05 | *** | 1.56 | 0.05 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -0.72 | 0.03 | *** | -0.73 | 0.03 | *** | -0.70 | 0.03 | *** |
| Other Housing | -0.23 | 0.07 | *** | -0.28 | 0.07 | *** | -0.24 | 0.07 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.07 | 0.03 | *** | -0.09 | 0.03 | *** | -0.09 | 0.03 | *** |
| Only lives with Father | -0.24 | 0.05 | *** | -0.20 | 0.05 | *** | -0.19 | 0.05 | *** |
| Other Household | -0.90 | 0.08 | *** | -0.89 | 0.08 | *** | -0.89 | 0.08 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 0.31 | 0.02 | *** | 0.33 | 0.02 | *** | 0.27 | 0.02 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 0.33 | 0.02 | *** | 0.34 | 0.02 | *** | 0.30 | 0.02 | *** |
| Constant | 0.32 | 0.05 | *** | 0.17 | 0.06 | *** | -1.75 | 0.07 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| **Pseudo $R^2$** | 0.15 | | | 0.15 | | | 0.15 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

GCSE attainment can also be expressed as an ordinal relationship by comparing whether a young person gains 5 or more, 1-4 or 0 A*-C GCSE passes. The associations with family background factors, identified within table 2, are consistent when ordinal logistic regression models are fitted for NS-SEC, RGSC and CAMSIS (see appendix table A7.1).

Having considered threshold measures of attainment, in table 7.4 the dependent variable in the models is the number of A*-C passes a young person achieves, fitted using a linear (Ordinary Least Squares) regression. There are a large number of young people who do not gain any A*-C passes and a sizeable proportion who gain 1-4 (demonstrated by figure 6.3 in chapter 6). All factors included in the models are highly significant and the proportion of variance explained ($R^2$) by these family background factors is consistently above 0.25, implying that over a quarter of the variation in GCSE attainment can be explained by the family background factors included in the models (see Treiman 2009: Ch. 6). Given other unmeasured potential sources of variation in the life circumstances of young people, this is remarkably high, and contrary to the detraditionalisation arguments made by Giddens (1990) and Beck and Beck-Gernsheim (2002).

The principle focus of analysis has so far demonstrated the consistency in the influence of parental occupation upon GCSE attainment, irrespective of the functional form of the outcome or the measure of parental occupation used. However, other stratification measures consistently influence attainment. It is clear that (irrespective of parental occupation, parental education and family structure) being a boy, being black or Pakistani, attending a comprehensive (till16) or secondary modern school and living in rented or other accommodation were

associated with lower GCSE attainment. Other things being equal Indian and Other Asian young people and those attending Independent or Grammar schools do better than the reference group (Nuttall et al. 1989; Mortimore, Sammons and Thomas 1994; Sammons 1995; Croxford 2000).

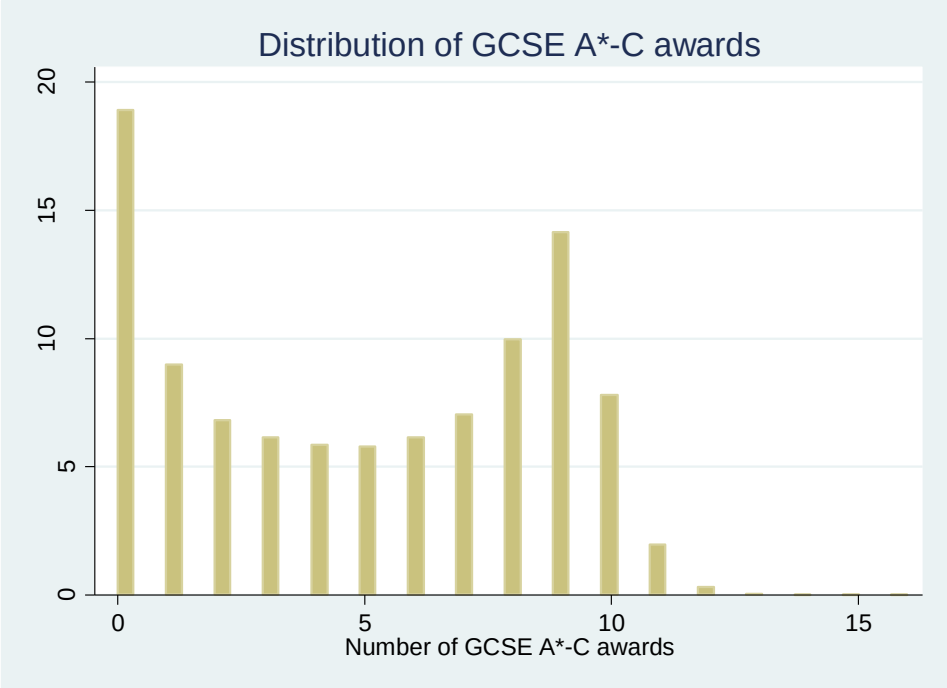



**Figure 7.1 Distribution of GCSE A*-C awards, YCS cohorts 5-10 (weighted)**

The number of A*- C passes is a count variable, with distribution shown in figure 7.1, therefore poisson regression models for count data were fitted (see Cameron and Trivedi 1998). However, the model fails the goodness of fit statistic (at the 0.01 level) for the poisson distribution (due to a large group of young people who gain a higher number of A*-C passes) and a negative binomial distribution may be a better fit (confirmed by the alpha statistic reported in the negative binomial regression models). In both instances, regardless of the assumptions made about the form of the outcome (GCSE attainment) the relationship with the predictor variables are consistent across the models (see appendix tables A7.2 and A7.3 for details of these models).

## Table 7.4 OLS Regression models, number of A*-C GCSE passes in year 11

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 0.30 | 0.02 | *** | 0.33 | 0.02 | *** | -0.20 | 0.03 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.41 | 0.11 | *** | | | | | | |
| NS-SEC = 2 | -0.61 | 0.08 | *** | | | | | | |
| NS-SEC = 3 | -1.06 | 0.09 | *** | | | | | | |
| NS-SEC = 4 | -1.66 | 0.09 | *** | | | | | | |
| NS-SEC = 5 | -1.58 | 0.10 | *** | | | | | | |
| NS-SEC = 6 | -1.85 | 0.09 | *** | | | | | | |
| NS-SEC = 7 | -1.99 | 0.10 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.03 | 0.04 | | | | | | | |
| NS-SEC 2 * Cohort | -0.02 | 0.03 | | | | | | | |
| NS-SEC 3 * Cohort | -0.02 | 0.03 | | | | | | | |
| NS-SEC 4 * Cohort | -0.09 | 0.03 | *** | | | | | | |
| NS-SEC 5 * Cohort | -0.18 | 0.04 | *** | | | | | | |
| NS-SEC 6 * Cohort | -0.18 | 0.03 | *** | | | | | | |
| NS-SEC 7 * Cohort | -0.23 | 0.03 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.43 | 0.08 | *** | | | |
| RGSC = 3.1 | | | | -0.91 | 0.09 | *** | | | |
| RGSC = 3.2 | | | | -1.56 | 0.09 | *** | | | |
| RGSC = 4 | | | | -1.48 | 0.10 | *** | | | |
| RGSC = 5 | | | | -1.62 | 0.14 | *** | | | |
| RGSC = 1 * Cohort | | | | | | | | | |
| RGSC = 2 * Cohort | | | | -0.03 | 0.03 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.10 | 0.03 | *** | | | |
| RGSC = 3.2 * Cohort | | | | -0.19 | 0.03 | *** | | | |
| RGSC = 4 * Cohort | | | | -0.24 | 0.03 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.27 | 0.05 | *** | | | |
| Parents CAMSIS | | | | | | | 0.04 | 0.00 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.01 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.70 | 0.02 | *** | -0.70 | 0.02 | *** | -0.69 | 0.02 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.50 | 0.11 | *** | -0.51 | 0.10 | *** | -0.49 | 0.10 | *** |
| Indian | 0.42 | 0.08 | *** | 0.31 | 0.08 | *** | 0.36 | 0.08 | *** |
| Pakistani | -0.41 | 0.11 | *** | -0.59 | 0.11 | *** | -0.52 | 0.11 | *** |
| Bangladeshi | 0.25 | 0.19 | | 0.23 | 0.19 | | 0.19 | 0.19 | |
| Other Asian | 0.94 | 0.14 | *** | 0.89 | 0.13 | *** | 0.91 | 0.13 | *** |
| Other Ethnicity | -0.10 | 0.15 | | -0.17 | 0.15 | | -0.20 | 0.15 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.20 | 0.03 | *** | -0.20 | 0.03 | *** | -0.19 | 0.03 | *** |
| Grammar | 3.20 | 0.05 | *** | 3.25 | 0.05 | *** | 3.19 | 0.05 | *** |
| Secondary Modern | -1.23 | 0.06 | *** | -1.25 | 0.06 | *** | -1.26 | 0.06 | *** |
| Independent | 2.63 | 0.05 | *** | 2.64 | 0.05 | *** | 2.54 | 0.05 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -1.18 | 0.03 | *** | -1.19 | 0.04 | *** | -1.14 | 0.04 | *** |
| Other Housing | -0.41 | 0.10 | *** | -0.48 | 0.10 | *** | -0.41 | 0.10 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.12 | 0.04 | *** | -0.15 | 0.04 | *** | -0.15 | 0.04 | *** |
| Only lives with Father | -0.41 | 0.07 | *** | -0.37 | 0.08 | *** | -0.35 | 0.08 | *** |
| Other Household | -1.29 | 0.09 | *** | -1.28 | 0.09 | *** | -1.29 | 0.09 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 0.58 | 0.03 | *** | 0.60 | 0.03 | *** | 0.50 | 0.03 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 0.59 | 0.04 | *** | 0.60 | 0.04 | *** | 0.53 | 0.03 | *** |
| Constant | 5.08 | 0.07 | *** | 4.83 | 0.08 | *** | 1.68 | 0.09 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| **$R^2$** | 0.26 | | | 0.25 | | | 0.25 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

Table 7.5 reports the relative contribution to deviance by the independent variables within the model by separating these into the blocks of dummy variables used in the model specification. The largest contribution to explained variation is made by the measure of school type (as might be expected from the polarisation in attainment between those attending selective schools and those who do not). Parental occupation (measured using NS-SEC including change during the 1990s and interaction with this change by occupational group) and parental education also contribute heavily. Family structure has a much lower impact. What is immediately clear when compared with the descriptive statistics in chapter 6 is that the influence of family background remains key net of the other relevant sources of variation.

**Table 7.5 Goodness of fit summary (based on model 1 in table 7.6)**

| Block | Independent Variables | Deviance Explained (F) | Block D.F. | $R^2$ | Change in R2 |
|---|---|---|---|---|---|
| 1 | NS-SEC, NS-SEC & Cohort interactions, Cohort | 827.92 | 15 | 0.156 | |
| 2 | Sex | 578.50 | 1 | 0.163 | 0.007 |
| 3 | Ethnicity | 42.30 | 6 | 0.166 | 0.003 |
| 4 | School Type | 1726.58 | 4 | 0.236 | 0.070 |
| 5 | Housing Type | 844.79 | 2 | 0.257 | 0.021 |
| 6 | Family Structure | 95.72 | 3 | 0.260 | 0.003 |
| 7 | Father's Education | 579.45 | 1 | 0.267 | 0.007 |
| 8 | Mother's Education | 314.92 | 1 | 0.271 | 0.004 |

NB: Design degrees of freedom is 88973 for the model with all variables significantly contributing to the variance

Table 7.6 reports regression models on GCSE total points score (as opposed to the number of passes in table 7.4). This measure is a cumulative points score measure based upon grade attained, where a GCSE at grade G is equivalent to 1 point and grade A*/A is equivalent to 7 points (grade A* was introduced in 1994). There are a small number of outliers who gain a very high number of GCSE passes. These have been truncated within the models in table 7.6, where the points have been limited to 84 (12 passes at grade A*/A). The effect does not significant change the relationships with the dependent variables and reduces the influence of these outliers upon the mean (see table 7.9).[27] From this point onwards, all regression models upon GCSE points score use the truncated measure due to the improvement of estimation upon the mean points score attained.

From models 1 and 2 in table 7.6, it appears that the gap has widened between young people with parents in higher professional occupations (NS-SEC 1.2) and those with lower supervisory and technical (NS-SEC 5), semi-routine (NS-SEC 6) or routine occupations (NS-SEC 7). The effect of the father or mother of a young person having A-levels (or above) is associated with higher GCSE points score (net of parental occupation and other controls, see table 7.6) and this is consistent with other literature (Blau 1999; Rice 1999; Ermisch and Francesconi 2001a; ONS 2002b; Bradley and Taylor 2004). Living in a mother only, father only or in another household forms carries a penalty in GCSE attainment net of other differences compared to living with both parents.

[27] For consistency, the non-limited points score measure has been modelled. The magnitude and direction of the coefficients are similar (see appendix table A7.4).

**Table 7.6 OLS Regression models, GCSE points score (truncated) in year 11**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 1.91 | 0.11 | *** | 2.05 | 0.11 | *** | 0.56 | 0.15 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -2.08 | 0.53 | *** | | | | | | |
| NS-SEC = 2 | -3.10 | 0.40 | *** | | | | | | |
| NS-SEC = 3 | -5.32 | 0.43 | *** | | | | | | |
| NS-SEC = 4 | -8.87 | 0.44 | *** | | | | | | |
| NS-SEC = 5 | -8.34 | 0.51 | *** | | | | | | |
| NS-SEC = 6 | -9.98 | 0.46 | *** | | | | | | |
| NS-SEC = 7 | -11.30 | 0.51 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.20 | 0.17 | | | | | | | |
| NS-SEC 2 * Cohort | 0.01 | 0.13 | | | | | | | |
| NS-SEC 3 * Cohort | 0.10 | 0.14 | | | | | | | |
| NS-SEC 4 * Cohort | -0.05 | 0.14 | | | | | | | |
| NS-SEC 5 * Cohort | -0.38 | 0.18 | ** | | | | | | |
| NS-SEC 6 * Cohort | -0.32 | 0.15 | ** | | | | | | |
| NS-SEC 7 * Cohort | -0.51 | 0.17 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -2.44 | 0.39 | *** | | | |
| RGSC = 3.1 | | | | -4.67 | 0.42 | *** | | | |
| RGSC = 3.2 | | | | -8.73 | 0.43 | *** | | | |
| RGSC = 4 | | | | -8.75 | 0.51 | *** | | | |
| RGSC = 5 | | | | -10.16 | 0.75 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.03 | 0.13 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.29 | 0.14 | ** | | | |
| RGSC = 3.2 * Cohort | | | | -0.40 | 0.14 | *** | | | |
| RGSC = 4 * Cohort | | | | -0.66 | 0.16 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.80 | 0.26 | *** | | | |
| Parents CAMSIS Male Scale, Dom. | | | | | | | 0.24 | 0.01 | *** |
| Parents CAMSIS * Cohort | | | | | | | 0.02 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -3.38 | 0.12 | *** | -3.32 | 0.12 | *** | -3.30 | 0.12 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -2.44 | 0.54 | *** | -2.44 | 0.54 | *** | -2.33 | 0.53 | *** |
| Indian | 2.26 | 0.39 | *** | 1.79 | 0.40 | *** | 2.00 | 0.40 | *** |
| Pakistani | -2.02 | 0.57 | *** | -2.71 | 0.57 | *** | -2.44 | 0.57 | *** |
| Bangladeshi | 0.99 | 1.00 | | 0.89 | 0.98 | | 0.65 | 0.99 | |
| Other Asian | 4.31 | 0.70 | *** | 4.18 | 0.68 | *** | 4.31 | 0.68 | *** |
| Other Ethnicity | -0.48 | 0.78 | | -0.86 | 0.77 | | -1.01 | 0.77 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -1.11 | 0.14 | *** | -1.10 | 0.14 | *** | -1.04 | 0.14 | *** |
| Grammar | 14.23 | 0.25 | *** | 14.50 | 0.26 | *** | 14.20 | 0.26 | *** |
| Secondary Modern | -6.14 | 0.31 | *** | -6.22 | 0.30 | *** | -6.27 | 0.30 | *** |
| Independent | 11.87 | 0.26 | *** | 11.90 | 0.26 | *** | 11.42 | 0.26 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -6.53 | 0.19 | *** | -6.42 | 0.19 | *** | -6.25 | 0.19 | *** |
| Other Housing | -2.15 | 0.49 | *** | -2.38 | 0.50 | *** | -2.09 | 0.50 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.86 | 0.20 | *** | -1.01 | 0.20 | *** | -1.05 | 0.20 | *** |
| Only lives with Father | -2.17 | 0.38 | *** | -2.00 | 0.39 | *** | -1.96 | 0.39 | *** |
| Other Household | -8.43 | 0.52 | *** | -8.41 | 0.52 | *** | -8.48 | 0.52 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 2.88 | 0.17 | *** | 2.94 | 0.17 | *** | 2.44 | 0.16 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 3.01 | 0.17 | *** | 3.03 | 0.17 | *** | 2.66 | 0.17 | *** |
| Constant | 38.77 | 0.36 | *** | 37.66 | 0.38 | *** | 19.66 | 0.47 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| **$R^2$** | 0.27 | | | 0.26 | | | 0.27 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

Figures 7.2, 7.3, 7.4 and 7.5 expand upon the differences associated with parental occupational measures in the models in table 7.6. These figures are the plotted coefficients from the models in table 7.6 using quasi-variances (see Gayle and Lambert 2007). Quasi-variances are used to overcome the reference category problem so that comparison between categories can be made beyond direct comparison with the reference category. The reference category problem occurs when there is no confidence interval fitted around the reference category when using dummy variables in conventional regression models (Gayle and Lambert 2007). Through publishing quasi-variances, it is possible to display comparison intervals around all of the categorical measures used in a set (i.e. both the reference category and all the dummy variables).

Comparison intervals constructed using quasi-variances also make it easier to see an apparent pattern that young people with the least advantaged parental occupational classifications fare similarly in GCSE attainment. The data suggests that there is a divide between non-manual and manual occupations with greater gradient in the non-manual occupations and greater similarity in the manual occupations. Moreover, the gap in attainment associated with parental occupational measures appears to be widening through the 1990s, as shown in table 7.6 by the significant interaction terms for the relationship between the occupational measures and cohort, which are expected graphically in figures 7.3 and 7.5, which short the interaction effects between the cohort and class measures.

Figure 7.3 however shows that when quasi-variances are plotted for the NS-SEC by cohort interactions the trend is supported but that the comparison intervals overlap. This indicates that whilst the trend is for the gap to have widened, it is not

significantly different from the reference category (NS-SEC 1.2). Similarly this is demonstrated in figure 7.5 using RGSC. It is possible that the NS-SEC trends could be a function of the measurement schema, for instance due to increased proportions of children over time having parents in the expanding higher levels of the NS-SEC scheme. Alternatively, it is possible that the RGSC is not sufficiently capable of measuring parent's occupational differences and therefore misses an important trend in background effects. This sensitivity analysis and use of quasi-variance estimates does however suggest that there is persisting stratification in GCSE attainment by parental occupation and that this has not reduced during the 1990s.

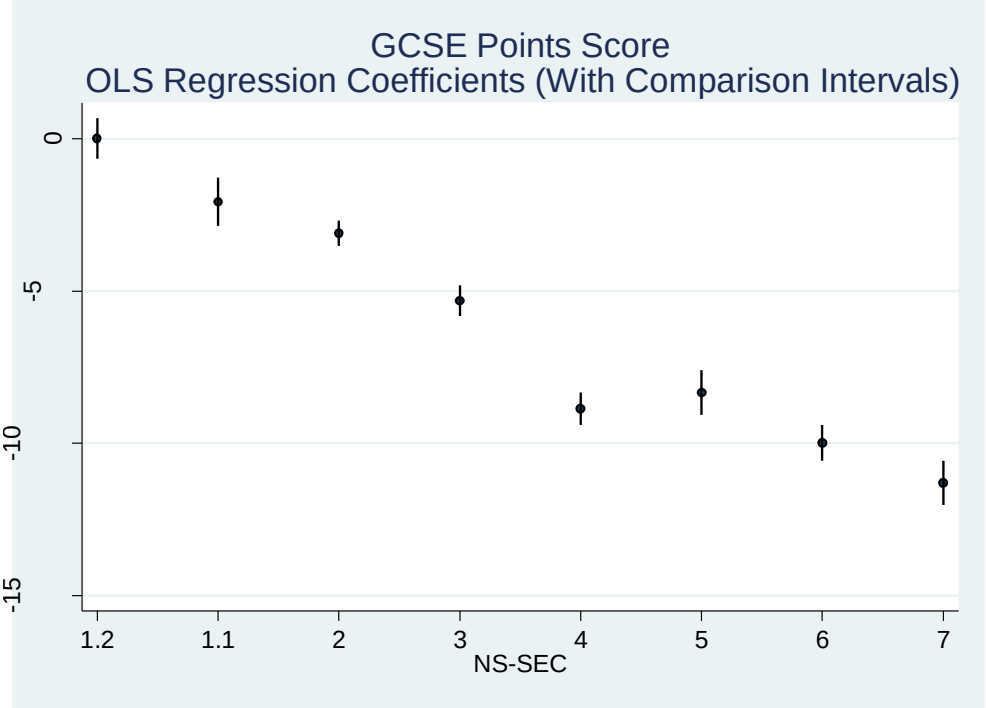



**Figure 7.2** **Estimates of GCSE Points Score, NS-SEC main effects**

Other variables: NS-SEC by cohort interactions, gender, ethnicity, school type, housing, household type, parental education (see model 1 in table 7.6)
Source: SN5765 & SN3532, n=88974, $R^2$= 0.27, 1990s YCS Cohorts.

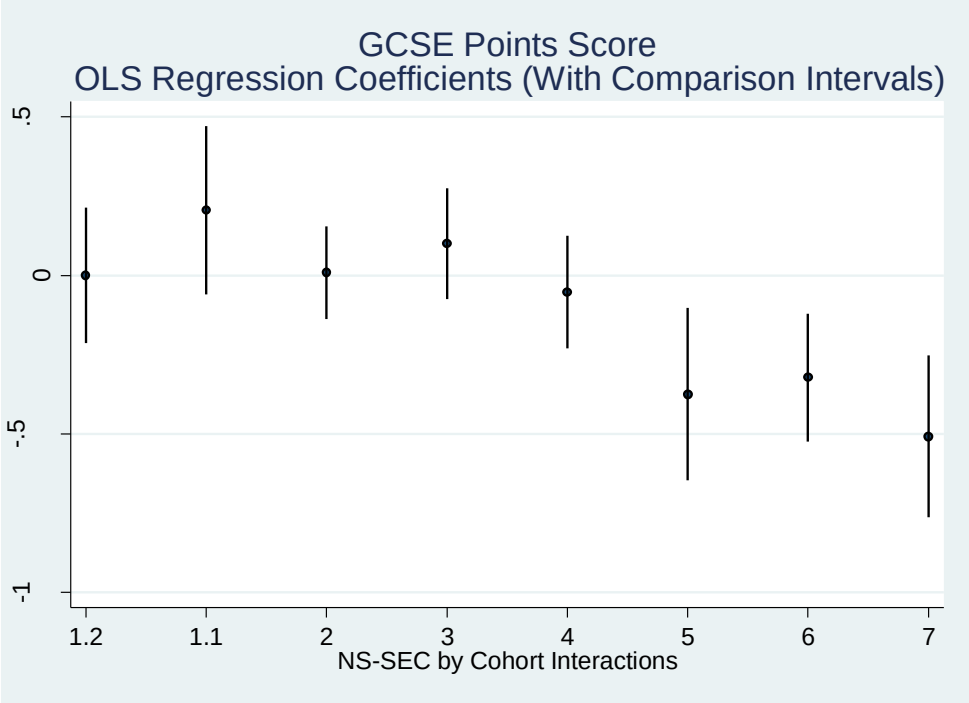



**Figure 7.3** **Estimates of GCSE Points Score, NS-SEC Time varying effects**

Other variables: NS-SEC, gender, ethnicity, school type, housing, household type, parental education (see model 1 in table 7.6)
Source: SN5765 & SN3532, n=88974, $R^2$= 0.27, 1990s YCS Cohorts.

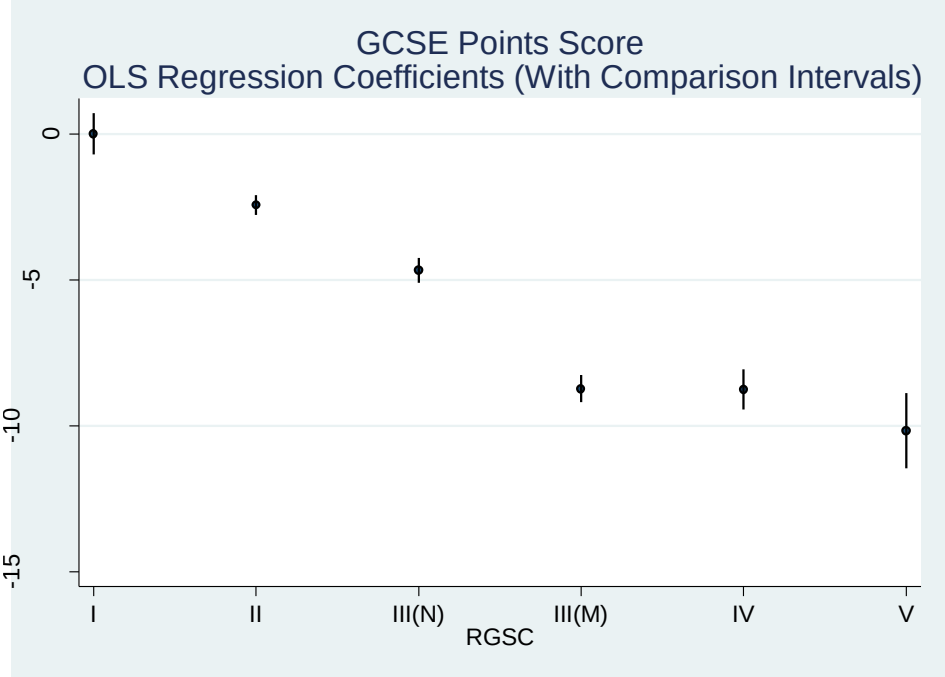



**Figure 7.4** **Estimates of GCSE Points Score, RGSC main effects**

Other variables: RGSC by cohort interactions, gender, ethnicity, school type, housing, household type, parental education (see model 2 in table 7.6)
Source: SN5765 & SN3532, n=89115, $R^2$=0.26; 1990s YCS Cohorts.

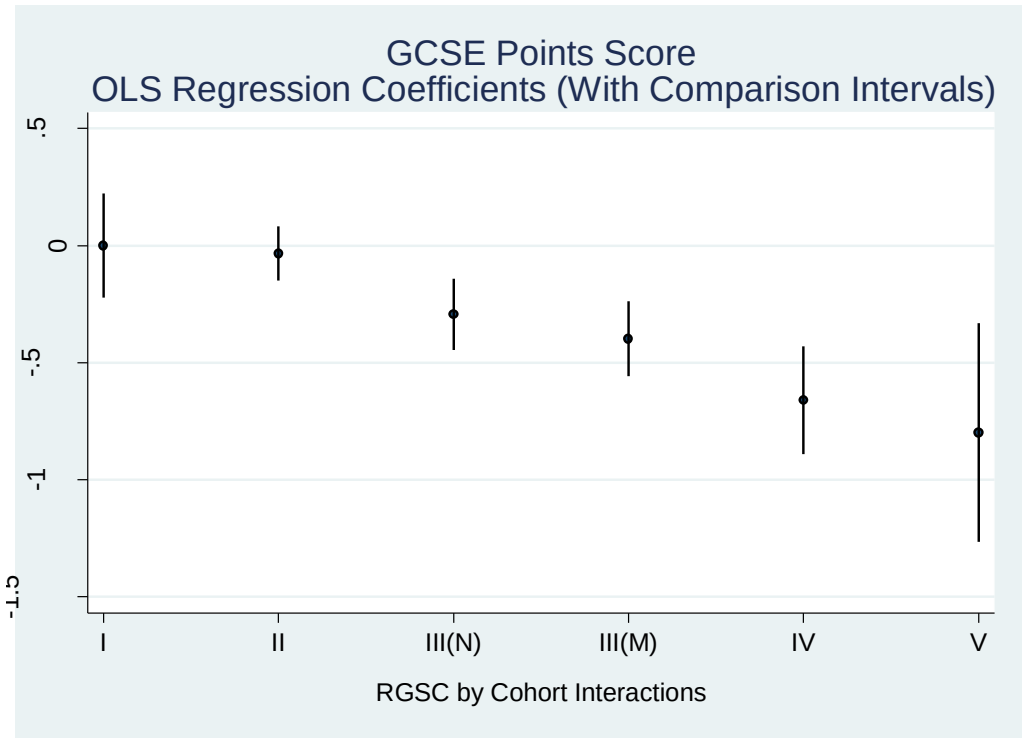



**Figure 7.5** **Estimates of GCSE Points Score, RGSC Time varying effects**

Other variables: RGSC, gender, ethnicity, school type, housing, household type, parental education (see model 2 in table 7.6)
Source: SN5765 & SN3532, n=89115, $R^2$=0.26; 1990s YCS Cohorts.

To verify the time varying coefficients between parental occupation measures and cohort within the models in table 7.6, individual models have been fitted for each cohort of the YCS during the 1990s (1990, 1992, 1993, 1995, 1997 and 1999). For each cohort, there is an increase in the outcome (rising grades). Table 7.7 contains three models of GCSE points score in 1990 including NS-SEC, RGSC & CAMSIS. Table 7.8 contains the same models for young people in 1999[28]. During this period, the constant has risen from 38 to 48 points (using NS-SEC) for the same reference category. This is equivalent to a grade per GCSE taken (assuming a student had taken 10 GCSEs).

[28] For details of the same models for 1992, 1993, 1995, 1997, please tables A7.5 - A7.8 in the appendix.

**Table 7.7 OLS Regression models, GCSE points score (truncated) in year 11 in 1990**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.73 | 0.72 | | | | | | | |
| NS-SEC = 2 | -2.47 | 0.55 | *** | | | | | | |
| NS-SEC = 3 | -5.37 | 0.60 | *** | | | | | | |
| NS-SEC = 4 | -8.90 | 0.62 | *** | | | | | | |
| NS-SEC = 5 | -8.86 | 0.72 | *** | | | | | | |
| NS-SEC = 6 | -10.99 | 0.66 | *** | | | | | | |
| NS-SEC = 7 | -13.70 | 0.71 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -2.43 | 0.54 | *** | | | |
| RGSC = 3.1 | | | | -5.90 | 0.59 | *** | | | |
| RGSC = 3.2 | | | | -10.58 | 0.62 | *** | | | |
| RGSC = 4 | | | | -11.47 | 0.70 | *** | | | |
| RGSC = 5 | | | | -13.13 | 0.93 | *** | | | |
| Parents CAMSIS | | | | | | | 0.34 | 0.01 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -3.42 | 0.27 | *** | -3.32 | 0.27 | *** | -3.33 | 0.27 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -4.85 | 1.51 | *** | -4.27 | 1.53 | *** | -3.91 | 1.46 | *** |
| Indian | 0.95 | 1.08 | | 0.16 | 1.09 | | 0.79 | 1.12 | |
| Pakistani | 0.41 | 1.39 | | -0.15 | 1.39 | | 0.00 | 1.46 | |
| Bangladeshi | 1.40 | 2.35 | | 0.49 | 2.35 | | 0.60 | 2.31 | |
| Other Asian | 3.64 | 1.74 | ** | 3.53 | 1.69 | ** | 3.50 | 1.64 | ** |
| Other Ethnicity | 0.00 | 0.00 | | 0.00 | 0.00 | | 0.00 | 0.00 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -1.18 | 0.32 | *** | -1.15 | 0.32 | *** | -1.00 | 0.32 | *** |
| Grammar | 14.54 | 0.74 | *** | 14.37 | 0.75 | *** | 14.38 | 0.75 | *** |
| Secondary Modern | -5.79 | 0.66 | *** | -5.84 | 0.66 | *** | -6.04 | 0.65 | *** |
| Independent | 14.49 | 0.47 | *** | 14.15 | 0.47 | *** | 13.57 | 0.47 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -7.57 | 0.39 | *** | -7.75 | 0.39 | *** | -7.35 | 0.39 | *** |
| Other Housing | -2.90 | 0.81 | *** | -3.41 | 0.81 | *** | -2.96 | 0.82 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | 0.03 | 0.47 | | -0.23 | 0.47 | | -0.34 | 0.47 | |
| Only lives with Father | -2.70 | 0.93 | *** | -2.34 | 0.93 | ** | -2.47 | 0.93 | *** |
| Other Household | -5.37 | 0.95 | *** | -5.03 | 0.94 | *** | -5.15 | 0.93 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 3.14 | 0.40 | *** | 3.51 | 0.40 | *** | 2.83 | 0.38 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 3.14 | 0.43 | *** | 3.14 | 0.43 | *** | 2.77 | 0.42 | *** |
| Constant | 37.82 | 0.56 | *** | 37.47 | 0.57 | *** | 13.07 | 0.69 | *** |
| **Observations** | 12632 | | | 12630 | | | 12632 | | |
| **$R^2$** | 0.31 | | | 0.31 | | | 0.32 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

**Table 7.8 OLS Regression models, GCSE points score (truncated) in year 11 in 1999**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.58 | 0.71 | | | | | | | |
| NS-SEC = 2 | -2.97 | 0.56 | *** | | | | | | |
| NS-SEC = 3 | -4.34 | 0.64 | *** | | | | | | |
| NS-SEC = 4 | -7.74 | 0.63 | *** | | | | | | |
| NS-SEC = 5 | -8.43 | 0.87 | *** | | | | | | |
| NS-SEC = 6 | -9.81 | 0.71 | *** | | | | | | |
| NS-SEC = 7 | -11.83 | 0.84 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -2.00 | 0.53 | *** | | | |
| RGSC = 3.1 | | | | -5.17 | 0.63 | *** | | | |
| RGSC = 3.2 | | | | -8.70 | 0.64 | *** | | | |
| RGSC = 4 | | | | -9.07 | 0.75 | *** | | | |
| RGSC = 5 | | | | -10.31 | 1.51 | *** | | | |
| Parents CAMSIS | | | | | | | 0.31 | 0.01 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -4.18 | 0.32 | *** | -4.10 | 0.32 | *** | -4.02 | 0.32 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.99 | 1.26 | | 0.01 | 1.21 | | -0.17 | 1.21 | |
| Indian | 4.31 | 0.97 | *** | 3.70 | 0.95 | *** | 3.96 | 0.97 | *** |
| Pakistani | -2.43 | 1.27 | * | -3.56 | 1.29 | *** | -3.68 | 1.29 | *** |
| Bangladeshi | 3.40 | 3.10 | | 2.62 | 2.65 | | 2.11 | 2.61 | |
| Other Asian | 5.97 | 1.42 | *** | 5.41 | 1.46 | *** | 5.82 | 1.39 | *** |
| Other Ethnicity | -0.90 | 1.58 | | -1.43 | 1.62 | | -1.42 | 1.59 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.99 | 0.37 | *** | -0.91 | 0.37 | ** | -0.91 | 0.36 | ** |
| Grammar | 14.01 | 0.62 | *** | 14.37 | 0.62 | *** | 13.87 | 0.61 | *** |
| Secondary Modern | -5.36 | 0.86 | *** | -5.54 | 0.88 | *** | -5.48 | 0.86 | *** |
| Independent | 9.88 | 0.58 | *** | 9.97 | 0.59 | *** | 9.09 | 0.59 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -6.88 | 0.48 | *** | -7.18 | 0.47 | *** | -6.86 | 0.46 | *** |
| Other Housing | -0.92 | 1.98 | | -1.37 | 2.05 | | -1.56 | 2.00 | |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.50 | 0.47 | | -0.71 | 0.47 | | -0.65 | 0.46 | |
| Only lives with Father | -2.07 | 0.96 | ** | -2.39 | 0.96 | ** | -1.95 | 0.95 | ** |
| Other Household | -7.24 | 1.44 | *** | -7.51 | 1.43 | *** | -7.42 | 1.44 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 3.07 | 0.41 | *** | 3.47 | 0.41 | *** | 2.83 | 0.41 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 2.87 | 0.41 | *** | 2.93 | 0.41 | *** | 2.51 | 0.40 | *** |
| Constant | 48.18 | 0.59 | *** | 47.32 | 0.60 | *** | 25.55 | 0.86 | *** |
| **Observations** | 11902 | | | 12036 | | | 12036 | | |
| **$R^2$** | 0.26 | | | 0.26 | | | 0.27 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

**Table 7.9 National average, weighting and standardised scores**

| Cohort | Mean (National)[29] | Mean GCSE Points Score unweighted | Mean GCSE Points Score weighted | n GCSE Points Score | Mean Standardised GCSE Points Score unweighted | Standard Deviation Standardised GCSE Points Score unweighted | Mean Standardised GCSE Points Score weighted | Standard Deviation Standardised GCSE Points Score weighted |
|---|---|---|---|---|---|---|---|---|
| **1990** | | 32.89 | 29.26 | 14269 | 0.00 | 1.00 | 0.21 | 1.00 |
| **1992** | | 34.09 | 33.82 | 24350 | 0.00 | 1.00 | 0.02 | 0.99 |
| **1993** | | 37.48 | 34.14 | 17769 | 0.00 | 1.00 | 0.18 | 0.97 |
| **1995** | 34.80 | 42.17 | 37.10 | 15761 | 0.00 | 1.00 | 0.27 | 0.95 |
| **1997** | 35.90 | 40.62 | 36.57 | 14662 | 0.00 | 1.00 | 0.21 | 0.96 |
| **1999** | 38.10 | 46.21 | 40.78 | 13101 | 0.00 | 1.00 | 0.31 | 0.92 |

Whilst the previous models prior to this point have included an indicator of the cohort number (i.e. the progress of time) this may not entirely control for the increase in attainment over time. Therefore, to control for this, a standardised measure has been constructed (see table 7.9). This standardised measure was constructed for the GCSE points score measure to fix the outcome around the mean and standard deviation for each cohort (the z-score). The distance between a young person's score and the mean score in absolute terms is given in points. If we divide by the standard deviation in that year, the result is the standardised difference in terms of standard deviations from the mean. To use a fictitious example, if a young person gained 35 points in a hypothetical year where the mean was 32 points and the standard deviation 3 points, they would be 3 points above the mean (35-32=3) or 1 standard deviation above the mean (3/3=1). The corresponding standardised score for each young person gives their position relative to the mean and standard deviation within each cohort.

