THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Young people and school General Certificate of Secondary Education attainment

Citation for published version:

Gayle, V, Murray, SJ & Connelly, R 2016, 'Young people and school General Certificate of Secondary Education attainment: Looking for the 'missing middle'', *British Journal of Sociology of Education*, vol. 37, no. 3, pp. 350-370. <https://doi.org/10.1080/01425692.2014.935292>

Digital Object Identifier (DOI):

[10.1080/01425692.2014.935292](https://doi.org/10.1080/01425692.2014.935292)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

British Journal of Sociology of Education

Publisher Rights Statement:

This is an Accepted Manuscript of an article published by Taylor & Francis in British Journal of Sociology of Education on 15 August 2014, available online:
<https://www.tandfonline.com/doi/full/10.1080/01425692.2014.935292>

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Young People and School GCSE Attainment: Looking for the 'Missing Middle'

Abstract

In Britain educational qualifications gained at school continue to play an important and central role in young people's educational and employment pathways. Recently there has been a growing interest in documenting the lives of 'ordinary' young people. In this paper we analyse the Youth Cohort Study of England and Wales in order to better document the experiences of those with 'middle' levels of school GCSE attainment.

The overall pattern of school GCSE attainment is one of increasing levels of performance. GCSE attainment is still highly stratified. Girls performed better than boys, and there were some marked differences in attainment for pupils from the main minority ethnic groups. Most notably parental socioeconomic positions are the most important factor.

The analyses fail to persuade us that there are clear boundaries that demark a 'middle' category of school GCSE attainment. We conclude that sociologists should study 'ordinary' young people, however school GCSE attainment is best understood as a continuum and measures such as the number of GCSEs, or point scores are preferable.

Keywords: Youth transitions; sociology of youth; educational attainment; GCSE; missing middle; Youth Cohort Study of England and Wales.

Introduction

The idea of a ‘marginalised mainstream’ or ‘missing middle’ is currently popular within the sociology of youth, and it has received special attention in a recent volume of *Sociological Research Online*¹. Brown (1987) colourfully describes this group as ordinary pupils who neither leave their names engraved on the school honours board, nor gouge them into their desktops. In three relatively recent papers Roberts (2011; 2012, 2013) makes an appeal to youth researchers to concentrate more analytical attention on ordinary young people. Roberts (2011) also petitions researchers to better document the experiences of this group through the secondary analysis of large-scale datasets to establish their social characteristics and how well qualified they might be. We respond to this appeal and use large-scale data to analyse school GCSE attainment.

Introduced in the mid-1980s the General Certificate of Secondary Education (GCSE) is the standard qualification undertaken by pupils in England and Wales at the end of year 11 (age 15-16) (Department of Education 1985; Mobley *et al.* 1986; North 1987). School GCSEs are worthy of sociological attention because they are public examinations and mark the first major branching point in a young person’s educational career. School GCSE results send clear signals to pupils, their peers, their parents, teachers and potential employers. Because of the progressive structure of the British education system poor school GCSE attainment is a considerable obstacle which often precludes young people from pursuing more advanced educational courses. School GCSEs are the first step in providing access to further study or employment and results are used as a yardstick to judge pupil’s abilities².

School GCSE attainment is strongly related to participation in post-compulsory education (Payne 1995; 2000; 2001; 2003). Leckie and Goldstein (2009) remind us that for young people who choose to leave education at the minimum age, their GCSEs are often their only educational qualifications. Rice (1999) observes a clear relationship between poor school GCSE performance, unemployment and participation in further education. Babb (2005) concludes that young people’s experiences at school and their attainment at GCSE level are strong determinants of their future success in both education and employment. Through the detailed

¹ Sociological Research Online, 2013, Volume 18.

² This point is made clearly on the Edexcel website
<http://www.edexcel.com/i-am-a/student/qualifications/Pages/GCSEs.aspx>.

examination of panel data, Murray (2011) reports that the negative effects of poor GCSE attainment follow young people into early adulthood. Jones *et al.* (2003) clearly illustrate that overall workers with poor school level qualifications (e.g. GCSEs) generally have less favourable labour market outcomes.

In contemporary Britain young people have grown up in shifting educational and economic circumstances (Furlong and Cartmel 2007). These include dramatic changes to the organisation and management of schools and revisions to educational qualifications. The overall pattern of school GCSE attainment is one of improvement (see DfES 2007). Given these changes it is theoretically plausible that a distinctive ‘middle’ group of young people have emerged, who obtain ‘middle’ (i.e. moderate) levels of GCSE attainment at school.

We are mindful of the methodological prescription issued by Merton (1987) which cautions that before sociologists proceed to explain or to interpret a phenomenon, it is advisable to establish that the phenomenon actually exists, and that it is enough of a regularity to require and to allow explanation. Therefore in this paper we address the rudimentary question, ‘is there a ‘middle’ group of young people that attain ‘middle’ or moderate levels of school GCSEs?’

This paper is a replication study, using specialised youth data in order to extend and augment earlier work. In a previous paper we began by exploring the mid-ground between what can broadly be termed as the educationally successful and the unsuccessful (Connelly, Murray and Gayle 2013). We undertook secondary analyses of the British Household Panel Survey (BHPS) to characterise members of this ‘middle’ group. The BHPS is a major longitudinal data resource which tracks young people within households and facilitates analyses of educational attainment, and activities in young adult life. We identified a group which might reasonably be called a ‘middle’ group and examined their activities in early adulthood. We compared their education and economic activities with the activities of their more and less educationally accomplished peers. The ‘middle’ group differed in their economic activities in early adulthood compared with the other two groups. Notably the ‘middle’ group made the transition from education into employment earlier. Overall, however, our results indicated that there was not a clearly defined ‘middle’ group and we were therefore cautious of making extended claims about this group without further exploration.

