Ethnic disproportionality in the identification of Special Educational Needs (SEN) in England:

Extent, causes and consequences

Steve Strand

Ariel Lindorff

University of Oxford

Department of Education

20 December 2018

Table of contents

Acknowledgments

This research was funded by the Department for Education (DFE) through the DFE highlight to the Economic and Social Research Council (ESRC) Secondary Data Analysis Initiative (SDAI): Grant Number ES/P000991/1.

Any enquiries regarding this publication should be sent to:
steve.strand@education.ox.ac.uk

This document is available for download at:
http://www.education.ox.ac.uk/research/the-unequal-representation-of-ethnic-minorities-in-special-education-in-england-extent-causes-and-consequences

Executive Summary

Background

For some considerable time there has been concern over the process of special education referral and the differential representation of ethnic minority groups with Special Educational Needs (SEN) both in the US (Dunn, 1968) and in England (Coard, 1971). *Ethnic disproportionality* exists when an ethnic group is significantly more, or significantly less, likely to be identified with SEN compared to the ethnic majority. A recent major review concluded that disproportionate identification of Black pupils with SEN is "among the most long-standing and intransigent issues in the field" (Skiba et al, 2008, p264).

Extensive research with nationally representative data in the US has established that Black pupils are substantially more likely to be identified with Special Educational Needs (SEN) than other ethnic groups, with the odds of being identified with Intellectual Disabilities 2.8 times higher, and the odds of being identified with Emotional Disturbance 2.3 times higher, than White pupils. In England there have been only two nationally representative studies on disproportionality in the last 25 years (Strand & Lindsay, 2009; 2012) but these also revealed the odds for Black Caribbean and Pakistani pupils being identified with Moderate Learning Difficulties (MLD) were 1.5 times higher than for White British pupils, and the odds for Black Caribbean and Mixed White and Black Caribbean (MWBC) pupils being identified with Social, Emotional and Mental Health (SEMH[1]) Needs were twice those for White British pupils.

MLD and SEMH are the highest frequency SEN, together accounting for nearly half of all identified SEN, but disproportionality is not limited to MLD and SEMH. Studies in the US on the identification of Autistic Spectrum Disorders (ASD) show mixed results for Black pupils but consistent under-representation for Hispanic pupils, with reported prevalence rates among 8-year olds Hispanic pupils of 0.59% compared to 0.90% for White pupils, a relative risk ratio of 0.66 (Travers et al, 2011; Sullivan, 2013). In England, Strand & Lindsay's (2009) analysis reveals substantial under-representation of Asian pupils with ASD, with the odds of identification for Indian, Pakistani and Bangladeshi pupils about half the odds for White British pupils. Thus the under-representation of some ethnic groups is just as important to understand as the over-representation of others, as it may indicate barriers to accessing services and provision.

In sum, there are positive outcomes of being identified with SEN, such as access to specialist resources and additional support. However, there are also possible negative outcomes, particularly for needs such as MLD and SEMH, which might include an inappropriate or narrowed curriculum, restriction of opportunities because

[1]. Prior to September 2014 the analogous category was Behavioural, Emotional and Social Difficulties (BESD).

of lowered expectations, or feelings of stigmatisation/labelling on the part of identified pupils. There is a danger that ethnic disproportionality, if not addressed, may through inadequate or inappropriate provision perpetuate the same unequal outcomes in the future.

This issue is increasingly salient as the minority ethnic population in England continues to grow. In the 2016 National School Census ethnic minority groups accounted for almost one-third (30%) of pupils of compulsory school age (aged 5-16) in England, more than double the 14.2% recorded in 2003 (DFE, 2016).

Causes of disproportionality

Some forms of SEN have a clear biological basis, for example sensory impairments, physical needs, or profound and multiple learning difficulties. These categories are often contrasted with categories like SEMH/MLD which are more socially constructed, in the sense that they rely on pupils' behaviour/performance being interpreted in terms of expected patterns or norms. A frequently proposed explanation for the over-representation of Black pupils with SEMH/MLD is inappropriate interpretation of ethnic and cultural differences including teacher racism, low expectations and a failure of schools to provide quality instruction or effective classroom management (e.g. Artiles et al, 2010; Waitoller et al, 2010).

However, an alternative hypothesis is that disproportionality reflects the fact that ethnic minority pupils are more at risk of SEMH/MLD because of the substantially greater socio-economic disadvantage they experience relative to the White majority. For example, in England in 2016, 14% of White British pupils are eligible for a Free School Meal (FSM) but this doubles to 25% of Black African, 28% of Black Caribbean and 29% of Mixed White and Black Caribbean pupils (Strand & Lindorff, this report). Some recent longitudinal studies in the US have even claimed that when further controls are included, for example for educational achievement and teacher's ratings of pupils' behaviour at Kindergarten entry, Black pupils, rather than being over-represented relative to White pupils, are actually under-represented (Hibel et al, 2010; Morgan et al, 2015, 2017).

The need for this study

Many of the studies in this area have methodological limitations. Most studies in England, with the exception of Strand & Lindsay, have been small scale and unrepresentative. Many of the US studies are large but typically based on aggregate district or school level data rather than pupil level data, or are based on longitudinal studies that are hampered by small samples of ethnic minority pupils with SEN. In contrast, the England National Pupil Database (NPD) offers a complete census of recorded SEN for all pupils in the population (over 6 million pupils each year), is collected at pupil level and is contemporary, not historic, data.

The project addressed the following research questions:

- Considering the most recent (2016) national data, what is the current picture of ethnic disproportionality in England? For which ethnic groups and which types of SEN does disproportionality exist?
- There has been substantial change in the ethnic composition of the population of England over the last decade or so, as well as changes in the rates of identification of SEN. Have patterns of ethnic disproportionality changed between 2005 and 2016?
- To what extent can ethnic disproportionality in different types of SEN be accounted for by age, sex, and socio-economic disadvantage?
- How does disproportionality develop dynamically as children progress through school over time? What can we learn by tracking a primary cohort from Reception to Y6, and a secondary cohort from Y6 to Y11? Can academic achievement or development on-entry to school account for disproportionality?
- What is the variability across schools in disproportionality? Do school variables (such as the socio-economic and ethnic composition of the school) have any additional association with disproportionality when these are modelled alongside pupil level variables?
- What is the variability across Local Authorities (LA) in disproportionality? What data on disproportionality might be reported to LAs to assist in highlighting local issues and needs?

What we did

Main analyses

The research had six main strands:

1. **A comprehensive analysis of the 2016 NPD data** for all pupils aged 5-16 to determine the current extent of ethnic disproportionality, and whether age, sex, and socio-economic factors such as poverty and neighbourhood deprivation can account for ethnic over- and under-representation, looking at all types of SEN;
2. **An exploration of trends over the last 12 years** through analysis of previous NPD datasets back to the 2005 dataset reported on by Strand & Lindsay, to identify trends in (a) the prevalence of MLD, SEMH and ASD over the period, and (b) the level of ethnic disproportionality for these SEN over time;
3. **Two longitudinal analyses, each of over 500,000 pupils, one tracked from age 5 to age 11 and the second from age 11 to age 16,** to assess the emergence of SEN over time using survival analysis, whilst further accounting for attainment and development on-entry to primary and secondary school respectively, as well as pupil background characteristics (age, sex, and socio-economic variables);

4. **Estimation of the relative influence of the pupil, school and Local Authority (LA)** in accounting for variability in SEN identification and in ethnic disproportionality. We consider specifically what role mainstream schools play in the identification of SEN, and conduct analyses separately for the primary (Y1-Y6) and secondary (Y7-Y11) phases using multilevel logistic regression models.
5. **Cross-validation of the NPD analysis** using the Second Longitudinal Study of Young People in England (LSYPE2), which contains a wider range of socio-economic and family background data.
6. **Calculating ethnic disproportionality indicators for each LA**, to help in identifying needs and issues in local areas.

The data

Level of SEN: Around 3% of pupils have a formal statement of SEN or Education and Health Care (EHC) Plan. This means a legal document is in place that sets out the child's needs and the extra help they should receive. However, the majority of pupils with SEN are identified at School Support (12.8% of the school population and 81% of all those with an identified SEN). These pupils also receive provision that is additional to or different from that made generally for others of the same age, and which goes beyond the differentiated approaches and learning arrangements normally provided as part of high quality, personalised teaching. We combined these groups in the majority of our analyses.

Type of SEN: Schools are asked to record the primary need of SEN pupils from one of twelve specific types of need. In the initial stages of our analysis we look at disproportionality for all twelve types of need, we later focus in depth on three types:

Moderate Learning Difficulties (MLD): this is the most frequently identified SEN, accounting for 4.0% of pupils aged 5-16. These pupils may "learn at a slower pace than their peers, even with appropriate differentiation" (DFE, 2015, p97)

Social, Emotional and Mental Health (SEMH): the second most frequently identified type of SEN, accounting for 2.8% of pupils aged 5-16. These difficulties "may include becoming withdrawn or isolated, as well as displaying challenging, disruptive or disturbing behaviour. These behaviours may reflect underlying health difficulties… [or] disorders such as attention deficit disorder, attention deficit hyperactive disorder or attachment disorder (DFE, 2015, p98)

Autistic Spectrum Disorders (ASD): this is the most commonly identified need among pupils who have a statement of SEN, and is also a rapidly growing need, increasing from 0.5% of the population in 2005 to 1.3% in 2016. "Pupils with ASD, including Asperger's syndrome and Autism, are likely to have particular difficulties with social interaction. They may also experience difficulties with language, communication and imagination, which can impact on how they relate to others" (DFE, 2015, p97)

Ethnicity: Pupils' ethnic group is recorded in 18 categories that were introduced in 2002/03 and are standard throughout education administrative databases in England and also used in the national decennial census. We use White British as the (majority) reference group and compare each of the ethnic minority groups to White British.

Measuring disproportionality

The key measure we employ is the Odds Ratio (OR), which represents the odds of identification for a particular ethnic minority group relative to the odds of identification for the White British majority group. Thus, an OR of 2.0 indicates twice the odds of being identified compared to White British pupils, an OR of 1.0 means the same odds of being identified as White British pupils, and an OR of 0.50 means half the odds of being identified compared to White British pupils. We considered the size of ORs in relation to the following cut-offs:

OR <= 0.67 "substantially under-represented"
OR <= 0.75 "under-represented"
OR >= 1.33 "over-represented"
OR >= 1.50 "substantially over-represented".

We avoid emphasis on results for very small ethnic groups as ORs for these groups are more volatile (e.g. Irish and Roma Traveller groups); these results are, however, included in tables in the full report.

Evaluation strategy

We first examine results that only take ethnic group into account (described as 'unadjusted' ORs). We then compute statistical models that control for other pupil background factors including year group, birth season (autumn/spring/summer), sex, eligibility for a Free School Meal (FSM) and home neighbourhood deprivation (Income Deprivation Affecting Children Index; IDACI), to produce 'adjusted' ORs for ethnic minority groups independent of the effects of those other background factors.

We follow the same strategy for our longitudinal analyses, looking first at ethnicity only, then adding age, sex and socio-economic factors, and then prior attainment / development at age 5 or age 11 respectively. In all our models we have a final step that includes consideration of school level variables, such as the percentage of pupils in the school entitled to FSM or the percentage of pupils from different ethnic minority groups.

Key Findings

There is marked disproportionality for the following ethnic groups and SEN:

- Black Caribbean and Pakistani pupils are over-represented for MLD, Indian and Chinese pupils are under-represented;
- Black Caribbean and Mixed White & Black Caribbean pupils are substantially over-represented for SEMH;
- All Asian Groups (Indian, Pakistani, Bangladeshi and Other Asian) are substantially under-represented for SEMH and for ASD.

The over-representation for MLD can be accounted for by socio-economic factors, but the ethnic disproportionalities for SEMH and ASD remain substantial even after pupil background controls for age, sex and socio-economic deprivation. This is not because of the limited socio-economic measures available in the NPD, as we found the same results after control for parental social class, parental education and family income using the Second Longitudinal Study of Young People in England (LSYPE2).