[29] See http://www.bstubbs.co.uk/5a-c.htm#table2 - *Source: Department for Education & Employment, London* accessed 15th July 2009

This is important because, table 7.9 shows that without standardising or using other controls, the same GCSE points score in 1990 and 1999 would reflect a different position within the distribution of attainment within each cohort. For example, using the national mean above, a young person gaining 37 points would be 2.2 points above the 1995 average of 34.8 points but 1.1 point below the 1999 average of 38.1 points. This young person would be of above average attainment in 1995 but below average attainment in 1999.

The standardisation was based on unweighted averages, and applying the weighting to the standardised scores increases the means unevenly (as can be seen by comparing the middle two columns with the final two columns in table 7.9). Nevertheless, the estimates produced using the standardised outcomes should provide an indication of whether the influence of the predictors of educational attainment has changed during the 1990s (net of the overall rise in attainment) and this is a valuable exercise. By standardising a score relative to the cohort average and standard deviation, this forms a control for increasing attainment over time.

## Table 7.10 OLS Regression models, GCSE points score (truncated) in year 11, standardised within cohort

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | -0.01 | 0.01 | | 0.00 | 0.01 | | -0.06 | 0.01 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.13 | 0.03 | *** | | | | | | |
| NS-SEC = 2 | -0.19 | 0.02 | *** | | | | | | |
| NS-SEC = 3 | -0.31 | 0.02 | *** | | | | | | |
| NS-SEC = 4 | -0.48 | 0.02 | *** | | | | | | |
| NS-SEC = 5 | -0.49 | 0.03 | *** | | | | | | |
| NS-SEC = 6 | -0.58 | 0.03 | *** | | | | | | |
| NS-SEC = 7 | -0.65 | 0.03 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.01 | 0.01 | | | | | | | |
| NS-SEC 2 * Cohort | 0.01 | 0.01 | | | | | | | |
| NS-SEC 3 * Cohort | 0.01 | 0.01 | | | | | | | |
| NS-SEC 4 * Cohort | 0.00 | 0.01 | | | | | | | |
| NS-SEC 5 * Cohort | -0.01 | 0.01 | | | | | | | |
| NS-SEC 6 * Cohort | -0.01 | 0.01 | | | | | | | |
| NS-SEC 7 * Cohort | -0.01 | 0.01 | | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.15 | 0.02 | *** | | | |
| RGSC = 3.1 | | | | -0.28 | 0.02 | *** | | | |
| RGSC = 3.2 | | | | -0.51 | 0.02 | *** | | | |
| RGSC = 4 | | | | -0.52 | 0.03 | *** | | | |
| RGSC = 5 | | | | -0.60 | 0.04 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | 0.00 | 0.01 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.01 | 0.01 | | | | |
| RGSC = 3.2 * Cohort | | | | -0.01 | 0.01 | | | | |
| RGSC = 4 * Cohort | | | | -0.02 | 0.01 | ** | | | |
| RGSC = 5 * Cohort | | | | -0.02 | 0.01 | * | | | |
| Parents CAMSIS | | | | | | | 0.01 | 0.00 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.00 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.19 | 0.01 | *** | -0.18 | 0.01 | *** | -0.18 | 0.01 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.14 | 0.03 | *** | -0.14 | 0.03 | *** | -0.13 | 0.03 | *** |
| Indian | 0.12 | 0.02 | *** | 0.10 | 0.02 | *** | 0.11 | 0.02 | *** |
| Pakistani | -0.12 | 0.03 | *** | -0.15 | 0.03 | *** | -0.14 | 0.03 | *** |
| Bangladeshi | 0.05 | 0.06 | | 0.04 | 0.05 | | 0.03 | 0.06 | |
| Other Asian | 0.24 | 0.04 | *** | 0.23 | 0.04 | *** | 0.24 | 0.04 | *** |
| Other Ethnicity | -0.03 | 0.04 | | -0.05 | 0.04 | | -0.05 | 0.04 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.06 | 0.01 | *** | -0.06 | 0.01 | *** | -0.06 | 0.01 | *** |
| Grammar | 0.79 | 0.01 | *** | 0.80 | 0.01 | *** | 0.79 | 0.01 | *** |
| Secondary Modern | -0.34 | 0.02 | *** | -0.35 | 0.02 | *** | -0.35 | 0.02 | *** |
| Independent | 0.66 | 0.01 | *** | 0.66 | 0.01 | *** | 0.64 | 0.01 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -0.37 | 0.01 | *** | -0.36 | 0.01 | *** | -0.35 | 0.01 | *** |
| Other Housing | -0.12 | 0.03 | *** | -0.13 | 0.03 | *** | -0.12 | 0.03 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.05 | 0.01 | *** | -0.05 | 0.01 | *** | -0.06 | 0.01 | *** |
| Only lives with Father | -0.12 | 0.02 | *** | -0.11 | 0.02 | *** | -0.11 | 0.02 | *** |
| Other Household | -0.43 | 0.03 | *** | -0.43 | 0.03 | *** | -0.43 | 0.03 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 0.16 | 0.01 | *** | 0.16 | 0.01 | *** | 0.14 | 0.01 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 0.17 | 0.01 | *** | 0.17 | 0.01 | *** | 0.15 | 0.01 | *** |
| Constant | 0.46 | 0.02 | *** | 0.40 | 0.02 | *** | -0.65 | 0.03 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| **$R^2$** | 0.25 | | | 0.24 | | | 0.24 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

Using these standardized scores, models 1 and 2 in table 7.10 report similar findings to the models 1 and 2 in table 7.6, suggesting that net of general increasing attainment in the 1990s, the GCSE points score of a young person is heavily influence by their parental occupation. The parental occupation measures contains a time invariant element (the main effect of NS-SEC and RGSC, see figures 7.6 and 7.8) and time variant element (for NS-SEC and RGSC by cohort, see figures 7.7 and 7.9) which implies that young people from less advantaged parental occupations have on average lower attainment and that this gap has slightly widened over time (net of increasing attainment and all the controls previously mentioned). As for the non-standardised GCSE points score measure (see figures 7.3 and 7.5) the comparison intervals overlap, suggesting that the coefficients for occupation group are not significantly different from the reference categories respectively (NS-SEC 1.2 and RGSC 1). This implies that the gap between those young people with parents in lower grade occupations and those with parents in higher grade occupations had not reduced and slightly increased during the 1990s.

A potential implication from the models constructed is that any changes to the relative contribution of parental occupation over time upon GCSE attainment may be conflated with the increasing propensity of parents being in more advantaged occupations over time. Being a metric scale, the CAMSIS measure has the potential advantage of controlling for changes to the occupation structure, since it can be standardised within a time period (as demonstrated in chapter 6 figure 6.7). It therefore represents a method of controlling for such changes (see Prandy 1990). As can be seen in table 7.6 and 7.10 (using the GCSE points score and the standardised points score) the differential by CAMSIS is persistent.

Furthermore by the end of the 1990s young people with parents in advantaged occupations continued to do better than those in less advantaged occupations than was the case at the beginning of the 1990s.

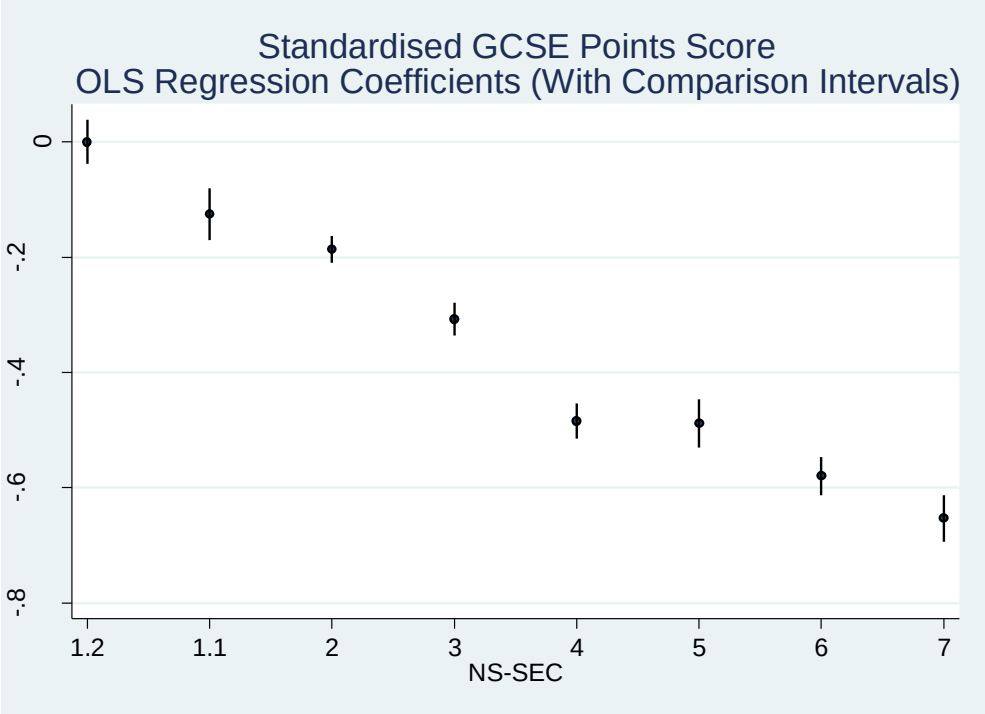



**Figure 7.6 Estimates of Standardised GCSE Points Score, NS-SEC main effects**

Other variables NS-SEC X cohort interactions, gender, ethnicity, school type, housing, household type, parental education (see model 1 in table 7.10)
Source: SN5765 & SN3532, n=88974, $R^2$=0.25, 1990s YCS Cohorts.

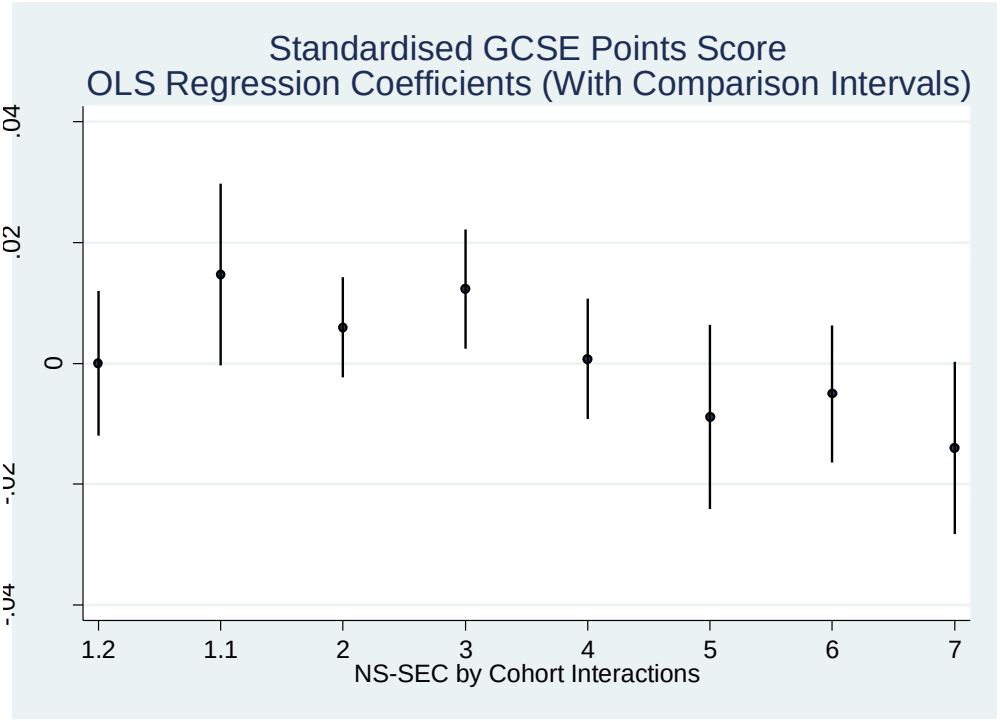



**Figure 7.7 Estimates of Standardised GCSE Points Score, NS-SEC Time varying effects**

Other variables NS-SEC, gender, ethnicity, school type, housing, household type, parental education (see model 1 in table 7.10)
Source: SN5765 & SN3532, n=88974, r2=0.25, 1990s YCS Cohorts.

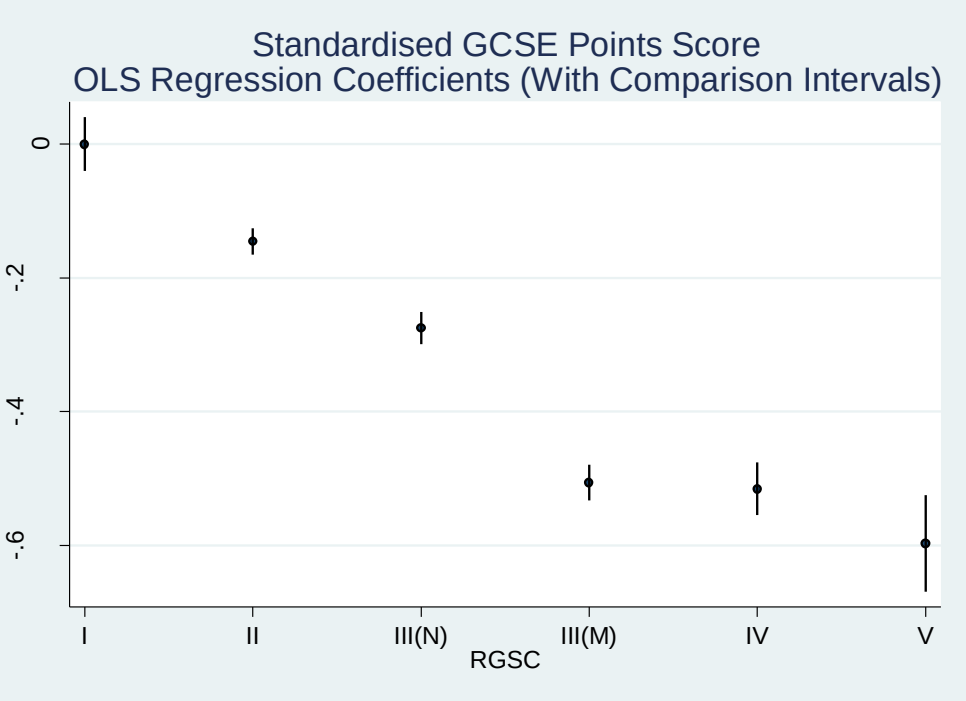



**Figure 7.8 Estimates of Standardised GCSE Points Score, RGSC main effects**

Other variables RGSC X cohort interactions, gender, ethnicity, school type, housing, household type, parental education (see model 2 in table 7.10)
Source: SN5765 & SN3532, n=89115, $R^2$=0.24, 1990s YCS Cohorts.

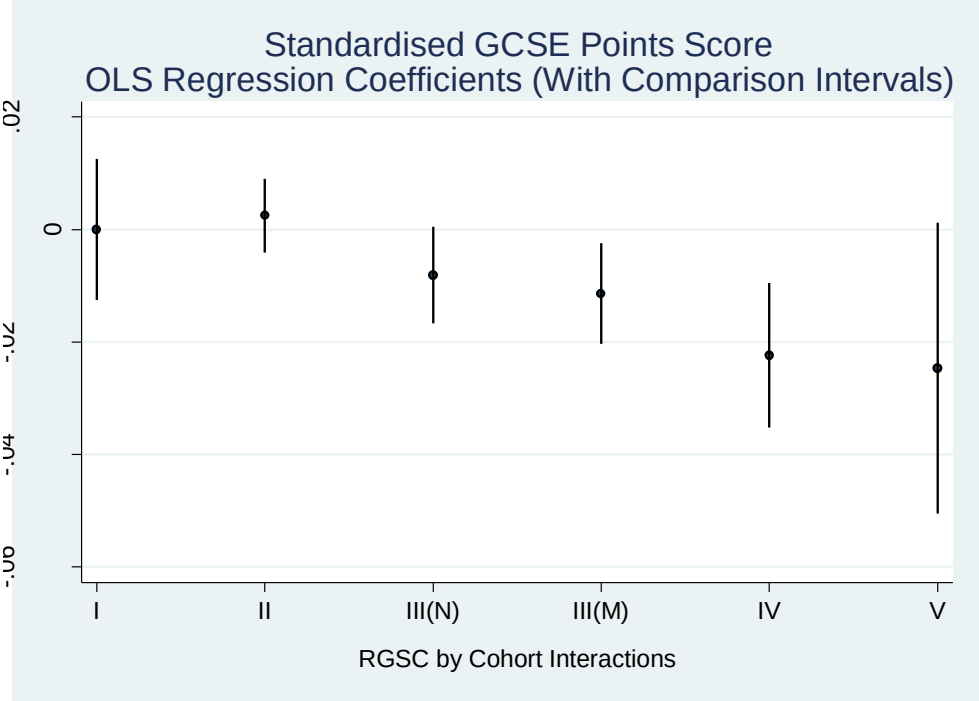



**Figure 7.9 Estimates of Standardised GCSE Points Score, RGSC Time varying effects**

Other variables RGSC, gender, ethnicity, school type, housing, household type, parental education (see model 2 in table 7.10)
Source: SN5765 & SN3532, n=89115, $R^2$=0.24, 1990s YCS Cohorts.

## School Type

The strong influence of school type within the previous models suggests that it may be worthwhile to exclude the elements of prior selection and fee-paying schools from the analysis. Tables 7.11 and 7.12 report models including only comprehensive school pupils. By considering comprehensive schools only, this removes the element of selection by attainment that occurs in independent and grammar schools (Goldstein and Thomas 1996) and verifies that the effects in these models are consistent with those including controls for school type. The findings are similar to the models in tables 7.3 and 7.6. This is also consistent with the models of Drew (Drew, Gray and Sime 1992; Drew 1995).

**Table 7.11 Logistic regression models, gaining five or more GCSE passes at grades A*-C in year 11 for comprehensive school pupils**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 0.16 | 0.02 | *** | 0.17 | 0.02 | *** | -0.13 | 0.02 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.25 | 0.08 | *** | | | | | | |
| NS-SEC = 2 | -0.34 | 0.06 | *** | | | | | | |
| NS-SEC = 3 | -0.62 | 0.07 | *** | | | | | | |
| NS-SEC = 4 | -1.03 | 0.07 | *** | | | | | | |
| NS-SEC = 5 | -0.94 | 0.08 | *** | | | | | | |
| NS-SEC = 6 | -1.12 | 0.07 | *** | | | | | | |
| NS-SEC = 7 | -1.21 | 0.08 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | -0.01 | 0.03 | | | | | | | |
| NS-SEC 2 * Cohort | -0.04 | 0.02 | * | | | | | | |
| NS-SEC 3 * Cohort | -0.02 | 0.02 | | | | | | | |
| NS-SEC 4 * Cohort | -0.05 | 0.02 | ** | | | | | | |
| NS-SEC 5 * Cohort | -0.09 | 0.03 | *** | | | | | | |
| NS-SEC 6 * Cohort | -0.10 | 0.02 | *** | | | | | | |
| NS-SEC 7 * Cohort | -0.12 | 0.03 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.22 | 0.06 | *** | | | |
| RGSC = 3.1 | | | | -0.54 | 0.06 | *** | | | |
| RGSC = 3.2 | | | | -0.93 | 0.07 | *** | | | |
| RGSC = 4 | | | | -0.83 | 0.08 | *** | | | |
| RGSC = 5 | | | | -1.00 | 0.12 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.04 | 0.02 | ** | | | |
| RGSC = 3.1 * Cohort | | | | -0.06 | 0.02 | ** | | | |
| RGSC = 3.2 * Cohort | | | | -0.09 | 0.02 | *** | | | |
| RGSC = 4 * Cohort | | | | -0.14 | 0.03 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.15 | 0.04 | *** | | | |
| Parents CAMSIS | | | | | | | 0.03 | 0.00 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.00 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.38 | 0.02 | *** | -0.38 | 0.02 | *** | -0.38 | 0.02 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.33 | 0.08 | *** | -0.35 | 0.08 | *** | -0.33 | 0.08 | *** |
| Indian | 0.27 | 0.06 | *** | 0.20 | 0.06 | *** | 0.24 | 0.06 | *** |
| Pakistani | -0.31 | 0.08 | *** | -0.44 | 0.08 | *** | -0.41 | 0.08 | *** |
| Bangladeshi | 0.06 | 0.15 | | 0.04 | 0.15 | | 0.03 | 0.15 | |
| Other Asian | 0.68 | 0.10 | *** | 0.63 | 0.10 | *** | 0.65 | 0.10 | *** |
| Other Ethnicity | 0.15 | 0.10 | | 0.08 | 0.10 | | 0.07 | 0.10 | |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -0.73 | 0.03 | *** | -0.74 | 0.03 | *** | -0.70 | 0.03 | *** |
| Other Housing | -0.27 | 0.07 | *** | -0.33 | 0.07 | *** | -0.28 | 0.07 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.07 | 0.03 | ** | -0.09 | 0.03 | *** | -0.09 | 0.03 | *** |
| Only lives with Father | -0.23 | 0.05 | *** | -0.20 | 0.06 | *** | -0.19 | 0.06 | *** |
| Other Household | -0.83 | 0.08 | *** | -0.82 | 0.08 | *** | -0.83 | 0.08 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 0.30 | 0.02 | *** | 0.31 | 0.02 | *** | 0.26 | 0.02 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 0.33 | 0.02 | *** | 0.34 | 0.02 | *** | 0.30 | 0.02 | *** |
| Constant | 0.30 | 0.06 | *** | 0.13 | 0.06 | ** | -1.74 | 0.07 | *** |
| **Observations** | 74709 | | | 74835 | | | 74837 | | |
| **Pseudo $R^2$** | 0.09 | | | 0.08 | | | 0.09 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Table 7.12 OLS regression models, GCSE points score (truncated) in year 11 for comprehensive school pupils

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 2.13 | 0.14 | *** | 2.21 | 0.14 | *** | 0.08 | 0.17 | |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -2.62 | 0.62 | *** | | | | | | |
| NS-SEC = 2 | -3.61 | 0.48 | *** | | | | | | |
| NS-SEC = 3 | -5.96 | 0.50 | *** | | | | | | |
| NS-SEC = 4 | -9.40 | 0.51 | *** | | | | | | |
| NS-SEC = 5 | -8.79 | 0.57 | *** | | | | | | |
| NS-SEC = 6 | -10.31 | 0.52 | *** | | | | | | |
| NS-SEC = 7 | -11.73 | 0.57 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.19 | 0.21 | | | | | | | |
| NS-SEC 2 * Cohort | -0.12 | 0.16 | | | | | | | |
| NS-SEC 3 * Cohort | -0.00 | 0.17 | | | | | | | |
| NS-SEC 4 * Cohort | -0.30 | 0.17 | * | | | | | | |
| NS-SEC 5 * Cohort | -0.59 | 0.20 | *** | | | | | | |
| NS-SEC 6 * Cohort | -0.60 | 0.17 | *** | | | | | | |
| NS-SEC 7 * Cohort | -0.75 | 0.19 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -3.01 | 0.47 | *** | | | |
| RGSC = 3.1 | | | | -5.35 | 0.49 | *** | | | |
| RGSC = 3.2 | | | | -9.00 | 0.50 | *** | | | |
| RGSC = 4 | | | | -8.99 | 0.57 | *** | | | |
| RGSC = 5 | | | | -10.49 | 0.82 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.03 | 0.16 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.36 | 0.16 | ** | | | |
| RGSC = 3.2 * Cohort | | | | -0.62 | 0.17 | *** | | | |
| RGSC = 4 * Cohort | | | | -0.89 | 0.19 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.98 | 0.28 | *** | | | |
| Parents CAMSIS | | | | | | | 0.24 | 0.01 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.03 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -3.76 | 0.13 | *** | -3.72 | 0.13 | *** | -3.68 | 0.13 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -2.47 | 0.58 | *** | -2.47 | 0.57 | *** | -2.38 | 0.56 | *** |
| Indian | 2.38 | 0.44 | *** | 1.88 | 0.45 | *** | 2.21 | 0.46 | *** |
| Pakistani | -2.06 | 0.57 | *** | -2.84 | 0.57 | *** | -2.45 | 0.58 | *** |
| Bangladeshi | 0.92 | 1.06 | | 0.80 | 1.04 | | 0.65 | 1.05 | |
| Other Asian | 4.93 | 0.80 | *** | 4.65 | 0.77 | *** | 4.82 | 0.76 | *** |
| Other Ethnicity | 0.71 | 0.76 | | 0.20 | 0.75 | | 0.07 | 0.75 | |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -6.66 | 0.19 | *** | -6.54 | 0.20 | *** | -6.27 | 0.20 | *** |
| Other Housing | -2.62 | 0.55 | *** | -2.92 | 0.56 | *** | -2.55 | 0.55 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.92 | 0.22 | *** | -1.10 | 0.22 | *** | -1.18 | 0.22 | *** |
| Only lives with Father | -2.27 | 0.41 | *** | -2.08 | 0.42 | *** | -2.00 | 0.42 | *** |
| Other Household | -8.31 | 0.53 | *** | -8.34 | 0.53 | *** | -8.44 | 0.52 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 2.95 | 0.19 | *** | 3.01 | 0.19 | *** | 2.52 | 0.18 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 3.18 | 0.19 | *** | 3.21 | 0.19 | *** | 2.82 | 0.19 | *** |
| Constant | 38.85 | 0.43 | *** | 37.72 | 0.44 | *** | 19.52 | 0.51 | *** |
| **Observations** | 74709 | | | 74835 | | | 74837 | | |
| **$R^2$** | 0.186 | | | 0.178 | | | 0.184 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Interaction effects

A series of models were fitted examining interaction effects to identify whether there was variation in the influence of some variables over time or in relation to family background measures. The results from interactions 1-4 are in the appendix (see tables A7.9 – A7.16). The main effects of previously identified characteristics remain significant which suggests that there is consistency in the effects that family background characteristics have upon GCSE attainment. In the majority of cases, GCSE attainment by ethnic group does not vary differentially for males as opposed to females (see tables A7.9 and A7.10 in appendix). Males of other ethnicity are more likely to gain 5 or more A*-C than White females (the reference category). Indian and Pakistani males gain higher grades points score than White females. For Indians, this reinforces a positive effect whereas for Pakistanis it moderates a negative effect. It also is moderated by the main effect of lower points score for males.

There is no difference in gaining 5 or more GCSE passes at grades A*-C for males from females within particular parental occupational classifications beyond the main effects previously described (see tables A7.11 and A7.12 in appendix). Males are less likely to 5 or A*-C grades (when modelled using RGSC) but this is not confirmed when modelled using NS-SEC or CAMSIS

GCSE attainment by parental occupation for ethnic groups mostly does not vary beyond the main effects (see tables A7.13 and A7.14 in appendix). Other Asians with parents in lower grade occupations are more likely to gain 5 or more A*-C grades than equivalent White young people. However, the findings are mixed and the interaction between parental occupation and ethnicity is not consistent across measures of parental occupation.

The interaction between parental occupation and family structure is largely non-significant but living with mother only has a significant interaction with CAMSIS and NS-SEC when compared to RGSC (see tables A7.15 and A7.16 in appendix).[30]

The interaction between parental occupation classifications and parental education is reported in tables 7.13 through to 7.16. There is an unusual effect, namely that higher parental education in lower grade occupations tends to reduce attainment. This may be an artefact of the variables included in the model (parental occupation, parental education and time varying interactions), and is therefore a suspected suppressor effect. The alternative explanation is that better qualified parents in lower grade explanations leads to lower educational attainment for their children. It is also possible that there is collinearity between parental occupation and parental education measures, whereby it would be expected that parents in higher grade occupations would be better educated, and this result is an artefact of that (as appears to be case for father's education measures in model 3 using CAMSIS in tables 7.13 and 7.15). The non-significance of the majority interaction effects suggests that, for the most part, the main effects of gender, ethnicity, parental occupation, parental education and family structure do not vary by other variables within the model.

[30] A propensity score matching analysis was undertaken to compare whether young people in lone mother families had lower GCSE attainment than dual parent families (ceteris paribus). However, it was not possible to match the two sub-samples due to the different characteristics by other variables within the models. The models produced included fewer covariates than the regression models and reached similar conclusions. As a consequence, the propensity score matching has not been reported. This did however provide an insight into the method and emphasised the advantages of regression models when analysing such data.