The young people within the BHPS are drawn from a representative sample of British households, and this means that the overall coverage of young people is smaller than it would be in a dedicated youth survey. In this current paper we present secondary analysis of the Youth Cohort Study of England and Wales (YCS) which is particularly suitable because it is a set of large-scale nationally representative samples of young people. The YCS also contains more detailed measures of educational qualifications. The analyses are organised so that they are comparable with our previous empirical work.

Youth Cohort Study of England and Wales

The Youth Cohort Study of England and Wales (YCS) is a suitable choice of dataset, and has been successfully used to explore educational attainment (Drew *et al.* 1992; Drew 1995; Demack *et al.* 2000; Gayle *et al.* 2003; Connolly 2006; Gayle *et al.* 2009; Sullivan *et al.* 2011). The YCS is a major longitudinal study that began in the mid - 1980s. It is a large-scale nationally representative survey funded by the government. The YCS is designed to monitor the behaviour of young people from when they reach the minimum school leaving age and either remain in education or enter the labour market. The survey collects detailed information on the young person's experiences of education, their qualifications, and information on employment and training. A limited amount of information is collected on the young person's personal characteristics, their family and their circumstances at home.

The YCS samples are nationally representative of year 11 pupils in England and Wales. A large sample from an academic year group (a cohort) is contacted in the spring following year 11. The young people are usually age 16-17 when they are first contacted. The main data collection instrument is a postal questionnaire. The cohorts are usually re-contacted on at least two occasions.

The YCS is primarily a monitoring tool although over its life-course it has been used for social science research. The survey is organised by school leaving cohorts and although cross-cohort comparisons are feasible, in practice there are a number of difficult practical challenges. Over the lifespan of the YCS there have been a number of major changes in education. These changes include amendments to qualifications and the curriculum, and alterations to the structure, organisation, management and financing of schools. These changes add substantially to the

complexity of comparing YCS cohorts. The survey has also been collected by different survey agencies. At a more practical level, there have been changes to questions within the surveys and changes in measurements and coding. This adds substantially to the challenge of developing comparable measures. Over the life-cycle of the YCS different government departments have been in charge of the survey and the structure and timings of data collection have varied between cohorts. A major obstacle is that the documentation for the YCS data is extremely poor.

The YCS has been conducted at irregular intervals. In the current paper we analyse data from eight YCS samples, which span the period from 1990 to 2006. We do not include any analyses of data prior to 1990 as young people in earlier YCS cohorts either did not undertake GCSE examinations, or did not have appropriate parental occupational information collected for comparable measures to be derived. We confine the analyses to young people who attended comprehensive schools in year 11. These pupils were at non-fee paying State funded schools. We undertake single-level analyses because there are no school-level or Local Authority-level indicators deposited with the public access datasets.

Exploring ‘Middle’ Levels of School GCSE Attainment

The attainment of five or more GCSEs at grades A*-C is a standard benchmark, for example it is used in school performance league tables (see Leckie and Goldstein 2009). This measure is routinely employed in a wide variety of social science applications (e.g. Gayle *et al.* 2003; Connolly 2006; Tunstall *et al.* 2011; Sullivan *et al.* 2011). We begin our analysis by exploring the idea that there might be a ‘middle’ group of young people with moderate levels of GCSE attainment. We theorise that this ‘middle’ group are neither well qualified nor completely unqualified. This ‘middle’ group may have obtained some GCSEs at grades A*-C (i.e. 1- 4), but they have not achieved the standard benchmark of five or more GCSEs at grades A*-C.

Note: 2003 and 2006 include GNVQ equivalents.

Figure 1: Proportion of Pupils Attaining School GCSEs at Grades A*-C by YCS School Year Cohort

The overall patterns of school GCSE attainment are depicted in Figure 1. The recognized pattern of improvement in performance by pupils in more recent school years is evident and the proportions gaining the benchmark of five or more GCSEs at grades A*-C are consistent with official figures (see DfES 2007).

Table 1: Multinomial Logistic Model (Survey Weighted) of School GCSE Attainment in Year 11 – Log Odds (significant variables highlighted in bold)

		5+ vs. 1-4				0 vs. 1-4			
		1990-1999	2001	2003^a	2006^a	1990-1999	2001	2003^a	2006^a
YCS Cohort	1990	0.00				0.00			
	1993	0.35				-0.02			
	1995	0.48				0.01			
	1997	0.61				0.14			
	1999	0.67				-0.14			
Gender	Girls	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Boys	-0.31	-0.37	-0.35	-0.31	0.49	0.46	0.45	0.38
Ethnicity	White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Black	-0.53	-0.39	-0.76	-0.38	-0.19	-0.65	-0.41	-0.59
	Indian	0.21	0.27	0.42	0.55	-0.31	-0.39	-1.35	-0.19
	Pakistani	-0.41	-0.08	-0.30	0.23	-0.21	0.03	0.01	-0.10
	Bangladeshi	0.18	0.32	-0.17	0.74	-0.62	-0.54	-1.42	-0.42
	Other Asian	0.61	1.10	0.40	0.84	-0.67	-0.11	0.40	0.26
	Other	-0.01	-0.08	0.14		-0.43	-0.29	-0.45	
	New Other Category ^b				-0.07				-0.57
Housing Tenure	Owned / Mortgage	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Renters	-0.69	-0.62	-0.79	-0.60	0.39	0.41	0.40	0.45
	Others	-0.22	-0.44	-1.01	-0.56	0.13	0.12	0.36	0.06
Household Type^c	Mother and Father	0.00	0.00	0.00	-	0.00	0.00	0.00	-
	Mother Only	-0.06	-0.18	-0.03	-	0.06	-0.07	0.07	-
	Father Only	-0.27	-0.65	-0.70	-	0.15	0.04	0.21	-
	Other Household	-0.60	-0.65	-0.24	-	0.38	0.22	0.91	-
Parental Education	Non-graduates	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Graduates	0.46	0.36	0.63	0.64	-0.11	-0.05	0.05	-0.18

Table 1 (continued): Multinomial Logistic Model (Survey Weighted) of School GCSE Attainment in Year 11 – Log Odds (significant variables highlighted in bold)