Prior attainment/development also does not account for the ethnic disproportionality in SEMH and ASD. Literacy and mathematics measures from the Early Years Foundation Stage Profile at age 5 were strongly predictive of the likelihood of subsequent identification of MLD, and the Personal, Social and Emotional Development (PSED) measure was highly predictive of subsequent identification of SEMH and ASD. However, this did not remove the ethnic disproportionality for SEMH and ASD which remained substantial. The findings for the secondary cohort, accounting for age 11 English and mathematics attainment on-entry to secondary school, led to the same conclusion.

Local Authorities (LAs) account for little (2%-6%) of the variation in the identification of SEN. Patterns of disproportionality vary little in direction across LAs e.g. of 113 LAs with sufficient data for SEMH calculations, 84 show over-representation of Black Caribbean/Mixed White & Black Caribbean pupils, none show under-representation. Similarly, of 94 LAs with sufficient data for ASD calculations, 79 show under-representation of Asian pupils, only three show over-representation. This consistency suggests that variation in LA policy and practice plays a limited role in the over-representation of Black Caribbean/Mixed White & Black Caribbean pupils with SEMH or the under-representation of Asian pupils with ASD.

There is variation between schools in the frequency with which they identify SEN, but schools play a limited role in accounting for ethnic disproportionality, with the notable exception of identification of SEMH in secondary school. In null models, around one-fifth of the variance in MLD is between schools (22%-25%) somewhat less for SEMH (13%-15%) and much less for ASD (11%-12%). Some of

this variation can be explained by the socio-economic composition of the pupil intake, and by factors like school size and type (e.g. Grammar schools had very few SEN pupils). However, differences between schools played little role in accounting for ethnic disproportionality, with the notable exception of SEMH in secondary schools. Differences between secondary schools account for a substantial part of the over-representation of Black Caribbean and Mixed White and Black Caribbean pupils with SEMH. i.e. their over-representation occurs much more in some secondary schools than it does in others.

Longitudinal studies, even with large representative samples, can often be under-powered to detect relatively low incidence outcomes like type of SEN for ethnic minority groups. Results from sample studies need to be interpreted with caution and more population level studies, like those reported here, are required.

Detailed findings

Moderate Learning Difficulties (MLD)

Pakistani pupils (OR= 1.36) and Black Caribbean pupils (OR= 1.38) were over-represented for MLD relative to White British pupils. Indian (OR= 0.56) and Chinese (OR= 0.30) were substantially under-represented.

Despite changes in prevalence, the extent of the ethnic disproportionality noted above has not changed notably since 2005. Prevalence rates for MLD increased from 2.6% in 2005 to 4.0% in 2016. In part this reflects the fact that from 2015 onwards type of SEN was requested for all pupils on School Support, not just those on the former School Action Plus, so more pupils are recorded as having a specific type of need. However, the change in prevalence did not alter the extent of ethnic disproportionality.

The over-representation of Pakistani and Black Caribbean pupils could be accounted for by socio-economic factors. Pupils were more likely to be identified with MLD if they were entitled to a Free School Meal (OR= 2.4), lived in a deprived neighbourhood (OR= 1.9), were boys (OR= 1.7) and were young for their year group (summer-born pupils OR= 1.8). After controlling for these factors, Pakistani and Black Caribbean pupils were no more likely to be identified than White British pupils with similar characteristics.

Accounting for attainment and/or social development at the start of school made little difference to the results by ethnic group. Literacy and mathematics measures from the Early Years Foundation Stage Profile (EYFSP) at age 5 were very strong predictors of a pupil's likelihood of subsequent identification of MLD during primary school. However, it did not change the pattern of ethnic group difference, with many ethnic groups (particular Black African, Indian and Bangladeshi pupils) less likely to be identified with MLD than White British pupils with the same prior attainment and socio-economic background. Similar conclusions apply when

accounting for reading and mathematics test scores at age 11 for the secondary longitudinal cohort.

Differences between LAs and schools made little contribution to ethnic disproportionality for MLD. Local Authorities (LAs) account for very little (5%-6%) of the variation in identification. There is more variation at the school level (22%-26%), with some schools more likely to have pupils identified with MLD than others, and this partly reflects the characteristics of the pupils attending the school (e.g. more pupils identified in small schools and those with more deprived intakes). Importantly though, accounting for differences between schools did not materially alter the ethnic coefficients for under/over-representation**,** either at primary or secondary phases.

Social, Emotional and Mental Health (SEMH)

Black Caribbean (OR= 2.29) and Mixed White and Black Caribbean (OR= 1.94) pupils were substantially over-represented relative to White British pupils. Asian groups were all substantially under-represented, as was the White Other group (OR= 0.57).

The extent of ethnic disproportionality for the above groups has remained constant since 2005. Prevalence rates increased from 1.9% in 2005 to 2.8% in 2016, although as stated earlier this partly reflects the increase since 2015 in the number of pupils for whom data on type of need is requested. Importantly, though, the ethnic disproportionality identified above has not altered with the change in terminology from BESD to SEMH. This is perhaps not surprising since displaying "challenging, disruptive or disturbing behaviour" remains central to the description of SEMH (DFE, 2015), whatever the putative drivers of such behaviour.

Demographic and socio-economic variables had very strong associations with identification of SEMH, but controlling for these factors did not account for the ethnic over-representation. The odds of being identified with SEMH needs were much higher for boys than girls (OR= 3.2); for pupils entitled to FSM (OR= 3.1), for pupils from disadvantaged neighbourhoods (OR= 1.9) and for pupils in secondary school, particularly Y10 and Y11 (OR= 2.1 and OR= 2.4 compared to Y1). Controlling for these factors attenuated but did not eliminate the over-representation of Black Caribbean (OR= 1.43) and Mixed White & Black Caribbean (OR= 1.38) pupils.

Similarly, controlling for prior attainment/development at the start of school did not account for Black Caribbean and Mixed White & Black Caribbean over-representation. Literacy and mathematics scores at age 5 had little association with subsequent identification of SEMH, but a below average Personal, Social and Emotional Development (PSED) score at age 5 raised the odds substantially (HR= 2.54). The mean PSED scores for Black Caribbean and Mixed White & Black Caribbean pupils were lower than the national average, but even after adjusting for

this Black Caribbean (HR= 1.42) and Mixed White & Black Caribbean (HR= 1.46) pupils were still over-represented. The findings for the secondary cohort, accounting for English and mathematics national test scores at age 11 on subsequent identification of SEMH during secondary school led to the same conclusion, with Black Caribbean (OR= 1.37) and Mixed White & Black Caribbean (OR= 1.53) pupils remaining over-represented.

Secondary schools seem to account for a significant part of the over-representation of Black Caribbean and Mixed White & Black Caribbean pupils with SEMH. In secondary schools the ORs for Black Caribbean and Mixed White & Black Caribbean pupils reduced substantially between single-level and multi-level models, from OR= 1.47 to 1.14 and from OR=1.47 to 1.29 respectively. This indicates that differences between schools play a part in the over-representation of these two specific ethnic groups. Our longitudinal analyses indicate that over-representation was reduced when account was taken of school composition factors, particularly in secondary schools. For example, schools in the top two quintiles of %FSM, and in the top two quintiles for % Black Caribbean pupils, had significantly raised odds of identification, and allowing for this did reduce the Black Caribbean and Mixed White & Black Caribbean over-representation. This suggests a particular focus on the context of, and processes occurring within, schools serving high deprivation communities and with large proportions of Black Caribbean and Mixed White & Black Caribbean pupils. What drives these associations is unknown, and could include unmeasured factors associated with high deprivation (e.g. high levels of crime, violence or gang culture), negative peer effects (such as disaffection or disengagement) or school policies (e.g. pre-emptive or zero tolerance disciplinary strategies).

Variation between LAs is minimal, accounting for <2% of variation in identification of SEMH. Of 113 LAs with sufficient data for SEMH calculations, 84 show over-representation for the combined Black Caribbean/MWBC group, none show under-representation. Nevertheless, there is a range in the risk ratios for 2016 from 0.77 in Newham to 3.15 in Barnsley. Data should be monitored annually to determine if any consistent LA patterns emerge.

Care needs to exercised in generalisations about 'Black' pupils. Black African pupils represent 3.7% of all pupils in England, a much larger group than either Black Caribbean (1.2%) or Mixed White and Black Caribbean (1.5%) pupils. They experience similar levels of socio-economic disadvantage yet they are not over-represented for SEMH, and are actually under-represented in the adjusted ORs, both in relation to socio-economic disadvantage and to prior attainment. This indicates that in the England context, care needs to be exercised in generalisations about 'Black' pupils. Similar differences have been reported for other outcomes such as exclusion from school; attitudes, aspiration and motivation; and academic achievement, and may be related to recency of migration (e.g. Strand, 2011, 2012).

Autistic Spectrum Disorders (ASD)

There was substantial ethnic disproportionality for ASD. Black Caribbean and Black Other pupils were over-represented (both ORs= 1.34) compared to White British pupils. Asian groups were under-represented, particularly Indian (OR= 0.46) and Pakistani pupils (OR= 0.54) where the odds of identification were half those for White British pupils. White Other pupils (OR= 0.60) were also under-represented.

There was more variation in ethnic disproportionality over time than was the case for other SEN. Black Caribbean pupils were not over-represented 2005-2009 but have been consistently over-represented since 2011 (OR= 1.12 in 2005 to OR= 1.34 by 2016). White Other groups were not under-represented 2005-2009 but have been consistently under-represented since 2011 (OR= 0.96 in 2005 to OR= 0.60 by 2016). On a positive note, the under-representation of Bangladeshi pupils has steadily declined (from OR= 0.38 in 2005 to OR= 0.79 by 2016).

Demographic and socio-economic variables had strong associations with identification of ASD. Controlling for these factors accounted for the over-representation of Black Caribbean and Black-Other groups, but did not account for the under-representation of Asian pupils. The odds of being identified with ASD were much higher for boys than girls (OR= 5.4) and for pupils entitled to FSM (OR= 2.3), and were slightly raised for pupils from more disadvantaged neighbourhoods (OR= 1.2). Controlling for these factors had little or no impact on the under-representation of Asian pupils, who were still about half as a likely as White British pupils to have an identification of ASD. On the other hand, Black Caribbean and Black Other pupils were no longer over-represented (OR= 1.12 and OR= 1.13), suggesting that their over-representation was largely attributable to socio-economic factors.

Controlling for prior attainment/development at the start of school did not change the pattern of ethnic disproportionality. Below average Personal, Social and Emotional Development (PSED) scores at age 5 were associated with substantially increased odds of ASD identification (HR= 3.2), and the mean PSED scores for Pakistani and Bangladeshi pupils were below the national average, but after adjusting for these scores Asian groups remained under-represented. Similarly, for the secondary cohort, higher English and mathematics national test scores at age 11 were associated with lower odds of identification with ASD, but controlling for prior attainment did not alter the Asian under-representation.

LA and school variability was small, and school composition variables had little impact on ethnic disproportionality. Around 4% of variance was at the LA level and 11%-12% at the school level, much lower than for MLD or SEMH. Generally, school level factors had little impact on ethnic disproportionality. However, both longitudinal cohorts suggested that pupils were somewhat more likely to be

identified in schools in the top two quintiles for % Asian pupils, particularly among secondary schools, indicating that a high concentration of Asian pupils slightly moderated the effect, but overall Asian pupils remained substantially under-represented compared to White British pupils.

Parental education qualifications may be an important factor in identification of ASD. Some, predominantly US, research has suggested that high socio-economic families are *more* likely to receive an ASD diagnosis (e.g. Durkin et al, 2010) while our data indicates the opposite. Our NPD measures are of socio-economic *dis*advantage, which may be blunt in differentiating at the more advantaged end of the SES range. However, our analysis of LSYPE2, using parental occupation, educational qualifications and family income, broadly confirmed the NPD results, with pupils from low SES homes (parents in routine and semi-routine occupations) more likely to have an ASD identification than those in managerial and professional households (OR= 3.0 and 2.6 respectively). It may be that in England the NHS provides more equitable access to services with fewer financial barriers than in the US, and our study is based in schools where all children can be assessed rather than in clinics or other settings. Nevertheless, we note that once parental occupation was controlled, the odds of ASD identification were twice as high in homes where one or more parent held a degree compared to similar homes where parents' highest educational qualifications were below degree level. This does suggest that parental awareness and access to resources may be an issue.