## Table 7.13 Logistic regression models, gaining five or more GCSE passes at grades A*-C in year 11: Interaction 5 – Parental Occupation by Parental Education (part 1)

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 0.13 | 0.02 | *** | 0.16 | 0.02 | *** | -0.09 | 0.02 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.15 | 0.08 | * | | | | | | |
| NS-SEC = 2 | -0.23 | 0.07 | *** | | | | | | |
| NS-SEC = 3 | -0.37 | 0.07 | *** | | | | | | |
| NS-SEC = 4 | -0.81 | 0.07 | *** | | | | | | |
| NS-SEC = 5 | -0.69 | 0.08 | *** | | | | | | |
| NS-SEC = 6 | -0.91 | 0.08 | *** | | | | | | |
| NS-SEC = 7 | -1.03 | 0.08 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.01 | 0.03 | | | | | | | |
| NS-SEC 2 * Cohort | -0.02 | 0.02 | | | | | | | |
| NS-SEC 3 * Cohort | 0.00 | 0.02 | | | | | | | |
| NS-SEC 4 * Cohort | -0.02 | 0.02 | | | | | | | |
| NS-SEC 5 * Cohort | -0.07 | 0.03 | *** | | | | | | |
| NS-SEC 6 * Cohort | -0.06 | 0.02 | *** | | | | | | |
| NS-SEC 7 * Cohort | -0.10 | 0.03 | *** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.09 | 0.07 | | | | |
| RGSC = 3.1 | | | | -0.27 | 0.07 | *** | | | |
| RGSC = 3.2 | | | | -0.70 | 0.08 | *** | | | |
| RGSC = 4 | | | | -0.64 | 0.09 | *** | | | |
| RGSC = 5 | | | | -0.77 | 0.12 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.04 | 0.02 | ** | | | |
| RGSC = 3.1 * Cohort | | | | -0.05 | 0.02 | ** | | | |
| RGSC = 3.2 * Cohort | | | | -0.07 | 0.02 | *** | | | |
| RGSC = 4 * Cohort | | | | -0.11 | 0.02 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.14 | 0.04 | *** | | | |
| Parents CAMSIS Male Scale, Dom. | | | | | | | 0.02 | 0.00 | *** |
| Parents CAMSIS * Cohort | | | | | | | 0.00 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.38 | 0.02 | *** | -0.37 | 0.02 | *** | -0.37 | 0.02 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.31 | 0.08 | *** | -0.33 | 0.08 | *** | -0.32 | 0.08 | *** |
| Indian | 0.28 | 0.06 | *** | 0.21 | 0.06 | *** | 0.23 | 0.06 | *** |
| Pakistani | -0.28 | 0.08 | *** | -0.41 | 0.08 | *** | -0.41 | 0.08 | *** |
| Bangladeshi | 0.08 | 0.14 | | 0.06 | 0.14 | | 0.02 | 0.14 | |
| Other Asian | 0.66 | 0.10 | *** | 0.61 | 0.10 | *** | 0.63 | 0.09 | *** |
| Other Ethnicity | 0.05 | 0.11 | | -0.01 | 0.10 | | -0.03 | 0.11 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.09 | 0.02 | *** | -0.09 | 0.02 | *** | -0.09 | 0.02 | *** |
| Grammar | 2.37 | 0.08 | *** | 2.38 | 0.08 | *** | 2.35 | 0.08 | *** |
| Secondary Modern | -0.78 | 0.05 | *** | -0.79 | 0.05 | *** | -0.80 | 0.05 | *** |
| Independent | 1.60 | 0.05 | *** | 1.59 | 0.05 | *** | 1.55 | 0.05 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -0.72 | 0.03 | *** | -0.73 | 0.03 | *** | -0.71 | 0.03 | *** |
| Other Housing | -0.23 | 0.07 | *** | -0.28 | 0.07 | *** | -0.25 | 0.07 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.07 | 0.03 | *** | -0.09 | 0.03 | *** | -0.09 | 0.03 | *** |
| Only lives with Father | -0.23 | 0.05 | *** | -0.19 | 0.05 | *** | -0.20 | 0.05 | *** |
| Other Household | -0.91 | 0.08 | *** | -0.88 | 0.08 | *** | -0.89 | 0.08 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 0.48 | 0.07 | *** | 0.48 | 0.07 | *** | -0.48 | 0.12 | *** |

## Table 7.14 Logistic regression models, gaining five or more GCSE passes at grades A*-C in year 11: Interaction 5 – Parental Occupation by Parental Education (part 2)

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| NS-SEC 1.2 * Father does not have A-level | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Father has A-level (mod) | 0.01 | 0.10 | | | | | | | |
| NS-SEC 2 * Father has A-level (mod) | -0.02 | 0.08 | | | | | | | |
| NS-SEC 3 * Father has A-level (mod) | -0.34 | 0.09 | *** | | | | | | |
| NS-SEC 4 * Father has A-level (mod) | -0.35 | 0.09 | *** | | | | | | |
| NS-SEC 5 * Father has A-level (mod) | -0.47 | 0.12 | *** | | | | | | |
| NS-SEC 6 * Father has A-level (mod) | -0.38 | 0.11 | *** | | | | | | |
| NS-SEC 7 * Father has A-level (mod) | -0.22 | 0.14 | | | | | | | |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 0.67 | 0.07 | *** | 0.61 | 0.07 | *** | -0.53 | 0.13 | *** |
| NS-SEC 1.2 * Mother does not have A-level | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Mother has A-level (mod) | -0.29 | 0.11 | *** | | | | | | |
| NS-SEC 2 * Mother has A-level (mod) | -0.22 | 0.08 | *** | | | | | | |
| NS-SEC 3 * Mother has A-level (mod) | -0.63 | 0.09 | *** | | | | | | |
| NS-SEC 4 * Mother has A-level (mod) | -0.35 | 0.10 | *** | | | | | | |
| NS-SEC 5 * Mother has A-level (mod) | -0.67 | 0.14 | *** | | | | | | |
| NS-SEC 6 * Mother has A-level (mod) | -0.68 | 0.11 | *** | | | | | | |
| NS-SEC 7 * Mother has A-level (mod) | -0.36 | 0.16 | ** | | | | | | |
| RGSC 1 * Father does not have A-level | | | | 0.00 | - | | | | |
| RGSC 2 * Father has A-level (mod) | | | | 0.00 | 0.08 | | | | |
| RGSC 3.1 * Father has A-level (mod) | | | | -0.33 | 0.08 | *** | | | |
| RGSC 3.2 * Father has A-level (mod) | | | | -0.58 | 0.10 | *** | | | |
| RGSC 4 * Father has A-level (mod) | | | | -0.41 | 0.13 | *** | | | |
| RGSC 5 * Father has A-level (mod) | | | | -0.80 | 0.28 | *** | | | |
| RGSC 1 * Mother does not have A-level | | | | 0.00 | - | | | | |
| RGSC 2 * Mother has A-level (mod) | | | | -0.20 | 0.08 | *** | | | |
| RGSC 3.1 * Mother has A-level (mod) | | | | -0.51 | 0.09 | *** | | | |
| RGSC 3.2 * Mother has A-level (mod) | | | | -0.51 | 0.10 | *** | | | |
| RGSC 4 * Mother has A-level (mod) | | | | -0.38 | 0.12 | *** | | | |
| RGSC 5 * Mother has A-level (mod) | | | | -0.26 | 0.28 | | | | |
| Parents CAMSIS * Father has A-level (mod) | | | | | | | 0.01 | 0.00 | *** |
| Parents CAMSIS * Mother has A-level (mod) | | | | | | | 0.01 | 0.00 | *** |
| Constant | 0.16 | 0.06 | ** | -0.03 | 0.07 | | -1.49 | 0.07 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| **Pseudo $R^2$** | 0.16 | | | 0.15 | | | 0.15 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Table 7.15 OLS Regression models, GCSE points score (truncated) in year 11: Interaction 5 – Parental Occupation by Parental Education (part 1)

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 1.82 | 0.11 | *** | 2.02 | 0.11 | *** | 0.68 | 0.15 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -1.42 | 0.60 | ** | | | | | | |
| NS-SEC = 2 | -2.77 | 0.47 | *** | | | | | | |
| NS-SEC = 3 | -4.06 | 0.49 | *** | | | | | | |
| NS-SEC = 4 | -7.93 | 0.50 | *** | | | | | | |
| NS-SEC = 5 | -7.02 | 0.56 | *** | | | | | | |
| NS-SEC = 6 | -8.74 | 0.51 | *** | | | | | | |
| NS-SEC = 7 | -10.26 | 0.56 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.27 | 0.18 | | | | | | | |
| NS-SEC 2 * Cohort | 0.02 | 0.13 | | | | | | | |
| NS-SEC 3 * Cohort | 0.23 | 0.14 | * | | | | | | |
| NS-SEC 4 * Cohort | 0.05 | 0.14 | | | | | | | |
| NS-SEC 5 * Cohort | -0.28 | 0.18 | | | | | | | |
| NS-SEC 6 * Cohort | -0.22 | 0.15 | | | | | | | |
| NS-SEC 7 * Cohort | -0.42 | 0.17 | ** | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -1.72 | 0.51 | *** | | | |
| RGSC = 3.1 | | | | -3.14 | 0.52 | *** | | | |
| RGSC = 3.2 | | | | -7.17 | 0.53 | *** | | | |
| RGSC = 4 | | | | -7.30 | 0.59 | *** | | | |
| RGSC = 5 | | | | -8.64 | 0.82 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.06 | 0.13 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.21 | 0.14 | | | | |
| RGSC = 3.2 * Cohort | | | | -0.31 | 0.14 | ** | | | |
| RGSC = 4 * Cohort | | | | -0.60 | 0.16 | *** | | | |
| RGSC = 5 * Cohort | | | | -0.78 | 0.26 | *** | | | |
| Parents CAMSIS | | | | | | | 0.22 | 0.01 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.02 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -3.38 | 0.12 | *** | -3.32 | 0.12 | *** | -3.30 | 0.12 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -2.31 | 0.54 | *** | -2.31 | 0.53 | *** | -2.29 | 0.53 | *** |
| Indian | 2.28 | 0.39 | *** | 1.83 | 0.40 | *** | 1.89 | 0.40 | *** |
| Pakistani | -2.03 | 0.57 | *** | -2.75 | 0.57 | *** | -2.60 | 0.57 | *** |
| Bangladeshi | 0.93 | 1.00 | | 0.71 | 0.98 | | 0.44 | 0.99 | |
| Other Asian | 4.37 | 0.70 | *** | 4.21 | 0.68 | *** | 4.29 | 0.68 | *** |
| Other Ethnicity | -0.42 | 0.77 | | -0.78 | 0.76 | | -1.02 | 0.76 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -1.09 | 0.14 | *** | -1.08 | 0.14 | *** | -1.04 | 0.14 | *** |
| Grammar | 14.16 | 0.25 | *** | 14.38 | 0.26 | *** | 14.14 | 0.26 | *** |
| Secondary Modern | -6.09 | 0.31 | *** | -6.14 | 0.30 | *** | -6.22 | 0.30 | *** |
| Independent | 11.69 | 0.26 | *** | 11.68 | 0.26 | *** | 11.24 | 0.26 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -6.55 | 0.18 | *** | -6.45 | 0.19 | *** | -6.34 | 0.19 | *** |
| Other Housing | -2.15 | 0.49 | *** | -2.38 | 0.50 | *** | -2.17 | 0.50 | *** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.85 | 0.20 | *** | -1.02 | 0.20 | *** | -1.04 | 0.20 | *** |
| Only lives with Father | -2.15 | 0.38 | *** | -1.96 | 0.39 | *** | -1.99 | 0.38 | *** |
| Other Household | -8.42 | 0.52 | *** | -8.36 | 0.52 | *** | -8.44 | 0.52 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 3.60 | 0.43 | *** | 3.92 | 0.47 | *** | -0.60 | 0.81 | |

## Table 7.16 OLS Regression models, GCSE points score (truncated) in year 11: Interaction 5 – Parental Occupation by Parental Education (part 2)

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| NS-SEC 1.2 * Father does not have A-level | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Father has A-level (mod) | -0.36 | 0.68 | | | | | | | |
| NS-SEC 2 * Father has A-level (mod) | 0.63 | 0.51 | | | | | | | |
| NS-SEC 3 * Father has A-level (mod) | -1.97 | 0.61 | *** | | | | | | |
| NS-SEC 4 * Father has A-level (mod) | -2.03 | 0.65 | *** | | | | | | |
| NS-SEC 5 * Father has A-level (mod) | -2.48 | 0.84 | *** | | | | | | |
| NS-SEC 6 * Father has A-level (mod) | -3.12 | 0.79 | *** | | | | | | |
| NS-SEC 7 * Father has A-level (mod) | -1.40 | 0.99 | | | | | | | |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 4.45 | 0.41 | *** | 4.34 | 0.42 | *** | -1.73 | 0.86 | ** |
| NS-SEC 1.2 * Mother does not have A-level | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Mother has A-level (mod) | -1.01 | 0.69 | | | | | | | |
| NS-SEC 2 * Mother has A-level (mod) | -0.61 | 0.49 | | | | | | | |
| NS-SEC 3 * Mother has A-level (mod) | -3.41 | 0.59 | *** | | | | | | |
| NS-SEC 4 * Mother has A-level (mod) | -0.80 | 0.66 | | | | | | | |
| NS-SEC 5 * Mother has A-level (mod) | -4.62 | 0.99 | *** | | | | | | |
| NS-SEC 6 * Mother has A-level (mod) | -3.90 | 0.76 | *** | | | | | | |
| NS-SEC 7 * Mother has A-level (mod) | -3.73 | 1.11 | *** | | | | | | |
| RGSC 1 * Father does not have A-level | | | | 0.00 | - | | | | |
| RGSC 2 * Father has A-level (mod) | | | | 0.24 | 0.52 | | | | |
| RGSC 3.1 * Father has A-level (mod) | | | | -2.48 | 0.61 | *** | | | |
| RGSC 3.2 * Father has A-level (mod) | | | | -4.04 | 0.67 | *** | | | |
| RGSC 4 * Father has A-level (mod) | | | | -3.85 | 0.94 | *** | | | |
| RGSC 5 * Father has A-level (mod) | | | | -5.39 | 1.75 | *** | | | |
| RGSC 1 * Mother does not have A-level | | | | 0.00 | - | | | | |
| RGSC 2 * Mother has A-level (mod) | | | | -0.61 | 0.48 | | | | |
| RGSC 3.1 * Mother has A-level (mod) | | | | -3.17 | 0.58 | *** | | | |
| RGSC 3.2 * Mother has A-level (mod) | | | | -3.81 | 0.71 | *** | | | |
| RGSC 4 * Mother has A-level (mod) | | | | -2.67 | 0.92 | *** | | | |
| RGSC 5 * Mother has A-level (mod) | | | | -2.92 | 1.94 | | | | |
| Parents CAMSIS * Father has A-level (mod) | | | | | | | 0.05 | 0.01 | *** |
| Parents CAMSIS * Mother has A-level (mod) | | | | | | | 0.07 | 0.01 | *** |
| Constant | 37.94 | 0.42 | *** | 36.45 | 0.48 | *** | 20.79 | 0.49 | *** |
| **Observations** | 88974 | | | 89115 | | | 89118 | | |
| $R^2$ | 0.27 | | | 0.27 | | | 0.27 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Summary

Increasing GCSE attainment during the 1990s, particularly among the children of parents from more advantaged socio-economic groups is consistently reported in the models within this chapter. There is clear evidence for persisting and relatively stable differentials in GCSE attainment according to parental occupation-based measures. Through the 1990s these do not appear to have lessened and the evidence from these models suggests that this gap has widened. This emphasises the continued importance of parental occupational stratification upon the GCSE attainment of young people. The three measures of parental occupation used in this analysis (NS-SEC, RGSC and CAMSIS) consistently report this finding.

This also has implications for understanding the decision to remain in education post-16, as described in chapter four. The strong association between parental occupation and GCSE attainment contributes to the trends initially reported by Courtenay and McAleese (1993a; 1993b; 1994) and by Park (1994) that the differences in staying-on rates between young people by parental occupation are considerably wider by age 18-19. The models in this chapter report similar findings to those fitted to 1980s YCS data by Jesson, Gray and Sime (1991), who claim that parental occupation influences staying-on via grades attained and the subsequent decision making processes after compulsory education ceases. It helps explain how parental occupation influences both directly and indirectly the likelihood of first attaining the requisite GCSE attainment to stay-on and then the likelihood of doing so. In order to gain access to more advantaged occupations, educational attainment is the first step. Young people whose parents are in more advantaged occupations are more likely to do well at school. Unsurprisingly, their experiences are more likely to be similar to their parents than dissimilar.

Park (1994) notes the influence of parental education appears to be less than that for parental occupation although it is not possible to entirely distinguish between these two effects as parental occupation is partially a consequence of parental education (i.e. more educated parents tend to be in more advantaged occupations). What is needed are data that measure the changing circumstances of people over their lifespan so that the processes of education, finding employment, marriage and raising a family and the early school years of their children can be fully investigated.

Parental education influences GCSE performance but the effect of parental occupation is stronger in multivariate models. Parental occupation and parental education are associated (as would be expected from studies of social mobility) but it is informative to consider their relative effects independently. This is especially the case in the absence of better data or information that might clarify the association between parental occupation and parental education. It is also of substantive interest to consider how parental education might influence a young person's attainment when the parent possessing A-levels was not necessarily a pre-requisite for their occupation (as might reasonably be considered the case for parents growing up in a context of lower educational attainment). This research suggests that parental education continues to have a positive effect on filial attainment net of parental occupation.

Family structure continues to have a role in educational attainment with young people living with their mother only doing less well than those living with both parents. This effect is consistent regardless of model form, although is more vulnerable to being reduced in significance when the sub sample is reduced (as is the case in tables 7.7 and 7.8).

From observing the differentials in attainment according to parental occupation, question five (in chapter five) examines whether the multiple sources of variation in GCSE attainment (manifested in different subject combinations) also influence the likelihood of success in certain subjects. This is important as some subjects are relevant to gaining access to post-compulsory educational courses and may represent some of the reason behind the low take up of post-compulsory education by disadvantaged social groups (net of prior attainment). It is also likely that young people from different social backgrounds may also have a tendency to do better or worse in different subjects. This is the subject of chapter ten.

The next chapter considers the influence that missing data may have on the estimates produced in the models in this chapter. This is important as a means of further verifying and examining the findings of the models in this chapter.

# 8 Missing data in the Youth Cohort Study

In this chapter, the issue of missing data the Youth Cohort Study is discussed and approaches to overcome these problems are introduced. The question being addressed is: *"What influence does missing data potentially have upon the estimates in models fitted (particularly with regard to parental occupation)?"* The different forms of missing data, assumptions about the nature of missing data and further detail of missing data are provided. Bayesian multiple imputation methods are applied and evaluated with regard to the YCS. Finally, missing data in parental educational and family structure measures are considered.

There are two principle types of missing data: unit non-response and item non-response (see Serfling 2004). Unit non-response occurs when a respondent does not respond to the survey or subsequently drops out of the survey (also referred to as attrition) and is typically addressed with by weighting the data. Weighting approaches attempt to correct for these differential response rates by sub-populations. The post-survey adjustments are made at the parameter estimation stage (Biemer and Christ 2006). No additional information is required but there is a tendency to inflate standard errors where the variance in the weights is larger (Sturgis 2004). Weights are dependent upon the analysis being performed and must be appropriate to the outcome variable. Different weights may be required depending upon the analysis conducted (Goldstein 2009).

Item non-response occurs when a respondent does not answer specific questions within the survey, either for structural reasons (for example, orphans not supplying parental occupational information due to not ever having known this information) or for unknown reasons (Shoemaker, Eichholz and Skewes 2002). Whether item non-

response may bias estimates produced during modelling of the data depends upon the conditions by which the data is assumed to be missing. If the response to a question is Missing Completely At Random (MCAR) then there should be no systematic error when estimating parameters, the sample mean being BLUE (the Best Linear Unbiased Estimator for the population mean). In such an instance, the missingness of the data is not dependent upon the question being asked (Lynn et al. 2005; Carpenter and Kenward 2008).

More commonly the data for a particular question is missing dependent upon other factors which have been measured. In such a scenario, it is said to be Missing At Random (MAR). These are instances where the responses to questions within the survey are non-randomly missing (i.e. Item non-response varies according to the characteristics of respondents). Sometimes this is due to differential attrition rates between groups of respondents (see for example Hawkes and Plewis 2006) but the subject of interest of this chapter is whether there is differential response rates according to parental occupation in the initial sweep of the YCS 1990s cohorts. It is important to note that when a variable is MAR, this is conditional upon other information; given the observed data, the factors influencing whether the variable are independent of (i.e. do not depend upon) the missing values of the missing variable itself (Carpenter and Kenward 2008). If this is not the case and the missingness of the variable depends upon the value of the variable then it is said to be Missing Not At Random (MNAR) or Non Ignorable (NI, see Treiman 2009).

## Missing parental occupation data

To establish the condition under which the data can be considered missing for a given variable, it is first necessary to observer the patterns of missing data by question in the survey. The first of these to be addressed is missing parental

occupational data. Father's and mother's employment details are recorded in the following manner in the YCS questionnaire.

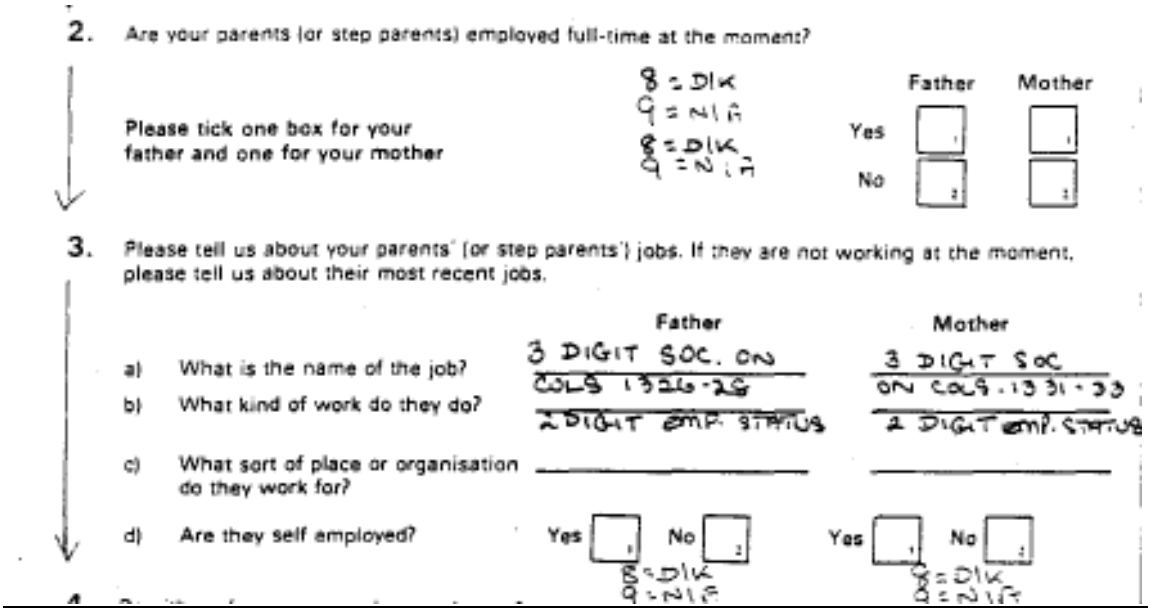



**Figure 8.1 Extract from YCS6 Questionnaire – Parental Occupation Questions**

The parental occupation data is based upon the SOC (Standard Occupation Codes) using the 1990 classification (see Rose 1995). Parental SOC is based upon the most recent job (see figure 9.1). Of the 101,713 young people in the YCS 1990s cohorts, Father's SOC is missing for 14,836 cases and Mother's for 20,201 (see table 8.1).

**Table 8.1 Information provided: Father's SOC by Mother's SOC**

| | Mother's SOC Reported | Missing | Total |
|---|---|---|---|
| Father's SOC Reported | 75,158 | 11,719 | 86,877 |
| Missing | 6,354 | 8,482 | 14,836 |
| *Total* | *81,512* | *20,201* | *101,713* |

The justification for using the "dominance" measure of parental occupation is largely as a consequence of the large amount of missing data on parental occupations (this can be seen in table 6.2 in chapter six). It is also because the highest NS-SEC, RGSC or CAMSIS position of parents is more likely to reflect the material circumstances of the family. Furthermore, there is less dispersion in female occupations (Prandy 1990; Lambert, Prandy and Bottero 2007). A large number of respondents do not report their parent's job when their parent is not employed full-time, particularly for mother's job (see tables 9.2 and 9.3). It is distinctly possible that young people (in the absence of advice, as the YCS is a postal survey) may decide not to respond if their parents are not working (even with the clarifying statement about recent jobs). Given that the first question asks about full-time work, the young person may decide that their parents' part-time job or jobs are not to be included.

**Table 8.2 Father's SOC provided by whether Father Employed**

| | *Is your father employed full-time at the moment?* | | | |
|---|---|---|---|---|
| | -9. not answered | 1. yes | 2. no | Total |
| Father's SOC Reported | 2,606 | 74,306 | 9,965 | 86,877 |
| Missing | 6,323 | 4,270 | 4,243 | 14,836 |
| Total | 8,929 | 78,576 | 14,208 | 101,713 |

**Table 8.3 Mother's SOC provided by whether Mother Employed**

| | *Is your mother employed full-time at the moment?* | | | |
|---|---|---|---|---|
| | -9. not answered | 1. yes | 2. no | Total |
| Mother's SOC Reported | 2,487 | 47,097 | 31,928 | 81,512 |
| Missing | 3,366 | 2,318 | 14,517 | 20,201 |
| Total | 5,853 | 49,415 | 46,445 | 101,713 |

## Missing parental education data

There is a greater quantity of missing data for parental education measures. Knowledge of parental education appeared to vary according to the respondent's own education. Using data from 1990 (YCS5), Park (1994) observed that 22% of those with an A or AS level and 49% of those without qualifications did not answer or were unsure of their parent's education. It is highly likely that, as for parental occupation, reporting of parental education varies by family background measures considered (Kreuter et al. 2010). The parental education measure used in the analysis used the parental degree variable to increase the A-level variable, as it is highly likely that those with a degree took the A-level route. Even if they did not, the qualification held is higher than an A-level so it can be considered that they are of A-level or above.

**Table 8.4 Frequency of missing data for parental education variables**

| *Did your father obtain one or more a levels?* | *Frequency* | *Percent* |
|---|---|---|
| 1. yes | 19,308 | 18.98 |
| 2. no | 38,450 | 37.80 |
| 3. other response | 16,237 | 15.96 |
| Missing | 27,718 | 27.25 |
| *Total* | *(101,713)* | *(100)* |

| *Did your father obtain a degree?* | *Freq.* | *Percent* |
|---|---|---|
| -9. not answered | 15,482 | 15.22 |
| 1. yes | 16,524 | 16.25 |
| 2. no | 54,087 | 53.18 |
| 3. other response | 15,620 | 15.36 |
| *Total* | *(101,713)* | *(100)* |

| *Did your mother obtain one or more a levels?* | *Freq.* | *Percent* |
|---|---|---|
| 1. yes | 17,995 | 17.69 |
| 2. no | 42,592 | 41.87 |
| 3. other response | 15,700 | 15.44 |
| Missing | 25,426 | 25.00 |
| *Total* | *(101,713)* | *(100)* |

| *Did your mother obtain a degree?* | *Freq.* | *Percent* |
|---|---|---|
| -9. not answered | 13,164 | 12.94 |
| 1. yes | 11,000 | 10.81 |
| 2. no | 62,205 | 61.16 |
| 3. other response | 15,344 | 15.09 |
| *Total* | *(101,713)* | *(100)* |

Using the modified A-level measure (including all those whose parents had degrees as having at least an A-level) increases the figure for fathers to 24,708 and mothers to 21,286. I.e. the figures in figure 6.11 (chapter 6) are the highest qualification gained[31]. As previously mentioned in chapter 6, the post compulsory education indicator was not useful for England and Wales as this should be equivalent to whether the parents had A-levels or above.

[31] Highest qualification: *16,524 Degrees + 8,184 A-levels = 24,708 with A-level or higher*

The other response includes the options *"not sure"* and *"prefer not to say"* and also *"don't know"* and *"not applicable"*. It seems inconsistent that the "prefer not to say" option should be included for parental educational questions but not for parental occupation questions. This is a pity as it decreases the amount of useful responses. In YCS6 the parental A-level question was not asked (only the degree) and for both YCS 5 and YCS 6 (as examples) the layout of the question on the page is confusing. The following is an extract from the YCS 5 documentation (page 14 of the sweep 1 survey).

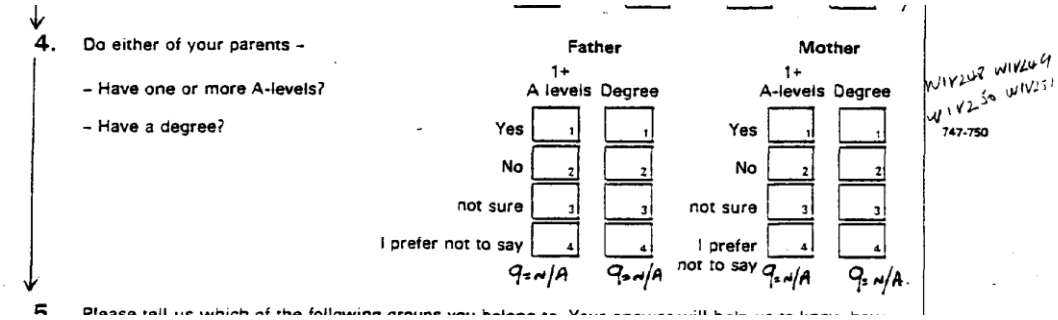



**Figure 8.2 Extract from YCS6 Questionnaire – Parental Education Questions**

Family structure has little missing data but the variable could be recorded with more information, such as marital status (see table 6.11 in chapter 6).

## Multiple Imputation

Another possible source of error may be the influence of differentially missing data upon the estimates. Tables 8.5 and 8.6 contain models of complete cases (i.e. with no missing data), to examine whether the missing data may influence the estimates. However, the coefficients are consistent with other models in chapter 7 which suggests that the sample weighting is providing some control for this. If the coefficients were different compared with all the models using sample weights (i.e. all the models in chapter 7), this would suggest a potential source of bias in the weighting. However, there is a penalty to using the complete case analysis in terms of reduced sample size and the potential for the data to be Missing At Random (MAR) or Missing Not At Random (MNAR) meaning that certain groups may have different propensities to respond to each question in the survey, from which the variables are derived (Treiman 2009).

**Table 8.5 Logistic regression models, gaining five or more GCSE passes at grades A*-C in year 11 for complete cases**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 0.12 | 0.02 | *** | 0.13 | 0.02 | *** | 0.03 | 0.03 | |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -0.26 | 0.09 | *** | | | | | | |
| NS-SEC = 2 | -0.41 | 0.07 | *** | | | | | | |
| NS-SEC = 3 | -0.72 | 0.08 | *** | | | | | | |
| NS-SEC = 4 | -1.16 | 0.08 | *** | | | | | | |
| NS-SEC = 5 | -1.16 | 0.09 | *** | | | | | | |
| NS-SEC = 6 | -1.44 | 0.09 | *** | | | | | | |
| NS-SEC = 7 | -1.53 | 0.10 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.00 | 0.03 | | | | | | | |
| NS-SEC 2 * Cohort | -0.01 | 0.02 | | | | | | | |
| NS-SEC 3 * Cohort | 0.01 | 0.02 | | | | | | | |
| NS-SEC 4 * Cohort | -0.00 | 0.02 | | | | | | | |
| NS-SEC 5 * Cohort | -0.02 | 0.03 | | | | | | | |
| NS-SEC 6 * Cohort | -0.01 | 0.03 | | | | | | | |
| NS-SEC 7 * Cohort | -0.03 | 0.03 | | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -0.30 | 0.07 | *** | | | |
| RGSC = 3.1 | | | | -0.68 | 0.07 | *** | | | |
| RGSC = 3.2 | | | | -1.20 | 0.08 | *** | | | |
| RGSC = 4 | | | | -1.19 | 0.09 | *** | | | |
| RGSC = 5 | | | | -1.42 | 0.15 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | -0.02 | 0.02 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.02 | 0.02 | | | | |
| RGSC = 3.2 * Cohort | | | | -0.02 | 0.02 | | | | |
| RGSC = 4 * Cohort | | | | -0.04 | 0.03 | | | | |
| RGSC = 5 * Cohort | | | | -0.02 | 0.05 | | | | |
| Parents CAMSIS | | | | | | | 0.04 | 0.00 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.00 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -0.46 | 0.02 | *** | -0.45 | 0.02 | *** | -0.45 | 0.02 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -0.33 | 0.10 | *** | -0.29 | 0.10 | *** | -0.27 | 0.10 | *** |
| Indian | 0.30 | 0.07 | *** | 0.22 | 0.07 | *** | 0.26 | 0.07 | *** |
| Pakistani | -0.30 | 0.09 | *** | -0.40 | 0.09 | *** | -0.39 | 0.10 | *** |
| Bangladeshi | 0.41 | 0.17 | ** | 0.34 | 0.17 | ** | 0.35 | 0.18 | ** |
| Other Asian | 0.76 | 0.12 | *** | 0.69 | 0.12 | *** | 0.72 | 0.12 | *** |
| Other Ethnicity | 0.07 | 0.13 | | 0.04 | 0.13 | | 0.03 | 0.13 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.08 | 0.02 | *** | -0.08 | 0.02 | *** | -0.07 | 0.02 | *** |
| Grammar | 2.41 | 0.10 | *** | 2.43 | 0.10 | *** | 2.40 | 0.10 | *** |
| Secondary Modern | -0.78 | 0.05 | *** | -0.78 | 0.05 | *** | -0.78 | 0.05 | *** |
| Independent | 1.60 | 0.06 | *** | 1.58 | 0.06 | *** | 1.53 | 0.06 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -0.79 | 0.03 | *** | -0.81 | 0.03 | *** | -0.77 | 0.03 | *** |
| Other Housing | -0.16 | 0.08 | * | -0.21 | 0.08 | ** | -0.17 | 0.09 | * |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.06 | 0.04 | | -0.05 | 0.04 | | -0.06 | 0.04 | * |
| Only lives with Father | -0.23 | 0.07 | *** | -0.21 | 0.07 | *** | -0.21 | 0.07 | *** |
| Other Household | -0.95 | 0.09 | *** | -0.94 | 0.09 | *** | -0.94 | 0.09 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 0.33 | 0.03 | *** | 0.36 | 0.03 | *** | 0.29 | 0.03 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 0.35 | 0.03 | *** | 0.35 | 0.03 | *** | 0.30 | 0.03 | *** |
| Constant | 0.49 | 0.07 | *** | 0.34 | 0.07 | *** | -2.25 | 0.09 | *** |
| **Observations** | 63912 | | | 63912 | | | 63912 | | |
| **Pseudo $R^2$** | 0.17 | | | 0.16 | | | 0.17 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Table 8.6 OLS regression models, GCSE points score (truncated) in year 11 for complete cases

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Cohort Number | 1.68 | 0.12 | *** | 1.83 | 0.12 | *** | 1.35 | 0.17 | *** |
| NS-SEC = 1.2 | 0.00 | - | | | | | | | |
| NS-SEC = 1.1 | -2.01 | 0.60 | *** | | | | | | |
| NS-SEC = 2 | -3.34 | 0.46 | *** | | | | | | |
| NS-SEC = 3 | -5.84 | 0.49 | *** | | | | | | |
| NS-SEC = 4 | -9.91 | 0.51 | *** | | | | | | |
| NS-SEC = 5 | -9.74 | 0.61 | *** | | | | | | |
| NS-SEC = 6 | -11.59 | 0.54 | *** | | | | | | |
| NS-SEC = 7 | -13.97 | 0.61 | *** | | | | | | |
| NS-SEC 1.2 * Cohort | 0.00 | - | | | | | | | |
| NS-SEC 1.1 * Cohort | 0.24 | 0.19 | | | | | | | |
| NS-SEC 2 * Cohort | 0.12 | 0.14 | | | | | | | |
| NS-SEC 3 * Cohort | 0.26 | 0.15 | * | | | | | | |
| NS-SEC 4 * Cohort | 0.19 | 0.15 | | | | | | | |
| NS-SEC 5 * Cohort | 0.02 | 0.20 | | | | | | | |
| NS-SEC 6 * Cohort | 0.01 | 0.17 | | | | | | | |
| NS-SEC 7 * Cohort | 0.17 | 0.19 | | | | | | | |
| RGSC = 1 | | | | 0.00 | - | | | | |
| RGSC = 2 | | | | -2.78 | 0.45 | *** | | | |
| RGSC = 3.1 | | | | -5.65 | 0.49 | *** | | | |
| RGSC = 3.2 | | | | -10.39 | 0.51 | *** | | | |
| RGSC = 4 | | | | -10.89 | 0.61 | *** | | | |
| RGSC = 5 | | | | -12.97 | 0.90 | *** | | | |
| RGSC = 1 * Cohort | | | | 0.00 | - | | | | |
| RGSC = 2 * Cohort | | | | 0.07 | 0.14 | | | | |
| RGSC = 3.1 * Cohort | | | | -0.07 | 0.15 | | | | |
| RGSC = 3.2 * Cohort | | | | -0.05 | 0.15 | | | | |
| RGSC = 4 * Cohort | | | | -0.15 | 0.19 | | | | |
| RGSC = 5 * Cohort | | | | 0.05 | 0.31 | | | | |
| Parents CAMSIS | | | | | | | 0.30 | 0.01 | *** |
| Male Scale, Dom. | | | | | | | | | |
| Parents CAMSIS * Cohort | | | | | | | 0.01 | 0.00 | *** |
| Female | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Male | -3.88 | 0.14 | *** | -3.83 | 0.14 | *** | -3.81 | 0.14 | *** |
| White | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Black | -3.05 | 0.68 | *** | -2.80 | 0.68 | *** | -2.69 | 0.67 | *** |
| Indian | 2.64 | 0.45 | *** | 2.05 | 0.45 | *** | 2.23 | 0.45 | *** |
| Pakistani | -1.86 | 0.67 | *** | -2.58 | 0.68 | *** | -2.42 | 0.68 | *** |
| Bangladeshi | 3.32 | 1.25 | *** | 2.85 | 1.22 | ** | 2.78 | 1.23 | ** |
| Other Asian | 4.87 | 0.79 | *** | 4.47 | 0.78 | *** | 4.60 | 0.76 | *** |
| Other Ethnicity | -0.16 | 0.91 | | -0.34 | 0.90 | | -0.44 | 0.89 | |
| Comprehensive (till 18) | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Comprehensive (till 16) | -0.90 | 0.16 | *** | -0.92 | 0.16 | *** | -0.86 | 0.16 | *** |
| Grammar | 13.69 | 0.27 | *** | 13.93 | 0.27 | *** | 13.65 | 0.27 | *** |
| Secondary Modern | -6.07 | 0.35 | *** | -6.11 | 0.35 | *** | -6.17 | 0.34 | *** |
| Independent | 11.14 | 0.28 | *** | 11.06 | 0.28 | *** | 10.54 | 0.28 | *** |
| Parents Own House | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Rented | -6.90 | 0.22 | *** | -6.98 | 0.23 | *** | -6.82 | 0.22 | *** |
| Other Housing | -1.29 | 0.59 | ** | -1.60 | 0.59 | *** | -1.38 | 0.58 | ** |
| Lives with both parents | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Only lives with Mother | -0.66 | 0.27 | ** | -0.57 | 0.27 | ** | -0.67 | 0.27 | ** |
| Only lives with Father | -2.30 | 0.50 | *** | -2.15 | 0.51 | *** | -2.11 | 0.50 | *** |
| Other Household | -8.60 | 0.57 | *** | -8.48 | 0.58 | *** | -8.53 | 0.57 | *** |
| Father does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Father has A-level (mod) | 2.99 | 0.18 | *** | 3.14 | 0.19 | *** | 2.62 | 0.18 | *** |
| Mother does not have A-level | 0.00 | - | | 0.00 | - | | 0.00 | - | |
| Mother has A-level (mod) | 3.05 | 0.19 | *** | 3.03 | 0.19 | *** | 2.64 | 0.18 | *** |
| Constant | 40.14 | 0.41 | *** | 39.13 | 0.43 | *** | 17.05 | 0.57 | *** |
| **Observations** | 63912 | | | 63912 | | | 63912 | | |
| **$R^2$** | 0.29 | | | 0.29 | | | 0.29 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

One potential solution might be to impute the mode or mean value for parental occupation (i.e. SOC) from the observed cases to the unobserved/missing values. However, modal or mean imputation tend to lead to reduced dispersion (i.e. underestimated standard errors or over-fitting, see Treiman 2009). Multiple Imputation is an attempt to gain unbiased estimates of parameters and to estimate more realistic standard errors (Allison 2000). Multiple Imputation (MI) methods by contrast attempt to model the missingness mechanism for the missing data in order to estimate values for and the distribution of the missing data using the data provided. Essentially, a new dataset or series of datasets are produced upon which complete case analyses are conducted with no post-estimation correction (Gelman, King and Liu 1998). The independent variables used to predict the "Model of Missingness" may be present within the dataset or auxiliary variables collected purely to predict missingness. Such a model is distinct from the Model of Interest (MOI) and whilst all variables in the MOI should be in the imputation model (Carpenter and Kenward 2008) it is likely that the additional predictors of missingness in the imputation model are quite different, for example interviewer or neighbourhood characteristics (Goldstein 2009).