		5+ vs. 1-4				0 vs. 1-4			
		1990-1999	2001	2003 ^a	2006 ^a	1990-1999	2001	2003 ^a	2006 ^a
Parents' Social Classification (NS-SEC)	1.1 Large Employers and Higher Managerial Occupations	0.40	0.34	0.39		-0.24	-0.43	0.33	
	1.2 Higher Professional Occupations	0.64	0.89	0.62		-0.35	-0.41	0.23	
	2 Lower Managerial and Professional Occupations	0.25	0.28	0.33		-0.09	-0.22	0.33	
	3 Intermediate Occupations	0.00	0.00	0.00		0.00	0.00	0.00	
	4 Small Employers and Own Account Workers	-0.41	-0.19	-0.22		0.26	-0.13	0.49	
	5 Lower Supervisory and Technical Occupations	-0.48	-0.58	-0.31		0.29	0.04	0.67	
	6 Semi-routine Occupations	-0.66	-0.60	-0.45		0.41	-0.07	0.66	
	7 Routine Occupations	-0.75	-0.64	-0.61		0.61	0.31	0.81	
Parents' Social Classification (Modified NS-SEC^d)	1 Higher Managerial and Professional Occupations				0.35				-0.27
	2 Lower Managerial and Professional Occupations				0.15				-0.05
	3 Intermediate Occupations				0.00				0.00
	4 Small Employers and Own Account Workers				-0.21				0.19
	5 Lower Supervisory and Technical Occupations				-0.36				0.42
	6 Semi-routine Occupations				-0.34				0.38
	7 Routine Occupations				-0.41				0.55
Constant		0.17	0.93	1.13	1.11	-0.64	-0.35	-1.12	-0.84

Table 1(continued): Multinomial Logistic Model (Survey Weighted) of School GCSE Attainment in Year 11 – Log Odds (significant variables highlighted in bold)

N	54236	12934	10269	5960
log likelihood (unweighted)	-49001	-10526	-7078	-5232
pseudo r squared (unweighed)	0.09	0.08	0.09	0.07
BIC	98547	21451	14543	10760

^a For the 2003 and 2006 cohorts the outcome variable includes other qualifications such as GNVQs.

^b The ethnicity question and its coding varies across the YCS cohorts. Ethnicity can be appropriately re-coded into seven main categories in the 1990 to 1999 cohorts (see Croxford *et al.* 2007). We were able to harmonise the data for the 2001 and 2003 cohorts but not for the 2006 cohort. Therefore a distinct ‘new other’ category has been explicitly reported for the 2006 cohort.

^c There is no comparable measure of household type deposited with the 2006 cohort data.

^d The parental socioeconomic information available in the 2006 cohort varies from the measures deposited in the previous cohorts. Usually the full eight class version of the National Statistics Socio-Economic Classification (NS-SEC) is deposited. However, in the 2006 cohort the first two categories of the eight class NS-SEC have been combined by the survey depositors. We envisage that this may lead to a minor loss of the theoretical clarity of this measure and it is therefore not fully comparable with previous cohorts (for a fuller discussion see Rose and Pevalin 2005).

We estimated a multinomial logistic regression model and report the results in Table 1. The outcome variable has three categories, five or more GCSEs at grades A*-C, 1-4 GCSEs at grades A*-C, and zero GCSEs at grades A*-C. We restrict our analyses to a set of established explanatory variables that are identified as being substantively important in previous studies of school GCSE attainment (for example Drew *et al.* 1992; Drew 1995; Demack *et al.* 2000; Gayle *et al.* 2003; Connolly 2006; Gayle *et al.* 2009; Sullivan *et al.* 2011). The 1990-1999 YCS cohorts are combined into a single cross-cohort analysis because we are able to construct comparable measures. The 2001, 2003 and 2006 cohorts are analysed separately because it is not possible to construct a full set of comparable variables.

Gender was significant net of the other variables included in the model across all eight of the YCS cohorts. Males had decreased odds of gaining five or more GCSEs at grades A*-C rather than being in the 'middle' category of 1-4 GCSEs at grades A*-C. Males had increased odds of gaining zero GCSEs at grades A*-C rather than being in the 'middle' category of 1-4 GCSEs at grades A*-C. This result chimes with other analyses of gender and educational attainment for this period (for example Warrington and Younger 2000; Gayle *et al.* 2003; Burgess *et al.* 2004; Younger and Warrington 2005; Connolly 2006).

The results for ethnicity show a familiar mixed pattern. It is well observed that there are differing levels of attainment and participation in post-compulsory education across ethnic groups (see Drew *et al.* 1992; Drew 1995; Biggart and Furlong 1996; Demack *et al.* 2000; Gillborn and Mirza 2000; Bhattacharyyal *et al.* 2003; Wilson *et al.* 2006). Young people of black ethnic origin had decreased odds of gaining five or more GCSE at grades A*-C rather than being in the 'middle' category of 1-4 GCSEs at grades A*-C. By contrast young people of Indian origin had increased odds of gaining five or more GCSE at grades A*-C rather than being in the 'middle' category of 1-4 GCSEs at grades A*-C. Although there is not a clear pattern for the recent YCS cohorts, young people in the 1990s cohorts from most of the minority ethnic groups had decreased odds of achieving zero GCSEs, compared with their white counterparts.

Housing tenure and household composition were associated with school GCSE attainment across all of the YCS cohorts. The offspring of renters had decreased odds of attaining five or more GCSEs at grades A*-C, but increased odds of attaining zero GCSEs at grades A*-C, compared to being in the 'middle' group. Household composition has a small effect on school GCSE attainment. Young people

in mother only households were generally not significantly different to their counterparts living in households with both parents. Those in father only, and in other households, tended to performed less well at GCSE.

The offspring of more educated parents had increased odds of gaining five or more GCSEs at grades A*-C. These results are consistent with other studies and earlier YCS results (Drew *et al.* 1992; Drew 1995; Gayle *et al.* 2003). Taken together they point towards the overall effect of having a more advantaged home background on school GCSE attainment.