The causes of ethnic disproportionality in identification of ASD are likely to be varied. Less extreme needs on the autistic spectrum can be subtle, identified by nuances in the use of language for social communication. These may be more difficult to identify if the first language of the assessor and pupil are not congruent, as might be the case for many pupils of Asian heritage. It may also be that these are communities with lower awareness of autism, parents' rights and relevant services; where having a child with SEN is particularly stigmatizing; where cultural or linguistic barriers impede access to services; or where the services available do not meet their needs (Corbett & Perapa, 2007). In any event, there is a need to raise awareness of ASD among Asian communities, improve outreach and review the extent to which services are configured appropriately.

Implications for policy and practice

- LAs, multi-academy trusts (MATs) and schools must have due regard to the Public Sector Equality Duty (PSED) requirements, and should monitor ethnic disproportionality in the identification of SEN.
- LAs or MATs with high levels of disproportionality should further investigate practices in their areas/trusts. Schools should identify priorities for the partnerships within which they work, so they can pool resources and develop effective responses.

- OFSTED should incorporate data on ethnic disproportionality into pre-inspection reports for LA SEND inspections, and include the issue of ethnic disproportionality within the LA SEND inspection.
- The original detailed guidance on data collection by type of SEN (DFE, 2005) no longer exists following the new SEND Code of Practice. The DFE should consider new guidance on definitions and criteria for defining different types of SEN.
- Secondary schools in particular should review their processes around the identification of SEMH needs, given variability between schools is a strong component of ethnic disproportionality in this domain.
- LAs and schools need to raise awareness of ASD among Asian communities, improve outreach and review the extent to which the services are configured appropriately for access by ethnic minority groups.
- Teachers need to be aware of the significant over-identification of summer born pupils for MLD and to consider carefully whether they are making sufficient allowance for the age of the child when forming their judgements.

Introduction

For some considerable time there has been concern over the process of special education referral and the differential representation of ethnic minority groups with Special Educational Needs (SEN) both in the US (Dunn, 1968) and in England (Coard, 1971). *Ethnic disproportionality* exists when an ethnic group is significantly more, or significantly less, likely to be identified with SEN compared to the ethnic majority. A recent major review concluded that disproportionate identification of Black students with SEN is "among the most long-standing and intransigent issues in the field" (Skiba et al, 2008, p264).

Ethnic disproportionality in the US

There is substantial nationally representative evidence from the US for ethnic disproportionality with respect to SEN (e.g. Donovan & Cross, 2002). Recent national data (US Department of Education, 2014) reports that Black students - relative to the combined data for all other ethnic groups - are 2.8 times more likely to be identified with Intellectual Disabilities (ID), 2.3 times more likely to be identified with Emotional Disturbance (ED) and 1.5 times more likely to be identified with Specific Learning Disabilities (SLD). Studies on the identification of Autistic Spectrum Disorders (ASD) show mixed results for Black pupils. Some studies have indicated Black pupils are over-represented relative to White pupils (Donovan & Cross, 2002); some have indicated no differences (CDC, 2009; Morrier & Hess, 2012); some that Black pupils are under-represented (Kogan et al, 2009; Mandell et al, 2009); and others that disproportionality varies by time, declining in more recent studies (Travers et al. 2011; 2014) or varies by the severity of the ASD (Jarquin et al, 2011). In a recent review the results have been judged “inconclusive” (Newschaffer, 2008, p240). However there does seem to be consistent under-representation for Hispanic pupils, with reported prevalence rates among 8-year olds Hispanic pupils of 0.59% compared to 0.90% for White pupils, a relative risk ratio of 0.66 (CDC, 2009; Travers et al, 2011; Sullivan, 2013).

Such data have led to a strong regulatory framework in US federal law in the form of the Individuals with Disabilities Education Act (IDEA) re-authorised in 1997 and 2004. States “must have in effect policies and procedures designed to prevent the inappropriate over-identification or disproportionate representation by race and ethnicity of children as children with disabilities”, should “collect, examine and publically report data on disproportionality in special education“, and, if disproportionality is found, must “pprovide for the review and revision of policies, procedures and practices used in identification or placement of children” (US Dept. of Education, 2014, 20 USC 1412(a)(24)).

Ethnic disproportionality in England

There have been only two nationally representative studies of disproportionality in England in the last 25 years. Strand & Lindsay (2009) report a detailed analysis of

the 2005 National School Census (SC) covering over 6.4 million students aged 5-16 years in all maintained schools in England. The analysis revealed that the odds of Black Caribbean and Pakistani students being identified with Moderate Learning Difficulties (MLD) were respectively 1.32 and 1.46 times higher than the odds for White British students. Even more strikingly the odds of Black Caribbean and Mixed White & Black Caribbean students being identified with Social Emotional and Mental Health (SEMH[2]) needs were 2.3 and 2.0 times higher than the odds for White British students. Importantly several other ethnic minority groups, including Black African, Mixed White & Black African, Indian, Bangladeshi and Chinese students, were substantially *under* represented relative to White British, particularly for SEMH. MLD and SEMH are the highest frequency SEN, together accounting for nearly half of all identified SEN, but disproportionality is not limited to MLD/SEMH. Strand & Lindsay (2012) analysed the 2011 school census and report over-representation of Black Other and Black Caribbean pupils for ASD, with odds 1.35 and 1.36 times higher than the odds for White British pupils. They also report substantial under-representation for Asian groups, with the odds of identification of ASD for Indian, Pakistani, Bangladeshi and other Asian pupils half the size of the odds for White British pupils.

Does ethnic disproportionality matter?

There are positive outcomes of SEN identification, including the identification and clarification of students' educational needs; individual action plans to address these needs; and access to specialist input and resources. There may be negative consequences if such needs are not identified because of cultural or other barriers, and the under-representation of particular ethnic groups is important to understand. However, there are also possible negative outcomes associated with SEN identification, particularly for needs such as MLD and SEMH, which might include: restriction of opportunities because of lowered expectations, an inappropriate or narrowed curriculum, and feelings of stigmatisation or labelling on the part of identified students. In these cases the over-representation of particular ethnic groups may also be a concern. In any event there is a danger that ethnic disproportionality, if not addressed, may through inadequate or inappropriate provision perpetuate unequal outcomes in the future. This issue is increasingly salient as the minority ethnic population in England continues to grow. In the 2016 National School Census ethnic minority groups accounted for almost one-third (30%) of pupils of compulsory school age (aged 5-16) in England, more than double the 14.2% recorded in 2003 (DFE, 2016).

These results have important implications for Government, Local Authorities and schools as the Equalities Act (2010) places a general duty on all public bodies to eliminate unlawful discrimination and advance equal opportunity. In England LAs have statutory responsibility for identifying and addressing SEN but, in contrast to

[2]. Prior to September 2014 the analogous category was Behavioural, Emotional and Social Difficulties (BESD).

the US, there is no requirement for LAs to collate, examine and report data on disproportionality in SEN. There is therefore a substantial vacuum between the general duties of the Equalities Act (2010) and the analysis needed to support effective action.

Explanations of disproportionality

Teacher bias or socio-economic disadvantage?

Some forms of SEN have a clear biological basis, for example sensory impairments, physical needs, or Profound and Multiple Learning Difficulties (PMLD). These categories are often contrasted with categories like MLD and BESD which are defined in terms of the student's actions within a context, mainly the school and classroom. These needs are socially constructed in the sense that students' behaviour is interpreted in terms of expected patterns or norms. 'Judgemental' categories like MLD and SEMH are not the only SEN evidencing disproportionality as we saw above, but they are those where the disproportionate identification of Black students is greatest (Skiba et al, 2008, p269). Consequentially it is no surprise that the main explanation forwarded for the over-representation of Black students with SEN is the long history of ethnic stratification within education. Disproportionality is assumed to reflect inappropriate interpretation of ethnic and cultural differences including teacher racism, low expectations and a failure of schools to provide quality instruction or effective classroom management (e.g. Coutinho & Oswald, 2000; Skiba et al, 2008; Artiles et al, 2010; Waitoller et al, 2010).

However an alternative hypothesis is that disproportionality reflects the fact that ethnic minority students are more at risk of SEN because of the greater socio-economic disadvantage (SED) they experience relative to the White majority. For example, in England in 2016, 14% of White British pupils are eligible for a Free School Meal (FSM) but this doubles to 25% of Black African, 28% of Black Caribbean and 29% of Mixed White and Black Caribbean pupils (Strand & Lindorff, this report). Most research in the US concludes that poverty is only weakly related to disproportionality (e.g. Skiba et al, 2005). However, in nearly all large scale US studies the data are not the ethnicity, SEN and poverty of individual students, but rates of occurrence of these variables at the district level (Hosp & Reschly, 2004; Skiba et al, 2005). The difficulty with using aggregated data to explore phenomena at an individual level is the 'ecological fallacy', i.e. that relationships that hold for groups necessarily hold for individuals (Robinson, 1950). Only student level data can adequately address the question of the role of socio-economic disadvantage in accounting for ethnic disproportionality. Strand & Lindsay (2009) report that controlling for socio-economic disadvantage reduced the extent of disproportionality for some ethnic groups for some types of SEN, although the over-representation of Black Caribbean and Mixed White and Black Caribbean pupils with SEMH remained.

It may also be that Socio-economic Status (SES) plays a role in the under-identification of Asian groups for ASD. The average neighbourhood deprivation score

of every Asian group was higher than the White British average, and levels of eligibility for FSM were particularly high for Pakistani (19%) and Bangladeshi (23%) pupils (Strand & Lindorff, this report). Some, predominantly US, research has suggested that pupils from disadvantaged backgrounds are less likely to receive an ASD diagnosis, while pupils from high socio-economic families have the awareness and resources to seek out and receive an ASD diagnosis (e.g. Bhasin & Schendel, 2007; Croen et al, 2002; Durkin et al, 2010; Thomas et al, 2012). The results are mixed, however, and other studies have suggested no relationship between SES and ASD or raised identification for low SES groups (e.g. Larson et al, 2005, Delobel-Ayoub et al, 2015, Rai et al, 2012). Studies that have attempted to control directly for SES also have mixed results. For example, Mandell et al (2009) have reported that the under-representation of Black and Hispanic groups remained after control for gender, IQ, birth weight and maternal education, and Strand & Lindsay (2012) report that controls for pupils entitlement to Free School Meals (FSM) and neighbourhood deprivation increased the relative under-representation of Asian groups. The current study will let us throw light on this question using national population data drawn from school settings.

Recent US longitudinal studies

Very recently a few studies have begun to emerge from the US based upon longitudinal surveys. Hibel et. al. (2010) analysed the US Early Childhood Longitudinal Study-Kindergarten (ECLS-K) tracking a sample of 11,000 students from age 3 to age 8/9. They report that Black and Latino students were actually under-represented for SEN after control for educational achievement and teacher's ratings of pupils' behaviour at Kindergarten entry. Morgan et. al. (2015) followed the same sample to age 11/12 and report the same result. In addition, Morgan et. al. (2017) review results across what they deem eight 'gold standard' studies and for five SEN types, and conclude Black students are under-represented relative to White students after adjusting for prior attainment and behaviour. Not all research concurs. Shifrer et. al. (2011) analysed data on 10,847 students aged 15/16 from the US Educational Longitudinal Study (ELS) and report that Black students were over-represented among those identified with Learning Disabilities (LD), although this disappeared after control for SES. Sullivan & Ball (2013) analysed data for 18,000 students attending 39 schools from one mid-western state and report that controls for student SES reduced, but did not eliminate, the SEN over-representation of Black students. The results are therefore mixed.