Multiple Imputation models are particularly useful when the item non-response is Missing At Random (MAR). The conditions under MAR are described by Allison (2000: 302): *"First, the data must be missing at random, meaning that the probability of missing data on a particular variable Y can depend on other observed variables, but not on Y itself (controlling for other observed variables). Second, the model used to generate the imputed values must be 'correct' in some sense. Third, the model used for the analysis must match up, in some sense, with the model used in the imputation."* If these conditions hold then we can get a valid estimate of the

distribution from the observed data (Rubin 1987). This is achieved by firstly modelling the missingness mechanism, then by chained equations to generate matrices of imputed values for the missing data based on known relationships between the observed data (based on the missingness mechanism).

Critics of multiple imputation suggest that other techniques may be more suitable, such as Inverse Probability Weighting (IPW). Vansteelandt, Carpenter and Kenward (2010) argue that IPW is relatively inefficient compared with MI because the weighting is based on fully observed individuals (rather than partially observed for MI). Further IPW models of increasing complexity have been developed referred to as doubly robust estimators which seek to combine models of response to variables of interest, models for the probability of observing variables and joint models of the mean of partially observed variables given fully observed variables (Carpenter, Kenward and Vansteelandt 2006; Vansteelandt, Carpenter and Kenward 2010).

## Modelling missing data with multiply imputed datasets

A number of models were fitted predict the likelihood of parental occupational data being missing according to the characteristics of the respondent. Parental SOC was more likely to be missing in later cohorts, for males than females, for all other ethnic groups compared to white young people, for some school types, for those living in rented accommodation, for those who did not live with both parents, for those with parents who did not have A-levels and for those who did not gain 5 or more A-C GCSE passes. However, stronger predictors of missing parental SOC were whether the respondent's father or mother had a full-time job and whether they truanted or not. The decision was made to fit chained equations[32] for GCSE attainment (using the points score measure) using the following independent variables: cohort, sex, ethnicity, RGSC, housing tenure type, family structure, father's A-level indicator, mother's A-level indicator, whether father is employed full-time, whether mother is employed full-time and frequency of truancy.

Therefore, the model of missingness contained the substantive variables of interest (similar to those included in the models in chapter 7) plus employment indicators for father and mother and the frequency of truancy. In the absence of auxiliary information, this was the best combination of predictors to estimate whether parental occupation data was missing or not. The chained equations then impute a value to insert into the missing item for a respondent based upon the associations predicted by the missingness model. The chained element of this is that values are imputed for a range of missing data items, not just the missing data variable of interest. Essentially, the procedure is repetitive with multiple equations being fitted to each of the missing data items to establish a likely value for the missing item

[32] In Stata this is performed using the 'ice' routine, see Royston P. (2004) Multiple imputation of missing values, *Stata Journal* 4(3) 227-241.

based upon the values supplied by respondents. Once this has concluded a complete dataset will have been produced with likely values imputed into the missing data items conditional upon the model of missingness which is used to estimate the likelihood of data being missing according to the characteristics of respondents.

To improve the robustness of the final estimates, this procedure is repeated a number of times to produce a series of complete datasets which will vary slightly in the values estimates for the missing data items. It is from this stage that the substantive model of interest is fitted to the multiply imputed datasets. The coefficients reported in the substantive model being the average of estimates in models fitted to each of the multiply imputed datasets.

In tables 8.7 and 8.8, the complete cases (model 1) then 5 datasets (model 2) and then 10 datasets (model 3) were produced from the multiple imputation for the missing data of each variable used. The survey weights used in all analysis were also included in the models.

**Table 8.7** **OLS regression on GCSE points score (truncated) in year 11 for comprehensive school pupils: comparison of weighted models with weighted MI models**

| | Model 1 | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|
| | B | SE | B | SE | FMI | B | SE | FMI |
| Cohort Number | 1.88 | 0.04 | 1.83 | 0.04 | 0.06 | 1.83 | 0.04 | 0.03 |
| RGSC = 1 | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| RGSC = 2 | -3.31 | 0.30 | -3.26 | 0.26 | 0.04 | -3.30 | 0.27 | 0.06 |
| RGSC = 3.1 | -6.83 | 0.33 | -6.33 | 0.29 | 0.10 | -6.35 | 0.29 | 0.05 |
| RGSC = 3.2 | -11.35 | 0.34 | -10.38 | 0.32 | 0.23 | -10.41 | 0.31 | 0.14 |
| RGSC = 4 | -12.15 | 0.39 | -11.23 | 0.37 | 0.29 | -11.29 | 0.35 | 0.20 |
| RGSC = 5 | -13.63 | 0.60 | -12.50 | 0.46 | 0.17 | -12.60 | 0.47 | 0.21 |
| Female | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Male | -4.10 | 0.15 | -3.54 | 0.13 | 0.06 | -3.54 | 0.13 | 0.04 |
| White | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Black | -1.59 | 0.73 | -1.52 | 0.45 | 0.01 | -1.56 | 0.47 | 0.06 |
| Indian | 2.68 | 0.50 | 1.75 | 0.42 | 0.13 | 1.73 | 0.41 | 0.08 |
| Pakistani | -2.76 | 0.67 | -3.73 | 0.47 | 0.16 | -3.72 | 0.45 | 0.08 |
| Bangladeshi | 3.50 | 1.24 | -0.22 | 0.69 | 0.15 | -0.17 | 0.67 | 0.11 |
| Other Asian | 6.31 | 0.90 | 5.32 | 0.69 | 0.07 | 5.19 | 0.70 | 0.09 |
| Other Ethnicity | 1.60 | 0.82 | 0.92 | 0.65 | 0.05 | 0.92 | 0.65 | 0.06 |
| Parents Own House | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Rented | -6.80 | 0.23 | -6.69 | 0.18 | 0.09 | -6.70 | 0.18 | 0.06 |
| Other Housing | -2.00 | 0.67 | -2.80 | 0.50 | 0.02 | -2.76 | 0.50 | 0.04 |
| Lives with both parents | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Only lives with Mother | -0.29 | 0.29 | -1.17 | 0.21 | 0.18 | -1.15 | 0.20 | 0.10 |
| Only lives with Father | -1.71 | 0.56 | -1.83 | 0.39 | 0.08 | -1.79 | 0.38 | 0.07 |
| Other Household | -7.53 | 0.58 | -8.42 | 0.44 | 0.12 | -8.41 | 0.42 | 0.07 |
| Father has A-level | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Father does not have A-level | -2.57 | 0.22 | -2.44 | 0.23 | 0.35 | -2.44 | 0.21 | 0.25 |
| Father Other response (A-level) | -5.92 | 0.30 | -5.62 | 0.33 | 0.52 | -5.57 | 0.30 | 0.39 |
| Mother has A-level | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Mother does not have A-level | -2.93 | 0.22 | -2.86 | 0.22 | 0.35 | -2.91 | 0.22 | 0.32 |
| Mother Other response (A-level) | -4.71 | 0.31 | -4.68 | 0.32 | 0.46 | -4.71 | 0.29 | 0.33 |
| Constant | 46.44 | 0.32 | 44.84 | 0.29 | 0.15 | 44.89 | 0.29 | 0.12 |
| **Observations (min)** | 52728 | | 86400 | | | 86400 | | |
| **$R^2$** | 0.217 | | | | | | | |
| *Number of Imputed Datasets* | - | | 5 | | | 10 | | |
| *Minimum Degrees of Freedom* | | | 18.4 | | | 64.9 | | |

Tables 8.7 and 8.8 use the weights provided with the YCS[33]. Fraction of Missing Information (FMI) is reported, which is the "relative increase in variance due to non-response" (Li, Raghunathan and Rubin 1991). There is little difference in the betas and standard errors between models 1, 2 and 3 in each instance which suggests that the multiple imputation of missing data has not greatly improved the models from the weighted models used through this thesis.

[33] Using Stata's xi: svy: mim: routine

**Table 8.8** **Logistic regression for gaining five or more GCSE passes at grades A*-C in year 11 for comprehensive school pupils: comparison of weighted models with weighted MI models**

| | Model 1 | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|
| | B | SE | B | SE | FMI | B | SE | FMI |
| Cohort Number | 0.12 | 0.01 | 0.11 | 0.00 | 0.01 | 0.11 | 0.01 | 0.02 |
| RGSC = 1 | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| RGSC = 2 | -0.38 | 0.04 | -0.34 | 0.04 | 0.09 | -0.34 | 0.04 | 0.07 |
| RGSC = 3.1 | -0.78 | 0.04 | -0.68 | 0.04 | 0.14 | -0.68 | 0.04 | 0.07 |
| RGSC = 3.2 | -1.29 | 0.05 | -1.13 | 0.04 | 0.18 | -1.13 | 0.04 | 0.10 |
| RGSC = 4 | -1.36 | 0.05 | -1.18 | 0.05 | 0.19 | -1.18 | 0.04 | 0.10 |
| RGSC = 5 | -1.53 | 0.09 | -1.34 | 0.07 | 0.06 | -1.35 | 0.06 | 0.06 |
| Female | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Male | -0.43 | 0.02 | -0.36 | 0.02 | 0.01 | -0.36 | 0.02 | 0.02 |
| White | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Black | -0.11 | 0.10 | -0.25 | 0.07 | 0.01 | -0.25 | 0.07 | 0.03 |
| Indian | 0.26 | 0.07 | 0.17 | 0.05 | 0.05 | 0.16 | 0.05 | 0.06 |
| Pakistani | -0.42 | 0.09 | -0.48 | 0.07 | 0.06 | -0.48 | 0.07 | 0.04 |
| Bangladeshi | 0.36 | 0.17 | -0.11 | 0.10 | 0.09 | -0.10 | 0.10 | 0.05 |
| Other Asian | 0.87 | 0.12 | 0.69 | 0.09 | 0.05 | 0.69 | 0.09 | 0.03 |
| Other Ethnicity | 0.21 | 0.12 | 0.13 | 0.09 | 0.01 | 0.13 | 0.09 | 0.02 |
| Parents Own House | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Rented | -0.80 | 0.03 | -0.75 | 0.03 | 0.05 | -0.75 | 0.02 | 0.03 |
| Other Housing | -0.22 | 0.09 | -0.30 | 0.07 | 0.08 | -0.29 | 0.07 | 0.03 |
| Lives with both parents | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Only lives with Mother | -0.00 | 0.04 | -0.09 | 0.03 | 0.03 | -0.09 | 0.03 | 0.02 |
| Only lives with Father | -0.13 | 0.07 | -0.17 | 0.05 | 0.04 | -0.17 | 0.05 | 0.03 |
| Other Household | -0.76 | 0.09 | -0.87 | 0.07 | 0.06 | -0.87 | 0.07 | 0.03 |
| Father has A-level | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Father does not have A-level | -0.24 | 0.03 | -0.23 | 0.02 | 0.04 | -0.23 | 0.02 | 0.10 |
| Father Other response (A level) | -0.68 | 0.04 | -0.63 | 0.04 | 0.26 | -0.62 | 0.04 | 0.35 |
| Mother has A-level | 0.00 | - | 0.00 | - | - | 0.00 | - | - |
| Mother does not have A-level | -0.32 | 0.03 | -0.30 | 0.03 | 0.26 | -0.30 | 0.03 | 0.22 |
| Mother Other response (A level) | -0.58 | 0.04 | -0.55 | 0.04 | 0.46 | -0.56 | 0.04 | 0.36 |
| Constant | 1.14 | 0.05 | 0.94 | 0.04 | 0.15 | 0.94 | 0.04 | 0.09 |
| **Observations (min)** | 52728 | | 84717 | | | 84717 | | |
| *Number of Imputed Datasets* | | | 5 | | | 10 | | |
| *Minimum Degrees of Freedom* | | | 23.4 | | | 75.7 | | |

It is possible to see in table 6.6 in chapter six, that the GCSE performance of those with missing parental occupation most closely resembles the lower RGSC categories. This would suggest that they are under-represented in the dataset. To explain further, non-respondents similarity in terms of GCSE attainment to those with parents in lower status occupations suggests they may be more likely to also have parents in lower status occupations but to not have responded to the question in the survey. This would suggest that perhaps the non-response is differential according to RGSC. This is supported by the increasing FMI for lower RGSC categories (reported in tables 8.7 and 8.8). This creates problems as this appears to support the MNAR condition (Missing Not At Random) whereby the likelihood of an

item being missing depends on the value of the item. In this instance, it is necessary to model both the response of interest and the missingness mechanism (Carpenter and Kenward 2008). This in part appears to be attributable to the difficulty in establishing a strong missingness model that is distinct from the MOI. As was noted earlier, only three variables were identified which helped explain missingness other than those included in the substantive model of interest. To overcome this, what is needed are additional data to predict the missingness mechanism. As has been demonstrated, unfortunately there is a shortage of independent auxiliary variables that would explain item non-response in the YCS.

## Evaluation of Multiple Imputation

*"All who have been faced with missing data know that the uncomfortable truth is that, while we may have some knowledge about why data are missing we do not usually know for certain."*

(Carpenter and Kenward 2008: 8).

The multiple imputation in this chapter has used information provided from the observed data to generate information for the missing data, under the condition of being MAR given the model of missingness fitted (Carpenter and Kenward 2008). This is of course conditional upon the information about the non-missing providing accurate explanation of the patterns of missingness (the missingness mechanism). As Rubin (2004: 298) notes, the Bayesian theory behind multiple imputation implies that: "...(1) the missing data has a distribution given the observed data (the predictive distribution) and (2) this distribution depends on assumption that have been made about the model." If these assumptions do not hold true then this is problematic. That is, if the missingness of a particular variable depends upon the value of the variable, then it is said to be MNAR (Missing Not At Random). This

appears to be fairly likely with respect to parental occupation information. We might expect that the SOC for children of parents from less advantaged occupations would be less likely to complete the survey. A potential reason for this is that the literacy and prevalence of survey completion is lower for such groups, particularly as the YCS is a postal survey. There is also the need for extra information to the survey data provided to explain this but not currently used in the substantive models. If such information were available, then it might be possible to fit a Heckman model predicting item non-response to parental occupation questions then a model of interest. This is a complicated procedure though and in any case, the information is not currently available.

What is suggested is that greater attention needs to be paid to data collection and survey design. The use of weighting is currently the best means to attempt to correct for differential probability of selection with the YCS. However, it must be noted that there still remains the likelihood that there is over-representation of more advantaged young people in the YCS (based on the work within this chapter). What this suggests is that the estimates produced when analysing disadvantage by family background will tend to be conservative.

## Summary and Conclusion

This chapter has demonstrated the difficulties of addressing issues with missing data. Analysis of secondary data is heavily reliant on the quality of data collection and avoiding systematic bias as a consequence of sample design. Analysing only those respondents who have answered all questions eliminates missing data but raises the question of whether certain groups are over or under represented within the dataset. As was described in chapter five, use of sample weights can reduce this effect (Lynn 2006).

However, Halse (2007) identified that for the YCS, there was little information for sweep 1 non-response weighting and that there were large differentials in non-response weights leading to large design effects and reduced sample efficiency. The issue of differential non response remains an issue. Through the use of multiple imputation, the data has been fully examined to see whether missing data has influenced the results in a disproportionate way for certain groups of respondents. What this has shown is that the technique is heavily reliant upon the assumptions made about missing data and the auxiliary information available that may explain why the data happen to be missing.

In defence of the YCS data collection, Drew (1995) argued that there was no systematic difference in the reporting of GCSE results. Further to this, Evans and Kelley (2004) analysed non respondents and respondents using postal survey data and did not observe systematic difference between the two groups. Use of multiple imputation provides a further method to examine whether this is the case for the YCS and is effectively an extension of the sensitivity analysis for parental occupation measures conducted in chapter seven.

This analysis has demonstrated that in the absence of predictors of missing data, multiple imputation provides similar substantive conclusions to results of a complete case analyses. This is a consequence of the assumptions made about the nature of missing data. The problem concerning the missing at random (MAR) assumption is highlighted by Allison (2000: 302) when considering the multiple imputation models within this chapter. *"The problem is that it is easy to violate these conditions in practice. There are often strong reasons to suspect that the data are not missing at random. Unfortunately, not much can be done about this."* Carpenter, Kenward and Vansteelandt (2006; 2008) argue that this is not the case but it is rather difficult when the data are missing not at random (MNAR), which appears to be the case for the YCS.

In the case of the YCS (and more generally) it is more important to concentrate on learning from the mistakes in data recording and collecting the data appropriately and thoroughly in the first instance to avoid missing data. Details of the LSYPE (successor to the YCS) is to be found in Noble, Lynn and Smith (2005) which has recommendations for how these might be overcome through improved sample design.

# 9 Latent Class Analysis of GCSE subject choice

This chapter addresses the fifth and final research question: how do the subjects studied and grades attained vary between young people from different family backgrounds? The modelling in chapter seven demonstrated the persisting stratification in GCSE attainment throughout the 1990s and the particular importance of parental occupation. The logistic regression models reported that gaining 5 or more A* to C passes was more likely for those young people with parents in advantaged occupations. What is also of substantive interest is those young people who are able to pass 1-4 GCSEs (at A* to C). Those young people who narrowly miss out on gaining 5 or more passes at these grades may have specific subjects which they find more difficult and it would be valuable to identify what subjects these are and the background of these young people. There may also be combinations or types of subject that young people from particular backgrounds do better at than those from other backgrounds. There is also an interest in GCSE subject-attainment on the basis of debate about the perceived difficulty of certain subjects (Coe et al. 2008). Newton (1997) argues that in practice, comparison of the difficulty of subjects is difficult to measure. There is further debate around rising attainment and gender differences in attainment in particular GCSE subjects (Elwood 1995; Reed 1999).

As has been previously identified, threshold categorical outcomes matter, particularly the 5+ A*-C measure in terms of admission to A-level courses and signals to potential employers. More recently, 5 or more A*-C passes including English and Maths is a widely used benchmark (Leckie and Goldstein 2009).

The likelihood of young people taking or passing particular combinations of subjects may be influenced by their family background. Payne, Cheng and Witherspoon (1996) identified that A-level subject choice was associated with gender. Using data from YCS cohort 5, they noted that girls were more likely to study English and social sciences & humanities whilst boys were more likely to study mathematics and physical sciences. In a series of multilevel logistic regression models for subject choice at A-level, it was clear that prior attainment in the subjects selected at GCSE was the key factor. More recently, White (2007) describes the large gender differences in choice of particular GCE A-level subjects in 2001/02. For example, 70% of those studying A-level Physics, 30% of those studying English Literature and 30% of those studying Social Studies are male. In part, it is likely that this is due to prior attainment at GCSE in these subjects.

For some subjects, the link between grade attained at GCSE and likelihood of going on to study the subject at A-level is particularly strong (Payne, Cheng and Witherspoon 1996). The likelihood of studying mathematics or physical sciences was greatly reduced for those young people who did not gain an A grade in GCSE Mathematics (Payne, Cheng and Witherspoon 1996). Controlling for prior attainment, Asian young people were significantly more likely than white young people to select mathematics and physical sciences. Girls were significantly less likely to study A-levels in mathematics or physical sciences than boys. Such choices are often cited as an underlying source of gender segregation in occupations and other related social inequalities (McQuaid and Bond 2004) which suggests the importance of understanding patterns of attainment by GCSE subject. The technique selected to do this is latent class analysis.

## Latent Class Modelling

Latent Class models are models of conditional independence, conditional on a latent or unmeasured variable (McCutcheon 1987; McCutcheon 1996; Becker and Yang 1998; McCutcheon 2002). The essence of the analysis is due to the interest in an underlying unmeasured latent property, potentially 'educational attainment.' Latent class analysis is designed to investigate latent (unobservable) properties through multiple manifest (observable) variables such as the grade achieved in particular GCSE subjects. The concept of local independence is that if we had been able to directly measure the latent variable (educational attainment), the manifest variables (GCSE subject grades) would be independent of one another conditional on their latent class membership. Were we able to directly measure and control for educational attainment, the associations between attainment in GCSE subjects would be removed. This is because the manifest variables are highly correlated with one another (i.e. performance in one GCSE is likely to be strongly correlated with performance in another). This is expressed below:

$$\pi_{ijk}^{ABC} = \sum_{t=1}^{T} \pi_t^X \, \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X}$$

Source: (McCutcheon 1987; McCutcheon 1996; McCutcheon 2002)

*"Where the latent class probability ($\pi_t^X$) is the probability that a randomly selected observation in the sample is located in the latent class t, and the conditional probabilities (e.g. $\pi_{it}^{A|X}$) are the probabilities that a member of latent class t will be at a specified level of an observable indicator variable"* (McCutcheon 2002: 58). For example, if the latent variable is a measure of educational attainment (t=1, High

level; t=2, Low level), the first indicator variable (Ai) could be a GCSE pass in English (i=1, A-C pass; i=2, D-G pass; i=3, no pass). The conditional probability $\pi_{11}^{A|X}$ is therefore the probability that a randomly selected high level attaining young person (i.e. latent class 1) would report gaining an A-C GCSE pass.

It is recognised that the distribution of attainment among subjects will not be unassociated. *"Latent class analysis is frequently used when the researcher has a set of categorically scored observed measures that are highly interrelated"* (McCutcheon 2002: 56). It is to be expected that those young people who do well in some subjects will also do well in other subjects. I wish to examine which groupings do well in certain academic subjects. Latent class analysis is superior to k-clusters means approaches for this analysis because it fits an associated probability distribution with the model, meaning that the conditional probability of gaining a pattern of subject level attainment which can be examined and questioned, whereas nearest neighbour techniques do not permit this inquiry (McCutcheon 1987; Goodman 2002). The modelling process investigates the best fit from the data in terms of groupings within the data. One would expect to see some young people in one group having a high probability of passing all the GCSE subjects they attempt. Similarly, I would also expect to see some young people in another group with a low probability of passing any GCSE subject they attempt. These groups might be considered High and Low attainment respectively. What is of particular interest are the potential probabilistic groupings in between who may be likely to pass some subjects but not others. These are substantively interesting. In addition, I would like to see whether there are distinctive groupings of young people who gain A-C passes in some subjects and D-G passes in other GCSE subjects. Using maximum likelihood

estimation, the procedure estimates the best fit by various grouping numbers of this latent property.

Latent class models are assortative techniques to factor analysis for insofar as they model conditional dependence upon a latent variable but for categorical as opposed to continuous variables (i.e. using scale rather than categorical measures, for example Carroll and Schweiker 1951; Merrifield 1974). There were many instances of the use of factor analysis in educational research. More recently, Jæger and Holm (2007) examined the Danish Youth Longitudinal Survey (DYLS) using confirmatory factor analysis to ascertain whether parental economic, social and cultural capital explained social class differences in educational attainment. Models including latent structures are also analysed by psychologists and educationalists; for example Frederickson and Petrides (2008) fitted Multiple Indicators Multiple Causes (MIMC) models to data from a questionnaire of 901 respondents, with two latent variables (Key Stage 3 test score and GCSE attainment). They identified that Pakistani pupils had lower educational attainment in both measures controlling for IQ and socio-economic background. However, no existing latent class analysis of educational attainment (i.e. using categorical measures) could be found.

The use of latent class analysis in this thesis provides the opportunity to identify groups of young people in terms of attainment in particular subjects. This is of value in identifying middle attainment young people and which subjects. Finally, logistic regression models will be introduced to predict the likelihood of a young person being in a latent class (grouping) by family background and other independent variables considered in chapter eight.

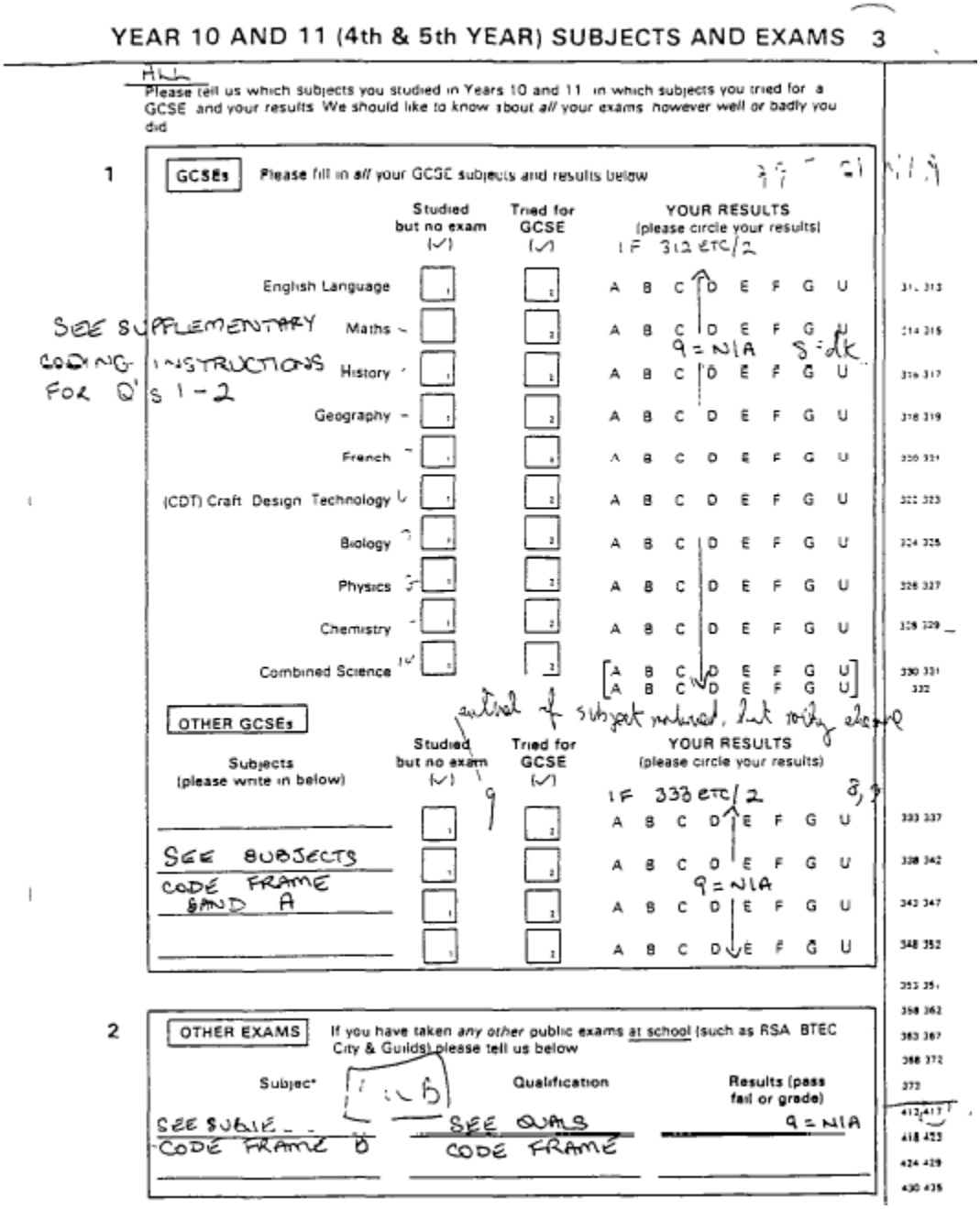
**GCSE Subjects**

As was described in chapter five, extensive work was conducted with harmonising the measures in this cohort with the dataset of Croxford, Ianelli and Shapira (2007). In producing the harmonised measures, subject and grade variables were recoded. This assisted in the preparatory work required to produce recoded forms of the GCSE subject information for this analysis of YCS6. One of the key constraints upon the model (as for other forms of modelling) is the ratio of the number of cases to the number of parameters to be estimated. Models should be parsimonious (i.e. efficient) in this respect. To include all the GCSE subjects attained would lead to the creation of a matrix of potential combinations with numerous cell counts of zero. This "sparseness" leads to problems in the estimation of parameters as the frequencies are converted into conditional probabilities. Any ratio that contains a zero cannot be estimated.

The data is recorded in the YCS questionnaire as a series of 11 subjects which are labelled accordingly and the grades attained and 7 "free format" subjects where the subject and grade attained are recorded in additional fields. This had to be recoded so that the data was produced in terms of subjects rather than in terms of GCSE results. Table 1 is a simplified 18 subject list of the subject data provided within YCS 6. As can be seen, whilst certain core subjects are studied by the majority of the sample (frequently compulsory at most schools), other subjects are optional with a high proportion of missing data because only certain students study them.

A decision was taken to simplify these subject groupings into 5 main subjects. English, Mathematics and Sciences are core subjects so were recoded into individual variables. The final two categories were humanities and other subjects. The full recode is detailed in table 2. From the 24 922 cases collected from year 11 students

in 1992, 18 053 cases were used which consisted of all those gaining at least 1 pass of G or more. The cases used were those who passed a GCSE (A to G) with distinction made between A-C and D-G (the grade A* did not exist in 1992). The software used was LEM (Vermunt 1997).