Parental socioeconomic classifications are central to explaining patterns of school GCSE attainment³. Compared with young people who have parents in Intermediate Occupations (NS-SEC 3), those with parents who are either Large Employers or Higher Managers (NS-SEC 1.1), or Higher Professionals (NS-SEC 1.2) have increased odds of gaining five or more GCSEs at grades A*-C. By contrast young people who have parents in NS-SEC categories four to seven (these categories include own account workers, routine and manual occupations) have decreased odds of gaining five or more GCSEs at grades A*-C. In the 1990s cohorts, young people with parents in NS-SEC categories four to seven have increased odds of gaining zero GCSEs at grades A*-C rather than being in the ‘middle’ category of 1-4 GCSEs. The effect is similar in the 2003 and 2006 cohorts but not in the 2001 cohort. The overall relationship between parental socioeconomic position and GCSE attainment accords with the well documented view that those from more occupationally advantaged family backgrounds perform better (Drew *et al.* 1992; Drew 1995; Demack *et al.* 2000; Gayle *et al.* 2003; Connolly 2006; Gayle *et al.* 2009).

These initial results intimate that there is a ‘middle’ group of young people with moderate levels of GCSE attainment. These young people are not unqualified,

³ The National Statistics Socio-Economic Classification (NS-SEC) scheme is the official UK measure. A thorough methodological description of this measure is provided by Rose and Pevalin (2005). A practical guide to the use of this measure is available at <http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html>. Here are list of examples of an occupation for each of the NS-SEC categories to assist a reader that is unfamiliar with this classification scheme. 1.1 Large Employers and Higher Managerial Occupations – General Managers; 1.2 Higher Professional Occupations – University Teachers; 2 Lower Managerial and Professional Occupations – Secondary School Teachers; 3 Intermediate Occupations – Medical Secretaries; 4 Small Employers and Own Account Workers – Builders / Building Contractors; 5 Lower Supervisory and Technical Occupations – Bricklayers; 6 Semi-routine Occupations – Store Keepers; 7 Routine Occupations – Cleaners.

they have some GCSEs at the higher grades (A*-C), but they have not reached the benchmark of five or more GCSEs. This 'middle' group are more likely to be male, and be from a lower attaining minority ethnic group. It is less likely that their parents are homeowners or graduates and they are more likely to come from less occupationally advantaged families. These initial findings lend support to the conception of a 'middle' group of ordinary young people with moderate levels of GCSE attainment.

Further Exploring the 'Middle' Group

In this section of the paper we go beyond these initial results and provide more comprehensive exploratory data analyses of the concept of a 'middle' level of school GCSE attainment. The question of how to measure school qualifications, or indeed what 'measure' means, raises interesting issues since there is no agreed standard way of categorising educational qualifications (Prandy, Unt and Lambert 2004; Schneider 2011). GCSE subjects are assessed separately and a subject specific GCSE is awarded. It is usual for pupils in Year 10 and 11 to study for about nine subjects, which will include core subjects (e.g. English, Mathematics and Science) and non-core subjects.

Each GCSE subject is awarded an alphabetical grade. Historically the highest grade was A, and the lowest grade was G. From 1994 the higher grade of A* was introduced (Yang and Woodhouse 2001). Because GCSEs are taken as diet of many individually graded subjects, there is no obvious single, or agreed, measure of school GCSE attainment. The formal benchmark of five or more GCSEs at grades A-C (and more recently A*-C) has been ubiquitous however.

The overall limitation of the benchmark measure is that it treats an A* in English, a B in Sociology and a C in Mathematics similarly in determining whether or not a pupil has five GCSEs at grades A*-C (Gorard and Taylor 2002). For quite some time the Government league tables have also included a measure of the percentage of pupils gaining five or more GCSEs at grades A*-C including Mathematics and English (Taylor 2011). The addition of achieving grades A*-C in Mathematics and English does not however overcome the more general obstacle of how best to suitably combine individual GCSE subject results. Analysing the number of GCSE passes at grades A*-C is a plausible analytical strategy.

Figure 2 illustrates quite clearly that there has been improvement in the number of GCSEs attained over the decade from 1990-1999. A smaller proportion of young people failed to attain any GCSEs at grades A*-C in later years. At the other end of the continuum we also see improved performance by pupils in the later years, and both the median and the mean number of GCSEs at grades A*-C rose over the decade. From a synoptic examination of Figure 2 it is not noticeable that there are clear clusters of GCSE attainment (at grades A*-C), with the exception of the spike at zero. This leads us to question the validity of the theoretical idea of a ‘middle’, or moderate, level of attainment.

Note: n=54,236

Figure 2: Number of GCSEs A*- C Attained in Year 11 by School Year Cohort, 1990-1999.

To formally investigate the ordinality of GCSE attainment we estimated a stereotype logistic regression model (see Anderson 1984; Lunt 2001)⁴. A model with the number of GCSEs at grades A*-C as the outcome was estimated for the 1990-

⁴ We do not report the full output of this model, which contains identical explanatory variables to the multinomial logistic regression models reported above. The full output is available from the authors by request.

1999 cohorts. The model included YCS cohort, gender, ethnicity, housing tenure, household type, parental education and parental socioeconomic classification as explanatory variables. This can be considered as a formal test of whether the linear predictor best discriminates the outcomes of the dependent variable. A parameter ϕ_k provides a measure of the distinguishability of categories in relation to the predictors. If the ϕ parameters for two categories are similar it is likely that the categories are indistinguishable (see Lunt 2001). The results from the model show a general monotonic decline in ϕ for each additional GCSE at grades A*-C, and this is positive evidence of ordinality. We formally tested adjacent levels of GCSE attainment (i.e. the number at grades A*-C). The adjacent levels of attainment were all significantly different, with the exception of 4 GCSEs and 5 GCSEs at grades A*-C. We therefore conclude that the number of year 11 school GCSEs attained at grades A*-C is appropriately considered as being ordinal, and there is no persuasive evidence of any clear clusters of attainment.