It is notable that Hibel et. al. (2010) and Morgan et. al. (2015) are unusual in that they report no significant over-representation of Black students for ID, ED and LD even *before* any adjustment for behaviour or attainment covariates, out of line with all other studies. The ECLS-K is representative and at 20,000 students is a large sample, but given the US national incidence of ID at 0.70% and ED at 0.60%, and African Americans at 15% of the school population, this represents just 21 and 18 students respectively in their sample, before considering any splits by gender, SES

etc. While the incidence of these SEN are somewhat higher in England (e.g. 2.5% with MLD and 2.4% with BESD/SEMH, Strand & Lindsay, 2012) the relatively small numbers in longitudinal surveys remain a substantial obstacle to accurate determination of disproportionality in UK longitudinal surveys. Further, in longitudinal studies the SEN data is also often derived from self-report by parents and retrospectively, rather than drawing on contemporaneous administrative records. Studies based on population data, such as the NPD, are therefore urgently needed.

School composition effects

Very few studies have modelled school effects using individual pupil data within an appropriate multi-level structure. However, there are good reasons to anticipate these may exist. Teacher judgements of acceptable student achievement or behaviour are based on the norms for a reference group which naturally consists of the other children in the school. A child with a given level of achievement or problem behaviour may be seen as more problematic in a school where levels of achievement are high and behaviour is good, compared to how the same student might be viewed in a low achieving school with more problem behaviours. This may be exacerbated in low performing and high ethnic minority schools where the special education resources available are often lower in the first place, as well as more severely strained by a large number of referrals. Hibel et al (2010) do indeed report such effects with pupils less likely to be identified with SEN in schools where academic behaviour was rated low and in schools with a high proportion of ethnic minority students. However, in the Sullivan & Ball (2013) study described previously there were no significant school effects. The mixed and inconclusive results may again reflect data and methodological limitations. For example, there are problems in obtaining reliable school composition measures in longitudinal surveys when the number of observations per school is low (e.g. on average there were just 11 students per school in Hibel et al, 2010) and in independently estimating school and student effects in a model where the school level variables are direct aggregates from the student level data.

Aims and Research Questions

- Considering the most recent (2016) national data, what is the current picture of ethnic disproportionality in England? For which ethnic groups and which types of SEN identification does disproportionality exist?
- There has been substantial change in the ethnic composition of the population of England over the last decade or so, as well as changes in the rates of identification of SEN. Have patterns of ethnic disproportionality changed between 2005 and 2016?
- To what extent can ethnic disproportionality in different types of SEN identification be accounted for by age, sex, and socio-economic disadvantage?

- How does disproportionality develop dynamically as children progress through school over time? What can we learn by tracking a primary cohort from Reception to Y6, and a secondary cohort from Y6 to Y11? Can academic achievement or development on-entry to school account for disproportionality?
- What is the variability across schools in disproportionality? Do school variables (such as the socio-economic and ethnic composition of the school) have any additional association with disproportionality when these are modelled alongside pupil level variables?
- What is the variability across Local Authorities (LA) in disproportionality? What data on disproportionality might be usefully reported to LAs?

The National Pupil Database (NPD)

This research uses data from the England School Census (SC), part of the England National Pupil Database (NPD). This dataset is ideal to use in this context since it collects pupil level data and is a complete census of all children in state funded schools in England[3], ensuring the results are general and not specific to a particular sample of the population. The NPD has substantial strengths for this research:

- Allows a focus on specific types of SEN, rather than a single heterogeneous SEN category
- Avoids crass (simple Black/White) ethnic categorisation
- Collects pupil level data (avoids the ecological fallacy, Robson, 1950)
- National populations>samples (avoids low incidence issues)
- Based on administrative records, not parent report / retrospective recall
- Controls for confounding variables (SES, prior attainment/development)
- Longitudinal data, allowing us to model dynamic development of SEN over time
- Recent, not historical, data

The particular data items and variables employed and the analyses completed are reported in the relevant results sections, as outlined below.

Structure of the report

There are four substantial sections in the reporting of the results.

Part 1: Pupil level results - All students aged 5-16

This part analyses the data for all 6 million students age 5-16 in state funded schools (primary, secondary and special) in England in January 2016. We complete a multinomial logistic regression to contrast the odds of being identified with each type of SEN against the odds of having no identified SEN for each ethnic minority group compared to White British students. Control variables at the individual student level

3. The data do not include children in private schools. State schools account for 93% of all students in England.

are then introduced including: gender, year group, month of birth (to give age within year group), eligibility for Free School Meals (FSM) and neighbourhood deprivation. We report the association of each of these variables with SEN, and how the ethnic groups ORs change after accounting for the controls. This section also considers trends in disproportionality over time by also completing the above analyses with the 2005, 2007, 2009, 2011, 2013, 2014 and 2015 national datasets.

Part 2: Multi level results - School and LA level effects

This part proceeds to multilevel logistic models to test whether school and LA composition variables have any additional impact over and above student level variables. These models allow partitioning of the variance across the student (level 1), school (level 2) and LA (level 3) levels to explore the relative influence of school and LA context. An important difference from Part 1 is that we run these analyses separately for primary and secondary schools, and for maintained mainstream schools only. This is School variables include school type (LA maintained, sponsored academy, converter academy, special schools/PRU); academic selectivity (grammar, comprehensive, secondary modern); size (number FTE students on roll), economic disadvantage (%FSM, mean IDACI score) and ethnic composition (% ethnic minority students).

Part 3: Longitudinal analyses of NPD for cohorts over time

We also explore how SEN identification develops over the course of a pupils' school career. For this purpose we track the cohort who enter Reception class at age 5 in 2008/09 and follow them through to age 11 in January 2015, recording their history of SEN identification at each January census. We do the same for the cohort entering Y7 in 2010/11 following them through the Y11 in January 2015. We use survival analysis (sometimes also called event history analysis or logit hazard modelling) to identify the cumulative likelihood of a pupils' first MLD, SEMH and ASD identification over time. This more accurately reflects the likelihood of SEN identification which is not a single-point in time event but instead occurs over time as children age. We are also able to include variables such as achievement and development on-entry to school attendance during the course of their first year of school (or their last year of primary school for the secondary cohort) as additional controls to test some of the conclusions from the US longitudinal studies. We are aware of no other study that has undertaken such an analysis of SEN with national population data.

Part 4: Cross-validation with Longitudinal Surveys

Despite the limitations in terms of sample size and data sources, longitudinal surveys do collect rich data on socio-economic circumstances compared to that available in administrative datasets. For example, the Second Longitudinal Study of Young People in England (LSYPE2) collects data on 13,000 pupils including the socio-economic classification of the home, parental education qualifications and family income. LSYPE2 covers the same cohort as our NPD secondary longitudinal

analysis, so we match in the SEN data from the NPD. This allows us to triangulate the results of the NPD analyses using the wider socio-economic variables available in LSYPE2.

Part 1: Pupil-level results (all pupils age 5-16)

Summary

This chapter investigates ethnic disproportionality in SEN identification amongst pupils of compulsory school-age (ages 5-16, Y1-11) in England in 2016, using data from the National Pupil Database. We also analyse trends over time, comparing back to similar data and analyses from previous years going back to 2005.

The first part of the chapter provides description of the data used and the analytical approach taken.

The second part of the chapter presents results regarding ethnic disproportionality. We report **unadjusted** estimates of over- and under-representation based only on ethnic group and then **adjusted** estimates that account for the impact of other aspects of pupil background such as socioeconomic deprivation, gender, and age (both in terms of year group and birth season within year group). We additionally present results accounting for interactions between ethnic group and FSM (as a proxy for socioeconomic deprivation) and between ethnic group and gender, to better understand whether any observed disproportionality differs within minority ethnic groups on these bases.

We report results for all types of SEN for the main analysis of unadjusted and adjusted disproportionality, but in our key findings and for more fine-grained results we focus on three types: Moderate Learning Difficulty (MLD), as it is the most prevalent SEN type; Social, Emotional and Mental Health (SEMH), as it is also one of the most prevalent types of need and is additionally of interest because of its high-judgment nature; and Autistic Spectrum Disorder (ASD), which is of particular interest because it has increased rapidly in prevalence over recent years. The focus on these three types of SEN carries through to the subsequent chapters of the report.

Key findings:

- Although the total prevalence of identified SEN has decreased over time, trends vary by type of SEN. ASD, for example, has more than doubled in prevalence. However ethnic disproportionality has remained a stable feature of the data, regardless of prevalence rates.
- Pakistani and Black Caribbean pupils are over-represented for MLD, but this can be accounted for by socio-economic factors. Indian and Bangladeshi pupils are under-represented.
- Black Caribbean and Mixed White & Black Caribbean pupils are over-represented for SEMH, while Asian pupils are broadly under-represented. These findings hold true even after accounting for other pupil background characteristics. Among the most disadvantaged pupils (entitled to FSM) the risks are equally raised for White British and Black Caribbean, the disproportionality is primarily among those not entitled to FSM.

Key findings (Cont.):

- Asian (Indian, Pakistani & Bangladeshi) pupils are under-represented for ASD, and this remains true in unadjusted and adjusted analyses.

What we did

Data source

The data used in this chapter come from the Department for Education (DFE) January 2016 School Census, which contains data on all pupils in state schools in England. The pupil-level dataset contains a variety of pupil background characteristics including entitlement to a Free School Meal (FSM), pupil neighbourhood income deprivation affecting children index (IDACI), whether a pupil has English as an Additional Language (EAL), pupil ethnic group, and – particularly of interest for the purposes of this project – Special Educational Needs (SEN) identification details including specific SEN type and level. The School Census does not contain information on pupils in independent schools, but these constitute only a small percentage of all pupils in England (under 7%).

Filtering of data

The analyses in this chapter are restricted to data on pupils in Y1 to Y11 (ages 5-16) in the January 2016 School Census.[4] Records for children in Reception - who would have been 4 years old at the start of the 2015-16 school year - have been excluded, as at the time of the January 2016 School Census these children had only have been attending school for a short period of time (4 months) and compared to other year groups relatively few are identified with SEN. Records of young people continuing into Y12 or higher were also excluded, as post-compulsory education is inherently selective and we wanted comprehensive population data.

There were missing values for only one of the variables used for the analysis in this chapter (neighbourhood IDACI) and as the number of records missing IDACI scores was very small (0.2% of the total Y-11 population), these were also excluded. After applying these filters, there were 6,490,615 pupil records included in the main analysis.

Measures

Level of SEN

[4] *A very small number of records (less than 0.001%) had duplicate Pupil Matching Reference numbers; these were also excluded.*

Around 3% of pupils have a formal statement of SEN or Education and Health Care (EHC) Plan. This means a legal document is in place that sets out the child's needs and the extra help they should receive. However, the majority of pupils with SEN are identified at School Support (12.8% of the school population and 81% of all those with an identified SEN). These pupils also receive provision that is additional to or different from that made generally for others of the same age, and which goes beyond the differentiated approaches and learning arrangements normally provided as part of high quality, personalised teaching. We combined these groups in the majority of our analyses.

Type of SEN

Schools are asked to record the primary need of SEN pupils from one of twelve specific types of need. In the initial stages of our analysis we look at disproportionality for all twelve types of need which are:

- No identified SEN
- Specific Learning Difficulty (SpLD)
- Moderate Learning Difficulty (MLD)
- Severe Learning Difficulty (SLD)
- Profound and Multiple Learning Difficulty (PMLD)
- Social, Emotional and Mental Health (SEMH)[5]
- Speech, Language and Communication Needs (SLCN)
- Autistic Spectrum Disorder (ASD)
- Hearing Impairment (HI)
- Visual Impairment (VI)
- Multi-sensory Impairment (MSI)
- Physical Disability (PD)
- Other (unspecified) type of need
- SEN support but no specialist assessment (NSA)

We later focus in depth on three types:

- *Moderate Learning Difficulties (MLD):* this is the most frequently identified SEN, accounting for 4.0% of pupils aged 5-16. These pupils may "learn at a slower pace than their peers, even with appropriate differentiation" (DFE, 2015, p97).
- *Social, Emotional and Mental Health (SEMH):* the second most frequently identified type of SEN, accounting for 2.8% of pupils aged 5-16. These difficulties "may include becoming withdrawn or isolated, as well as displaying challenging, disruptive or disturbing behaviour. These behaviours may reflect

[5] *This largely corresponds to the category Behavioural, Emotional and Social Difficulties (BESD) from previous years; this shift was one of several changes in SEN-related policy in 2014 and apparent in School Census data from 2015. Any tables including results relevant to SEN type from 2014 and earlier will include BESD rather than SEMH.*

underlying health difficulties… [or] disorders such as attention deficit disorder, attention deficit hyperactive disorder or attachment disorder (DFE, 2015, p98).