YEAR 10 AND 11 (4th & 5th YEAR) SUBJECTS AND EXAMS 3

Please tell us which subjects you studied in Years 10 and 11, in which subjects you tried for a GCSE and your results. We should like to know about *all* your exams, however well or badly you did

1 GCSEs — Please fill in *all* your GCSE subjects and results below

| | Studied but no exam (✓) | Tried for GCSE (✓) | YOUR RESULTS (please circle your results) |
|---|---|---|---|
| English Language | 1 | 2 | A B C D E F G U |
| Maths | | | A B C D E F G U |
| History | | | A B C D E F G U |
| Geography | | | A B C D E F G U |
| French | | | A B C D E F G U |
| (CDT) Craft Design Technology | | | A B C D E F G U |
| Biology | | | A B C D E F G U |
| Physics | | | A B C D E F G U |
| Chemistry | | | A B C D E F G U |
| Combined Science | | | A B C D E F G U / A B C D E F G U |

OTHER GCSEs

| Subjects (please write in below) | Studied but no exam (✓) | Tried for GCSE (✓) | YOUR RESULTS (please circle your results) |
|---|---|---|---|
| | | | A B C D E F G U |
| SEE SUBJECTS CODE FRAME BAND A | | | A B C D E F G U |
| | | | A B C D E F G U |
| | | | A B C D E F G U |

2 OTHER EXAMS — If you have taken *any other* public exams at school (such as RSA, BTEC, City & Guilds) please tell us below

| Subject* | Qualification | Results (pass, fail or grade) |
|---|---|---|
| SEE SUBJECT CODE FRAME D | SEE QUALS CODE FRAME | 9 = N/A |
| | | |

**Figure 9.1 YCS6 Questionnaire**

Source: YCS 6 Questionnaire section 3 – see Courtenay (1996b)

## Table 9.1 YCS 6 GCSE Attainment by subject

| | | A-C GCSE grade | D-G GCSE grade | No Pass | Missing | ROW TOTAL % |
|---|---|---|---|---|---|---|
| **English** | Freq. | 14,630 | 7,898 | 1,205 | 1,189 | |
| | Percent | 58.7 | 31.7 | 4.8 | 4.8 | 100.0 |
| **Maths** | Freq. | 11,443 | 9,958 | 1,113 | 2,408 | |
| | Percent | 45.9 | 40.0 | 4.5 | 9.7 | 100.0 |
| **History** | Freq. | 5,231 | 3,658 | 390 | 15,643 | |
| | Percent | 20.99 | 14.68 | 1.56 | 62.77 | 100 |
| **Geography** | Freq. | 6,192 | 5,085 | 517 | 13,128 | |
| | Percent | 24.9 | 20.4 | 2.1 | 52.7 | 100.0 |
| **French** | Freq. | 6,409 | 6,350 | 382 | 11,781 | |
| | Percent | 25.7 | 25.5 | 1.5 | 47.3 | 100 |
| **CDT** | Freq. | 4,694 | 4,454 | 618 | 15,156 | |
| | Percent | 18.8 | 17.9 | 2.5 | 60.8 | 100 |
| **Biology** | Freq. | 3,325 | 3,105 | 331 | 18,161 | |
| | Percent | 13.3 | 12.5 | 1.3 | 72.9 | 100 |
| **Physics** | Freq. | 3,399 | 2,479 | 220 | 18,824 | |
| | Percent | 13.6 | 10.0 | 0.9 | 75.5 | 100 |
| **Chemistry** | Freq. | 3,331 | 2,578 | 245 | 18,768 | |
| | Percent | 13.4 | 10.3 | 1.0 | 75.3 | 100 |
| **Science (D)** | Freq. | 5,197 | 6,412 | 735 | 12,578 | |
| | Percent | 20.9 | 25.7 | 3.0 | 50.5 | 100 |
| **Other Sciences** | Freq. | 1,397 | 1,284 | 108 | 22,133 | |
| | Percent | 5.6 | 5.2 | 0.4 | 88.8 | 100 |
| **Other Hum.** | Freq. | 4,012 | 2,616 | 220 | 18,074 | |
| | Percent | 16.1 | 10.5 | 0.9 | 72.5 | 100 |
| **Other Languages** | Freq. | 3,684 | 2,094 | 74 | 19,070 | |
| | Percent | 14.8 | 8.4 | 0.3 | 76.5 | 100 |
| **Religious Education** | Freq. | 2,663 | 1,860 | 95 | 20,304 | |
| | Percent | 10.7 | 7.5 | 0.4 | 81.5 | 100 |
| **Arts** | Freq. | 6,202 | 3,673 | 198 | 14,849 | |
| | Percent | 24.9 | 14.7 | 0.8 | 59.6 | 100 |
| **Physical Education** | Freq. | 673 | 601 | 47 | 23,601 | |
| | Percent | 2.7 | 2.4 | 0.2 | 94.7 | 100 |
| **Other** | Freq. | 2,593 | 2,612 | 237 | 19,480 | |
| | Percent | 10.4 | 10.5 | 1.0 | 78.2 | 100 |

**Table 9.2 Subject recoding into simplified form**

| Subject Group | Subject Coding |
|---|---|
| English | English |
| Maths | Maths |
| Science | Biology<br>Physics<br>Chemistry<br>Science (D)<br>Other Science |
| Humanity | History<br>Geography<br>Other Humanity<br>RE |
| Other Subject | French<br>CDT<br>Other Language<br>Arts<br>PE<br>Other |

In selecting the appropriate number of latent classes, several models were fitted and the model fit statistics compared. As can be seen in table 9.3, the four class model was the best fit given the number of parameters estimated (see AIC & BIC). Whilst the five class model had a closer overall fit, it was less parsimonious and so the four class model was chosen.

**Table 9.3 Comparison of different Latent Class models**

| Model | $X^2$ | $L^2$ | AIC | BIC | DF |
|---|---|---|---|---|---|
| 2 Class | 1518.4 | 1365.9 | 1325.9 | 1169.8 | 20 |
| 3 Class | 245.9 | 212.9 | 184.9 | 75.7 | 14 |
| **4 Class** | **6.3** | **6.4** | **-9.6** | **-72.0** | **8** |
| 5 Class | 0.9 | 0.9 | -3.1 | -18.7 | 2 |

The top half of table 9.4 describes the four category unrestricted latent class model based upon YCS6 (1992) data. In this model there are five manifest variables (subject groupings English, Maths, Science, Humanity and Other) with two outcomes each (D-G pass or A-C pass) which are listed in the first column. There is one latent variable (a proxy for "attainment") the four classes of which are listed in the next four columns (with the headings X1 – X4). The titles given to each of the latent classes are self-defined labels to broadly describe the properties of the latent class.

The first row contains the mixing proportion for the overall probability of a young person being in a particular latent class, based upon the proportion of respondents fitting the patterns listed by the link function. The link function provides the characteristics of the particular latent classes, in terms of the associated probability of a young person gaining an A-C (or D-G) pass in the subject listed in the first column given that they are in the latent class. For example in this model, approximately half (49.1%) of young people can pass nearly all the subject groups (henceforth referred to as "subjects") they study for at grades A-C. The latter part of the sentence is deduced by observing the associated probability of gaining an A-C pass for members of latent class 4. These are 98.5% for English, 95.0% for Maths, 91.7% for Science, 95.5% for Humanities and 93.8% for any other subject.

By contrast, those in latent class 2 (20.7% of respondents) are very unlikely to gain A-C passes. Latent class 1 (10.2% of respondents) are of middle attainment and better at Science whilst latent class 3 (the remaining 20.1%) are of middle attainment and better at English and Other subjects.

The lower half of the table contains measures of overall attainment for each latent class within the table (by volume and mean measures rather than probability of attainment in manifest subjects). Latent class 4 has the highest percentage passing 5 or more GCSEs at grades A-C (with and without English and Maths). For all measures used, this group has the highest level of attainment and is therefore labelled "Pass all". Correspondingly, latent class two (20% of young people) consists of those with the lowest individual subject probabilities and overall percentage chance of gaining A-C grades at GCSE and is labelled "Poor grades."[34] Perhaps the greatest substantive interest lies in latent classes one and three, as these are intermediate groups with mean number of A-C passes of 3.68 and 3.49 respectively. Latent class one, consisting of 10% of the young people in the model, is described as "Science" as this group has a much higher percentage chance of gaining an A-C pass in Science than in the other subjects. Young people with patterns of attainment described in latent class three do better in "Arts" subjects and struggle with maths and science. This is important, as the large majority of young people in latent classes 1 (science) and 3 (arts) gain 1-4 A-C GCSE passes. They have quite different requirements if we are to ask the question: "What subjects would these students need to do better in to gain 5 or more GCSE passes at A-C?" For the Science group this would be English and humanities. For the Arts group (the larger of the two) this would be maths and science.

[34] For further details of the subject combinations that have been assigned to each latent class, see appendix.

**Table 9.4 Latent Class Mixing Proportion and Link Function (YCS 6)**

| Latent Class | X1 Science | X2 Poor Grades | X3 Arts | X4 Pass All | ALL |
|---|---|---|---|---|---|
| **Mixing Proportion** | **10.2%** | **20.7%** | **20.1%** | **49.1%** | |
| **Link Function** | | | | | |
| A1 English D-G | 49.2% | 83.8% | 17.7% | 1.5% | |
| A2 English A-C | 50.8% | 16.2% | 83.3% | 98.5% | |
| B1 Maths D-G | 43.0% | 94.7% | 70.9% | 5.0% | |
| B2 Maths A-C | 57.0% | 5.3% | 29.1% | 95.0% | |
| C1 Science D-G | 29.0% | 94.6% | 84.9% | 8.3% | |
| C2 Science A-C | 71.0% | 5.4% | 15.1% | 91.7% | |
| D1 Humanity D-G | 52.1% | 91.2% | 42.7% | 4.5% | |
| D2 Humanity A-C | 47.9% | 8.8% | 57.3% | 95.5% | |
| E1 Other D-G | 37.3% | 73.0% | 29.7% | 6.2% | |
| E2 Other A-C | 62.7% | 27.0% | 70.3% | 93.8% | |
| *N* | *1227* | *3981* | *3317* | *9528* | ***18053*** |
| *% in class (modal category)* | *6.8%* | *22.1%* | *18.4%* | *52.8%* | |
| **Other Measures** | | | | | |
| **% 0** A-C Passes | 0% | 56.8% | 0% | 0% | **15.3%** |
| **% 1-4** A-C Passes | 77.1% | 43.1% | 81.9% | 2.9% | **33.7%** |
| **% 5+** A-C Passes | 22.9% | <0.1% | 18.1% | 97.1% | **50.1%** |
| % 1-4 A-C Passes (inc Eng & Maths) | 6.6% | 0% | 15.3% | 2.2% | **4.4%** |
| % 5+ A-C Passes (inc Eng & Maths) | 2.5% | 0% | 4.5% | 92.1% | **44.6%** |
| Total GCSE Points Score (A) | 34.6 | 23.2 | 34.8 | 51.1 | **39.4** |
| Mean Number of Exams Studied for (B) | 7.9 | 7.6 | 8.2 | 8.7 | **8.3** |
| Mean GCSE Points Score (A/B) | 4.4 | 3.0 | 4.2 | 5.9 | **4.8** |
| **Mode grade of mean GCSE score** | D | E | D | B | **D** |
| **Mean Number of A-C Passes** | 3.7 | 0.5 | 3.5 | 7.9 | **4.8** |
| **Mean Number of A-F Passes**[35] | 8.0 | 6.9 | 8.2 | 8.9 | **8.1** |

[35] NB: It is possible for the mean number of A-F passes to be higher than the mean number of exams studied for because these are based on two different variables within the YCS. The data has been preserved as recorded to ensure consistency with harmonised equivalents from other cohorts even though this is clearly an error.

**Model Outcomes**

The next empirical question is therefore, what is the likelihood of a young person being in a particular latent class according to their background? The reference category in these models is a young person with parents in NS-SEC1, a girl, white, attended comprehensive to 18, parents own their house, living with both parents, and whose parents do not have A-levels.

The predictors of being in the "pass all" group (latent class four) are virtually identical to those for overall attainment in chapter seven. The clear gradient by parental occupation can also be seen in the quasi-variance comparison interval plots[36] in figures 9.2 and 9.3 (Gayle and Lambert 2007). For latent classes 2 and 4, the patterns are similar to those in chapter 7. As would be expected, strong predictors of high attainment (membership of the "pass all" group) are higher grade parental occupations, using NS-SEC and RGSC, attending grammar or independent schools, whether the father or mother has A-levels and being other Asian ethnicity. Those less likely to pass all their GCSEs at grades A-C are boys, those attending comprehensive (till 16) or secondary modern schools, those living in rented accommodation or other housing type and those living with their father only.

Correspondingly, the predictors of consistently poor grades are the inverse of those that predict membership of the "pass all" group. The only notable difference is that those with fathers who gained A-levels are less likely to be in the poor grades class group whereas mothers with A-levels is not significantly associated with diminished likelihood. In this instance, mother's education increases the likelihood of being in the high attainment group but does not decrease the likelihood of being in the low

[36] Chapter 7 contains further details of interpreting and reading quasi-variance comparison intervals

attainment group. Father's education is significant in both instances. The quasi variance plots for NS-SEC and RGSC are below:

There is no significant association between parental NS-SEC/RGSC and being able to pass science but struggling with other subjects (membership of latent class 1). This can be seen below:

Those more likely to be in the science group are boys, Indians, those attending secondary modern schools, those living in rented accommodation and those living with their father only. Those less likely are Black, those attending grammar or independent schools, and those living with their mother only. There seems to be a masculine bias towards membership of this science group in terms of both the sex of the young person but also whether they live with their father only. As a note of caution, the sub-sample is small for this group.

Finally, those more likely to be in the arts group are Black or Pakistani (compared to White), attended secondary modern school or live with mother only. This reinforces the point above about a male-science and arts-female tendency which may appear to be stereotypical but is of potential interest. Boys, those attending grammar or independent schools and those with parents who have gained A-levels (fathers or mothers) are less likely to be in the group who can pass arts subjects but struggle with maths and science. The influence of NS-SEC or RGSC is mixed and suggests the likelihood of being in the arts groups is greatest for those whose parents are in middle grade occupations.

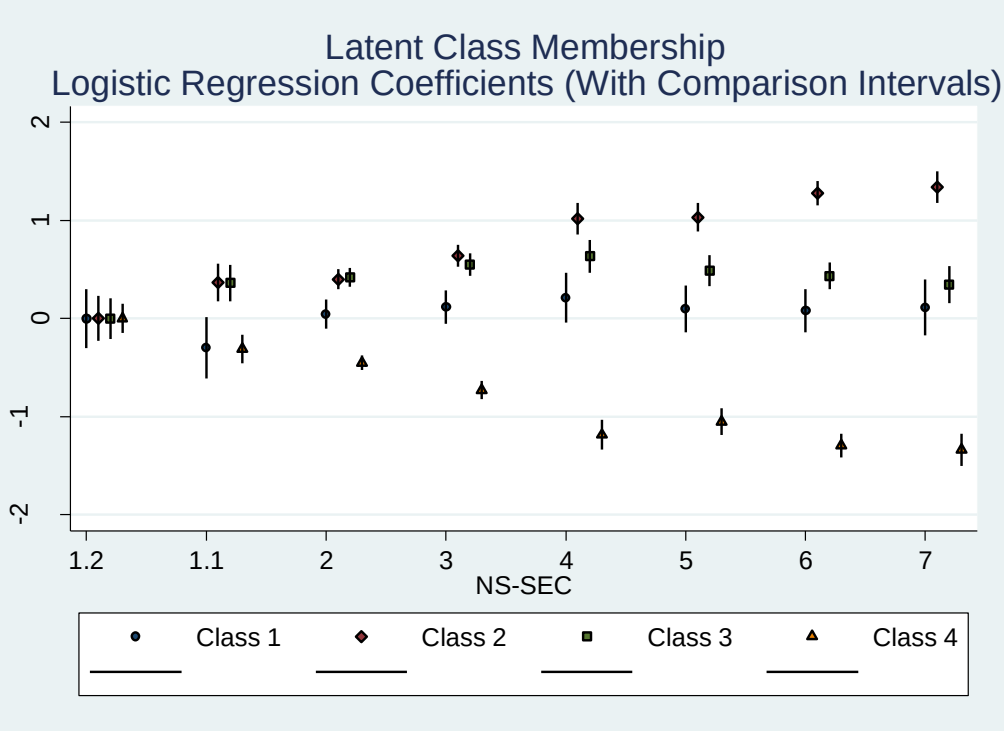



**Figure 9.2** **All Schools NS-SEC**

Other variables: gender, ethnicity, school type, housing, family structure, parental education (see also table 9.5)
Source: SN3532, n=16,637, YCS6.

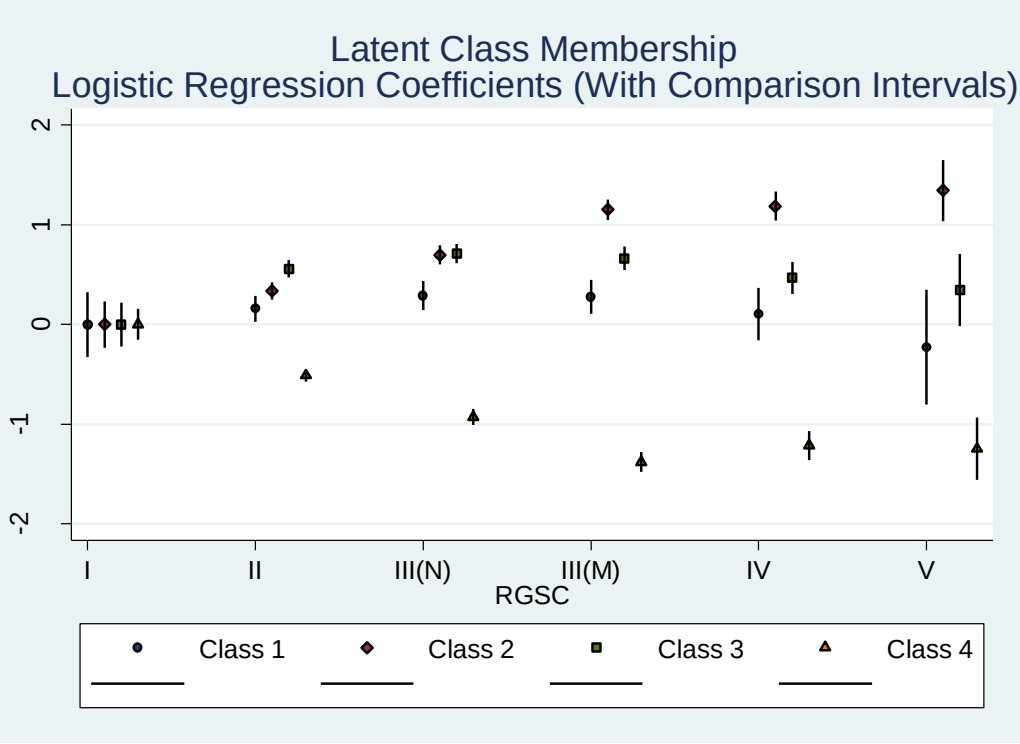



**Figure 9.3** **All Schools RGSC**

Other variables: gender, ethnicity, school type, housing, family structure, parental education (see also table 9.6)
Source: SN3532, n=16,637, YCS6.

## Table 9.5 Logistic Regression Models – All Schools - NS-SEC

| | Class 1 Science | | | Class 2 Poor Grades | | | Class 3 Arts | | | Class 4 Pass All | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NS-SEC = 1.1 | -0.30 | 0.22 | | 0.37 | 0.15 | ** | 0.36 | 0.14 | *** | -0.31 | 0.10 | *** |
| NS-SEC = 2 | 0.04 | 0.17 | | 0.40 | 0.12 | *** | 0.42 | 0.11 | *** | -0.45 | 0.08 | *** |
| NS-SEC = 3 | 0.12 | 0.18 | | 0.64 | 0.13 | *** | 0.55 | 0.12 | *** | -0.73 | 0.09 | *** |
| NS-SEC = 4 | 0.21 | 0.20 | | 1.02 | 0.14 | *** | 0.64 | 0.14 | *** | -1.19 | 0.11 | *** |
| NS-SEC = 5 | 0.10 | 0.20 | | 1.03 | 0.14 | *** | 0.49 | 0.13 | *** | -1.05 | 0.10 | *** |
| NS-SEC = 6 | 0.08 | 0.19 | | 1.28 | 0.13 | *** | 0.43 | 0.13 | *** | -1.30 | 0.10 | *** |
| NS-SEC = 7 | 0.11 | 0.21 | | 1.34 | 0.14 | *** | 0.35 | 0.15 | ** | -1.34 | 0.11 | *** |
| Boys | 1.25 | 0.09 | *** | 0.44 | 0.05 | *** | -0.53 | 0.05 | *** | -0.34 | 0.04 | *** |
| Black | -1.35 | 0.41 | *** | -0.00 | 0.21 | | 0.65 | 0.19 | *** | -0.34 | 0.18 | * |
| Indian | 0.65 | 0.23 | *** | -0.26 | 0.17 | | 0.03 | 0.18 | | -0.07 | 0.16 | |
| Pakistani | -0.05 | 0.40 | | -0.03 | 0.30 | | 0.55 | 0.26 | ** | -0.38 | 0.25 | |
| Bangladeshi | -0.14 | 0.64 | | -0.07 | 0.37 | | 0.66 | 0.47 | | -0.56 | 0.48 | |
| Other Asian | 0.06 | 0.44 | | -1.00 | 0.28 | *** | 0.04 | 0.26 | | 0.64 | 0.22 | *** |
| Other Ethnicity | 0.42 | 0.55 | | -0.65 | 0.61 | | 0.25 | 0.38 | | 0.14 | 0.37 | |
| Comprehensive (till 16) | 0.06 | 0.08 | | 0.16 | 0.05 | *** | -0.04 | 0.05 | | -0.13 | 0.05 | *** |
| Grammar | -2.15 | 0.53 | *** | -2.59 | 0.40 | *** | -0.98 | 0.19 | *** | 2.13 | 0.17 | *** |
| Secondary Modern | 0.45 | 0.22 | ** | 0.57 | 0.14 | *** | 0.32 | 0.15 | ** | -1.07 | 0.15 | *** |
| Independent | -0.84 | 0.26 | *** | -2.24 | 0.27 | *** | -0.92 | 0.17 | *** | 1.70 | 0.13 | *** |
| Rented | 0.25 | 0.12 | ** | 0.62 | 0.07 | *** | -0.00 | 0.08 | | -0.79 | 0.07 | *** |
| Other Housing | -0.12 | 0.34 | | 0.33 | 0.16 | ** | 0.33 | 0.19 | * | -0.50 | 0.16 | *** |
| Only lives with Mum | -0.34 | 0.15 | ** | -0.05 | 0.09 | | 0.25 | 0.09 | *** | -0.06 | 0.08 | |
| Only lives with Dad | 0.49 | 0.23 | ** | 0.33 | 0.16 | ** | -0.03 | 0.18 | | -0.50 | 0.16 | *** |
| Other Household | -0.22 | 1.15 | | 0.33 | 0.77 | | | | | 0.87 | 0.67 | |
| Dad has A-level (mod) | -0.15 | 0.13 | | -0.32 | 0.08 | *** | -0.23 | 0.08 | *** | 0.40 | 0.06 | *** |
| Mum has A-level (mod) | -0.25 | 0.14 | * | -0.10 | 0.08 | | -0.37 | 0.09 | *** | 0.35 | 0.06 | *** |
| Constant | -3.35 | 0.17 | *** | -1.90 | 0.13 | *** | -1.51 | 0.12 | *** | 0.62 | 0.08 | *** |
| **Observations** | 16637 | | | 16637 | | | 16628 | | | 16637 | | |
| **Pseudo $R^2$** | 0.07 | | | 0.12 | | | 0.04 | | | 0.16 | | |

## Table 9.6 Logistic Regression Models – All Schools - RGSC

| | Class 1 Science | | | Class 2 Poor Grades | | | Class 3 Arts | | | Class 4 Pass All | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RGSC = 2 | 0.16 | 0.18 | | 0.34 | 0.12 | *** | 0.56 | 0.12 | *** | -0.51 | 0.08 | *** |
| RGSC = 3.1 | 0.29 | 0.18 | | 0.70 | 0.13 | *** | 0.71 | 0.12 | *** | -0.93 | 0.09 | *** |
| RGSC = 3.2 | 0.28 | 0.19 | | 1.15 | 0.13 | *** | 0.66 | 0.13 | *** | -1.38 | 0.10 | *** |
| RGSC = 4 | 0.10 | 0.22 | | 1.19 | 0.14 | *** | 0.47 | 0.14 | *** | -1.21 | 0.11 | *** |
| RGSC = 5 | -0.23 | 0.34 | | 1.34 | 0.20 | *** | 0.35 | 0.22 | | -1.25 | 0.18 | *** |
| Boys | 1.24 | 0.09 | *** | 0.43 | 0.05 | *** | -0.54 | 0.05 | *** | -0.32 | 0.04 | *** |
| Black | -1.31 | 0.41 | *** | -0.01 | 0.21 | | 0.68 | 0.19 | *** | -0.36 | 0.18 | ** |
| Indian | 0.70 | 0.23 | *** | -0.25 | 0.17 | | 0.03 | 0.18 | | -0.09 | 0.16 | |
| Pakistani | 0.03 | 0.40 | | -0.04 | 0.30 | | 0.57 | 0.26 | ** | -0.41 | 0.25 | * |
| Bangladeshi | -0.04 | 0.63 | | -0.08 | 0.37 | | 0.67 | 0.47 | | -0.56 | 0.47 | |
| Other Asian | 0.04 | 0.43 | | -1.01 | 0.28 | *** | 0.00 | 0.26 | | 0.68 | 0.21 | *** |
| Other Ethnicity | 0.44 | 0.54 | | -0.72 | 0.62 | | 0.24 | 0.39 | | 0.19 | 0.37 | |
| Comprehensive (till 16) | 0.06 | 0.08 | | 0.16 | 0.05 | *** | -0.04 | 0.05 | | -0.14 | 0.04 | *** |
| Grammar | -2.16 | 0.54 | *** | -2.60 | 0.40 | *** | -0.98 | 0.19 | *** | 2.14 | 0.17 | *** |
| Secondary Modern | 0.46 | 0.22 | ** | 0.58 | 0.15 | *** | 0.34 | 0.15 | ** | -1.09 | 0.15 | *** |
| Independent | -0.84 | 0.26 | *** | -2.23 | 0.27 | *** | -0.91 | 0.17 | *** | 1.68 | 0.13 | *** |
| Rented | 0.29 | 0.12 | ** | 0.66 | 0.07 | *** | 0.01 | 0.08 | | -0.84 | 0.07 | *** |
| Other Housing | -0.10 | 0.34 | | 0.36 | 0.16 | ** | 0.35 | 0.19 | * | -0.54 | 0.16 | *** |
| Only lives with Mum | -0.33 | 0.15 | ** | -0.03 | 0.09 | | 0.25 | 0.09 | *** | -0.08 | 0.08 | |
| Only lives with Dad | 0.48 | 0.23 | ** | 0.28 | 0.16 | * | -0.04 | 0.18 | | -0.46 | 0.16 | *** |
| Other Household | -0.19 | 1.13 | | 0.33 | 0.70 | | | | | 0.79 | 0.66 | |
| Dad has A-level (mod) | -0.14 | 0.12 | | -0.35 | 0.08 | *** | -0.22 | 0.08 | *** | 0.41 | 0.06 | *** |
| Mum has A-level (mod) | -0.25 | 0.14 | * | -0.09 | 0.08 | | -0.36 | 0.09 | *** | 0.33 | 0.06 | *** |
| Constant | -3.49 | 0.18 | *** | -1.84 | 0.13 | *** | -1.65 | 0.12 | *** | 0.70 | 0.09 | *** |
| **Observations** | 16637 | | | 16637 | | | 16628 | | | 16637 | | |
| **Pseudo $R^2$** | 0.07 | | | 0.12 | | | 0.04 | | | 0.15 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Logistic Regression Models – only comprehensive schools

The following are logistic regression models predicting membership of the particular latent classes, restricted to those pupils who attended comprehensive schools (with admissions till 16 or till 18). The reason for doing so was largely concerning the unequal distribution of pupils according to school type into the latent classes (see table 9.7). Those at grammar and independent schools are largely in latent class 4 ("pass all"). Students at secondary modern and other state funded schools constitute a small percentage of the total number of young people and it is substantively interesting to look at the variation purely for those at comprehensive schools. The majority of young people do not attend selective schools (Goldstein and Thomas 1996).

**Table 9.7 Only comprehensive schools**

| | **Latent Class** | **1** | **2** | **3** | **4** | **Total** |
|---|---|---|---|---|---|---|
| **School type** | | | | | | |
| Comprehensive to 16 | *n* | 394 | 1,362 | 1,008 | 2,246 | 5,010 |
| | *Row %* | *7.86* | *27.19* | *20.12* | *44.83* | *100.00* |
| | | | | | | |
| Comprehensive to 18 | *n* | 768 | 2,465 | 2,114 | 5,521 | 10,868 |
| | *Row %* | *7.07* | *22.68* | *19.45* | *50.80* | *100.00* |
| | | | | | | |
| Grammar | *n* | 6 | 9 | 42 | 654 | 711 |
| | *Row %* | *0.84* | *1.27* | *5.91* | *91.98* | *100.00* |
| | | | | | | |
| Secondary modern | *n* | 34 | 116 | 87 | 101 | 338 |
| | *Row %* | *10.06* | *34.32* | *25.74* | *29.88* | *100.00* |
| | | | | | | |
| Other state funded | *n* | 5 | 13 | 12 | 24 | 54 |
| | *Row %* | *9.26* | *24.07* | *22.22* | *44.44* | *100.00* |
| | | | | | | |
| Independent | *n* | 20 | 16 | 54 | 982 | 1,072 |
| | *Row %* | *1.87* | *1.49* | *5.04* | *91.60* | *100.00* |
| | | | | | | |
| **Total** | **n** | **1,227** | **3,981** | **3,317** | **9,528** | **18,053** |
| | *Row %* | *6.80* | *22.05* | *18.37* | *52.78* | *100.00* |

It was not possible to do a likelihood ratio test as the models were non-nested when the non-comprehensive schools were removed. The AIC and BIC for the models produced using purely comprehensive schools was lower than for the other

two models due to the relative efficiency (parameters estimated relative to degrees of freedom remaining). This is to demonstrate that these are broadly comparable models and whilst a small degree of explanatory power is lost, the substantive justification for doing so is good as it is of interest to understand the variables purely in the comprehensive school environment. Once the non-comprehensive schools had been removed, 15 878 cases remained. Tables 9.9 and 9.10 provide details of the model coefficients.

**Table 9.8 Comparative Likelihood**

| Model | Obs | ll(null) | ll(model) | df | Deviance | AIC | BIC | Pseudo $R^2$ |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| NS-SEC All Schools | 16637 | -4063.984 | -3794.783 | 26 | 7589.566 | 7641.566 | 7842.27 | 0.0662 |
| NS-SEC All Schools (no School dummies) | 16637 | -4063.984 | -3848.610 | 22 | 7697.22 | 7741.220 | 7911.046 | 0.0530 |
| NS-SEC Just Comps | 14576 | -3768.24 | -3571.217 | 22 | 7142.434 | 7186.434 | 7353.351 | 0.0523 |

**Table 9.9 Logistic Regression Models – Just Comprehensives - NS-SEC**

| | Class 1 Science | | | Class 2 Poor Grades | | | Class 3 Arts | | | Class 4 Pass All | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NS-SEC = 1.1 | -0.24 | 0.22 | | 0.45 | 0.15 | *** | 0.25 | 0.14 | * | -0.31 | 0.11 | *** |
| NS-SEC = 2 | -0.01 | 0.17 | | 0.48 | 0.13 | *** | 0.28 | 0.12 | ** | -0.42 | 0.08 | *** |
| NS-SEC = 3 | 0.07 | 0.18 | | 0.76 | 0.13 | *** | 0.45 | 0.13 | *** | -0.76 | 0.09 | *** |
| NS-SEC = 4 | 0.21 | 0.20 | | 1.02 | 0.15 | *** | 0.62 | 0.14 | *** | -1.19 | 0.11 | *** |
| NS-SEC = 5 | 0.07 | 0.20 | | 1.17 | 0.14 | *** | 0.37 | 0.14 | *** | -1.09 | 0.11 | *** |
| NS-SEC = 6 | 0.08 | 0.19 | | 1.36 | 0.14 | *** | 0.39 | 0.13 | *** | -1.35 | 0.10 | *** |
| NS-SEC = 7 | 0.02 | 0.21 | | 1.47 | 0.15 | *** | 0.27 | 0.15 | * | -1.39 | 0.12 | *** |
| Boys | 1.30 | 0.09 | *** | 0.49 | 0.05 | *** | -0.60 | 0.05 | *** | -0.37 | 0.04 | *** |
| Black | -1.30 | 0.41 | *** | -0.00 | 0.21 | | 0.64 | 0.19 | *** | -0.33 | 0.19 | * |
| Indian | 0.67 | 0.23 | *** | -0.30 | 0.17 | * | -0.08 | 0.17 | | 0.04 | 0.15 | |
| Pakistani | -0.30 | 0.43 | | -0.02 | 0.29 | | 0.45 | 0.27 | | -0.24 | 0.24 | |
| Bangladeshi | -0.11 | 0.64 | | -0.10 | 0.37 | | 0.76 | 0.48 | | -0.66 | 0.55 | |
| Other Asian | 0.05 | 0.47 | | -0.98 | 0.29 | *** | 0.11 | 0.27 | | 0.61 | 0.23 | *** |
| Other Ethnicity | 0.45 | 0.55 | | -0.66 | 0.62 | | 0.39 | 0.40 | | 0.07 | 0.40 | |
| Rented | 0.26 | 0.12 | ** | 0.62 | 0.07 | *** | -0.05 | 0.08 | | -0.76 | 0.08 | *** |
| Other Housing | -0.25 | 0.32 | | 0.33 | 0.17 | * | 0.35 | 0.19 | * | -0.49 | 0.16 | *** |
| Only lives with Mum | -0.27 | 0.15 | * | -0.07 | 0.09 | | 0.31 | 0.09 | *** | -0.10 | 0.08 | |
| Only lives with Dad | 0.37 | 0.24 | | 0.35 | 0.16 | ** | -0.01 | 0.18 | | -0.48 | 0.15 | *** |
| Other Household | 0.44 | 1.09 | | 0.10 | 1.03 | | | | | 0.93 | 0.81 | |
| Dad has A-level (mod) | -0.07 | 0.12 | | -0.33 | 0.08 | *** | -0.21 | 0.08 | ** | 0.38 | 0.06 | *** |
| Mum has A-level (mod) | -0.29 | 0.14 | ** | -0.11 | 0.09 | | -0.39 | 0.09 | *** | 0.38 | 0.06 | *** |
| Constant | -3.34 | 0.17 | *** | -1.95 | 0.13 | *** | -1.41 | 0.12 | *** | 0.61 | 0.09 | *** |
| **Observations** | 14576 | | | 14576 | | | 14569 | | | 14576 | | |
| **Pseudo $R^2$** | 0.06 | | | 0.06 | | | 0.03 | | | 0.08 | | |

**Table 9.10 Logistic Regression Models – Just Comprehensives - RGSC**

| | Class 1 Science | | | Class 2 Poor Grades | | | Class 3 Arts | | | Class 4 Pass All | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RGSC = 2 | 0.13 | 0.18 | | 0.44 | 0.13 | *** | 0.38 | 0.12 | *** | -0.48 | 0.09 | *** |
| RGSC = 3.1 | 0.32 | 0.18 | * | 0.85 | 0.13 | *** | 0.58 | 0.13 | *** | -0.99 | 0.09 | *** |
| RGSC = 3.2 | 0.26 | 0.19 | | 1.31 | 0.14 | *** | 0.52 | 0.13 | *** | -1.42 | 0.10 | *** |
| RGSC = 4 | 0.13 | 0.22 | | 1.28 | 0.15 | *** | 0.34 | 0.15 | ** | -1.22 | 0.11 | *** |
| RGSC = 5 | -0.33 | 0.34 | | 1.50 | 0.20 | *** | 0.25 | 0.22 | | -1.30 | 0.19 | *** |
| Boys | 1.30 | 0.09 | *** | 0.48 | 0.05 | *** | -0.60 | 0.05 | *** | -0.36 | 0.04 | *** |
| Black | -1.26 | 0.41 | *** | -0.01 | 0.21 | | 0.67 | 0.19 | *** | -0.35 | 0.19 | * |
| Indian | 0.70 | 0.23 | *** | -0.29 | 0.17 | * | -0.07 | 0.17 | | 0.01 | 0.15 | |
| Pakistani | -0.23 | 0.43 | | -0.01 | 0.29 | | 0.47 | 0.28 | * | -0.27 | 0.25 | |
| Bangladeshi | -0.03 | 0.64 | | -0.09 | 0.38 | | 0.78 | 0.47 | | -0.68 | 0.53 | |
| Other Asian | 0.04 | 0.47 | | -0.98 | 0.29 | *** | 0.08 | 0.27 | | 0.65 | 0.22 | *** |
| Other Ethnicity | 0.44 | 0.55 | | -0.73 | 0.63 | | 0.37 | 0.40 | | 0.14 | 0.40 | |
| Rented | 0.30 | 0.12 | ** | 0.65 | 0.07 | *** | -0.03 | 0.08 | | -0.82 | 0.08 | *** |
| Other Housing | -0.23 | 0.32 | | 0.36 | 0.17 | ** | 0.37 | 0.19 | * | -0.54 | 0.16 | *** |
| Only lives with Mum | -0.27 | 0.15 | * | -0.04 | 0.09 | | 0.30 | 0.09 | *** | -0.12 | 0.08 | |
| Only lives with Dad | 0.36 | 0.24 | | 0.30 | 0.16 | * | -0.02 | 0.18 | | -0.42 | 0.15 | *** |
| Other Household | 0.44 | 1.09 | | 0.08 | 0.90 | | | | | 0.89 | 0.80 | |
| Dad has A-level (mod) | -0.05 | 0.12 | | -0.34 | 0.08 | *** | -0.20 | 0.08 | ** | 0.38 | 0.06 | *** |
| Mum has A-level (mod) | -0.29 | 0.14 | ** | -0.10 | 0.09 | | -0.39 | 0.09 | *** | 0.37 | 0.07 | *** |
| Constant | -3.51 | 0.18 | *** | -1.93 | 0.13 | *** | -1.50 | 0.12 | *** | 0.68 | 0.09 | *** |
| **Observations** | 14576 | | | 14576 | | | 14569 | | | 14576 | | |
| **Pseudo $R^2$** | 0.06 | | | 0.06 | | | 0.03 | | | 0.08 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

The removal of the non-comprehensives has not altered the relative influence of the remaining independent variables within the models. This is an important finding as it demonstrates the importance of family background factors in the academic attainment among those at comprehensive schools.