In the next stage of the analyses we estimated a model of the number of GCSEs attained at grades A*-C for the 1990-1999 cohorts. Standard linear regression analysis is not suitable for count data (Cameron and Trivedi 1998). Poisson regression models are routinely used but, as we have indicated there is an over-representation of zero counts (i.e. over 9,000 pupils with no GCSEs at grades A*-C in the 1990-1999 cohorts). This limitation of a poisson approach is elaborated upon by Long (1997). The zero inflated poisson (ZIP) model overcomes this obstacle by modelling a two-state process (see Lambert 1992). In the present context this involves a logistic model which estimates the attainment of no GCSEs at grades A*-C, followed by a poisson model of the number of GCSEs at grades A*-C.

The zero inflated poisson model, Table 2, indicates a pattern of inequality in year 11 school GCSE attainment. Boys are more likely to gain zero GCSEs, there are some ethnic differences, and young people from more advantaged home backgrounds are less likely to gain zero GCSEs at grades A*-C.

Table 2: Zero Inflated Poisson Regression Model (Survey Weighted) GCSE Attainment Year 11 (Number of GCSEs at Grades A*-C), 1990-1999 School Year Cohorts

		Zero GCSEs (A*-C)	Number of GCSEs (A*-C)
YCS Cohort	1990	0.00	0.00
	1993	-0.15	0.16
	1995	-0.19	0.23
	1997	-0.12	0.24
	1999	-0.43	0.28
Gender	Girls	0.00	0.00
	Boys	0.64	-0.10
Ethnicity	White	0.00	0.00
	Black	0.02	-0.17
	Indian	-0.41	0.05
	Pakistani	-0.04	-0.12
	Bangladeshi	-0.71	0.06
	Other Asian	-0.99	0.15
	Other	-0.43	-0.02
Housing Tenure	Owned / Mortgage	0.00	0.00
	Renters	0.67	-0.25
	Others	0.25	-0.07
Household Type	Mother and Father	0.00	0.00
	Mother Only	0.09	-0.02
	Father Only	0.28	-0.08
	Other Household	0.60	-0.22
Parental Education	Non-graduates	0.00	0.00
	Graduates	-0.39	0.13
Parents' Social Classification (NS-SEC)	1.1 Large Employers and Higher Managerial Occupations	-0.50	0.12
	1.2 Higher Professional Occupations	-0.79	0.16
	2 Lower Managerial and Professional Occupations	-0.24	0.07
	3 Intermediate Occupations	0.00	0.00
	4 Small Employers and Own Account Workers	0.47	-0.12
	5 Lower Supervisory and Technical Occupations	0.52	-0.15
	6 Semi-routine Occupations	0.71	-0.22
	7 Routine Occupations	0.93	-0.28
Constant		-1.49	1.60

Total observations = 54,236; Non zero observations = 44,862; Zero observations = 9,374; Vuong test (unweighted model) $z = 74$, $p < .001$; Log likelihood (unweighted model) -135485.

Given that a young person obtains some GCSEs at grades A*-C, the model suggests that the factors we have identified are important predictors of the number of GCSEs at grades A*-C that they attain. Females perform better than males and there is an ethnicity related pattern to attainment. Pupils from more advantaged home backgrounds and those with more educated parents perform better in year 11. Parental occupational position is important and pupils with parents in more advantaged occupations gain more GCSEs at grades A*-C.

The model reported a significant Vuong test, and we therefore have solid grounds for favouring the zero inflated poisson model over a standard poisson model (see Vuong 1989). The results from the zero inflated poisson model persuade us that year 11 school GCSE attainment is reasonably considered as being located on a continuum. We are progressively dissuaded that there are clear clusters or groups of GCSE attainment. Therefore we are increasingly sceptical that there is a 'middle' level of school attainment that is characteristically different.

In the next stage of the analyses we explore year 11 school attainment by examining a points based measure of GCSE attainment. There are an infinite number of possible scores that could be assigned to the alphabetical grades ascribed to the levels of GCSE attainment. The point score deposited with the data in the 1990-1999 cohorts was calculated by allocating 7 points for an A*/A, 6 points for a B, 5 points for a C, 4 points for a D, 3 points for an E, 2 points for a F, and 1 point for a G (Croxford *et al.* 2007 p.52). This scoring was in line with the Qualifications and Curriculum Authority (QCA) approach when the harmonised 1990-1990 dataset was constructed. Because the A* grade was introduced midway through the data series, a grade A and a grade A* are awarded the same score. Yang and Woodhouse (2001) adopt the same strategy to splice GCSE data spanning the introduction of the A* grade.

We use a measure of GCSE attainment capped at 84 points (i.e. the equivalent of twelve GCSEs at grade A*/A). We chose this approach to limit the effects of pupils achieving higher scores simply as a function of having taken more GCSEs. Webber and Butler (2007) used a similar approach on the advice of DfES officials.

The Qualifications and Curriculum Authority have more recently developed a scoring system which awards an A* 58 points, an A 52 points, a B 46 points, a C 40

points, a D 34 points, an E 28 points, a F 22 points, and a G 16 points⁵. It is not possible to recode the GCSE scores in these older YCS samples onto the new QCA scale. We suspect that because the new and old scores for each GCSE grade are similarly spaced, the overall substantive interpretations of analyses that use the new scoring system will not be dramatically altered⁶.

There was not an extreme spike at zero points, and this is because many of the pupils that fail to achieve any GCSEs at grades A*-C are awarded points for subjects for which they gain awards at grades D-G. Table 3 reports the results for a standard linear regression model for the year 11 GCSE points scores. We present analyses for YCS cohort 1990-1999 in a single model. The 2001 and 2003 YCS cohorts are modelled separately because an alternative point score was deposited with these datasets which include other qualification (e.g. GCSE short courses). We are unable to include an analysis of the 2006 cohort because a suitable measure was not deposited with the public access dataset.