- *Autistic Spectrum Disorders (ASD):* this is the most commonly identified need among pupils who have a statement of SEN, and is also a rapidly growing need, increasing from 0.5% of the population in 2005 to 1.3% in 2016. "Pupils with ASD, including Asperger's syndrome and Autism, are likely to have particular difficulties with social interaction. They may also experience difficulties with language, communication and imagination, which can impact on how they relate to others" (DFE, 2015, p97).

As a result of changes to the 2014 SEN code of practice, prevalence rates of SEN types from 2015 include pupils with different levels of need. That is, before 2015, levels of SEN were defined according to the categories of School Action, School Action Plus, and Statement, and the primary type of need was recorded in the School Census only for those pupils having School Action Plus or a Statement. A transition to a different set of categories began in 2015; these new categories included SEN support (which did away with the distinction between School Action and School Action Plus) and Education, Health and Care (EHC) Plan (which replaced the Statement category and was intended to represent a more comprehensive, multi-agency plan for support). The transition to the new set of levels was mostly complete by the January 2016 School Census, except for some remaining records of Statements not yet transitioned to EHC Plans. Thus, for the 2016 data, we define level of need using three categories: No SEN, SEN support, and Statement/EHC Plan).

Ethnicity

The DFE defines ethnicity as a personal awareness of a common cultural identity. Ethnicity relates to how a person feels and not how they are perceived by others. It is a subjective decision as to which category a person places themselves in and does not infer any other characteristics such as religion, country of origin etc. Ethnicity monitoring advice is available from the department's website. The DFE recommends that those with parental responsibility determine the ethnic background for children at primary school. The Information Commissioner advises that pupils aged 11-15 are sufficiently mature to determine their own ethnic background, and the DFE recommends that decisions are best made with the support and knowledge of those with parental responsibility in a family context. Pupils aged 16 and over can make their own decisions (DFE, School Census Guide 2015-16, p41).

The ethnicity codeset reflects categories used in the 2001 national population census, with additional categories for Travellers of Irish heritage and pupils of Gypsy/Roma heritage. Local Authorities may make use a DFE approved extended ethnic code set to reflect local needs and priorities. All codes aggregate to 18 categories that are standard throughout education administrative databases in England (see Table 1.1). Where ethnicity has not yet been collected this is recorded

as ‘NOBT’ (information not yet obtained). If a pupil or parent has refused to provide ethnicity, ‘REFU’ (refused) is recorded and returned.

We use White British as the (majority) reference group and compare each of the ethnic minority groups to White British. Table 1.1 lists the ethnic main codes in the order they are typically listed, and then by the percentage of the total Y1-Y11 population constituted by each ethnic group.

Table 1.1: Frequencies and percentages of ethnic groups (2016, Y1-11)

Ethnic group (in order of reporting)	**Ethnic group** (by size of population)	*N*	*Percent*
White British (reference group)	White British	4502558	69.4
White Irish	White other groups	367017	5.7
Irish Traveller	Pakistani	275269	4.2
Gypsy/Roma	Black African	235333	3.6
White other groups	Indian	179111	2.8
Mixed White and African	Any other mixed background	122534	1.9
Mixed White and Caribbean	Any other ethnic group	111023	1.7
Mixed White and Asian	Any other Asian	110319	1.7
Any other mixed background	Bangladeshi	108478	1.7
Indian	Mixed White & Caribbean	96033	1.5
Pakistani	Black Caribbean	79909	1.2
Bangladeshi	Mixed White & Asian	78940	1.2
Any other Asian ethnic group	Unknown	60484	0.9
Black African	Black other groups	46924	0.7
Black Caribbean	Mixed White & African	45042	0.7
Black other	Chinese	25993	0.4
Chinese	Traveller Gypsy/Roma	21735	0.3
Any other ethnic group	White Irish	19044	0.3
Unknown (not obtained or Refused)	Traveller Irish	4869	0.1
	Total	6490615	100.0

Other variables

Other pupil background characteristics including:

- *Year group* (with Year 1 as the reference group)
- *Birth season* (with ‘autumn born’ as the reference group). Autumn born (Sep-Dec) will be the oldest pupils in the year group, spring born (Jan-Apr) will be the middle third, and summer born (May-Aug) the youngest pupils in the year group.
- *Sex* (with ‘Girl’ as the reference group) so that coefficients reflect the effect of being a boy.
- *Entitlement to a FSM* (with ‘not entitled’ as the reference group). A Free School Meal (FSM) is a statutory benefit available to school-aged children from families who receive other qualifying benefits and who have been through the relevant

registration process[6]. The qualifying benefits are: Income Support; Income-based Jobseekers Allowance; Income-related Employment and Support Allowance; Support under Part VI of the Immigration and Asylum Act 1999; the guaranteed element of State Pension Credit; Child Tax Credit (provided they are not also entitled to Working Tax Credit and have an annual gross income of no more than £16,190); Working Tax Credit run-on - paid for 4 weeks after qualification for Working Tax Credit ends; Universal Credit. In January 2016, 15.2% of pupils aged 5-16 were recorded as eligible for FSM.

- *Income Deprivation Affecting Children Index (IDACI):* IDACI is produced by the Department for Communities and Local Government (DCLG). The index is based on 32,482 Super Output Areas (SOAs) in England, which are geographical regions of around 1,500 residents, designed to include those of similar social backgrounds. The IDACI score is the percentage of under-16s in the SOA living in income deprived households (primarily defined by being in receipt of certain benefits). In 2016 the average was 23.1%. This variable is highly skewed and so for the purpose of the current analysis the measure was normal score transformed to give a variable with a mean of 0 and SD of 1. A score above 0 indicate greater than average deprivation, and score below 0 indicate less than average deprivation, relative to the average for the population. Both 2001 and 2007 IDACI measures were available. The means of the two were nearly identical (24.7% and 25.7%) and they correlated r=0.97, so the more recent 2007 values were used. Further information about IDACI can be found at: https://www.gov.uk/government/publications/english-indices-of-deprivation-2010.

Descriptive statistics by ethnic group

Our interest in including controls for pupil background characteristics is in accounting for their impact on estimates of ethnic disproportionality. Therefore, it is important to first describe the distribution/frequency of these variables by ethnic group to identify any patterns and differences across groups. Table 1.2 presents descriptive statistics for entitlement to FSM, gender, IDACI score and birth season by ethnic group.

Gender and birth season are relatively consistently distributed across ethnic groups. However, there is large variation between ethnic groups in the socio-economic disadvantage measures. For example, 14% of White British pupils are eligible for a Free School Meal (FSM) but this rises to 19% of Pakistani, 23% of Bangladeshi, 25% of Black African, 28% of Black Caribbean and 29% of Mixed White and Black Caribbean pupils. The average neighbourhood deprivation (IDACI) score of White British pupils was 19%, but is higher than this for every Asian group, and indeed is more than half as large again for Mixed white and Black Caribbean (30%), Pakistani (35%), Black Caribbean and Black Other (39%) and Black African (41%) pupils.

[6]. Pupils who are only in receipt of a free school lunch due to the infant pupil universal entitlement are not recorded as FSM eligible and not eligible to receive pupil premium.

Table 1.2: Pupil descriptive statistics by ethnic group (2016, Y1-11, ages 5-16)

	Total	**FSM entitlement**		**Gender**		**IDACI score**		**Birth Season**					
		Entitled		*Boy*				*Autumn*		*Spring*		*Summer*	
Ethnic group	N	N	%	N	%	M	SD	N	%	N	%	N	%
White British	4502558	632217	14.0	2306046	51.2	19.1	15.6	1508244	33.5	1453833	32.3	1540481	34.2
White Irish	19044	2896	15.2	9758	51.2	21.4	17.3	6445	33.8	6097	32.0	6502	34.1
Traveller Irish	4869	3057	62.8	2529	51.9	30.5	19.4	1611	33.1	1645	33.8	1613	33.1
Traveller Gypsy/Roma	21735	6896	31.7	11059	50.9	32.2	16.9	7074	32.5	7104	32.7	7557	34.8
White Other	367017	32224	8.8	188083	51.2	28.7	18.2	119042	32.4	118381	32.3	129594	35.3
Mixed White & African	45042	10410	23.1	22574	50.1	29.4	18.5	14797	32.9	14462	32.1	15783	35.0
Mixed White & Caribbean	96033	27723	28.9	48350	50.3	29.8	18.3	31977	33.3	30981	32.3	33075	34.4
Mixed White & Asian	78940	12534	15.9	40379	51.2	21.7	16.8	26350	33.4	25592	32.4	26998	34.2
Mixed Other	122534	23591	19.3	62699	51.2	27.4	18.8	40211	32.8	39798	32.5	42525	34.7
Indian	179111	11109	6.2	92361	51.6	25.4	15.8	60427	33.7	58265	32.5	60419	33.7
Pakistani	275269	52313	19.0	141411	51.4	34.9	15.3	94061	34.2	89974	32.7	91234	33.1
Bangladeshi	108478	25396	23.4	54707	50.4	43.8	18.1	37163	34.3	35566	32.8	35749	33.0
Asian Other	110319	13528	12.3	57130	51.8	28.6	16.8	36978	33.5	35909	32.6	37432	33.9
Black African	235333	57666	24.5	118388	50.3	40.9	18.0	78638	33.4	76287	32.4	80408	34.2
Black Caribbean	79909	22151	27.7	40564	50.8	39.0	17.2	27505	34.4	25625	32.1	26779	33.5
Black Other	46924	12049	25.7	23989	51.1	38.8	18.3	15724	33.5	15223	32.4	15977	34.0
Chinese	25993	1767	6.8	13061	50.2	24.4	18.6	8691	33.4	8569	33.0	8733	33.6
Any Other	111023	25869	23.3	57474	51.8	36.0	19.5	36195	32.6	36430	32.8	38398	34.6
Unknown	60484	10515	17.4	31578	52.2	24.0	17.7	20122	33.3	19673	32.5	20689	34.2
Total (Overall)	6490615	983911	15.2	3322140	51.2	23.1	17.6	2171255	33.5	2099414	32.3	2219946	34.2

Approach to analysis

Our analytical strategy within this section of the report consists of two main stages. First, we conducted some descriptive analyses to facilitate understandings in context and to inform and justify decisions relevant to the main analysis. Second, for the main analysis within this section, we used multinomial logistic regression to investigate ethnic disproportionality in SEN identification, contrasting the odds of being identified with each SEN type against the odds of having no identified SEN for each ethnic minority group compared to White British students (i.e. taking the majority ethnic group as the reference group). Multinomial regression was first run with ethnic group as the only predictor, and then again with additional pupil predictors to investigate the extent to which ethnic disproportionality can be accounted for by other pupil background characteristics.

Interpretation

Because the dataset is very large, statistical significance is not a good measure of substantive importance (p-values are substantially inflated), so we report Odds Ratios (ORs) as the relevant effect size, with the added practical advantage that these have a fairly intuitive interpretation. The OR represents the odds of identification for a particular ethnic minority group relative to the odds of identification for the White British majority group. Thus, an OR of 2.0 indicates twice the odds of being identified compared to White British pupils, an OR of 1.0 means the same odds of being identified as White British pupils, and an OR of 0.50 means half the odds of being identified compared to White British pupils.

We considered the size of ORs in relation to the following cut-offs:

- **OR <= 0.67** "substantially under-represented"
- OR <= 0.75 "under-represented"
- OR >= 1.33 "over-represented"
- **OR >= 1.50** "substantially over-represented".

The more stringent cut-off values of 0.67 and 1.50 correspond to odds of 2:3 and 3:2 of being identified with a given type of SEN, respectively, as compared to the reference group (White British). The less stringent threshold values of 0.75 and 1.33 correspond to odds of 3:4 and 4:3 relative to the odds of identification for White British pupils.