## Multinomial logistic regression models for comprehensive school pupils only

These models have the form of predicting whether a young person is more likely or less likely to be in the following latent classes (comparison category) when compared with a base category. Model 1a predicts how much more likely a young person is to be in the Science latent class than the Pass All class (labelled as 1 vs. 4 Base).[37]

**Table 9.11 Multinomial comparison table**

| *Page 1* | **COMPARISON CATEGORY** | | | **BASE CATEGORY** | |
|---|---|---|---|---|---|
| Model 1a | Science | class 1 | vs. | Pass All | class 4 |
| Model 1b | Poor Grades | class 2 | vs. | Pass All | class 4 |
| Model 1c | Arts | class 3 | vs. | Pass All | class 4 |
| | | | | | |
| *Page 2* | | | | | |
| Model 2a | Science | class 1 | vs. | Arts | class 3 |
| Model 2b | Poor Grades | class 2 | vs. | Arts | class 3 |
| Model 3 | Science | class 1 | vs. | Poor Grades | class 2 |

These models have been completed for the comprehensive schools so that comparison can be made between pupils in a non-selective environment. See details on the following pages.

The first three models (1a, 1b & 1c) are all concerned with comparisons with the Pass All group (latent class 4). Young people in lower NS-SEC groups are more likely to be in the Science, Poor Grades or Arts group than the Pass All group. This is clearly demonstrated in the quasi variance plots on pages 26 and 27. The relationship is similar but perhaps less clearly demonstrated using RGSC.

[37] A series of multinomial logistic regression models were fitted predicting class membership on the basis of socio-economic characteristics of young people attending all school types. This analysis was conducted for the sake of completeness but has not been reported. The substantive findings are similar to the analysis of comprehensive schools pupils.

Boys are more likely than girls to be in the Science or Poor Grades groups but less likely to be in the Arts group, than in the Pass All group. Black young people are more likely than White young people to be in the Arts group and less likely to be in the Science group than in the Pass All group. Indians are more likely than Whites to be in the Science group. Other Asians are less likely to be in the Poor Grades group than the Pass All group (compared to Whites). Pakistanis are more likely than Whites to be in the Arts group than the Pass All group.

Living in rented accommodation is similar to the effect of being in a lower NS-SEC with consistently higher likelihood of being categories other than the Pass All group than those whose parents own their house. Living in Other Housing increases the likelihood of being in the Poor Grades or Arts group than the Pass All group.

Living with Father only increases the likelihood of being in the Science or Poor Grades groups than the Pass All group. Living with Mother only increases the likelihood of being in the Arts group than the Pass All group (but less likely to be in the Science group). Young people whose parents have A-levels are less likely to be in any group than the Pass All group.

These factors all appear consistent with the logistic regression models prior to this section. What predicts being in the Science group compared to being the in the Arts group? Model 2a suggests that boys are more likely than girls to be in the Science group. Black young people are more likely to be in the Arts group than White young people. Indians are more likely to also be in the Science group. Those living in rented accommodation are more likely to be in either the Science group or the Poor Grades group than the Arts group.

Model 2b suggests that young people living with Mother only are more likely than those living with both parents to be in the Arts group than the Poor Grades group. Finally, model 3 suggests that Boys and Indian young people are more likely (than girls and White young people) to be in the Science group than the Poor Grades group.

Not every permutation is discussed but further details can be seen in the models. It is also important to note that the likelihoods of one group versus another can be compared in the opposite direction by reversing the sign before the coefficient. These points may sound initially confusing but there are consistent patterns across these models. It is apparent that boys appear to be more often in the Science group when compared with all other groups. Contrastingly, Black young people appear to be more likely to be in all other groups than the Science group. In policy terms this is potentially important as there appears to be systemic issues for different groups of young people. There seems to be a male and Indian tendency towards the Sciences and a female and Black preference for the Arts subjects (among those not in the Pass All group). Furthermore, which parent you live with (if not both) emphasises this with living with Fathers only being more associated with Science and Mothers only being associated with Arts.

## Table 9.12 Multinomial Logistic Regression models – Comprehensive Schools only – NS-SEC

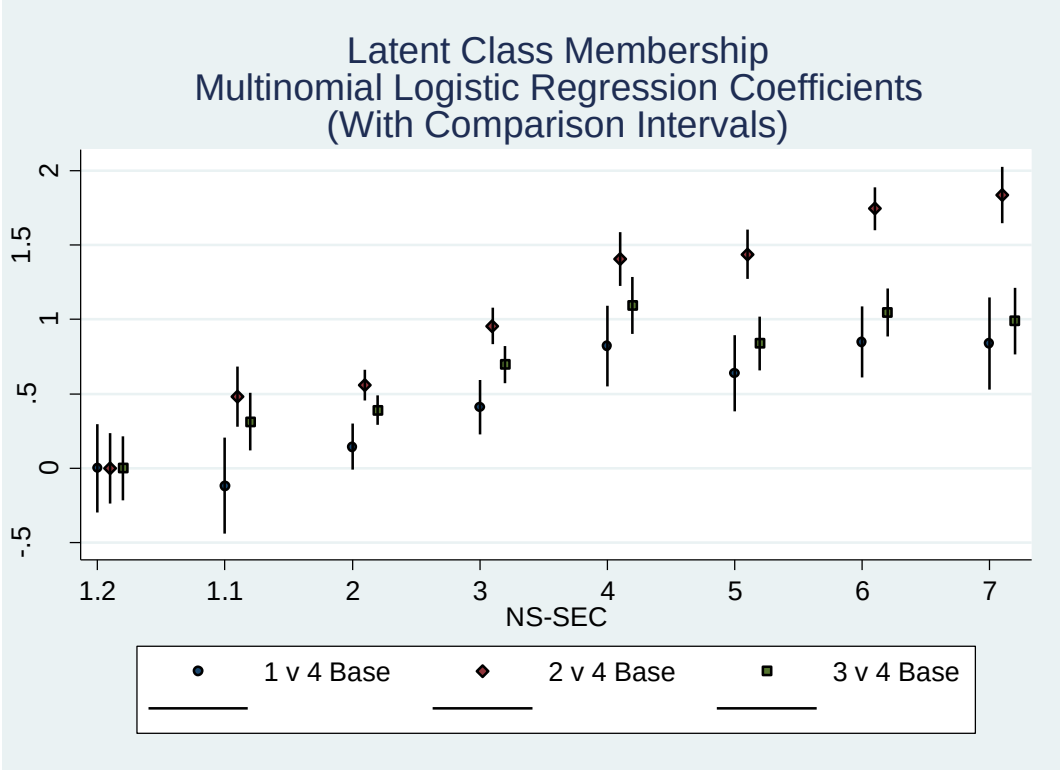
| | Model 1a 1 v Base 4 | | | Model 1b 2 v Base 4 | | | Model 1c 3 v Base 4 | | | Model 2a 1 v Base 3 | | | Model 2b 2 v Base 3 | | | Model 3 1 v Base 2 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NS-SEC = 1.1 | -0.12 | 0.22 | | 0.48 | 0.16 | *** | 0.31 | 0.14 | ** | -0.43 | 0.25 | * | 0.17 | 0.20 | | -0.60 | 0.26 | ** |
| NS-SEC = 2 | 0.15 | 0.17 | | 0.56 | 0.13 | *** | 0.39 | 0.12 | *** | -0.24 | 0.19 | | 0.17 | 0.16 | | -0.41 | 0.20 | ** |
| NS-SEC = 3 | 0.41 | 0.18 | ** | 0.95 | 0.14 | *** | 0.70 | 0.13 | *** | -0.29 | 0.21 | | 0.26 | 0.17 | | -0.54 | 0.21 | *** |
| NS-SEC = 4 | 0.82 | 0.21 | *** | 1.41 | 0.15 | *** | 1.10 | 0.15 | *** | -0.27 | 0.23 | | 0.31 | 0.19 | * | -0.58 | 0.24 | ** |
| NS-SEC = 5 | 0.64 | 0.20 | *** | 1.44 | 0.15 | *** | 0.84 | 0.14 | *** | -0.20 | 0.23 | | 0.60 | 0.18 | *** | -0.80 | 0.23 | *** |
| NS-SEC = 6 | 0.85 | 0.20 | *** | 1.74 | 0.14 | *** | 1.05 | 0.14 | *** | -0.20 | 0.22 | | 0.70 | 0.18 | *** | -0.89 | 0.22 | *** |
| NS-SEC = 7 | 0.84 | 0.22 | *** | 1.84 | 0.16 | *** | 0.99 | 0.16 | *** | -0.15 | 0.25 | | 0.85 | 0.19 | *** | -1.00 | 0.24 | *** |
| Boys | 1.44 | 0.09 | *** | 0.57 | 0.05 | *** | -0.25 | 0.06 | *** | 1.69 | 0.10 | *** | 0.82 | 0.06 | *** | 0.87 | 0.09 | *** |
| Black | -1.06 | 0.43 | ** | 0.20 | 0.24 | | 0.66 | 0.21 | *** | -1.72 | 0.44 | *** | -0.46 | 0.24 | * | -1.26 | 0.44 | *** |
| Indian | 0.57 | 0.25 | ** | -0.22 | 0.19 | | -0.09 | 0.19 | | 0.66 | 0.26 | ** | -0.13 | 0.21 | | 0.79 | 0.26 | *** |
| Pakistani | -0.17 | 0.45 | | 0.12 | 0.31 | | 0.48 | 0.29 | | -0.65 | 0.47 | | -0.36 | 0.35 | | -0.30 | 0.47 | |
| Bangladeshi | 0.37 | 0.77 | | 0.43 | 0.54 | | 1.06 | 0.65 | | -0.69 | 0.71 | | -0.63 | 0.46 | | -0.06 | 0.65 | |
| Other Asian | -0.32 | 0.48 | | -1.10 | 0.31 | *** | -0.25 | 0.28 | | -0.07 | 0.53 | | -0.85 | 0.36 | ** | 0.78 | 0.53 | |
| Other Ethnicity | 0.34 | 0.60 | | -0.54 | 0.66 | | 0.24 | 0.42 | | 0.09 | 0.62 | | -0.78 | 0.65 | | 0.88 | 0.73 | |
| Rented | 0.81 | 0.13 | *** | 0.92 | 0.09 | *** | 0.49 | 0.09 | *** | 0.31 | 0.14 | ** | 0.43 | 0.09 | *** | -0.11 | 0.13 | |
| Other Housing | 0.07 | 0.33 | | 0.54 | 0.19 | *** | 0.57 | 0.21 | *** | -0.50 | 0.36 | | -0.04 | 0.22 | | -0.47 | 0.35 | |
| Only lives with Mum | -0.21 | 0.15 | | 0.01 | 0.10 | | 0.29 | 0.09 | *** | -0.50 | 0.17 | *** | -0.28 | 0.11 | *** | -0.21 | 0.16 | |
| Only lives with Dad | 0.66 | 0.25 | *** | 0.55 | 0.18 | *** | 0.29 | 0.20 | | 0.36 | 0.28 | | 0.25 | 0.21 | | 0.11 | 0.26 | |
| Other Household | -0.11 | 1.15 | | -0.48 | 1.02 | | -44.94 | 0.00 | | 22.83 | . | | 22.45 | 1.28 | *** | 0.37 | 1.28 | |
| Dad has A-level (mod) | -0.26 | 0.12 | ** | -0.44 | 0.08 | *** | -0.35 | 0.08 | *** | 0.09 | 0.14 | | -0.09 | 0.11 | | 0.18 | 0.14 | |
| Mum has A-level (mod) | -0.45 | 0.14 | *** | -0.27 | 0.09 | *** | -0.50 | 0.09 | *** | 0.05 | 0.16 | | 0.23 | 0.12 | ** | -0.17 | 0.16 | |
| Constant | -2.97 | 0.17 | *** | -1.64 | 0.13 | *** | -1.21 | 0.12 | *** | -1.76 | 0.20 | *** | -0.43 | 0.16 | *** | -1.33 | 0.20 | *** |
| **Observations** | 14576 | | | 14576 | | | 14576 | | | 14576 | | | 14576 | | | 14576 | | |
| **Pseudo $R^2$** | 0.07 | | | 0.07 | | | 0.07 | | | 0.07 | | | 0.07 | | | 0.07 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01

## Table 9.13 Multinomial Logistic Regression models – Comprehensive Schools only – RGSC

| | Model 1a 1 v Base 4 | | | Model 1b 2 v Base 4 | | | Model 1c 3 v Base 4 | | | Model 2a 1 v Base 3 | | | Model 2b 2 v Base 3 | | | Model 3 1 v Base 2 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RGSC = 2 | 0.29 | 0.18 | * | 0.54 | 0.13 | *** | 0.49 | 0.12 | *** | -0.19 | 0.20 | | 0.05 | 0.16 | | -0.25 | 0.21 | |
| RGSC = 3.1 | 0.76 | 0.18 | *** | 1.13 | 0.14 | *** | 0.92 | 0.13 | *** | -0.16 | 0.21 | | 0.21 | 0.17 | | -0.37 | 0.22 | * |
| RGSC = 3.2 | 1.02 | 0.19 | *** | 1.72 | 0.14 | *** | 1.17 | 0.13 | *** | -0.14 | 0.22 | | 0.55 | 0.17 | *** | -0.70 | 0.22 | *** |
| RGSC = 4 | 0.77 | 0.22 | *** | 1.56 | 0.15 | *** | 0.89 | 0.15 | *** | -0.12 | 0.25 | | 0.67 | 0.19 | *** | -0.79 | 0.25 | *** |
| RGSC = 5 | 0.43 | 0.36 | | 1.74 | 0.23 | *** | 0.90 | 0.24 | *** | -0.48 | 0.39 | | 0.84 | 0.26 | *** | -1.32 | 0.36 | *** |
| Boys | 1.43 | 0.09 | *** | 0.56 | 0.05 | *** | -0.26 | 0.06 | *** | 1.69 | 0.10 | *** | 0.82 | 0.06 | *** | 0.87 | 0.09 | *** |
| Black | -1.02 | 0.43 | ** | 0.21 | 0.24 | | 0.69 | 0.21 | *** | -1.71 | 0.44 | *** | -0.48 | 0.24 | ** | -1.22 | 0.44 | *** |
| Indian | 0.61 | 0.25 | ** | -0.20 | 0.19 | | -0.07 | 0.19 | | 0.69 | 0.27 | *** | -0.13 | 0.21 | | 0.81 | 0.26 | *** |
| Pakistani | -0.09 | 0.45 | | 0.14 | 0.31 | | 0.51 | 0.29 | * | -0.61 | 0.47 | | -0.37 | 0.34 | | -0.24 | 0.47 | |
| Bangladeshi | 0.44 | 0.76 | | 0.44 | 0.54 | | 1.08 | 0.63 | * | -0.64 | 0.71 | | -0.64 | 0.47 | | 0.00 | 0.66 | |
| Other Asian | -0.35 | 0.48 | | -1.12 | 0.30 | *** | -0.29 | 0.28 | | -0.06 | 0.53 | | -0.82 | 0.36 | ** | 0.77 | 0.53 | |
| Other Ethnicity | 0.30 | 0.60 | | -0.63 | 0.67 | | 0.19 | 0.42 | | 0.11 | 0.62 | | -0.82 | 0.67 | | 0.92 | 0.74 | |
| Rented | 0.88 | 0.13 | *** | 0.98 | 0.09 | *** | 0.55 | 0.09 | *** | 0.34 | 0.14 | ** | 0.44 | 0.09 | *** | -0.10 | 0.13 | |
| Other Housing | 0.12 | 0.33 | | 0.59 | 0.19 | *** | 0.62 | 0.21 | *** | -0.50 | 0.36 | | -0.03 | 0.22 | | -0.47 | 0.34 | |
| Only lives with Mum | -0.18 | 0.15 | | 0.04 | 0.10 | | 0.31 | 0.09 | *** | -0.49 | 0.17 | *** | -0.26 | 0.11 | ** | -0.23 | 0.16 | |
| Only lives with Dad | 0.61 | 0.25 | ** | 0.48 | 0.18 | *** | 0.25 | 0.20 | | 0.36 | 0.28 | | 0.22 | 0.21 | | 0.14 | 0.26 | |
| Other Household | -0.12 | 1.16 | | -0.48 | 0.92 | | -29.85 | 0.51 | *** | 22.73 | . | | 22.37 | 1.21 | *** | 0.36 | 1.21 | |
| Dad has A-level (mod) | -0.24 | 0.12 | ** | -0.45 | 0.08 | *** | -0.34 | 0.08 | *** | 0.10 | 0.14 | | -0.10 | 0.11 | | 0.20 | 0.14 | |
| Mum has A-level (mod) | -0.43 | 0.14 | *** | -0.26 | 0.09 | *** | -0.49 | 0.09 | *** | 0.05 | 0.16 | | 0.23 | 0.11 | ** | -0.18 | 0.16 | |
| Constant | -3.15 | 0.18 | *** | -1.65 | 0.13 | *** | -1.31 | 0.12 | *** | -1.84 | 0.21 | *** | -0.34 | 0.17 | ** | -1.50 | 0.21 | *** |
| **Observations** | 14576 | | | 14576 | | | 14576 | | | 14576 | | | 14576 | | | 14576 | | |
| **Pseudo $R^2$** | 0.07 | | | 0.07 | | | 0.07 | | | 0.07 | | | 0.07 | | | 0.07 | | |

Standard errors in second column, * p<.10, ** p<.05, *** p<.01



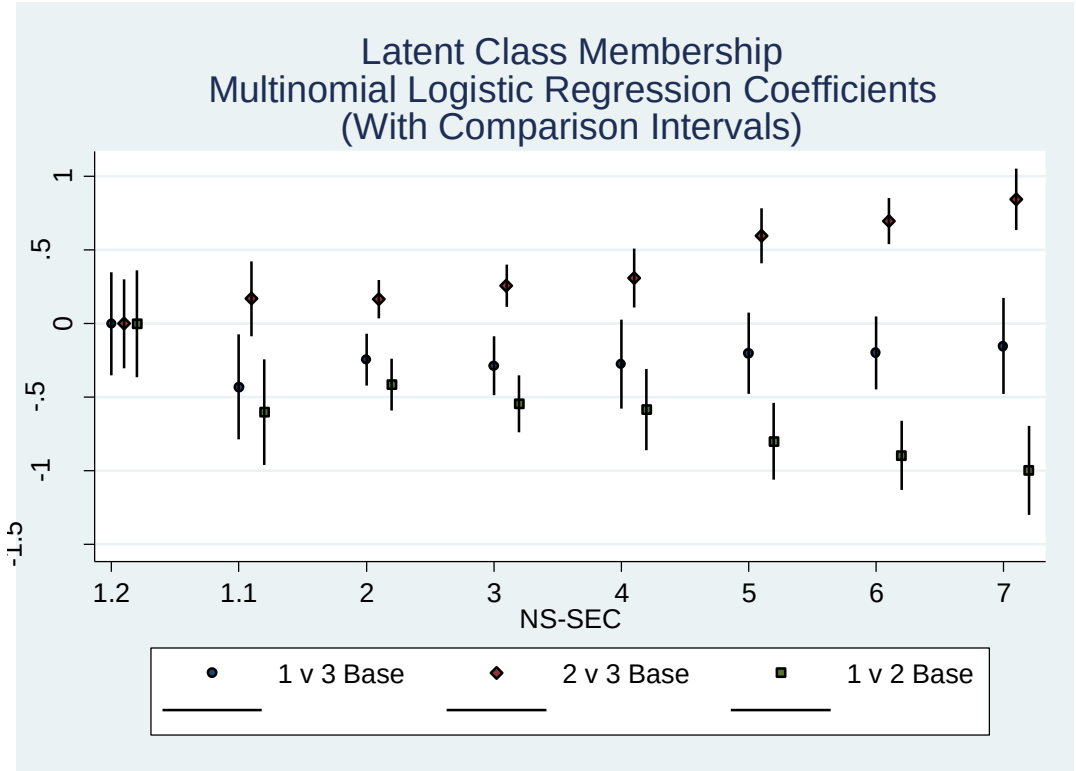
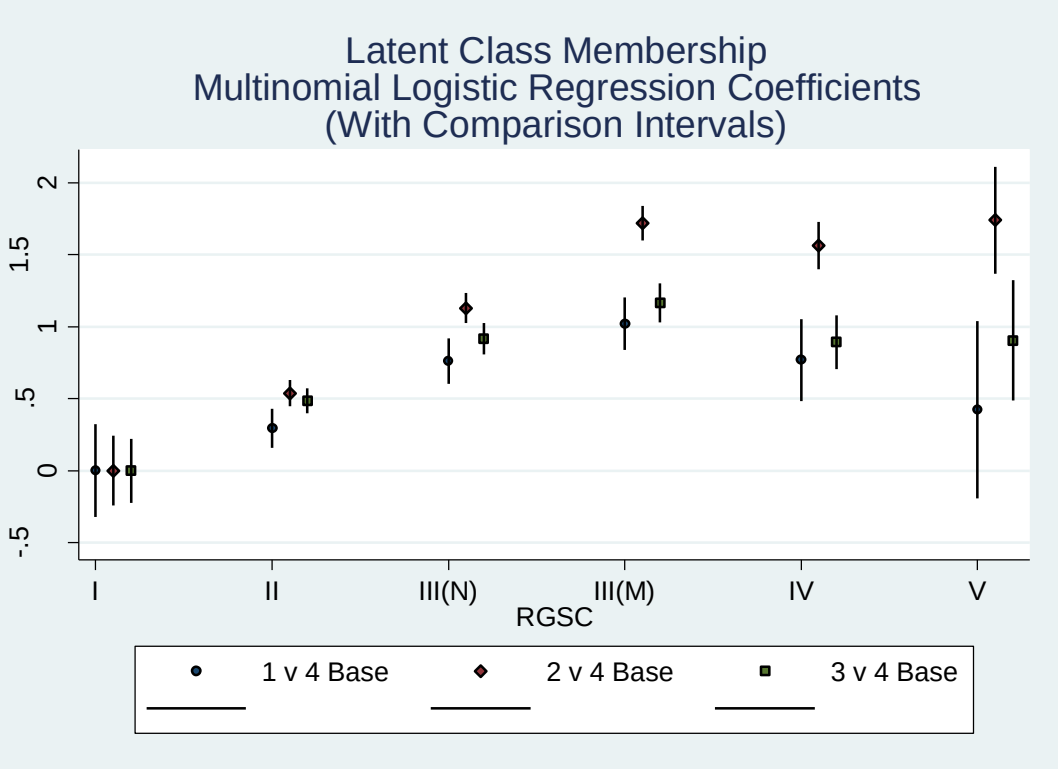




**Figure 9.4 Multinomial Logistic Regression – Comprehensive Schools only – NS-SEC**

Other variables: gender, ethnicity, school type, housing, family structure, parental education (see table 9.12)
Source: SN3532, n=14,576, YCS6.



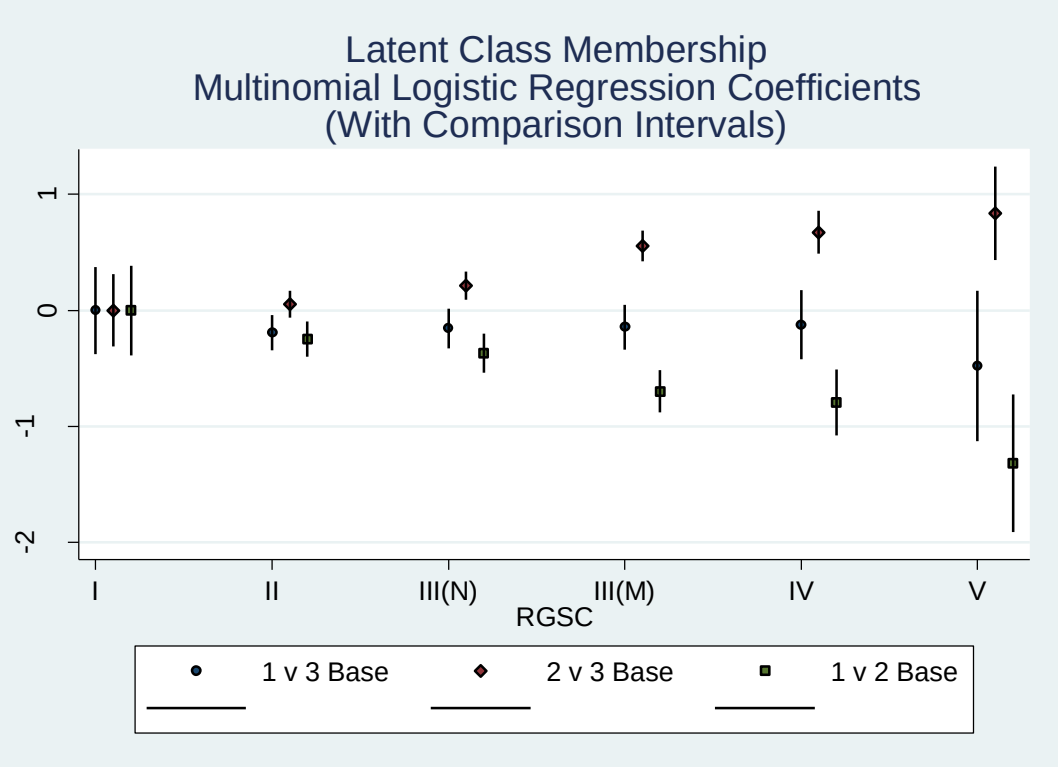




**Figure 9.5** **Multinomial Logistic Regression – Comprehensive Schools only – RGSC**

Other variables gender, ethnicity, school type, housing, family structure, parental education (see table 9.13)
Source: SN3532, n=14,576, YCS6.

## Summary

The latent attainment of pupils is manifested through grades achieved in different GCSE subjects. Within YCS cohort 6 (1992) there are clear clusters of attainment with different conditional probabilities of gaining A-C passes in GCSE subjects according to the latent class membership. As was identified in chapter two, this is of substantive interest given the 'liberalisation' of the national curriculum in 1991 to permit students to choose the subjects studied, with the exception of the core subjects of English, Maths and Science (Benn and Chitty 1997). Of the four classes identified, two groups are unambiguous, those who are highly likely to pass all the subjects they study for at grades A-C and those who are highly unlikely to do so, forming the classes "Pass All" and "Poor Grades" (consisting of 49% and 21% of the sample respectively). As was identified in chapter eight, membership of either of these classes is strongly associated with family background and in particular, parental occupation. This is important because this provides further evidence for the persistent influence of parental occupation. It is highly unlikely that young people in the "Poor Grades" group can improve sufficiently to gain 5 or more GCSEs at grades A-C to remain in education post-16 on A-level courses. Furthermore, the effects of family background remain clear and unambiguous once non-comprehensive schools are removed from the analysis (i.e. issues of prior selection on attainment accounted for).

In the two intermediate classes identified, who gain on average 3.5 GCSE passes at grades A-C, there seems to be a clear division between those who are better at Sciences (10% of the sample) and those who are better at Arts subjects (20%). This presents policy makers with two distinct target groups should they wish to concentrate resources in attempting to improve the grades of these young people. Being in the Science group does not appear to be related to parental occupation but is predicted more by gender (boys being much more likely to be in this group). Parental occupation

does predict likelihood of being in the Arts group but less clearly, with the children of parents in middle grade occupations being more likely to be in this group. Family structure appears to be important in these models with those living with just their father being more likely to be in the Science group and those living with just their mother being more likely to be in the Arts group. Parental education also appears to have differential effects. Whilst the subsample sizes may be small for some of these groups, it would be of interest in future to extend this work to look across a greater range of cohorts. This would permit the study of subject level variation in GCSE attainment by a wider range of combination and to further examine these findings.

# 10 Conclusions

This thesis addressed the principle question of the role of family background in the stratification of GCSE attainment in the 1990s, in terms of parental occupations, parental education and family structure. It has been identified that GCSE attainment is stratified by each of these factors but also gender, ethnicity and several other indicators of the circumstances in which a young person grows up. Numerous studies in chapter four using YCS data were highlighted which measured these associations with GCSE attainment, but most reported on bivariate associations between these variables and none contained the full set of 1990s data. Through the harmonisation of YCS cohort 6 with the other 1990s YCS cohorts, it has been possible to present a more comprehensive analysis of the period in order to study this key period of change. In particular, educational attainment was consistently associated with parental occupational advantage and this was demonstrated in chapter six in the descriptive statistics reported for all the 1990s cohorts of the YCS. From this foundation, multivariate models were introduced to investigate all the stratifying factors identified, to examine missing data and to explore variations in GCSE attainment by combinations of GCSE subject.

Of the specific questions addressed in chapter one (and reiterated in chapter five), the first question considered was: *"Has the association between parents' occupations and their children's educational attainment increased, decreased or remained stable through the 1990s?"* It has been consistently demonstrated by the models fitted in chapter seven that the association between parental occupation and young people's educational attainment has slightly increased through the 1990s. This is despite the overall increase in grades across all groups as the models include controls for rising attainment. Furthermore, these models included controls for variation in numerous

other potential explanatory variables. Despite the introduction of these additional variables, the influence of parental occupation remained as the strongest predictor of GCSE attainment after school type, which it has been acknowledged contains elements of selection by attainment prior to age 16 amongst independent schools (Goldstein and Thomas 1996).

The strength and significance of the association between parental occupation and GCSE attainment (among other factors considered) suggests that the detraditionalisation thesis (Heelas, Lash and Morris 1996) discussed in chapter three, does not account for such persisting differentials and does not reflect the empirical evidence. Instead, the analyses presented tend to support the continuation of processes of social stratification (Scott 1996; Morgan 1996). It is acknowledged that the contextual factors (discussed in chapter two) have changed markedly in the decades prior to the 1990s. However, the longer term trends in social mobility tend to be slower to change and more stable (Penn 2006; Lambert, Prandy and Bottero 2007; Goldthorpe 2007b). The persistence of differentials in educational attainment by family background continues to be important because of the different life chances experienced by young people according to this attainment (Morgan 1996; Payne, Cheng and Witherspoon 1996; Furlong and Cartmel 1997; White 2007).

What the detraditionalisation thesis has emphasised is that the 1990s are a crucial period for educational research. Opinion regarding the nature of youth transitions from school to work and family background may have divided sociologists of youth but there is consensus about the change in educational context that occurred (Gayle, Lambert and Murray 2009a) which suggests that the period requires further investigation. The decreased likelihood of young people entering work immediately post-school, rising

GCSE attainment and decline of traditional routes available to those with lower attainment suggests a growth in the importance of educational attainment if a young person wishes to continue into further and higher education: *"Young people entering today's workplace without academic credentials enter it disarmed"* (Bills 2004: 200)

The evidence that GCSE attainment is more strongly stratified by family background measures at the end of the 1990s than the start suggests widening gaps between young people. It also provides evidence for the failure of comprehensive school ideals, although it is acknowledged that these principles were never fully implemented. The promotion of market based ideals in education seems to have favoured those with parents in more advantaged occupations.

On the other hand, the evaluation of social policies with cohort data is problematic due to the non-experimental nature of secondary social survey research. The evidence within this thesis, using 100,000 respondents within the Youth Cohort Study, suggests that over the decade these gaps persist and cannot be ignored, but there is no counterfactual available to compare these patterns against. Nevertheless, if the political objective has been to reduce the difference between young people from advantaged and disadvantaged backgrounds, then this has not been observed during the 1990s using these data.

These analyses emphasise the persistence of inequality in GCSE attainment by parental occupation over a 10 year period. This supports and expands the findings of similar work using the YCS to analyse the 1990s cohorts (e.g. Lupton, Heath and Salter 2009; and Gayle, Lambert and Murray 2009b). Similarly, these findings develop the analysis of YCS data which studied early 1990s cohorts by including data for the remainder of

the decade (Courtenay and McAleese 1993a; Courtenay and McAleese 1993b; Payne 1995a; Payne 1995b).

## Measures of Family Background

The second question addressed was, *"How do different measures of parental occupation and parental education report this association? Are the different measures consistent?"* The models in chapter seven included an extended sensitivity analysis of measures of parental occupation. The three different measures of parental occupation, NS-SEC (National Statistics Socio-Economic Classification, see Rose and Pevalin 2003), RGSC (Registrar General's Social Class, see Rose 1995) and CAMSIS (Cambridge Social Interaction and Stratification Scale, see Prandy 1999; Prandy and Lambert 2003) report this association consistently. Analysis of the complete set of 1990s cohorts of the Youth Cohort Study has suggested the persistence of family background factors in the stratification of GCSE attainment. In particular, there are strong differentials in attainment by parental occupation and this is confirmed by the extended sensitivity analysis using RGSC, NS-SEC and CAMSIS. Strong differentials by parental occupation persist throughout the 1990s and do not diminish despite the overall context of rising attainment. This relationship remains net of the other factors listed, irrespective of the measure of parental occupation or the GCSE attainment outcome measure used.