⁵See:http://www.education.gov.uk/schools/performance/secondary_11/PointsScoreAllocation2011.pdf.

⁶ More recently some official statistics are capped at the level to the best eight GCSEs. Other alternative approaches have been employed for example Haque and Bell (2001) convert GCSE attainment into numerical scores (A*=8, A=7...U=0) and calculate a mean GCSE score for each pupil. They chose this approach because they believe that this helps to prevent discrimination against pupils who have taken fewer GCSEs as a result of their school's internal policy. Similarly we can envisage the use of other summary measures of overall attainment, for example median scores. Ideally, we would pursue sensitivity analyses of additional alternative GCSE attainment measures, but such measures cannot be derived from data deposited with the SN5765 dataset.

Table 3: Linear Regression Models (Survey Weighted) School GCSE Attainment Year 11 (GCSE Points Score) – Beta Values (significant variables highlighted in bold)

		1990-99	2001 ^a	2003 ^a
YCS Cohort	1990			
	1993	4.78		
	1995	7.95		
	1997	7.21		
	1999	10.88		
Gender	Girls	0.00	0.00	0.00
	Boys	-4.73	-5.01	-5.53
Ethnicity	White	0.00	0.00	0.00
	Black	-3.43	-1.19	-2.80
	Indian	3.00	4.87	8.25
	Pakistani	-2.01	0.75	-1.98
	Bangladeshi	3.28	7.92	4.77
	Other Asian	6.46	8.42	1.72
	Other	0.84	1.11	2.77
Housing Tenure	Owned / Mortgage	0.00	0.00	0.00
	Rented	-7.37	-7.69	-10.74
	Others	-2.67	-5.79	-15.99
Household Type	Mother and Father	0.00	0.00	0.00
	Mother Only	-1.19	-1.10	-2.00
	Father Only	-2.94	-6.21	-8.16
	Other Household	-7.98	-8.44	-10.01
Parental Education	Non-graduates	0.00	0.00	0.00
	Graduates	4.95	4.23	6.35
Parents' Social Classification (NS-SEC)	1.1 Large Employers and Higher Managerial	4.53	3.83	1.10
	1.2 Higher Professional Occupations	6.44	8.02	3.98
	2 Lower Managerial and Professional Occupations	2.43	2.70	1.31
	3 Intermediate Occupations	0.00	0.00	0.00
	4 Small Employers and Own Account Workers	-4.72	-2.78	-4.68
	5 Lower Supervisory and Technical Occupations	-5.09	-5.33	-6.77
	6 Semi-routine Occupations	-6.96	-5.22	-7.78
	7 Routine Occupations	-9.14	-7.69	-10.54
Constant		33.83	44.77	51.22
R Squared		0.24	0.18	0.21
n		54236	12934	10269

^a For the 2001 and 2003 school year cohorts an alternative point score was deposited with the data which include other qualifications (e.g. GCSE short courses).

The overall pattern was one of increasing levels of attainment over time. The now familiar, and dispiriting, pattern of educational inequality emerges. There was a significant gender gap in all of the YCS school cohorts, with boys performing less well than girls. There was a mixed pattern of attainment across the minority ethnic groups. Young people with more educated parents scored higher on average, and pupils from more advantaged home backgrounds also performed better. The effect of parental occupational positions was dramatic for all of the samples and those pupils with parents in less advantaged occupations performed significantly less well, *ceteris paribus*.

The linear regression models presented capture a reasonable proportion of the variance in school GCSE attainment and convince us that there are clear substantive benefits to modelling point scores. We also consider that there are methodological benefits. Altman (2006) contends that the categorisation of variables can provide enticing simplicity, but this is gained at a cost. The drawbacks of categorisation are demonstrated in the greater explanatory power that is usually associated with modelling continuous variables. By using categorical measures of school GCSE attainment the extent of the variation between groups might potentially be underestimated. Categorising school GCSE attainment also has the potential to imply a misleading degree of homogeneity within categories. Finally, Sauerbrei and Royston (2010) remind us that continuous measures are attractive as they typically lend themselves to more parsimonious forms of analyses (e.g. linear regression).

Conclusion

Educational qualifications gained at school continue to be a motor that propels young people along alternative pathways. Indeed Noah and Eckstein (1992) state that while particular examinations have come and gone during the past forty years, the underlying social and educational significance of school examinations has been preserved.

Official figures report that levels of school GCSE examination performance have increased (see DfES 2007). These improvements in performance are mirrored in the YCS samples. Despite overall improvements in results, GCSE attainment remained highly stratified. Given the many decades of comprehensive secondary

schooling and the removal of the earlier two-tier qualification system⁷ this overall message of inequality is particularly dismal. On average girls performed better than boys and there were some marked differences in attainment for pupils from the main minority ethnic groups. A striking result is the impact of parental socioeconomic positions and, to a lesser extent, the other variables associated with the young person's home background. This is particularly important as much of the popular discourse associated with differences in school attainment focus on gender differences rather than differences between pupils from dissimilar social backgrounds.

We began the analysis with an open mind and, following Merton (1987), we have attempted to establish whether or not there is a distinctive 'middle' level of school GCSE attainment. The evidence provided here persuades us that there are no crisp boundaries that demark a 'middle' category of moderate GCSE attainment. A sizeable proportion of young people failed to gain any GCSEs at grades A*-C. This is obviously far short of the benchmark target and is consequential because those without school level qualifications usually have fewer choices and chances than their better qualified counterparts.

We are persuaded that GCSE attainment is situated on a continuum. Whether measured by the number of GCSEs gained at grades A*-C or by a GCSE score, attainment was similarly stratified. With the exception of the sharp spike of young people that were unsuccessful in gaining any awards at grades A*-C, we fail to observe the presence of any clusters that indicate clear cohesive GCSE attainment groups. Therefore we suggest that researchers exercise a suitable degree of caution before making additional claims about the GCSE attainment of the 'missing middle'.