What we found

Prevalence of SEN: Types and levels of need

Prevalence of primary types of need

Table 1.3 shows the number and percentage of Y1-11 pupils identified with each type of primary SEN (overall and as a percentage of pupils with a primary need) as of January 2016, as well as the number and percent identified with each type

amongst only those having SEN support and amongst only those having Statements/EHC plans.

MLD is the most prevalent type of primary SEN overall (4.0%), and SEMH is the second most prevalent (2.8%); these constitute 25.5% and 17.7% of all pupils with a primary SEN, respectively. ASD is a fairly low-incidence type of primary need overall (1.3% of all Y1-11 pupils), but is the most common type of primary need amongst pupils with a Statement or EHC Plan, representing 26% of those with a statement/EHCP. For these reasons, although our multinomial regression analysis included all types of primary SEN, we emphasize these three SEN types (MLD, SEMH, ASD) in reporting detailed results in the subsequent multilevel analysis.

Although for the purposes of this initial descriptive information we report all types of SEN as recorded in the annual Census return, not all of these types are equally informative based on the information inherent to their descriptions. Specifically, the No Specialist Assessment (NSA) and Other categories do not provide information on the nature of a pupils' SEN; the distinction between these is, ostensibly, that NSA is essentially a placeholder category (noting that a pupil has SEN but had not received/completed specialist assessment to further determine the type of need by the Census date), while Other implies a need not falling under any of the other types of SEN listed.

Table 1.3: Prevalence by primary type of SEN, overall and by level of SEN, 2016 (Y1-11)

	SEN support		*Statement/EHC*		*Overall*		
Primary type of SEN	*N*	*% of those with SEN support*	*N*	*% of those with Statement/ EHC plan*	*N*	*% of all pupils*	*% of those with a primary need*
None	0	0.0	0	0.0	5470700	84.3	--
MLD	234129	28.3	26166	13.6	260295	4.0	25.5
SEMH	153252	18.5	27211	14.2	180463	2.8	17.7
SLCN	148817	18.0	27992	14.6	176809	2.7	17.3
SpLD	134167	16.2	8230	4.3	142397	2.2	14.0
ASD	37136	4.5	48882	25.5	86018	1.3	8.4
Other	45645	5.5	4626	2.4	50271	0.8	4.9
NSA	33558	4.1	235	0.1	33793	0.5	3.3
PD	16905	2.0	10845	5.7	27750	0.4	2.7
SLD	2525	0.3	21864	11.4	24389	0.4	2.4
HI	12835	1.5	5097	2.7	17932	0.3	1.8
VI	7174	0.9	2809	1.5	9983	0.2	1.0
PMLD	534	0.1	7337	3.8	7871	0.1	0.8
MSI	1430	0.2	514	0.3	1944	0.0	0.2
Total	828107	100.0	191808	100.0	6490615	100.0	--

Prevalence within year groups

Different types of primary SEN may be more likely to be identified at different ages. Figure 1-1 displays the prevalence of MLD, SEMH, ASD and a combined category including all sensory and physical types of primary SEN (Hearing, Visual, Physical, and Multi-Sensory) across Years 1-11 in the 2016 School Census.

It is apparent that the different focal SEN types have distinct profiles in terms of their prevalence across year groups. MLD identification increases in prevalence in the primary year groups (from 2.1% of pupils in Y1 to 5.4% in Y6) and then decreases in secondary (from 4.4% in Y7 to 3.5% in Y11). SEMH identification increases across primary year groups (from 1.7% of pupils in Y1 to 3.1% in Y7), drops very slightly in the transition to secondary (2.9% in Y7), and increases again across secondary year groups (to 3.7% in Y11). ASD increases very slightly across primary year groups (from 1.0% in Y1 to 1.4% in Y6) and remains constant across secondary Year groups (1.5%), suggesting that ASD may be more consistently identified earlier in pupils' school careers.

As a basis for comparison, Figure 1-1 also shows the prevalence of a combined Sensory and Physical category of primary SEN identification (including Hearing Impairment, Visual Impairment, Multi-Sensory Impairment, and Physical Disability), which has roughly equal prevalence across all year groups (0.9% of pupils); this may well be because Sensory and Physical needs are more visible and identification is a more medical, lower-judgment (more objective) process than for many of the other primary types of need.

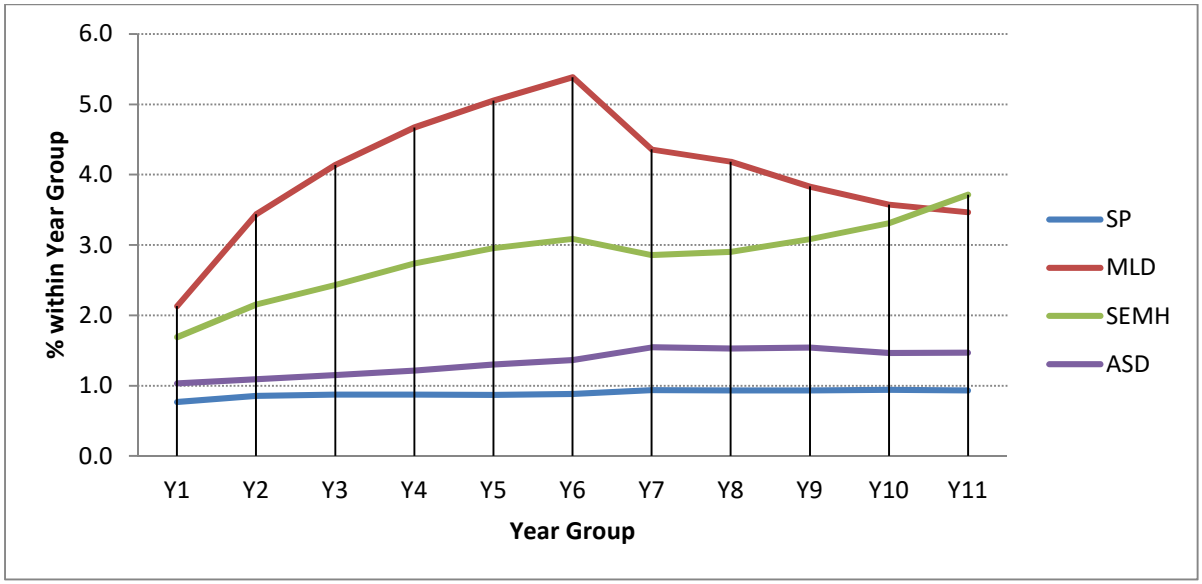


Figure 1-1: Prevalence of primary type of SEN (MLD, SEMH, ASD, and Sensory/Physical) by year group (2016)

It is worth noting that these patterns come from cross-sectional data, so that they provide information on prevalence across year groups but not over time as such. Part 3 of this report will address longitudinal patterns using survival/hazard analysis.

Prevalence of secondary types of need

Table 1.4 shows the prevalence of each secondary type of SEN, both in the population of all pupils and amongst those pupils identified with a primary type of need. Of those pupils identified with SEN, 78.1% are only identified as having a primary type of need. Of those with both a primary and a secondary type of need, SLCN is the most common secondary type of need (4.9% of those with a primary need), with SEMH (4.5%) and MLD (4.3%) close behind.

For those identified with MLD as their primary type of SEN, 18.3% are identified as having a secondary need, with the most frequently-identified secondary types being SLCN (6.5%) and SEMH (6.1%). Of those with SEMH as a primary type of need, 19.8% are identified as having a secondary type of need, with the most common secondary types being MLD (6.8%) and SLCN (4.3%). A greater percentage of those pupils identified with ASD have a secondary type of need identified (34.6%), with the most common secondary types including SLCN (11.2%), SEMH (8.0%), and MLD (4.9%).

Appendix A provides the full cross-tabulation of primary and secondary types of SEN from the 2016 School Census.

Table 1.4: Prevalence by secondary type of SEN

Secondary type of SEN	*N*	*% of all pupils*	*% of pupils with a primary need*
Only primary need identified	796099	--	78.1
SLCN	50478	0.8	4.9
SEMH	46229	0.7	4.5
MLD	43781	0.7	4.3
SpLD	20409	0.3	2.0
Other	18171	0.3	1.8
ASD	15391	0.2	1.5
PD	9936	0.2	1.0
SLD	5667	0.1	0.6
VI	4321	0.1	0.4
HI	4246	0.1	0.4
NSA	3195	0.0	0.3
MSI	1278	0.0	0.1
PMLD	714	0.0	0.1
Total	223816	15.7	100.0

For consistency, the analyses reported in later sections of this report focus on *primary* type of SEN, as this is recorded in the School Census for all pupils with SEN identified at any level and secondary type of need has extremely low rates overall.

Prevalence of SEN by ethnic group

Prevalence in 2016

Table 1.5 provides prevalence rates for all primary types of need within each ethnic group in 2016. The data in the table are the percentage of pupils within each ethnic group with each primary type of need. It is apparent from these base rates that the proportions identified with specific SEN types varies by ethnic group, for example:

- 6.3% of Black Caribbean pupils were identified with SEMH needs, compared to 3.0% of White British pupils;
- 5.6% of Pakistani pupils were identified with MLD, compared to 4.1% of White British pupils;
- 0.7% of Indian pupils were identified with ASD, compared to 1.4% of White British pupils.

We explore these differences in terms of odds of identification for each primary type of SEN using multinomial regression analysis in a later section.

Table 1.5: Prevalence (%) of each primary SEN type by ethnic group, 2016

	Cognition & Learning				Social, Emotional & Mental Health	Communication & Interaction		Sensory & Physical				Unspecified/ Other	
Ethnic Group	**SpLD**	**MLD**	**SLD**	**PMLD**	**SEMH**	**SLCN**	**ASD**	**HI**	**VI**	**MSI**	**PD**	**Other**	**NSA**
White British	2.6	4.1	0.4	0.1	3.0	2.5	1.4	0.3	0.2	0.0	0.5	0.8	0.5
White Irish	3.0	3.4	0.3	0.1	2.8	2.4	1.5	0.3	0.1	0.1	0.5	0.8	0.6
Traveller Irish	4.5	13.7	0.5	0.2	6.5	5.2	0.5	0.4	0.2	0.0	0.5	1.8	2.0
Traveller Gypsy/Roma	3.4	12.9	0.6	0.1	4.1	5.1	0.5	0.7	0.2	0.0	0.4	1.1	1.4
White other groups	1.4	3.3	0.3	0.1	1.8	3.3	0.9	0.2	0.1	0.0	0.3	0.7	0.6
Mixed White & African	1.9	3.4	0.4	0.1	3.6	3.0	1.3	0.2	0.1	0.0	0.3	0.7	0.5
Mixed White & Caribbean	2.5	4.8	0.3	0.1	5.5	2.8	1.5	0.2	0.2	0.0	0.4	0.9	0.7
Mixed White & Asian	1.4	2.9	0.3	0.1	2.2	2.4	1.2	0.2	0.1	0.0	0.3	0.6	0.4
Any other mixed background	1.8	3.2	0.4	0.2	3.2	2.8	1.5	0.2	0.1	0.0	0.4	0.8	0.5
Indian	0.8	2.5	0.3	0.1	0.8	2.3	0.7	0.2	0.1	0.0	0.3	0.5	0.3
Pakistani	1.0	5.6	0.6	0.3	1.5	3.4	0.8	0.6	0.3	0.0	0.5	0.8	0.7
Bangladeshi	0.9	3.6	0.4	0.2	1.4	4.1	1.1	0.4	0.2	0.0	0.3	0.7	0.7
Any Other Asian	0.9	2.6	0.4	0.2	1.0	3.0	0.9	0.3	0.1	0.0	0.3	0.6	0.4
Black African	1.2	3.5	0.5	0.2	2.5	4.1	1.6	0.2	0.1	0.0	0.3	0.7	0.6
Black Caribbean	2.4	5.2	0.4	0.1	6.3	4.1	1.7	0.2	0.1	0.0	0.4	1.0	0.7
Black other groups	1.5	4.1	0.6	0.2	3.8	4.1	1.8	0.2	0.1	0.0	0.3	0.8	0.7
Chinese	0.7	1.4	0.3	0.1	0.7	3.2	1.4	0.2	0.1	0.0	0.1	0.4	0.3
Any other ethnic group	1.3	3.8	0.4	0.2	1.9	3.8	1.0	0.3	0.1	0.0	0.3	0.8	0.8
Unclassified/Refused	2.4	3.7	0.5	0.1	3.3	2.6	1.7	0.2	0.2	0.0	0.4	0.9	0.7
All pupils	**2.2**	**4.0**	**0.4**	**0.1**	**2.8**	**2.7**	**1.3**	**0.3**	**0.2**	**0.0**	**0.4**	**0.8**	**0.5**

Time trends in prevalence 2005-2016: Any SEN

Table 1.6 shows the trends over time in the rate of SEN identification by ethnic group (as a percentage) from the 2005 to 2016 school Censuses. SEN identification is here defined as having any type of SEN at any level including School Action, School Action Plus, Statemented or EHC Plan.