Being a scale rather than a series of categories, CAMSIS is more finely graded than NS-SEC or RGSC, and it is important to see the influence of parental occupation on a more granular scale to examine whether the differences were an artefact of the categorical measurement scheme used. If CAMSIS had not been associated with GCSE attainment, whilst NS-SEC and RGSC had been, then it might have been possible that the categorical groupings were capturing a non linear effect of NS-SEC and RGSC.

However, this was not the case and provides strong evidence that parental occupation (however it is measured) is associated with filial GCSE attainment (for further evidence of the consistency of occupational measures see Lambert and Bihagen 2007).

Measures of parental education proved somewhat limited when compared with the effect of parental occupation. Whilst there was consistently higher GCSE attainment among the children of parents with A-levels or above, it is suggested that this measure was insufficiently detailed to fully evaluate using these data. Parental occupation measures consistently captured more of the variation which suggests that this is a better predictor of filial attainment, emphasising the caveat that with improved data it would be possible to examine levels of parental education in greater detail. Parental education remains an important part of understanding family background in predicting parental occupational advantage but also may indicate variation in expectations. This is a speculative point and it would be interesting to examine with data where it was possible to evaluate parental attitudes towards filial attainment in conjunction with measures of parental occupation and education.

It is also important to note the influence of the other variables beyond the three family background measures selected, with gender and ethnicity being sources of persistent stratification in GCSE attainment. The gap between boys' and girls' GCSE attainment continued to grow in this period. These forms of stratification were also considered in the latent class analysis in chapter nine, which is reflected upon at the end of this chapter.

## Lone Parenthood and GCSE attainment

The third question was, *"Have the differences in attainment between pupils from lone-parent and couple families increased or reduced through the 1990s?"* The difference in attainment between pupils from lone-parent and couple families has increased between 1990 and 1999 but was not significantly different from couple families in either year net of parental occupation. However, using the full set of YCS data from the 1990s, it is shown that this is an artefact of smaller sample size. Young people growing up with their mother only have lower GCSE attainment than those in couple families net of other family background factors. Whether the gap had widened over the course of the 1990s by family structure was inconclusive due to the smaller sub-samples by each year (when interaction effects were fitted). These findings support those of Ermisch and Francesconi (2001a; 2001b), that family structure has a distinct effect beyond parental occupation but that parental occupation has a greater influence than family structure.

It was not possible to distinguish between the subgroups of lone parents identified in chapter four (younger poorer single mothers and older more affluent divorcees) due to the limitations of the measures available. Improved measures of family structure over time, and an indicator of whether a young person's parents were married, would be valuable for future research. In addition, future studies of the role of lone parenthood and educational attainment require educational data which take greater account of the detail of the family structure, in particular in terms of the duration of the current structure and previous family history. This would assist in differentiating between young people who had grown up in a lone parent household for a longer period of time, and studying the potential implications this may have had upon their educational attainment. Further longitudinal data observing young people over a longer duration would be beneficial, particularly of attainment pre-16 and early development. The

Longitudinal Study of Young People in England and Wales (LSYPE) offers such an opportunity for future research.

## Missing Data Techniques

The YCS, as with other studies, is susceptible to missing data, both in terms of item and unit non-response. The YCS in particular may be more prone because it is a postal survey collected for different reasons by the DfEE. Chapter eight focused upon the fourth question which was, *"What influence does missing data potentially have upon the estimates in the models fitted? (particularly with regard to parental occupation)"* The evidence within this thesis suggests that a complete case analysis, prior to analysing patterns of missingness within the data, is sensible. However, it is theoretically advisable to undertake more sophisticated analysis. Therefore, weighting for unit non-response has been performed for all analyses within this thesis. Item non-response remains an area for concern in all surveys but is most salient in postal surveys such as the Youth Cohort Study.

The exploration of patterns of missing data (with emphasis on parental occupation data) in this thesis demonstrates the value of investigating missing data and has implications for analysis. The YCS provides the opportunity to survey young people between the ages of 16-18 but data collection methods could be improved. Postal surveys are problematic in the study of educational attainment and family background as it is likely that young people from poorer backgrounds will not respond. This is a difficult issue to correct for using post-survey methods (such as weighting or multiple imputation of missing data). This is because assumptions must be made about why these data are missing and the appropriateness of the technique to correct for potential bias in the estimates, as a consequence of the data being missing.

The technique of multiple imputation of missing data is conditional upon the assumptions by which the data are missing, being a more robust technique when the data are Missing At Random (MAR). Under this condition, the missing data may be imputed based on the patterns of response provided by those completing the survey (Carpenter and Kenward 2007). The strength of the imputation model depends upon the availability of data which predict the probability of missingness. In the absence of additional data which may explain why respondents did not supply occupational information, such procedures are limited.

When the response to a particular question is missing and the missing value is dependent upon the information contained in the question, it is said to be Missing Not At Random (MNAR). An example of data MNAR might be that respondents from lower occupational groups are less likely to respond to questions about parental occupation. The analysis in chapter eight suggests that this is the likely condition under which parental occupational data are missing in the YCS. It therefore remains highly desirable to minimise missing data at the data collection stage (Lynn 2006; Carpenter and Kenward 2007).

Improvements are noted in surveys such as the LSYPE where records are taken from the National Pupil Database (NPD), thus avoiding some of the problem of item non-response to some questions, such as educational attainment. The linkage of other data resources such as school records would provide more robust data. However, linkages such as these are not always possible and cannot currently assist with missing parental occupational data. This thesis has provided a useful demonstration of the techniques available to researchers after the data has been collected and also highlighted the advantages and disadvantages of these methods under different conditions. The

development of multiple imputation techniques have been desirable to both estimate this effect upon analyses and to provide robust reliable estimates. In practice this has demonstrated that these techniques are limited by the information available to the secondary data analyst. Nevertheless, the investigation is a reasonable pursuit given the development of multiple imputation models. Although these techniques have limitations, these approaches are becoming more widespread and this is a valid application of these techniques in further understanding YCS data. In the models presented, the influence of missing data does not influence the estimates significantly (in the absence of additional data predicting why these data were missing).

## GCSE attainment by combinations of subject

Educational attainment has been investigated in terms of the volume of qualifications but little has been done to investigate subject level variation in attainment by family background, and nothing particularly using the YCS. The use of latent class analysis as a method to investigate attainment in combinations of GCSE subjects is an original contribution to the study of GCSE attainment. Using data from 1992 it was possible to answer the question of: *"How do the subjects studied and grades attained vary between young people from different family backgrounds?"* By including the subject level variation and patterns of attainment into an outcome measure, it has been possible to identify four principle groups. Those with high and low levels of attainment in GCSE subjects were perhaps unsurprising but the difference between the middle attaining groups, notably in terms of performance in GCSE Science and English, is a key finding.

The differential in GCSE attainment of girls relative to boys has continued to grow since the introduction of GCSEs in 1988. Gender and ethnicity remain key sources of stratification in overall GCSE attainment, particularly among those with middle levels of

attainment (on average 3.5 GCSE passes at grades A*-C). Boys and black young people are more likely to gain an A*-C pass in English and Humanities and less likely to do so in Science. This represents a development of Drew's work on ethnicity in respect to educational attainment (Drew, Gray and Sime 1992; Drew 1995; Demack, Drew and Grimsley 2000). It may also help to direct future policy related research investigating the patterning of GCSE attainment by subject with respect to gender.

## The Youth Cohort Study and methods

The Youth Cohort Study provides a large scale nationally representative survey dataset of young people's GCSE results throughout the 1990s. Despite the practical limitations of the data, it has been demonstrated in this thesis that it contains valuable information regarding family background. The YCS is primarily an administrative data resource. However, it is the only large scale nationally representative survey dataset covering school leavers in this period with such a large sample size. It has been noted that there is a relative shortage of longitudinal quantitative research involving the analysis of young people in the 1990s (Payne 2003; White 2007). This thesis, through the examination of family background factors helps improve understanding of stratification in GCSE attainment, which is influential in the choices made by young people at age 16 (Payne, Cheng and Witherspoon 1996; Gayle, Berridge and Davies 2002).

The latent class analysis within chapter nine used data from YCS cohort 6 (1992) following the harmonisation of YCS cohort 6 data with the other 1990s YCS cohorts. It would be desirable to extend the analysis to include other 1990s cohorts but it is recognised that this represents a considerable amount of work. The preparation of the data for each cohort would be very time consuming as the GCSE subject performance

is not contained within the harmonised dataset of Croxford, Iannelli and Shapira (2007). An increased sample size would permit a greater range of subjects to be analysed and the testing of whether the four class model is consistent. The latent class analysis within this thesis was exploratory (i.e. no restrictions to the models were fitted) and the analysis concentrated on identifying probabilistic combinations of attainment by GCSE subject. Future work might include testing of hypotheses regarding particular combinations of GCSE subjects.

## Final Remarks

This work has identified the family background factors influencing GCSE attainment during the 1990s. It is hoped that this will inform future researchers and policy formulation and contribute to understanding the continuing stratification of attainment according to these factors, particularly with respect to parental occupation, gender and ethnicity. The understanding of attainment in particular subjects is also important in planning and understanding the supply of young people skilled in these subjects with respect to their future occupations.

Further work into the role of parental education and family structure will similarly require longitudinal data resources and the multivariate analysis of these datasets. It is recognised that these data resources are expensive but they are important in understanding the lives of young people with a broad and national representative sample (Payne 2003). Further empirical research is required into different trajectories of young people in the pre-16 educational process according to family background and prior attainment. Birth cohort data are ideally suited to studying these processes.

The analysis of stratification in educational attainment continues to be relevant to comprehending and potentially addressing the opportunities available to young people at the end of compulsory education. There is further work to done to identify how stratification in educational attainment may change in the early decades of the 21st century, however it is hoped that this thesis has contributed to advancing understanding of the persisting stratification of educational attainment during the 1990s.

# Appendix

Please go to http://www.staff.stir.ac.uk/c.j.playford/ for appendices.

# References


AINLEY, P. (1988) *From school to YTS: Education and training in England and Wales 1944-1987,* Milton Keynes, Open University Press.

AINLEY, P. (1990) *Vocational education and training,* London, Cassell Educational Ltd.

AINLEY, P. & RAINBIRD, H. (1999) *Apprenticeship: towards a new paradigm of learning,* London, Kogan Page.

ALLATT, P. (1993) Becoming Privileged : the role of family process. IN BATES, I. & RISEBOROUGH, G. (Eds.) *Youth and Inequality.* Buckingham, Open University Press.

ALLISON, P. D. (2000) Multiple Imputation for Missing Data. *Sociological Methods & Research,* 28**,** 301-309.

ARCHER, L., HUTCHINGS, M. & ROSS, A. (2003) *Higher education and social class: issues of exclusion and inclusion,* London, Routledge Falmer.

ARNOT, M., GRAY, J., JAMES, M. & JEAN, R. (1998) Recent Research on Gender and Educational Performance. London, Office for Standards in Education.

ARNOT, M. & MAC AN GHAILL, M. (2006) *The RoutledgeFalmer reader in gender and education,* London, Routledge.

ASHER, H. B. (1976) *Causal Modeling,* London, Sage Publications.

ASHFORD, S. & GRAY, J. (1993) Part-timers: part-time participation in education and training amongst 16-19 year olds in the late eighties. *Employment Department Research Series YCS Reports 24.* Sheffield, Employment Department.

ASHFORD, S., GRAY, J. & TRANMER, M. (1993) The introduction of GCSE exams and changes in post-16 participation. *Employment Department Research Series YCS Reports 23.* Sheffield, Employment Department.

ATKINSON, W. (2007) Beck, individualization and the death of class: a critique1. *The British Journal of Sociology,* 58**,** 349-366.

ATTEWELL, P. & LAVIN, D. E. (2007) *Passing the torch: Does Higher Education for the Disadvantaged Pay Off Across the Generations?,* New York, Russell Sage Foundation.

AUGHINBAUGH, A., PIERRET, C. R. & ROTHSTEIN, D. S. (2005) The Impact of Family Structure Transitions on Youth Achievement: Evidence from the Children of the NLSY79. *Demography,* 42**,** 447-468.

AUSTRIN, T. (1977) Industrial relations in the construction industry: some sociological considerations on wage contracts and trade unionism (1919-1973). University of Bristol.

BABB, P. (2005) A Summary of Focus on Social Inequalities. London, Office for National Statistics.

BALL, S. J. (2003) The more things change... Educational research, social class and 'interlocking inequalities'. London, Institute of Education, University of London.

BALL, S. J. (2006) *Education Policy and Social Class: The selected worked of Stephen J. Ball,* London, Routledge.

BARKER, B. (2008) School reform policy in England since 1988: relentless pursuit of the unattainable. *Journal of Education Policy,* 23**,** 669-683.

BARNES, M., WILLITTS, M., ANDERSON, T., CHAPLIN, J., DEBBIE, C., GROBEN, S., MORRIS, S., NOBLE, J., PHILLIPS, M. & SNEADE, I. (2004) Families and Children in Britain: Findings from the 2002 Families and Children Study (FACS). *DWP Research Report No. 206.* Leeds, Department for Work and Pensions.

BATES, I. & RISEBOROUGH, G. (1993) *Youth and Inequality,* Buckingham, Open University Press.

BECK, U. (1992) *Risk society: towards a new modernity,* London, Sage.

BECK, U. & BECK-GERNSHEIM, E. (2002) *Individualization: institutionalized individualism and its social and political consequences,* London, SAGE.

BECKER, G. S. (1991) *A Treatise on the Family,* Cambridge, Mass., Harvard University Press.

BECKER, G. S. (1993) *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education,* Chicago, University of Chicago Press.

BECKER, G. S. & TOMES, N. (1976) Child Endowments, and the Quantity of Children. *Journal of Political Economy,* 84**,** S143-S162.

BECKER, M. P. & YANG, I. (1998) Latent Class Marginal Models for Cross-Classifications of Counts. *Sociological Methodology,* 28**,** 293-325.

BELL, D. N. F. & BLANCHFLOWER, D. G. (2009) What should be done about rising unemployment in the UK? Bonn, IZA Discussion Paper 4040.

BELL, L. & STEVENSON, H. (2006) *Education Policy: Process, themes and impact,* London, Routledge.

BENN, C. & CHITTY, C. (1997) *Thirty Years On: Is comprehensive education alive and well or struggling to survive?,* London, Penguin Books.

BERTHOUD, R., MCKAY, S. & ROWLINGSON, K. (1999) Becoming a single mother. IN MCRAE, S. (Ed.) *Changing Britain: Families and households in the 1990s.* Oxford, Oxford University Press.

BIEMER, P. P. & CHRIST, S. L. (2006) Weighting Survey Data. IN DE LEEUW, E. D., HOX, J. J. & DILLMAN, D. A. (Eds.) *International Handbook of Survey Methodology.* New York, Lawrence Erlbaum Associates.

BIGGART, A. & FURLONG, A. (1996) Educating "discouraged workers" - Cultural diversity in the upper secondary school. *British Journal of Sociology of Education,* 17**,** 253-266.

BILLS, D. B. (2004) *The sociology of education and work* Oxford, Blackwell.

BLACKBURN, R. M. (2003) The concept of capital: sense and nonsense. *SSRG Conference.* Cardiff.

BLAIKIE, N. (2003) *Analyzing Quantitative Data,* London, Sage.

BLAU, D. M. (1999) The Effect of Income on Child Development. *The Review of Economics and Statistics,* 81**,** 261-276.

BLAU, P. M. & DUNCAN, O. D. (1967) *The American Occupational Structure,* New York, John Wiley & Sons.

BLOCK, F. (1990) *Post-industrial Possibilities: A Critique of Economic Discourse,* Berkeley, University of California Press.

BLOSSFELD, H.-P., KLIJZING, E., MILLS, M. & KURZ, K. (Eds.) (2008) *Globalization, Uncertainty and Youth in Society,* Abingdon, Routledge.

BLOSSFELD, H.-P. & TIMM, A. (2003) *Who Marries Whom? Educational Systems as Marriage Markets in Modern Societies,* Dordrecht, Kluwer.

BLUNDELL, R., DEARDEN, L., GOODMAN, A. & REED, H. (2000) The Returns to Higher Education in Britain: Evidence from a British Cohort. *The Economic Journal,* 110**,** F82-F99.

BOTTERO, W. (2005) *Stratification: Social division and inequality,* London, Routledge.

BOTTERO, W. & PRANDY, K. (2003) Social interaction distance and stratification. *British Journal of Sociology,* 54**,** 177-197.

BOURDIEU, P. & PASSERON, J. C. (1990) *Reproduction in education, society and culture,* London, Sage.

BRADLEY, S. & TAYLOR, J. (2002) The Effect of the Quasi–market on the Efficiency–equity Trade–off in the Secondary School Sector. *Bulletin of Economic Research,* 54**,** 295-314.

BRADLEY, S. & TAYLOR, J. (2004) Ethnicity, educational attainment and the transition from school. *The Manchester School,* 72**,** 317-346.

BRADSHAW, J. & MILLAR, J. (1991) Lone Parent Families in the UK. London, Department of Social Security.

BRANNEN, J. (2006) Cultures of intergenerational transmission in four-generation families. *Sociological Review,* 54**,** 133-154.

BRANNEN, J. & NILSEN, A. (2002) Young people's time perspectives: From youth to adulthood. *Sociology,* 36**,** 513-537.

BREEN, R. (Ed.) (2004) *Social Mobility in Europe,* Oxford, Oxford University Press.

BREEN, R. & GOLDTHORPE, J. H. (1997) Explaining Educational Differentials: Towards a Formal Rational Action Theory. *Rationality and Society,* 9**,** 275-305.

BREEN, R. & JONSSON, J. O. (2005) Inequality of Opportunity in Comparative Perspective: Recent research on Educational Attainment and Social Mobility. *Annual Review of Sociology,* 31**,** 223-243.

BROECKE, S. & HAMED, J. (2008) Gender Gaps in Higher Education Participation: *An Analysis of the Relationship between Prior Attainment and Young Participation by Gender, Socio-Economic Class and Ethnicity*. London, DIUS.

BURGESS, S., MCCONNELL, B., PROPPER, C. & WILSON, D. (2004) Girls Rock, Boys Roll: An Analysis of the Age 14–16 Gender Gap in English Schools. *Scottish Journal of Political Economy,* 51**,** 209-229.

CAMERON, A. C. & TRIVEDI, P. K. (1998) *Regression Analysis of Count Data,* Cambridge, Cambridge University Press.

CARPENTER, J. R. & KENWARD, M. G. (2007) Missing data in clinical trials - a practical guide. Birmingham, National Health Service Co-ordinating Centre for Research Methodology.

CARPENTER, J. R. & KENWARD, M. G. (2008) Missing data in clinical trials - a practical guide. Birmingham, National Health Service Co-ordinating Centre for Research Methodology.

CARPENTER, J. R., KENWARD, M. G. & VANSTEELANDT, S. (2006) A comparison of multiple imputation and doubly robust estimation for analyses with missing data. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 169**,** 571-584.

CARROLL, J. B. & SCHWEIKER, R. F. (1951) Factor Analysis in Educational Research. *Review of Educational Research,* 21**,** 368-388.

CASEY, B. & SMITH, D. (1995) Truancy and Youth Transitions. *Employment Department Research Series YCS Reports 34.* Sheffield, Employment Department.

CASSEN, R. & KINGDON, G. (2007) Tackling low educational achievement. York, Joseph Rowntree Foundation.

CASSON, M. (1979) *Youth unemployment,* London, Macmillan.

CHAPMAN, P. G. & TOOZE, M. J. (1987) *The youth training scheme in the United Kingdom,* Aldershot, Avebury.

CHERLIN, A. J., KIERNAN, K. E. & CHASE-LANSDALE, P. L. (1995) Parental Divorce in Childhood and Demographic Outcomes in Young Adulthood. *Demography,* 32**,** 299-318.

CHEUNG, S. Y. & EGERTON, M. (2007) Great Britain: Higher Education Expansion and Reform - Changing Educational Inequalities. IN SHAVIT, Y., ARUM, R. & GAMORAN, A. (Eds.) *Stratification in Higher Education.* Stanford, Stanford University Press.

CHEVALIER, A. (2004) Parental Education and Child's Education: A natural experiment. *IZA Discussion Paper,* 1153.

CHEVALIER, A. & LANOT, G. (2002) The Relative Effect of Family Characteristics and Financial Situation on Educational Achievement. *Education Economics,* 10**,** 165-181.

CHITTY, C. (2004) *Education Policy in Britain,* Basingstoke, Palgrave Macmillan.

CLARK, D., CONLON, G. & GALINDO-RUEDA, F. (2005) Post-Compulsory Education and Qualification Attainment. IN MACHIN, S. & VIGNOLES, A. (Eds.) *What's the Good of Education?* Oxford, Princeton University Press.

CLASEN, J. & CLEGG, D. (2003) Unemployment Protection and Labour Market Reform in France and Great Britain in the 1990s: Solidarity Versus Activation? *Journal of Social Policy,* 32**,** 361-381.

COE, R., SEARLE, J., BARMBY, P., JONES, K. & HIGGINS, S. (2008) Relative difficulty of examinations in different subjects. *Report for Score (Science Community Supporting Education).* Durham, Centre for Evaluation and Modelling, University of Durham.

COLEMAN, J. S. (1990) *Equality and Achievement in Education,* Boulder, San Francisco, Westview Press.

COLEMAN, J. S. & HOFFER, T. (1987) *Public and Private Schools: the impact of communities,* New York, Basic Books.

COLLINS, R. (1979) *The Credential Society: An historical sociology of education and stratification,* New York, Academic Press.

CONNOLLY, P. (2006) The effects of social class and ethnicity on gender differences in GCSE attainment: a secondary analysis of the Youth Cohort Study of England and Wales 1997–2001. *British Educational Research Journal,* 32**,** 3 - 21.

COURTENAY, G. (1996a) Youth Cohort Study of England and Wales, 1991-1993; Cohort Five, Sweep One to Three [computer file]. August 1996 ed., Colchester, Essex. UK Data Archive SN3531.

COURTENAY, G. (1996b) Youth Cohort Study of England and Wales, 1992-1994; Cohort Six, Sweep One to Three [computer file]. August 1996 ed., Colchester, Essex. UK Data Archive SN3532.

COURTENAY, G. & MCALEESE, I. (1993a) Cohort 5: aged 16-17 years old in 1991 Report on Sweep 1. *Employment Department Research Series YCS Reports 22.* Sheffield, Employment Department.

COURTENAY, G. & MCALEESE, I. (1993b) Cohort IV: Young people 16-17 years old in 1989 Report on Sweep 1. *Employment Department Research Series YCS Reports 21.* Sheffield, Employment Department.

COURTENAY, G. & MCALEESE, I. (1994) Cohort 4: Young people 17-18 years old in 1990 Report on Sweep 2. *Employment Department Research Series YCS Reports 27.* Sheffield, Employment Department.

COURTENAY, G. & MEKKELHOLT, P. (1996) England and Wales Youth Cohort Study : Handbook : the first ten years. London, Department for Education and Employment Research studies RS22.

CREEMERS, B. P. M. & KYRIAKIDES, L. (2010) Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice,* 21**,** 409 - 427.

CREGAN, C. (1999) *Young People in the Workplace,* London, Mansell.

CREGAN, C. (2001) What's happened to the labour market for early school-leavers in Britain? *Industrial Relations Journal,* 32**,** 126-135.

CROLL, P. (2009) EDUCATIONAL PARTICIPATION POST-16: A LONGITUDINAL ANALYSIS OF INTENTIONS AND OUTCOMES. *British Journal of Educational Studies,* 57**,** 400-416.

CROMPTON, R. (2008) *Class and Stratification,* Cambridge, Polity Press.

CROXFORD, L. (2000) Inequality in Attainment at Age 16: A 'Home International' Comparison. *Special CES Briefing No. 19.* Edinburgh, Centre for Educational Sociology.

CROXFORD, L. (2006a) Construction of Social Class Variables. *Working Paper 4, ESRC Research Project on Education and Youth Transitions in England, Wales and Scotland, 1984-2002.* Edinburgh, Centre for Educational Sociology.

CROXFORD, L. (2006b) The Youth Cohort Survey Data. *Working Paper 5, ESRC Research Project on Education and Youth Transitions in England, Wales and Scotland, 1984-2002.* Edinburgh, Centre for Educational Sociology.

CROXFORD, L. (2006c) The Youth Cohort Surveys - How Good is the Evidence? *Special CES Briefing No. 38.* Edinburgh, Centre for Educational Sociology.

CROXFORD, L., IANELLI, C. & SHAPIRA, M. (2007) Documentation of the Youth Cohort Time-Series Datasets. IN CENTRE FOR EDUCATIONAL SOCIOLOGY (Ed.) *UK Data Archive Study Number 5765 - Youth Cohort Time Series for England, Wales and Scotland, 1984-2002.* Edinburgh, UK Data Archive.

CROXFORD, L., IANNELLI, C. & SHAPIRA, M. (2007) Youth Cohort Time Series for England, Wales and Scotland, 1984-2002 [computer file]. Edinburgh, UK Data Archive [distributor].

DALE, A. & DAVIES, R. B. (Eds.) (1994) *Analyzing Social and Political Change,* London, Sage.

DAVID, M. E. (1995) The Education Policy Context 1976-1992. IN DAWTREY, L., HOLLAND, J., HAMMER, M. & SHELDON, S. (Eds.) *Equality and Inequality in Education Policy.* Clevedon, Multilingual Matters.

DAVIS, J. A. (1985) *The Logic of Causal Order,* London, Sage Publications.

DE GRAAF, N. D., DE GRAAF, P. M. & KRAAYKAMP, G. (2000) Parental Cultural Capital and Educational Attainment in the Netherlands: A Refinement of the Cultural Capital Perspective. *Sociology of Education,* 73**,** 92-111.

DEAKIN, B. M. (1996) *The youth labour market in Britain; the role of intervention,* Cambridge, Cambridge University Press.

DEMACK, S., DREW, D. & GRIMSLEY, M. (2000) Minding the Gap: Ethnic, Gender and Social Class Differences in Attainment at 16, 1988-95. *Race, Ethnicity and Education,* 3**,** 117-143.

DEPARTMENT FOR EDUCATION (2011) Youth Cohort Study & Longitudinal Study of Young People in England: The Activities and Experiences of 19 year olds: England 2010. Sheffield, Department for Education.

DEVINE, F. (2004) *Class practices : how parents help their children get good jobs* Cambridge, Cambridge University Press.

DFES (2006) Trends in Attainment Gaps: 2005. London, DfES.

DOLTON, P. & VIGNOLES, A. (2000) The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review,* 19**,** 179-198.

DREW, D. (1995) *'Race', education and work: the statistics of inequality,* Aldershot, Avebury.

DREW, D. & GRAY, J. (1990) The fifth-year examination achievements of black young people in England and Wales. *Educational Research,* 32**,** 107 - 117.

DREW, D., GRAY, J. & SIME, N. (1992) Against The Odds: The Education and Labour Market Experiences of Black Young People. *Youth Cohort Series 19 Employment Department Training Research and Development Series,* 68**,** 1-56.

ECCLES, J. S. & DAVIS-KEAN, P. E. (2005) Influences of parents' education on their children's educational attainments: the role of parent and child perceptions. *London Review of Education,* 3**,** 191-204.

EGERTON, M. & SAVAGE, M. (2000) Age Stratification and Class Formation: A Longitudinal Study of the Social Mobility of Young Men and Women, 1971-1991. *Work, Employment & Society,* 14**,** 23-49.

EKERT-JAFFE, O., JOSHI, H., LYNCH, K., MOUGIN, R., RENDELL, M. & SHAPIRO, D. (2002) Fertility Timing of Births and Socio-economic Status in France and Britain: Social Policies and Occupational Polarization. *Population,* 57**,** 475-507.

ELIAS, P., HOGARTH, T. & PIERRE, G. (2002) The wider benefits of education and training. London, Department for Work and Pensions, Research Report No. 178.

ELIAS, P. & PURCELL, K. (2004) Is Mass Higher Education Working? Evidence from the Labour Market Experiences of Recent Graduates. *National Institute Economic Review,* 190**,** 60-74.

ELLIOTT, J., DALE, A. & EGERTON, M. (2001) The Influence of Qualifications on Women's Work Histories, Employment Status and Earnings at Age 33. *European Sociological Review,* 17**,** 145-168.

ELWOOD, J. (1995) Undermining gender stereotypes: examination and coursework performance in the UK at 16. *Assessment in Education: Principles, Policy & Practice,* 2**,** 283-303.

EPSTEIN, D., ELWOOD, J., HEY, V. & MAW, J. (1998) *Failing boys? Issues in gender and achievement,* Buckingham, Open University Press.

ERIKSON, R. (1984) Social Class of Men, Women and Families. *Sociology,* 18**,** 500-514.

ERIKSON, R. & GOLDTHORPE, J. H. (1992) *The Constant Flux: A Study of Class Mobility in Industrial Societies,* Oxford, Clarendon Press.

ERIKSON, R., GOLDTHORPE, J. H., JACKSON, M., MEIR, Y. & COX, D. R. (2005) On class differentials in educational attainment. *PNAS,* 102.

ERMISCH, J. (1995) Pre-marital Cohabitation, Childbearing and the Creation of One-Parent Families. Colchester, ESRC Research Centre on Micro-social Change.

ERMISCH, J. & FRANCESCONI, M. (2000) Cohabitation in Great Britain: Not for Long, But Here to Stay. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 163**,** 153-171.

ERMISCH, J. & FRANCESCONI, M. (2001a) Family Matters: Impacts of Family Background on Educational Attainments. *Economica,* 68**,** 137-156.

ERMISCH, J. & FRANCESCONI, M. (2001b) Family structure and children's achievements. *Journal of Population Economics,* 14**,** 249-270.

ERMISCH, J., FRANCESCONI, M. & PEVALIN, D. J. (2001) Outcomes for children of poverty. Leeds, Corporate Document Services.

ERMISCH, J. & PEVALIN, D. J. (2005) Early Motherhood and later partnerships. *Journal of Population Economics,* 18**,** 469-489.

EVANS, M. D. R. & KELLEY, J. (2004) *Australian economy and society 2002: Religion, morality, and public policy in international perspective 1984-2002,* Sydney, Federation Press.

FEATHERMAN, D. L., LANCASTER JONES, F. & HAUSER, R. M. (1975) Assumptions of social mobility research in the U.S.: The case of occupational status. *Social Science Research,* 4**,** 329-360.

FEINSTEIN, L., DUCKWORTH, K. & SABATES, R. (2008) *Education and the Family: Passing success across the generations,* London, Routledge.

FIELD, J. (2008) *Social Capital,* London, Routledge.

FINCH, H., O'CONNOR, W., MILLAR, J., HALES, J., SHAW, A. & ROTH, W. (1999) New Deal for lone parents: Learning from the prototype areas. *DSS Research Report No. 92.* Leeds, Department of Social Security.

FINLAYSON, L., FORD, R., MARSH, A., MCKAY, S. & MUKHERJEE, A. (2000) The British Lone Parent Cohort 1991 to 1998. *DSS Research Report No. 128.* Leeds, Department of Social Security.

FINLAYSON, L. & MARSH, A. (1998) Lone parents on the margins of work. *DSS Research Report No. 80.* Leeds, Department of Social Security.

FITZ-GIBBON, C. T. (1996) *Monitoring education : indicators, quality, and effectiveness,* London, Cassell.

FORD, R., MARSH, A. & FINLAYSON, L. (1998) What happens to lone parents: A cohort study 1991-1995. *DSS Research Report No. 77.* London, Department of Social Security.

FORD, R., MARSH, A. & MCKAY, S. (1995) Changes in Lone Parenthood. *DSS Research Report No. 40.* London, Department of Social Security.

FORSYTH, A. & FURLONG, A. (2000) Socioeconomic disadvantage and access to higher education. Bristol, Joseph Rowntree Foundation.

FORSYTH, A. & FURLONG, A. (2003a) Access to Higher Education and Disadvantaged Young People. *British Educational Research Journal,* 29**,** 205-225.

FORSYTH, A. & FURLONG, A. (2003b) Socio-economic disadvantage and experience in higher education. Bristol, Joseph Rowntree Foundation.

FRANCIS, B. (2000) *Boys, girls and achievement: addressing the classroom issues,* London, RoutledgeFalmer.

FREDERICKSON, N. & PETRIDES, K. V. (2008) Ethnic, gender, and socio-economic group differences in academic performance and secondary school selection: A longitudinal analysis. *Learning and Individual Differences,* 18**,** 144-151.

FRONSTIN, P., GREENBERG, D. H. & ROBINS, P. K. (2001) Parental disruption and the labour market performance of children when they reach adulthood. *Journal of Population Economics,* 14**,** 137-172.

FULLER, A. & UNWIN, L. (2009) Change and continuity in apprenticeship: the resilience of a model of learning. *Journal of Education and Work,* 22**,** 405-416.

FULLER, A., UNWIN, L., AINLEY, P. & RAINBIRD, H. (1999) A sense of belonging: the relationship between community and apprenticeship. *Apprenticeship: towards a new paradigm of learning.* London, Kogan Page.

FURLONG, A. (Ed.) (2009) *Handbook of Youth and Young Adulthood: New perspectives and agendas,* Oxford, Routledge.

FURLONG, A. & CARTMEL, F. (1997) *Young people and social change: individualization and risk in late modernity,* Buckingham, Open University Press.

FURLONG, A., CARTMEL, F., BIGGART, A., SWEETING, H. & WEST, P. (2006) Social Class in an 'Individualised' Society. *Sociology Review,* 15**,** 28-32.

FURTHER EDUCATION FUNDING COUNCIL (2000) Citizenship for 16-19 Year Olds in Education and Training. Coventry, FEFC.

GALLIE, D. (2000) The Labour Force. IN HALSEY, A. H. (Ed.) *Twentieth-Century British Social Trends.* Basingstoke, MacMillan.

GALLIE, D., MARSH, C. & VOGLER, C. M. (1994) *Social change and the experience of unemployment,* Oxford, Oxford University Press.

GALTON, M., SIMON, B. & CROLL, P. (1980) *Inside the Primary Classroom,* London, Routledge & Kegan Paul.

GANZEBOOM, H. B. (2006) *Tools for Deriving Status Measures from ISKO-88 and ISCO-68,* Amsterdam, Amsterdam University.