Roberts (2011) calls for analyses of intermediate (or ordinary) groups that fall between the dualism of the successful and the unsuccessful groups. Our analysis convinces us that there are clear benefits to understanding school attainment on a continuum, and that measures which reflect the heterogeneity of GCSE performance as fully as possible (e.g. measured by the number of GCSEs gained at grades A*-C, or a GCSE score) should be preferred. In many social surveys only crude measures of educational attainment are available. For many analyses a categorical measure of school GCSE attainment will be adequate and functional. In such circumstances we recommend that categorical GCSE attainment measures should be understood as

⁷ General Certificate of Education Ordinary Levels, and Certificates of Secondary Education.

being more coarse groupings of a finer scale, rather than discrete substantively meaningful categories.

Theoretically the conception of a ‘middle’ group of ordinary young people was initially intellectually attractive. Replicating our previous analysis using YCS data has been insightful and has allowed us to further investigate the concept of ‘middle’ levels of educational attainment with larger samples with more detailed information on GCSE attainment. The idea that school GCSE attainment is best understood as a continuum rather than discrete clusters also emerges from our earlier analysis of the BHPS. The consistent finding leads us to believe that this is what Goldthorpe (2000) would term as an ‘empirical regularity’. Therefore we conclude that there is no persuasive evidence that there is a distinctive ‘middle’ level of overall school GCSE attainment.

In conclusion, and on wider reflection, we nevertheless believe that sociologists of youth should study ‘ordinary’ young people and moderate, or unspectacular, levels of educational attainment. There is much to be gained by understanding the educational experiences, characteristics and qualifications of all young people. That notwithstanding, researchers should avoid relying on overly simplistic categorisations and should seek to study the stratification of educational attainment across the fullest spectrum.

References

- Altman, D. 2006. "The cost of dichotomising continuous variables." *British Medical Journal* 332 (7549):1080.
- Anderson, John A. 1984. "Regression and ordered categorical variables." *Journal of the Royal Statistical Society. Series B. Methodological* 46 (1):1-30.
- Ashford, S., J. Gray, and M. Tranmer. 1993. *The introduction of GCSE exams and changes in post-16 participation*. Research Paper No. 23. Sheffield: Department of Employment.
- Babb, P. 2005. *A Summary of Focus on Social Inequalities*. London: Office for National Statistics.
- Bhattacharyya, G., L. Ison, and M. Blair. 2003. *Minority Ethnic Attainment and Participation in Education and Training: The evidence*. London: Department for Education and Skills.
- Biggart, A., and A. Furlong. 1996. "Educating 'discouraged workers' - Cultural diversity in the upper secondary school." *British Journal of Sociology of Education* 17 (3):253-266.
- Brown, P. 1987. *Schooling ordinary kinds; inequality, unemployment and the new vocationalism*. London: Tavistock.
- Burgess, S., B. McConnell, C. Propper, and D. Wilson. 2004. "Girls Rock, Boys Roll: An Analysis of the Age 14-16 Gender Gap in English Schools." *Scottish Journal of Political Economy* 51 (2):209-229.
- Cameron, A.C., and P.K. Trivedi. 1998. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- Connelly, R., S. Murray, and V. Gayle. 2013. "Young People and School GCSE Attainment: Exploring the 'Middle'." *Sociological Research Online* 18 (1).
- 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 (1):3-21.
- Croxford, L, C. Ianelli, and M. Shapira. 2007. Documentation of the Youth Cohort Time-Series Datasets, UK Data Archive Study Number 5765. Colchester: Economic and Social Data Service.
- Demack, S, D. Drew, and M. Grimsley. 2000. "Minding the Gap: Ethnic, gender and social class differences in attainment at 16, 1988-95." *Race Ethnicity and Education* 3 (2):117-143. doi: 10.1080/13613320050074005.
- Department for Education and Skills. 2001. *Statistics of Education: Public examinations GCSE/GNVQ and GCE/AGNVQ in England 2000*. Norwich: Her Majesty's Stationary Office.
- Department for Education and Skills. 2007. *Gender and Education: The evidence on pupils in England*. Nottingham: Department for Education and Skills.
- Department of Education. 1985. *General Certificate of Secondary Education: A general introduction*. London: Her Majesty's Stationary Office.
- Drew, D. 1995. *'Race', Education and Work: The Statistics of Inequality*. Aldershot: Avebury.
- Drew, D., J. Gray, and F Sime. 1992. *Against the Odds: The Education and Labour Market Experiences of Black Young People*. *Youth Cohort Series No. 19*. Sheffield: Training, Agency Research and Development Series.
- France, A. 2007. *Understanding youth in late modernity*. Buckingham: Open University Press.