These percentages are directly comparable over time, and show an increase from 20.4% of the population in 2005 to a high of 23.8% in 2009, followed by a gradual decrease back to 20.1% in 2014, and large drops to 15.7% in 2016.

Table 1.6: Time trends in the prevalence of any SEN identification by ethnic group, 2005-16

	Rate of SEN identification (any type and level*)							
Ethnic Group	*2005*	*2007*	*2009*	*2011*	*2013*	*2014*	*2015*	*2016*
White British	21.1	22.2	23.2	22.1	20.7	19.5	16.9	15.8
White Irish	57.0	58.2	61.4	59.8	53.3	49.9	40.9	36.0
Traveller Irish	50.9	52.7	53.6	48.8	42.5	39.3	33.0	30.4
Traveller Gypsy/Roma	20.2	21.5	23.3	23.2	19.9	18.2	14.5	12.9
White other groups	21.4	23.6	25.0	24.3	21.6	20.2	17.0	15.5
Mixed White & African	24.8	26.9	29.1	29.3	26.1	25.2	21.4	19.9
Mixed White & Caribbean	15.9	17.5	18.7	18.7	16.8	15.8	13.2	12.3
Mixed White & Asian	20.4	22.0	23.3	22.9	20.4	19.4	16.3	15.1
Any other mixed	14.1	15.1	16.1	15.3	13.0	11.8	9.8	8.9
Indian	24.5	26.6	28.0	26.8	22.8	21.2	17.5	16.0
Pakistani	21.6	23.3	25.1	23.8	19.8	18.1	15.2	14.0
Bangladeshi	15.6	16.8	18.1	17.6	15.3	14.3	11.8	10.6
Any Other Asian	23.7	25.9	28.0	26.5	22.6	20.9	17.0	15.5
Black African	29.7	31.9	33.5	33.6	29.8	28.8	24.7	22.8
Black Caribbean	26.7	28.6	30.7	29.7	25.7	24.3	20.3	18.3
Black other groups	12.4	13.5	14.5	13.9	12.2	11.6	9.5	8.8
Chinese	20.5	22.8	24.7	24.6	21.5	19.8	16.1	14.6
Any other	23.1	24.7	25.0	24.7	22.6	21.4	18.0	16.7
Unclassified/Refused	20.1	21.7	23.5	23.5	21.3	20.4	17.5	16.2
All students	20.4	22.1	23.8	23.7	21.2	20.1	17.1	15.7

*this includes pupils with any level of SEN, including those with School Action, School Action Plus, SEN support (in 2015 and 2016) and Statement/EHC plan

These changes over time seem consistent within each major ethnic group.

Time trends in prevalence 2005-2016: By type of SEN

The new SEN Code of Practice introduced from September 2014 contained changes to the way type of SEN was recorded. Up until 2014 type of SEN was recorded only for pupils with the higher levels of need, those on School Action Plus (SAP) or with a Statement. It was not required for pupils recorded at School Action (SA).

From 2015 onwards, the distinction between SA and SAP was removed, both groups were now referred to as School Support, and the type of need was recorded for any pupil with any level of SEN identification. Thus from 2015 onwards type of SEN was recorded for a much larger number of pupils. There were particular issues with 2015 as this represented a transition year. Schools were able to retain the School Action / School Action Plus codes or to shift to the new system at the school's discretion, so the School Census data for 2015 included a mixture of new and old SEN level classifications.

The number of pupils with a recorded primary SEN type in 2014 was **606,100** (9.6% of the age 5-16 roll); in 2015 it was **913,008** (14.3%), and in 2016 was **1,022,797** (15.8%). As a result identification rates for specific types of SEN are not directly comparable over time before and after the policy change.

Table 1.7 presents the trends in rate of identification (percentages of each ethnic group) for the focal types of SEN (MLD, SEMH, ASD) from 2005 to 2016. Given the mixed economy we describe above in 2015, we calculated both upper and lower bounds. The upper bound rates were based on all pupils recorded as having a primary SEN type, i.e. including all pupils in schools that had transitioned to the new system (over two-thirds of schools). It therefore provides an estimate based on all pupils we *know* to have had a primary type of need recorded in 2015. The lower bound rates were conservative and tried to retain parity with the previous year. We matched each pupil to their SEN status in the previous year, and excluded any pupils on School Support who were School Action in 2014, so that the estimate was highly likely to be for pupils at the previous SAP level. Together, these upper and lower bound estimates help to bridge the policy shift from 2014 to 2016, allowing for better-informed interpretations of trends across the relevant years (in particular, the extent to which we might attribute increases in incidence to changes in recording versus actual changes in incidence.

There is an apparent increase in overall rates of identification in 2015 and 2016 for all three focal SEN types, which is in keeping with the trends over earlier years for SEMH (previously BESD) and ASD, but not for MLD (which had decreased in overall prevalence in 2013 and 2014). Some of the increases in 2015 and 2016 may be attributed to the change in policy and consequently the availability of SEN type information on a greater number of pupils than in previous year.

We will look further at these data in term of Odds Ratios to show the extent of disproportionality, and how this might have changed over time, in the concluding section of this chapter. First we describe the how we calculate 'unadjusted' and 'adjusted' ORs for the 2016 data.

Table 1.7: Time trends in the prevalence of focal primary types of need (MLD, SEMH/BESD, ASD) by ethnic group 2005-2016

Ethnic Group	MLD									BESD (SEMH)									ASD								
	2005	*2007*	*2009*	*2011*	*2013*	*2014*	*2015 (L.B.)**	*2015 (U.B.)**	*2016*	*2005*	*2007*	*2009*	*2011*	*2013*	*2014*	*2015 (L.B.)**	*2015 (U.B.)**	*2016*	*2005*	*2007*	*2009*	*2011*	*2013*	*2014*	*2015 (L.B.)**	*2015 (U.B.)**	*2016*
White British	2.6	2.6	2.7	2.5	2.2	2.0	2.2	3.7	4.1	1.9	2.2	2.5	2.5	2.3	2.3	2.2	2.8	3.0	0.5	0.6	0.8	0.9	1.0	1.1	1.2	1.3	1.4
White Irish	2.6	2.6	2.6	2.1	1.9	1.7	2.0	3.1	3.4	2.1	2.4	2.9	2.5	2.2	2.2	1.9	2.4	2.8	0.6	0.7	0.8	1.0	1.2	1.2	1.4	1.5	1.5
Traveller Irish	13.5	13.2	13.9	12.8	9.4	7.5	8.5	13.5	13.6	5.4	5.9	6.5	6.2	5.8	5.7	5.1	6.2	6.5	0.1	0.2	0.2	0.3	0.4	0.5	0.6	0.6	0.5
Gypsy/Roma	12.0	11.3	11.4	8.6	6.3	5.6	6.7	10.7	12.9	4.4	5.2	5.1	4.3	3.6	3.7	3.3	4.2	4.1	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4
White other groups	2.0	1.9	2.0	1.9	1.5	1.4	1.7	3.0	3.3	1.4	1.6	1.8	1.8	1.4	1.3	1.2	1.7	1.8	0.5	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.9
Mixed White & BA	2.0	2.2	2.4	2.2	1.8	1.6	1.8	3.1	3.4	2.8	3.0	3.1	3.1	2.7	2.8	2.6	3.4	3.6	0.6	0.6	0.7	0.9	1.0	1.0	1.2	1.3	1.3
Mixed White & BC	2.5	2.8	3.0	3.0	2.6	2.4	2.6	4.3	4.8	3.9	4.4	4.9	4.8	4.3	4.5	4.0	5.1	5.5	0.5	0.6	0.8	1.0	1.1	1.1	1.3	1.4	1.5
Mixed White &Asian	1.7	1.8	1.8	1.7	1.5	1.4	1.5	2.6	2.9	1.3	1.6	1.7	1.9	1.7	1.7	1.5	2.0	2.2	0.5	0.7	0.8	0.9	0.9	1.0	1.1	1.1	1.2
Any other mixed	2.0	2.1	2.2	2.0	1.8	1.6	1.8	2.9	3.2	2.4	2.7	2.9	2.8	2.4	2.6	2.4	3.0	3.2	0.6	0.8	0.9	1.1	1.2	1.2	1.3	1.4	1.5
Indian	1.7	1.7	1.7	1.5	1.2	1.1	1.3	2.1	2.5	0.4	0.5	0.6	0.6	0.6	0.5	0.5	0.7	0.8	0.2	0.3	0.4	0.4	0.5	0.5	0.6	0.6	0.7
Pakistani	3.7	3.7	3.8	3.4	2.9	2.5	2.8	4.8	5.6	0.9	1.0	1.2	1.1	0.9	1.0	0.9	1.3	1.5	0.2	0.3	0.4	0.4	0.5	0.6	0.7	0.7	0.8
Bangladeshi	2.5	2.6	2.6	2.2	1.8	1.7	1.9	3.2	3.6	0.7	0.8	1.0	1.1	0.8	0.8	0.9	1.3	1.4	0.2	0.3	0.4	0.5	0.7	0.8	1.0	1.0	1.1
Any Other Asian	1.4	1.5	1.5	1.4	1.2	1.1	1.4	2.4	2.6	0.6	0.7	0.9	0.9	0.8	0.8	0.7	1.0	1.0	0.3	0.4	0.4	0.6	0.7	0.7	0.8	0.8	0.9
Black African	2.2	2.4	2.7	2.4	1.9	1.7	1.9	3.2	3.5	2.0	2.2	2.5	2.3	1.9	1.8	1.6	2.3	2.5	0.5	0.7	0.9	1.1	1.3	1.4	1.5	1.5	1.6
Black Caribbean	3.4	3.3	3.5	3.4	2.8	2.6	2.8	4.6	5.2	4.3	4.9	5.6	5.5	4.8	4.9	4.4	5.9	6.3	0.5	0.7	1.0	1.2	1.4	1.4	1.6	1.6	1.7
Black other groups	2.4	2.7	2.9	2.8	2.3	2.1	2.3	3.7	4.1	3.6	3.9	4.4	3.7	3.0	3.2	2.8	3.6	3.8	0.7	0.9	1.2	1.4	1.5	1.6	1.7	1.8	1.8
Chinese	0.8	0.8	0.9	0.7	0.7	0.6	0.7	1.2	1.4	0.3	0.5	0.6	0.5	0.5	0.4	0.5	0.7	0.7	0.5	0.6	0.6	0.7	1.0	1.1	1.1	1.2	1.4
Any other	2.1	2.2	2.5	2.4	2.0	1.8	2.1	3.5	3.8	1.2	1.3	1.6	1.7	1.4	1.4	1.3	1.7	1.9	0.3	0.3	0.5	0.6	0.7	0.8	0.9	0.9	1.0
Unclassified/Refused	2.8	2.8	2.6	2.4	2.1	1.9	2.2	3.5	3.7	2.7	3.0	3.1	2.9	2.5	2.7	2.5	3.2	3.3	0.5	0.7	0.8	1.1	1.3	1.3	1.4	1.4	1.7
All students	**2.6**	**2.6**	**2.7**	**2.5**	**2.1**	**2.0**	**2.1**	**3.6**	**4.0**	**1.9**	**2.2**	**2.4**	**2.4**	**2.2**	**2.2**	**2.0**	**2.6**	**2.8**	**0.5**	**0.6**	**0.7**	**0.9**	**1.0**	**1.0**	**1.1**	**1.2**	**1.3**

*Note: 'L.B.' ('lower bound') indicates results calculated after discounting SEN types in 2015 for those pupils recorded as having 'School Action' in 2014; was done to account for the fact that in previous years, the School Census requested SEN types for only those pupils with needs identified as 'School Action Plus' or above. 'U.B.' ('upper bound') indicates results calculated using all available SEN type records from the January 2015 School Census.