GANZEBOOM, H. B., LUIJKX, R. & TREIMAN, D. J. (1989) Intergenerational class mobility in comparative perspective. *Research in Social Stratification and Mobility,* 8**,** 3-84.

GANZEBOOM, H. B. G., TREIMAN, D. D. & ULTEE, W. C. (1991) COMPARATIVE INTERGENERATIONAL STRATIFICATION RESEARCH: Three Generations and Beyond. *Annual Review of Sociology,* 17**,** 277-302.

GAYLE, V. (2000) Descriptive Statistics. IN BURTON, D. (Ed.) *Research Training for Social Scientists.* London, Sage.

GAYLE, V. (2005) Youth Transitions. IN ERMISCH, J. & WRIGHT, R. E. (Eds.) *Changing Scotland: Evidence from the British Household Panel Survey.* Bristol, Policy Press.

GAYLE, V., BERRIDGE, D. & DAVIES, R. B. (2000) Young People's Routes To Higher Education: Exploring Social Processes With Longitudinal Data. *Higher Education Review,* 33**,** 47-64.

GAYLE, V., BERRIDGE, D. & DAVIES, R. B. (2002) Young People's Entry To Higher Education: Quantifying Influential Factors. *Oxford Review of Education,* 28**,** 5-20.

GAYLE, V., BERRIDGE, D. & DAVIES, R. B. (2003) Econometric analysis of the demand for higher education. Nottingham, DfES Publications.

GAYLE, V. & LAMBERT, P. (2009) Logistic Regression Models in Sociological Research. University of Stirling, Technical Paper 2009-1 of the Data Management through e-Social Science Research Node.

GAYLE, V., LAMBERT, P. & MURRAY, S. (2009a) Growing Up in the 1990s – An exploration of the educational experiences of cohorts of Rising 16s in the BHPS. *British Household Panel Survey Conference.* University of Essex.

GAYLE, V. & LAMBERT, P. S. (2006) Using Quasi-variance to Communicate Sociological Results from Statistical Models. *Working Paper 3 of 'Longitudinal Data Analysis for Social Science Researchers', ESRC Researcher Development Initiative training programme.* Stirling, University of Stirling.

GAYLE, V. & LAMBERT, P. S. (2007) Using Quasi-variance to Communicate Sociological Results from Statistical Models. *Sociology,* 41**,** 1191-1208.

GAYLE, V., LAMBERT, P. S. & MURRAY, S. (2009b) School-To-Work in the 1990s: Modelling Transitions with large-scale datasets. IN BROOKS, R. (Ed.) *Transitions from Education to Work: New Perspectives from Europe and Beyond.* Basingstoke, Palgrave Macmillan.

GAYLE, V., PLAYFORD, C. J. & LAMBERT, P. S. (2008) Scottish Social Survey Data, Past, Present and Future – Does Scotland Need its Own Data Strategy? *Radical Statistics,* 97**,** 82-97.

GELMAN, A., KING, G. & LIU, C. (1998) Not Asked and Not Answered: Multiple Imputation for Multiple Surveys. *Journal of the American Statistical Association,* 93**,** 846-857.

GEWIRTZ, S., BALL, S. J. & BOWE, R. (1995) *Markets, Choice and Equity in Education,* Buckingham, Open University Press.

GIDDENS, A. (1990) *The Consequences of Modernity,* Cambridge, Polity Press.

GILLARD, D. (2007) Education in England: a brief history. http://www.educationengland.org.uk/history/.

GILLBORN, D. & GIPPS, C. (1996) Recent Research on the Achievement of Ethnic Minority Pupils. London, Office for Standards in Education.

GINTHER, D. K. & POLLAK, R. A. (2004) Family Structure and Children's Educational Outcomes: Blended Families, Stylized Facts, and Descriptive Regressions. *Demography,* 41**,** 671-696.

GLENNERSTER, H. (2007) *British Social Policy: 1945 to the present,* Oxford, Blackwell.

GOLDSTEIN, H. (2009) Handling attrition and non-response in longitudinal data. *Longitudinal and Life Course Studies,* 1**,** 63-72.

GOLDSTEIN, H. (2011) *Multilevel Statistical Models,* Chichester, John Wiley & Sons.

GOLDSTEIN, H. & THOMAS, S. (1996) Using Examination Results as Indicators of School and College Performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 159**,** 149-163.

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

GOLDTHORPE, J. H. (1983) Women and class analysis: in defence of the conventional view. *Sociology,* 17**,** 466-488.

GOLDTHORPE, J. H. (1991) The Uses of History in Sociology: Reflections on Some Recent Tendencies. *The British Journal of Sociology,* 42**,** 211-230.

GOLDTHORPE, J. H. (1996) Class analysis and the reorientation of class theory: The case of persisting differentials in educational attainment. *British Journal of Sociology,* 47**,** 481-505.

GOLDTHORPE, J. H. (2007a) *On Sociology: Vol. 1 Critique and program,* Stanford, CA, Stanford University Press.

GOLDTHORPE, J. H. (2007b) *On Sociology: Vol. 2 Illustration and Retrospect,* Stanford, CA, Stanford University Press.

GOLDTHORPE, J. H. & JACKSON, M. (2007) Intergenerational class mobility in contemporary Britain: political concerns and empirical findings1. *The British Journal of Sociology,* 58**,** 525-546.

GOLDTHORPE, J. H. & LLEWELLYN, C. (1983) Class mobility in Britain: three theses examined. IN COSIN, B. & HALES, M. (Eds.) *Education, policy and society: theoretical perspectives, a reader.* London, Routledge and Kegan Paul.

GOLDTHORPE, J. H., LLEWELLYN, C. & PAYNE, C. (1980) *Social Mobility and the Class Structure in Modern Britain,* Oxford, Clarendon.

GOODMAN, L. A. (2002) Latent Class Analysis: The empirical study of latent types, latent variables, and latent structures. IN HAGENAARS, J. A. & MCCUTCHEON,

A. L. (Eds.) *Applied Latent Class Analysis.* Cambridge, Cambridge University Press.

GOSPEL, H. F. (1995) The decline of apprenticeship training in Britain. *Industrial Relations Journal,* 26**,** 32.

GRAY, J. (2004) School effectiveness and the 'other outcomes' of schooling: a reassessment of three decades of British research. *Improving Schools,* 7**,** 185-198.

GRAY, J., JESSON, D. & TRANMER, M. (1993) Boosting Post-16 Participation: a study of some key factors. *Employment Department Research Series YCS Reports 20.* Sheffield, Employment Department.

GRAY, J., JESSON, D. & TRANMER, M. (1994) Local labour market variations in post-16 participation: evidence from the end of the eighties. *Employment Department Research Series YCS Reports 26.* Sheffield, Employment Department.

GRAY, J. & SIME, N. (1990) Extended Routes and Delayed Transitions Amongst 16-19 Year Olds: National Trends and Local Contexts. *British Journal of Education and Work,* 3**,** 13-40.

GROVES, R. M. (2006) Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly,* 70**,** 646-675.

HALSE, J. (2007) Winning the War of Attrition? Sampling, response analysis and weighting using the National Pupil Database. *University of Bristol: Pupil Level Annual Schools Census (PLASC) User Group Workshop.* Bristol, Department of Children, Schools and Families.

HALSEY, A. H., HEATH, A. F. & RIDGE, J. M. (1980) *Origins and Destinations: Family, Class and Education in Modern Britain,* Oxford, Clarendon Press.

HASKEY, J. (1998) Families: Their Historical Context, and Recent Trends in the Factors Influencing their Formation and Dissolution. IN DAVID, M. (Ed.) *The Fragmenting Family: Does It Matter?* London, The IEA Health and Welfare Unit.

HASKEY, J. (2002) One-parent families - and the dependent children living in them - in Great Britain. *Population Trends,* 109**,** 46-57.

HAWKES, D. & PLEWIS, I. (2006) Modelling non-response in the National Child Development Study. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 169**,** 479-491.

HEELAS, P., LASH, S. & MORRIS, P. (Eds.) (1996) *Detraditionalization: Critical reflections on authority and identity,* Oxford, Blackwell.

HEINZ, W. R. (2009) Youth transitions in an age of uncertainty. IN FURLONG, A. (Ed.) *Handbook of Youth and Young Adulthood: New perspectives and agendas.* Oxford, Routledge.

HENZ, U. & MILLS, C. (2011) Social class origins and assortative mating in Britain, 1949-2005. *ISA RC28.* University of Essex.

HILLS, J., SEFTON, T. & STEWART, K. (Eds.) (2009) *Towards a more equal society?: Povery, inequality and policy since 1997,* Bristol, The Policy Press.

HODGESON, A. & SPOURS, K. (2001) Part-time work and full-time education in the UK: the emergence of a curriculum and policy issue. *Journal of Education and Work,* 14**,** 373-388.

HOPKINS, D. (2002) *Improving the quality of education for all,* London, David Fulton Publishers.

HOWIESON, C. & CROXFORD, L. (1996) Using the YCS to analyse the outcomes of careers education and guidance. London, Department for Education and Employment Research studies RS40.

HOY, C., BAYNE-JARDINE, C. C. & WOOD, M. (2000) *Improving quality in education* London, Falmer Press.

JACKSON, M. P. (1985) *Youth unemployment,* London, Croom Helm.

JÆGER, M. M. & HOLM, A. (2007) Does parents' economic, cultural, and social capital explain the social class effect on educational attainment in the Scandinavian mobility regime? *Social Science Research,* 36**,** 719-744.

JENKINS, A. & WOLF, A. (2003) Employers' Selection Decisions: the Role of Qualifications and Tests. IN DICKENS, R., GREGG, P. & WADSWORTH, J. (Eds.) *The labour market under New Labour.* London, MacMillan.

JESSON, D., GRAY, J. & SIME, N. (1991) Youth cohort study : England & Wales. Participation, progress and performance in post-compulsory education *Employment Department Research Series YCS Reports 15.* Sheffield, Employment Department.

JONATHAN, R. (1990) State education service or prisoner's dilemma: the "hidden hand" as source of educational policy. *British Journal of Educational Studies,* 38**,** 116-132.

JONES, K. (2003) *Education in Britain: 1944 to the present,* Cambridge, Polity.

JUDGE, H. (2006) The road not taken: Deconstructing the 1960s, the formative decade. IN HEWLETT, M., PRING, R. & TULLOCH, M. (Eds.) *Comprehensive Education: Evolution, achievement and new directions.* Northampton, University of Northampton Press.

KIECOLT, K. J. & NATHAN, L. E. (1985) *Secondary Analysis of Survey Data,* Thousand Oaks, Ca, Sage.

KIERNAN, K. (1998) Family Change: Issues and Implications. IN DAVID, M. (Ed.) *The Fragmenting Family: Does It Matter?* London, The IEA Health and Welfare Unit.

KIERNAN, K. (2004) Unmarried Cohabitation and Parenthood in Britain and Europe. *Law & Policy,* 26**,** 33-55.

KIERNAN, K., AGULINK, P., HILLS, J., LE GRAND, J. & PIACHAUD, D. (2002) Demography and Disadvantage: Chicken and Egg? *Understanding social exclusion.* Oxford, Oxford University Press.

KIERNAN, K. & ESTAUGH, V. (1993) Cohabitation: Extra-marital childbearing and social policy. London, Family Policy Studies Centre.

KIERNAN, K. & MUELLER, G. (1999) Who Divorces? IN MCRAE, S. (Ed.) *Changing Britain: Families and households in the 1990s.* Oxford, Oxford University Press.

KIERNAN, K. E. (1992) The Impact of Family Disruption in Childhood on Transitions Made in Young Adult Life. *Population Studies,* 46**,** 213-234.

KOGAN, M. (1993) The End of the Dual System? The Blurring of Boundaries in the British Tertiary Education System. IN GELLERT, C. (Ed.) *Higher Education in Europe.* London, Jessica Kingsley.

KORUPP, S. E., GANZEBOOM, H. B. & VAN DER LIPPE, T. (2002) Do Mothers Matter? A Comparison of Models of the Influence of Mothers' and Fathers' Educational and Occupational Status on Children's Educational Attainment. *Quality and Quantity,* 36**,** 17-42.

KREUTER, F., ECKMAN, S., MAAZ, K. & WATERMANN, R. (2010) Children's Reports of Parents' Education Level: Does it Matter Whom You Ask and What You Ask About? *Survey Research Methods,* 4**,** 127-138.

KYRIAKIDES, L. (2008) Testing the validity of the comprehensive model of educational effectiveness: a step towards the development of a dynamic model of effectiveness. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice,* 19**,** 429 - 446.

LAMBERT, P. & BIHAGEN, E. (2007) Concepts and Measures: Empirical evidence on the interpretation of ESeC and other occupation-based social classifications. *RC28 Summer meeting.* Montreal.

LAMBERT, P. S., PRANDY, K. & BOTTERO, W. (2007) By slow degrees: Two centuries of social reproduction and mobility in Britain. *Sociological Research Online,* 12.

LAMPARD, R. (2007) Is Social Mobility an Echo of Educational Mobility? Parents' Educations and Occupations and Their Children's Occupational Attainment. *Sociological Research Online,* 12.

LAMPARD, R., GALLIE, D., MARSH, C. & VOGLER, C. M. (1994) An examination of the relationship between marital dissolution and unemployment. *Social change and the experience of unemployment.* Oxford, Oxford University Press.

LARSON, P. D. & CHOW, G. (2003) Total cost/response rate trade-offs in mail survey research: impact of follow-up mailings and monetary incentives. *Industrial Marketing Management,* 32**,** 533-537.

LECKIE, G. & GOLDSTEIN, H. (2009) The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 172**,** 835-851.

LEEMING, A., UNELL, J. & WALKER, R. (1994) Lone Mothers. *DSS Research Report No. 30.* London, Department of Social Security.

LEUNG, G. M., JOHNSTON, J. M., SAING, H., TIN, K. Y. K., WONG, I. O. L. & HO, L.-M. (2004) Prepayment was superior to postpayment cash incentives in a randomized postal survey among physicians. *Journal of Clinical Epidemiology,* 57**,** 777-784.

LI, K. H., RAGHUNATHAN, T. E. & RUBIN, D. B. (1991) Large-Sample Significance Levels from Multiply Imputed Data Using Moment-Based Statistics and an F Reference Distribution. *Journal of the American Statistical Association,* 86**,** 1065-1073.

LODGE, P. & BLACKSTONE, T. (1982) *Educational Policy and Educational Inequality,* Oxford, Martin Robertson.

LUPTON, R., HEATH, N. & SALTER, E. (2009) Education: New Labour's top priority. IN HILLS, J., SEFTON, T. & STEWART, K. (Eds.) *Towards a more equal society?: Povery, inequality and policy since 1997.* Bristol, The Policy Press.

LYNCH, K. & O'RIORDAN, C. (1998) Inequality in higher education: a study of class barriers. *British Journal of Sociology of Education,* 19**,** 445-479.

LYNN, P. (2006) The problem of nonresponse. IN DE LEEUW, E. D., HOX, J. J. & DILLMAN, D. A. (Eds.) *International Handbook of Survey Methodology.* New York, Lawrence Erlbaum Associates.

LYNN, P., BUCK, N., BURTON, J., JÄCKLE, A. & LAURIE, H. (2005) A Review of Methodological Research Pertinent to Longitudinal Survey Design and Data Collection. *Working Papers of the Institute for Social and Economic Research, paper 2005-29.* Colchester, University of Essex.

LYNN, P. & CLARKE, P. (2002) Seperating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys. *The Statistician,* 51**,** 319-333.

LYNN, P., PURDON, S., HEDGES, B. & MCALEESE, I. (1994) The Youth Cohort Study: An assessment of alternative weighting strategies and their effects. *Employment Department Research Series YCS Reports 30.* Sheffield, Employment Department.

MACHIN, S. (2003) Wage inequality since 1975. IN DICKENS, R., GREGG, P. & WADSWORTH, J. (Eds.) *The labour market under New Labour.* London, MacMillan.

MACHIN, S. (2006) Social Disadvantage and Education Experiences. Paris, OECD.

MAGUIRE, M. & MAGUIRE, S. (1997) Young people and the labour market. IN MACDONALD, R. (Ed.) *Youth, the 'Underclass' and Social Exclusion.* London, Routledge.

MARE, R. D. (1981) Change and Stability in Educational Stratification. *American Sociological Review,* 46**,** 72-87.

MARSH, A. & MCKAY, S. (1993) Families, Work and Benefits. London, Policy Studies Institute.

MARSH, A. & VEGERIS, S. (2004) The British Lone Parent Cohort and their Children 1991 to 2001. *DWP Research Report No. 209.* Leeds, Department for Work and Pensions.

MATTHEWS, R. C. O., FEINSTEIN, C. H. & ODLING-SMEE, J. C. (1982) *British Economic Growth 1856-1973,* Oxford, Clarendon Press.

MCCUTCHEON, A. L. (1987) *Latent Class Analysis,* London, Sage Publications.

MCCUTCHEON, A. L. (1996) Models for the Analysis of Categorically-Scored Panel Data. IN ENGEL, U. & REINECKE, J. (Eds.) *Analysis of Change: advanced techniques in panel data analysis.* Berlin, de Gruyter.

MCCUTCHEON, A. L. (2002) Basic Concepts and Procedures in Single- and Multiple-Group Latent Class Analysis. IN HAGENAARS, J. A. & MCCUTCHEON, A. L. (Eds.) *Applied Latent Class Analysis.* Cambridge, Cambridge University Press.

MCGOVERN, P., HILL, S., MILLS, C. & WHITE, M. (2007) *Market, Class, and Employment,* Oxford, Oxford University Press.

MCINTOSH, S. (2003) Evidence on the Balance of Supply and Demand for Qualified Workers. IN DICKENS, R., GREGG, P. & WADSWORTH, J. (Eds.) *The labour market under New Labour.* London, MacMillan.

MCKAY, S. & MARSH, A. (1994) Lone Parents and Work: The Effects of Benefits and Maintenance. London, Department of Social Security.

MCQUAID, R. & BOND, S. (2004) *Gender Stereotyping in Career Choices,* Glasgow, Careers Scotland.

MENARD, S. (2001) *Applied Logistic Regression Analysis,* Berkeley, Ca., Sage.

MERRIFIELD, P. R. (1974) Factor Analysis in Educational Research. *Review of Research in Education,* 2**,** 393-434.

MERTON, R. K. (1957) *Social Theory and Social Structure,* Toronto, Collier-Macmillan.

MOBLEY, M., EMERSON, C., GODDARD, I., GOODWIN, S. & LETCH, R. (1986) *All about GCSE,* London, Heinemann Educational.

MORGAN, D. H. J. (1985) *The family, politics and social theory,* London, Routledge.

MORGAN, D. H. J. (1996) *Family connections: an introduction to family studies,* Cambridge, Polity Press.

MORGAN, P. (1998) An Endangered Species? IN DAVID, M. (Ed.) *The Fragmenting Family: Does It Matter?* London, The IEA Health and Welfare Unit.

MORGAN, P. (2000) *Marriage-Lite: The rise of cohabitation and its consequences,* London, Insititute for the Study of Civil Society.

MORTIMORE, P., SAMMONS, P. & THOMAS, S. (1994) School Effectiveness and Value Added Measures. *Assessment in Education: Principles, Policy & Practice,* 1**,** 315 - 332.

MURPHY, M. (2000) The Evolution of Cohabitation in Britain, 1960-95. *Population Studies,* 54**,** 43-56.

MURPHY, M. & WANG, D. (1999) Forecasting British families into the twenty-first century. IN MCRAE, S. (Ed.) *Changing Britain: Families and households in the 1990s.* Oxford, Oxford University Press.

NATIONAL CENTRE FOR SOCIAL RESEARCH (1999) Youth Cohort Study Cohort 9 Sweep 1 (C9S1): Technical Report. London, The National Centre for Social Research.

NEWTON, P. E. (1997) Measuring Comparability of Standards between Subjects: why our statistical techniques do not make the grade. *British Educational Research Journal,* 23**,** 433 - 449.

NOBLE, I., LYNN, P. & SMITH, P. (2005) A new and better method for sample designs with disproportionate sampling. *Annual meeting of the American Association For Public Opinion Association (AAPOR).* Fontainebleau Resort, Miami Beach, FL.

NOBLE, I. & MOON, N. (2008) 'Your Old Man's a Dustman?' The accuracy of young people's reports of parental occupations. *RC33 Logic and Methodology in Sociology, 7th International Conference on Social Science Methodology.* Naples.

NORTH, J. (1987) *The GCSE: an examination,* London, Claridge Press, in association with Education Research Trust.

NUTTALL, D. L., GOLDSTEIN, H., PROSSER, R. & RASBASH, J. (1989) Differential school effectiveness. *International Journal of Educational Research,* 13**,** 769-776.

O'DONOGHUE, C., THOMAS, S., GOLDSTEIN, H. & KNIGHT, T. (1997) 1996 DfEE study of value added for 16–18 year olds in England. *DfEE research series.* London, Department for Education and Employment.

OECD (1997) Is Job Insecurity on the Increase in OECD Countries? *Employment Outlook 1997.* Paris, OECD.

OLSSEN, M., CODD, J. & O'NEILL, A.-M. (2004) *Education Policy: Globablization, Citizenship & Democracy,* London, Sage.

ONS (2000a) Social Trends 30. London, Office for National Statistics.

ONS (2000b) The UK Service Sector. London, Office for National Statistics.

ONS (2001) Social Trends 31. London, Office for National Statistics.

ONS (2002a) Social Trends 33. London, Office for National Statistics.

ONS (2002b) Social Trends 34. London, Office for National Statistics.

ONS (2005) Focus on People and Migration. Basingstoke, Office for National Statistics.

ONS (2006a) Focus on Ethnicity and Identity. London, Office for National Statistics.

ONS (2006b) Social Trends 36. London, Office for National Statistics.

ONS (2007a) Focus on Families. Basingstoke, ONS.

ONS (2007b) Social Trends 37. London, Office for National Statistics.

ONS (2007c) Statistics of Education: The Characteristics of High Attainers. *National Statistics Bulletin June 2007.* London, Office for National Statistics.

ONS (2008) GCSE and Equivalent Examination Results in England 2006/07 (Revised). *Statistical First Release.* London, Office for National Statistics.

ONS (2009a) GCSE and Equivalent Results in England, 2008/09 (Provisional). *Statistical First Release.* London, Office for National Statistics.

ONS (2009b) Social Trends 39. London, Office for National Statistics.

PARK, A. (1994) Cohort 4: Young people 18-19 years old in 1991 Report on Sweep 3. *Employment Department Research Series YCS Reports 29.* Sheffield, Employment Department.

PARSONS, T. (1967) *The Structure of Social Action,* New York, Free Press.

PAYNE, J. (1995a) Qualifications between 16 and 18: a comparison of achievements on routes beyond compulsory schooling. *Employment Department Research Series YCS Reports 32.* Sheffield, Employment Department.

PAYNE, J. (1995b) Routes beyond compulsory schooling. *Employment Department Research Series YCS Reports 31.* Sheffield, Employment Department.

PAYNE, J. (2000) Young people not in education, employment or training : data from the England and Wales Youth Cohort Study *Research Report 201.* Nottingham, Department for Education and Employment.

PAYNE, J. (2001) Post-16 Students and Part-Time Jobs: Patterns and Effects. *Research Report 323.* London, Department for Education and Skills.

PAYNE, J. (2003) Choice at the end of compulsory schooling: A research review. *Research Report 414.* Sheffield, Department for Education and Skills.

PAYNE, J., CHENG & WITHERSPOON, S. (1996) Education and training for 16-18 year olds in England and Wales : individual paths and national trends London, Policy Studies Institute.

PAYNE, J. & RANGE, M. (1998) Lone parents' lives: an analysis of partnership, fertility, employment and housing histories in the 1958 British birth cohort. London, Stationery Office.

PENN, R. (2006) *Social Change and Economic Life in Britain,* Bologna, Homeless Books.

PLEWIS, I. (2009) Ethnic Group Differences in Educational Attainments and Progress Revisited. CCSR Working Paper 2009 - 01.

PORTES, A. (1998) Social Capital: Its origins and applications in modern sociology. *Annual Review of Sociology,* 24**,** 1-24.

PRANDY, K. (1990) The Revised Cambridge Scale of Occupations. *Sociology,* 24**,** 629-655.

PRANDY, K. (1999) The social interaction approach to the measurement and analysis of social stratification. *International Journal of Sociology and Social Policy,* 19**,** 204-236.

PRANDY, K. & LAMBERT, P. (2003) Marriage, Social Distance and the Social Space:: An Alternative Derivation and Validation of the Cambridge Scale. *Sociology,* 37**,** 397-411.

PUTNAM, R. D. (1996) The strange disappearance of civic America. *The American Prospect,* 24**,** 34-48.

RAFFE, D. (1988) *Education and the youth labour market: schooling and scheming,* London, Falmer P.

RAFFE, D. (1992) Educational participation amongst 16-18 year olds. *Briefing Paper.* London, National Commission on Education.

RAFFE, D. (2005) Devolution and divergence in education policy. IN ADAMS, J. & SCHMUEKER, K. (Eds.) *Devolution in practice 2006: public policy differences within the UK.* Newcastle, IPPR North.

RANSON, S. (2008) The Changing Governance of Education. *Educational Management Administration & Leadership,* 36**,** 201-219.

RAPHAEL REED, L. (1999) Troubling boys and disturbing discourses on masculinity and schooling: a feminist exploration of current debates and interventions concerning boys in school. *Gender and Education,* 11**,** 93–110.

REAY, D. (2006) The zombie stalking English schools: Social class and educational inequality. *British Journal of Educational Studies,* 54**,** 288-307.

REED, L. R. (1999) Troubling Boys and Disturbing Discourses on Masculinity and Schooling: A feminist exploration of current debates and interventions concerning boys in school. *Gender and Education,* 11**,** 93 - 110.

RICE, P. (1999) The impact of local labour markets on investment in further education: Evidence from the England and Wales youth cohort studies. *Journal of Population Economics,* 12**,** 287-312.

ROBERTS, K. (1977) *The fragmentary class structure* London, Heinemann Educational.

ROBERTS, K. (1984) *School-leavers and their prospects: youth in the labour market in the 1980s,* Milton Keynes, Open University Press.

ROBERTS, K. & CHADWICK, C. (1991) England & Wales youth cohort study: Transitions into the labour market. the new routes of the 1980s, a study of transitions 1984-87. *Employment Department Research Series YCS Reports 16.* Sheffield, Employment Department.

ROSE, D. (1995) Official Social Classificiations in the UK. *Social Research Update.* Guildford, University of Surrey.

ROSE, D. (1998) Once more Unto the Breach: In Defence of Class Analysis Yet Again. *Work, Employment & Society,* 12**,** 755-767.

ROSE, D. & HARRISON, E. (Eds.) (2009) *Social class in Europe : an introduction to the European Socio-economic Classification,* Abingdon, Oxon, Routledge.

ROSE, D. & PEVALIN, D. J. (2003) A Researchers Guide to the National Statistics Socio-Economic Classification. London, Sage.

ROTHON, C. (2007) Can achievement differentials be explained by social class alone? *Ethnicities,* 7**,** 306-322.

RUBIN, D. B. (1987) *Multiple Imputation for non-response in surveys,* Chichester, Wiley.

RUBIN, D. B. (2004) The Design of a General and Flexible System for Handling Nonresponse in Sample Surveys. *The American Statistician,* 58**,** 298-302.

SAMMONS, P. (1995) Gender, Ethnic and Socio-Economic Differences in Attainment and Progress: A Longitudinal Analysis of Student Achievement over 9 Years. *British Educational Research Journal,* 21**,** 465-485.

SANDERSON, M. (2007) Education and the Labour Market. IN CRAFTS, N. F. R., GAZELEY, I. & NEWELL, A. (Eds.) *Work and pay in twentieth-century Britain.* Oxford, Oxford University Press.

SAUNDERS, P. (1990) Social class and stratification. London, Routledge.

SAVAGE, M. (2000) *Class analysis and social transformation,* Buckingham, Open University Press.

SAVAGE, M. (2003) REVIEW ESSAY: A New Class Paradigm? *British Journal of Sociology of Education,* 24**,** 535 - 541.

SAVAGE, M. & EGERTON, M. (1997) Social Mobility, Individual Ability and the Inheritance of Class Inequality. *Sociology,* 31**,** 645-672.

SCHROEDER, A., MILES, A., SAVAGE, M., HALFORD, S. & TAMPUBOLON, G. (2008) Mobility, careers and inequalities: a study of work-life mobility and the returns from education. Manchester, Equality and Human Rights Commisson.

SCOTT, J. (1996) Stratification and power : structures of class, status and command Cambridge, Polity Press.

SCOTT, J. (2004) Family, Gender and Educational Attainment in Britain: A Longitudinal Study. *Journal of Comparative Family Studies,* 35**,** 565-89.

SEAMAN, P. & SWEETING, H. (2004) Assisting young people's access to social capital in contemporary families: a qualitative study. *Journal of Youth Studies,* 7**,** 173 - 190.

SERFLING, O. (2004) The interaction between unit and item nonresponse in view of the reverse cooperation continuum. *Working papers 04/01.* Faculty of Business and Economics, University of Basel.

SHAVIT, Y. & MÜLLER, W. (Eds.) (1998) *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations,* Oxford, Clarendon Press.

SHAVIT, Y., YAISH, M. & BAR-HAIM, E. (2007) The Persistence of Persistent Inequality. IN SCHERER, S., POLLAK, R., OTTE, G. & GANGL, M. (Eds.) *From Origin to Destination. Trends and Mechanisms in Social Stratification Research.* Frankfurst a. M. and New York, Campus.

SHOEMAKER, P. J., EICHHOLZ, M. & SKEWES, E. A. (2002) Item Nonresponse: Distinguishing between don't Know and Refuse. *International Journal of Public Opinion Research,* 14**,** 193-201.

SMART, C. & STEVENS, P. (2000) Cohabitation breakdown. London, Family Policy Studies Centre.

SMITH, G. (2000) Schools. IN HALSEY, A. H. (Ed.) *Twentieth-Century British Social Trends.* Basingstoke, MacMillan.

SMITH, T. & NOBLE, M. (1995) *Education Divides: Poverty and schooling in the 1990s,* London, CPAG.

SMITHERS, A. (2001) Education Policy. IN SELDON, A. (Ed.) *The Blair Effect - The Blair Government 1997-2001.* London, Little, Brown and Company.

SMITHERS, A. (2007) *Blair's education: An international perspective,* Buckingham, Centre for Education and Employment Research, University of Buckingham.

SOBOTKA, T. & TOULEMON, L. (2008) Overview Chapter 4: Changing family and partnership behaviour: Common trends and persistent diversity across Europe. *Demographic Research,* 19**,** 85-138.

STATACORP (2009) Stata Statistical Software: Release 11. College Station, TX, StataCorp LP.

STEELE, F., CONSTANTINOS, K., GOLDSTEIN, H. & JOSHI, H. (2005) The Relationship Between Childbearing and Transitions From Marriage and Cohabitation in Britain. *Demography,* 42**,** 647-673.

STOBART, G., ELWOOD, J. & QUINLAN, M. (1992) Gender Bias in Examinations: how equal are the opportunities? *British Educational Research Journal,* 18**,** 261 - 276.

STONEY, S. M. & LINES, A. (1987) *YTS: the impact on FE,* Windsor, NFER-Nelson.

STRAND, S. (1999) Ethnic Group, Sex and Economic Disadvantage: associations with pupils' educational progress from Baseline to the end of Key Stage 1. *British Educational Research Journal,* 25**,** 179 - 202.

STRAND, S. (2010) Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice,* 21**,** 289 - 314.

STRAND, S. (2011) The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal,* 37**,** 197-229.

STURGIS, P. (2004) Analysing Complex Survey Data: Clustering, Stratification and Weights. *Social Research Update,* 43.

SULLIVAN, A. (2002) Bourdieu and education: how useful is Bourdieu's theory for researchers? *Netherlands' Journal of Social Sciences,* 38**,** 144-166.

TAYLOR, S. & LYNN, P. (1997) The effect of time between contacts, questionnaire length, personalisation and other factors on response to the Youth Cohort Study. London, Department for Education and Employment Research Series RS8.

TEDDLIE, C. & REYNOLDS, D. (Eds.) (2000) *The international handbook of school effectiveness research,* London, Falmer Press.

THOMPSON, R. (2009) Social class and participation in further education: Evidence from the Youth Cohort Study of England and Wales. *British Journal of Sociology of Education,* 30**,** 29-42.

TINKLIN, T. & RAFFE, D. (1999) Entrants to Higher Education. Edinburgh, Centre for Educational Sociology.

TREIMAN, D. J. (2009) *Quantitative Data Analysis: Doing social research to test ideas,* San Francisco, Jossey-Bass.

UK DATA FORUM (2007) The National Strategy for Data Resources for Research in the Social Sciences. Warwick, University of Warwick.

VANSTEELANDT, S., CARPENTER, J. R. & KENWARD, M. G. (2010) Analysis of Incomplete Data Using Inverse Probability Weighting and Doubly Robust Estimators *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences,* 6**,** 37-48.

VEGERIS, S. & PERRY, J. (2003) Families and children 2001: Living standards and the children. *DWP Research Report No. 190.* Leeds, Department for Work and Pensions.

VERMUNT, J. K. (1997) LEM: A General Program for the Analysis of Categorical Data. Department of Methodology and Statistics, Tilburg University.

WALFORD, G. (1994) *Choice and Equity in Education,* London, Cassell.

WEST, A., BARHAM, E. & HIND, A. (2009) Secondary school admissions in England: policy and practice. London, LSE: Education Research Group.

WEST, A., HIND, A. & PENNELL, H. (2006) School Admissions and 'Selection' in Comprehensive Schools: Policy and practice. IN HEWLETT, M., PRING, R. & TULLOCH, M. (Eds.) *Comprehensive Education: Evolution, achievement and new directions.* Northampton, University of Northampton Press.

WHITE, P. (2007) *Education and Career Choice: A New Model of Decision Making* Basingstoke, Palgrave Macmillan.

WHITTY, G. (2002) Making Sense of Education Policy. London, Sage.

WITHERSPOON, S. (1995) Careers Advice and the Careers Service: The Experiences of Young People. *Employment Department Research Series YCS Reports 33.* Sheffield, Employment Department.

WOLF, A. (2002) *Does Education Matter? Myths about education and economic growth,* London, Penguin.

WOOLCOCK, M. (2001) The Place of Social Capital in Understanding Social and Economic Outcomes. *Isuma: Canadian Journal of Policy Research,* 2**,** 11-17.

YANG, M. & WOODHOUSE, G. (2001) Progress from GCSE to A and AS Level: institutional and gender differences, and trends over time. *British Educational Research Journal,* 27**,** 245 - 267.

YOUNGER, M. & WARRINGTON, M. (2005) Raising boys' achievement. London, Department for Education and Skills.