- Furlong, A., and F. Cartmel. 2007. *Young People and Social Change: new perspectives*. Berkshire: Open University Press/McGraw Hill.
- Gayle, V, D. Berridge, and R.B. Davies. 2003. *Econometric Analysis of the Demand for Higher Education. Research Report 472*. London: Department for Education and Skills Research Series.
- Gayle, V, P Lambert, and S Murray. 2009. "School-to-Work in the 1990s: Modelling Transitions with large-scale datasets." In *Transitions from Education to Work: New Perspectives from Europe and Beyond*, edited by R Brooks. London: Palgrave Macmillan.
- Gillborn, D, and H Mirza. 2000. Educational Inequality. Mapping Race, Class and Gender. A Synthesis of Research Evidence. London: Office for Standards in Education.
- Goldthorpe, J. 2000. *On Sociology: Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press.
- Gorard, S., and C. Taylor. 2002. "What is segregation? A comparison of measures in terms of 'strong' and 'weak' compositional invariance." *Sociology* 36 (4):875-895.
- Haque, Z., and J. Bell. 2001. "Evaluating Performances of Minority Ethnic Pupils in Secondary Schools." *Oxford Review of Education* 27 (3):357-368.
- Jenkins, R. 1983. *Lads, citizens, and ordinary kids: working class youth lifestyles in Belfast*. London: Routledge and Kegan Paul.
- Jones, J., M Joyce, and J Thomas. 2003. "Non-employment and labour availability." *Bank of England Quarterly Bulletin* Autumn:291-303.
- Lambert, D. 1992. "Zero-inflated Poisson regression, with an application to defects in manufacturing." *Technometrics* 34:1-34.
- Leckie, G., and H. Goldstein. 2009. "The limitations of Using School League Tables to Inform School Choice." *Centre for Market and Public Organisation Working Paper Series No. 09/208*.
- Long, J.S. 1997. *Regression Models for Categorical and Limited Dependent Variables, Advanced Quantitative Techniques in the Social Sciences Series 7*. London: Sage.
- Lunt, M. 2001. "Stereotype Ordinal Regression." *Stata Technical Bulletin* 10:298-307.
- Merton, R.K. 1987. "Three Fragments From a Sociologist's Notebook: Establishing the Phenomenon, Specified Ignorance, and Strategic Research Materials." *Annual Review of Sociology* 13:1-28.
- Mobley, M., C. Emerson, Y. Goddard, S. Goodwin, and R. Letch. 1986. *All about GCSE*. London: Heinemann.
- Murray, S.J. 2011. "Growing up in the 1990s: Tracks and trajectories of the 'Rising 16's': A longitudinal analysis using the British Household Panel Survey." PhD, School of Applied Social Science, PhD Thesis, University of Stirling.
- Noah, H., and M. Eckstein. 1992. "The Two Faces of Examinations: A Comparative and International Perspective." In *Examinations: Comparative and International Studies*, edited by M. Eckstein and H. Noah, 147-170. Oxford: Pergamon Press.
- North, J. 1987. *The GCSE: An Examination*. London: The Clarridge Press.
- Payne, J. 1995. *Routes Beyond Compulsory Schooling. Youth Cohort Report No. 31*. Employment Department Research Series.
- Payne, J. 2000. *Progress of Low Achievers after Age 16: An Analysis of Data from the England and Wales Youth Cohort Study. Research Report No. 185*. Department of Education and Employment.

- Payne, J. 2001. *Work-based Training for Young People: Data from the England and Wales Youth Cohort Study. Research Report No. 276.* Department for Education and Skills Research Series.
- Payne, J. 2003. *Vocational Pathways at Age 16-19. Research Report No. 501.* Department for Education and Skills Research Series.
- Playford, C.J. 2011. "The role of families in the stratification of attainment: Parental occupations, parental education and family structure in the 1990s." PhD, School of Applied Social Science, PhD Thesis, University of Stirling.
- Prandy, K., M. Unt, and P. Lambert. 2004. "Not by degrees: Education and social reproduction in twentieth-century Britain." International Sociological Association Research Committee 28 on Social Stratification and Mobility, Neuchatel, Switzerland, 7-9th May.
- Pye, J. 1988. *Invisible Children: Who are the real losers in school?* Oxford: Oxford University Press.
- Raftery, A.E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111-163.
- Rice, P. 1999. "The impact of local labour markets on investment in higher education: evidence from the England and Wales Youth Cohort Studies." *Journal of Population Economics* 12:287-312.
- Roberts, Kenneth. 2013a. "Education to Work Transitions: How the Old Middle Went Missing and Why the New Middle Remains Elusive." *Sociological Research Online* 18 (1):3.
- Roberts, S. 2011. "Beyond 'NEET' and 'tidy' pathways: considering the 'missing middle' of youth transition studies." *Journal of Youth Studies* 14 (1):21-39.
- Roberts, S. 2012. "I just got on with it: the educational experiences of ordinary, yet overlooked boys." *British Journal of Sociology of Education* 22 (2):203-221.
- Roberts, S., and R. MacDonald. 2013. "The Marginalised Mainstream: Making Sense of the 'Missing Middle' of Youth Studies." *Sociological Research Online* 18 (1):21.
- Roberts, Steven. 2013b. "Youth Studies, Housing Transitions and the 'Missing Middle': Time for a Rethink?" *Sociological Research Online* 18 (3):11.
- Rose, D., and D.J. Pevalin. 2005. *The National Statistics Socio-Economic Classification: Origins, Development and Use.* Colchester: University of Essex.
- Sauerbrei, W., and P. Royston. 2010. "Continuous Variables: To Categorize or to Model?" In *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics*, edited by C. Reading. Voorburg: International Statistical Institute.
- Schneider, S. 2011. Measuring Educational Attainment. Survey Question Bank: Topic Overview 6.
- Sullivan, A, H Joshi, and D. Leonard. 2011. "Single-sex schooling and labour market outcomes." *Oxford Review of Education* 37 (3):311-332.
- Taylor, J. 2011. *The English Baccalaureate: how not to measure school performance.* Lancaster: Lancaster University Management School.
- Treiman, D.J. 2009. *Quantitative Data Analysis: Doing Social Research to Test Ideas.* San Francisco: John Wiley & Sons.
- Tunstall, R., R. Lupton, D. Kneale, and A Jenkins. 2011. *Growing up in social housing in the new millennium: Housing, neighbourhoods and early outcomes for children born in 2000. CASE paper 142.* London: CASE.

- Vuong, Q.H. 1989. "Likelihood ratio tests for model selection and non-nested hypotheses." *Econometrica* 57:307-333.
- Warrington, M, and M Younger. 2000. "The Other Side of the Gender Gap." *Gender and Education* 12 (4):493-508.
- Webber, R., and T. Butler. 2007. "Classifying pupils by where they live: How well does this predict variations in their GCSE results?" *Urban Studies* 44 (7):1229-1253.
- Wilson, D., R. Burgess, and A Briggs. 2006. *The Dynamics of School Attainment of England's Minorities. CASE paper 105*. London: CASE.
- Yang, M., and G. Woodhouse. 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, and M Warrington. 2005. *Rising Boys' Achievement. Research Report 636*. London: Department for Education and Skills.