Ethnic disproportionality in SEN identification: Odds Ratios

In this section we report the results from multinomial logistic regression models. Table 1.8 presents the unadjusted ORs for each ethnic group, for each primary type of SEN. These ORs (highlighted according to the thresholds explained above) show whether and to what extent each minority ethnic group is over- or under-represented for each type of SEN, compared to White British pupils. Note that there is no implicit value judgment about whether disproportionality (over- or under-representation) is a positive, neutral or negative circumstance; these results only show whether and how much a particular ethnic group's odds of identification are higher or lower than White British pupils' odds of identification.

Unadjusted Odds Ratios: 'Raw' measures of ethnic disproportionality

Before adjusting for any additional pupil background characteristics, results for each SEN type indicated that, compared to White British pupils:

SpLD: Only the two Traveller groups had higher odds of identification, while pupils from the White other group had lower odds of identification. Most of the Black groups (Black African, Black other, Mixed White and African) had lower odds of identification, but the Black Caribbean and Mixed White and Caribbean groups had odds no higher or lower than those of White British pupils. Asian groups in general (Indian, Pakistani, Bangladeshi, Chinese, Mixed White and Asian, Asian other) had even lower odds of identification. [7]

MLD: The two Traveller groups were over-represented. Of the Black groups, most had odds of identification that were not substantially higher or lower than those of White British pupils, with the exception of the over-represented Black Caribbean group. Of the Asian groups, several were under-represented (Indian, Chinese, Asian other, Mixed White and Asian), but Pakistani pupils were over-represented.

SLD: The two Traveller groups were over-represented, as were Pakistani, Black African and Black other pupils. Chinese pupils were somewhat under-represented.

PMLD: The Irish Traveller group was over-represented, as were pupils from the Mixed White and African and Any other mixed ethnic groups. Pakistani, Bangladeshi, and the Asian other groups were also over-represented, as were Black groups (African, Caribbean, and Other). No groups were substantially under-represented.

SEMH: The two Traveller groups were over-represented, as were the Black Caribbean and Mixed White and Black Caribbean groups. Asian groups

[7]. The 'Any other ethnic group' category was also under-represented, but we refrain from making inferences about this category as it lacks substantive definition.

(Indian, Bangladeshi, Pakistani, other Asian), Chinese, White Other and Mixed White and Asian were all under-represented.

SLCN: Traveller groups were over-represented, as were Black (African, Caribbean, other), Pakistani and Bangladeshi groups. No minority ethnic groups were substantially under-represented.

ASD: Only the Black Caribbean and Black other groups had higher odds of identification than White British pupils. The Traveller, White other, Indian, Pakistani, and Asian other groups were under-represented.

HI: The two Traveller groups were over-represented, as were the Pakistani and Bangladeshi groups. Black African and Black other pupils were under-represented, along with White other, Mixed White and African, Mixed White and Caribbean, and Mixed White and Asian groups.

VI: The two Traveller groups were over-represented. Of the Asian groups, most had lower odds of identification (Indian, Chinese, Asian other), as did the Mixed White and Asian group, but Pakistani pupils were over-represented. Black African, Black other and Mixed White & African groups were under-represented, and so were the White other and Any other mixed groups.

MSI: As this SEN type had a low overall rate of identification and some ethnic groups constituted a small fraction of the total population as well, many of these results were not statistically significant (as indicated in the table). Irish Traveller and White Irish pupils had OR>1.50 but these results were not statistically significant. Asian and Black African groups were generally under-represented.

PD: Irish Traveller pupils were over-represented. All Asian groups except Pakistani pupils were under-represented, as were Mixed White & Asian pupils. White other, Mixed White and African, and pupils from the 'Any other mixed' groups were also all under-represented. As for MSI, because of low incidence and some ethnic groups making up a very small proportion of the total population, not all of these results were statistically significant.

Adjusted Odds Ratios: Ethnic disproportionality controlling for other pupil background characteristics

After adjusting for additional background characteristics including pupil FSM eligibility, gender, birth season, year group and neighbourhood IDACI, results indicated that, compared to White British pupils (Table 1.9):

SpLD: ORs – and therefore patterns of over- and under-representation – were largely unchanged relative to the unadjusted results for each ethnic group.

MLD: The White other, Mixed White & African, Any other mixed, Bangladeshi, Black African and Black other groups were under-represented, whereas they had

not been (according to even the less-stringent threshold of 0.75) in the unadjusted model. The Pakistani and Black Caribbean groups were no longer over-represented as they had been in the unadjusted models. Traveller groups were still over-represented, but considerably less so than in the unadjusted model, and the Mixed White and Asian, Indian, Chinese, and Asian other groups were still under-represented as they were in the unadjusted model.

SLD: The Irish Traveller group was no longer over-represented, while the Roma Traveller group still had higher odds of identification than White British pupils. The Mixed White and Caribbean group was somewhat under-represented, but had not been so in the unadjusted model. No other group remained substantially over- or under-represented.

PMLD: The Irish Traveller, Mixed White and African, and Black Caribbean groups were no longer over-represented compared to White British pupils (according to even the less-conservative 1.33 threshold). The odds of identification were slightly reduced for most other groups that had been over-represented in the unadjusted model, but the patterns of over- and under-representation were otherwise largely the same before and after adjusting for additional pupil background characteristics.

SEMH: The Roma Traveller group was no longer over-represented as in the unadjusted model. Black African pupils were under-represented only after adjusting for additional pupil background characteristics. Patterns of ethnic disproportionality in identification for the other minority ethnic groups were largely similar to those based on unadjusted ORs.

SLCN: Both Traveller groups were still over-represented, though somewhat less so. Of the other ethnic groups, only Black Caribbean pupils remained over-represented after adjusting, suggesting that the raw over-representation of other groups was to a great extent attributable e.g. deprivation, gender and age.

ASD: No group was over-represented after adjustment. The Bangladeshi group was substantially under-represented after adjusting, while other groups remained under-represented both before and after including additional pupil controls.

HI: The Roma Traveller group and Pakistani pupils were the only over-represented groups after adjustment for pupil background. The Bangladeshi group was no longer substantially over-represented. White other, Mixed White and African, Mixed White & Caribbean, and Mixed White and Asian groups remained under-represented both before and after adjusting for additional pupil background characteristics. The Black Caribbean group was under-represented only after adjusting for additional pupil background

characteristics, while Black African and Black other groups remained under-represented both before and after including additional controls.

VI: Pakistani pupils remained over-represented, as did the two traveller groups. Other results remained largely similar to those from the unadjusted model, with almost all other ethnic minority groups were under-represented.

MSI: The pattern of results was largely unchanged, most were not statistically significant.

PD: The Irish Traveller group was no longer over-represented after controlling for additional pupil background characteristics. Black Caribbean and Black other groups were under-represented only after including additional pupil controls, while other patterns of under-representation remained largely similar to those in the unadjusted model.

Figures 1-2A, 1-2B and 1-2C provide a visual comparison of unadjusted and adjusted ORs by ethnic group for MLD, SEMH and ASD respectively.

Table 1.8: Unadjusted Odds Ratios by ethnic minority group, 2016 (Y1-11)

	Cognition & Learning				Social, Emotional & Mental Health	Communication & Interaction		Sensory & Physical				Unspecified/ Other		
	SpLD	MLD	SLD	PMLD	SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA	N
White Irish	1.16 *	0.82 *	0.90	1.30	0.92	0.95	1.04	1.25	0.77	1.41	0.96	1.03	1.21 *	19044
Traveller Irish	***2.27*** *	***4.35*** *	***1.86*** *	***2.06*** *	***2.87*** *	***2.79*** *	***0.44*** *	***1.67*** *	***2.06*** *	***1.61***	1.44	***2.86*** *	***5.41*** *	4869
Traveller Gypsy/Roma	***1.59*** *	***3.78*** *	***2.10*** *	1.11	***1.64*** *	***2.48*** *	***0.39*** *	***3.15*** *	***1.59*** *	***0.33***	1.02	***1.71*** *	***3.39*** *	21735
White other groups	***0.53*** *	0.78 *	0.76 *	1.06	***0.57*** *	1.29 *	***0.60*** *	0.74 *	***0.61*** *	0.69 *	***0.54*** *	0.85 *	1.16 *	367017
Mixed White & African	0.74 *	0.82 *	1.00	1.35 *	1.19 *	1.23 *	0.94	***0.65*** *	***0.45*** *	0.73	0.70 *	0.88 *	1.04	45042
Mixed White & Caribbean	1.03	1.23 *	0.92	1.20	***1.94*** *	1.21 *	1.12 *	0.74 *	0.98	0.72	0.89 *	1.15 *	1.48 *	96033
Mixed White & Asian	***0.52*** *	0.68 *	0.87 *	1.01	0.72 *	0.94 *	0.84 *	0.74 *	0.69 *	***0.55*** *	0.70 *	0.69 *	0.80 *	78940
Mixed other	0.68 *	0.78 *	1.12 *	*1.50* *	1.07 *	1.14 *	1.06 *	0.80 *	0.74 *	0.90	0.74 *	0.96	1.04	122534
Indian	***0.28*** *	***0.56*** *	0.80 *	1.02	***0.24*** *	0.85 *	***0.46*** *	0.78 *	0.72 *	***0.42*** *	***0.56*** *	***0.60*** *	***0.60*** *	179111
Pakistani	***0.38*** *	1.36 *	***1.52*** *	***2.62*** *	***0.50*** *	1.39 *	***0.54*** *	***2.09*** *	***2.13*** *	0.75 *	1.13 *	0.96	***1.51*** *	275269
Bangladeshi	***0.36*** *	0.87 *	1.20 *	***1.94*** *	***0.46*** *	***1.62*** *	0.79 *	1.45 *	0.97	***0.54*** *	***0.63*** *	0.80 *	1.30 *	108478
Any other Asian	***0.31*** *	***0.59*** *	1.10 *	1.49 *	***0.31*** *	1.15 *	***0.63*** *	0.92	***0.62*** *	***0.36*** *	***0.61*** *	***0.67*** *	0.82 *	110319
Black African	***0.45*** *	0.84 *	1.41 *	***1.55*** *	0.83 *	***1.66*** *	1.15 *	***0.65*** *	0.73 *	***0.60*** *	***0.63*** *	0.87 *	1.23 *	235333
Black Caribbean	1.03	1.38 *	1.21 *	1.39 *	***2.29*** *	***1.82*** *	1.34 *	0.87	0.87	0.69	0.87 *	1.34 *	***1.63*** *	79909
Black other groups	***0.61*** *	1.03	***1.64*** *	***1.63*** *	1.31 *	***1.70*** *	1.34 *	0.67 *	***0.66*** *	0.85	0.76 *	1.05	1.39 *	46924
Chinese	***0.24*** *	***0.30*** *	0.72 *	0.81	***0.20*** *	1.20 *	0.91	0.82	***0.45*** *	***0.64***	***0.27*** *	***0.44*** *	***0.62*** *	25993
Any other ethnic group	***0.48*** *	0.91 *	1.03	***1.62*** *	***0.61*** *	***1.52*** *	0.73 *	1.09	0.73 *	***0.66*** *	***0.64*** *	0.95	***1.52*** *	111023
Unknown	0.94 *	0.90 *	1.30 *	1.23	1.11 *	1.04	1.24 *	0.90	1.05	1.15	0.90	1.14 *	1.37 *	60484
N	142397	260295	24389	7871	180463	176809	86018	17932	9983	1944	27750	50271	33793	6490615

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared = 0.011

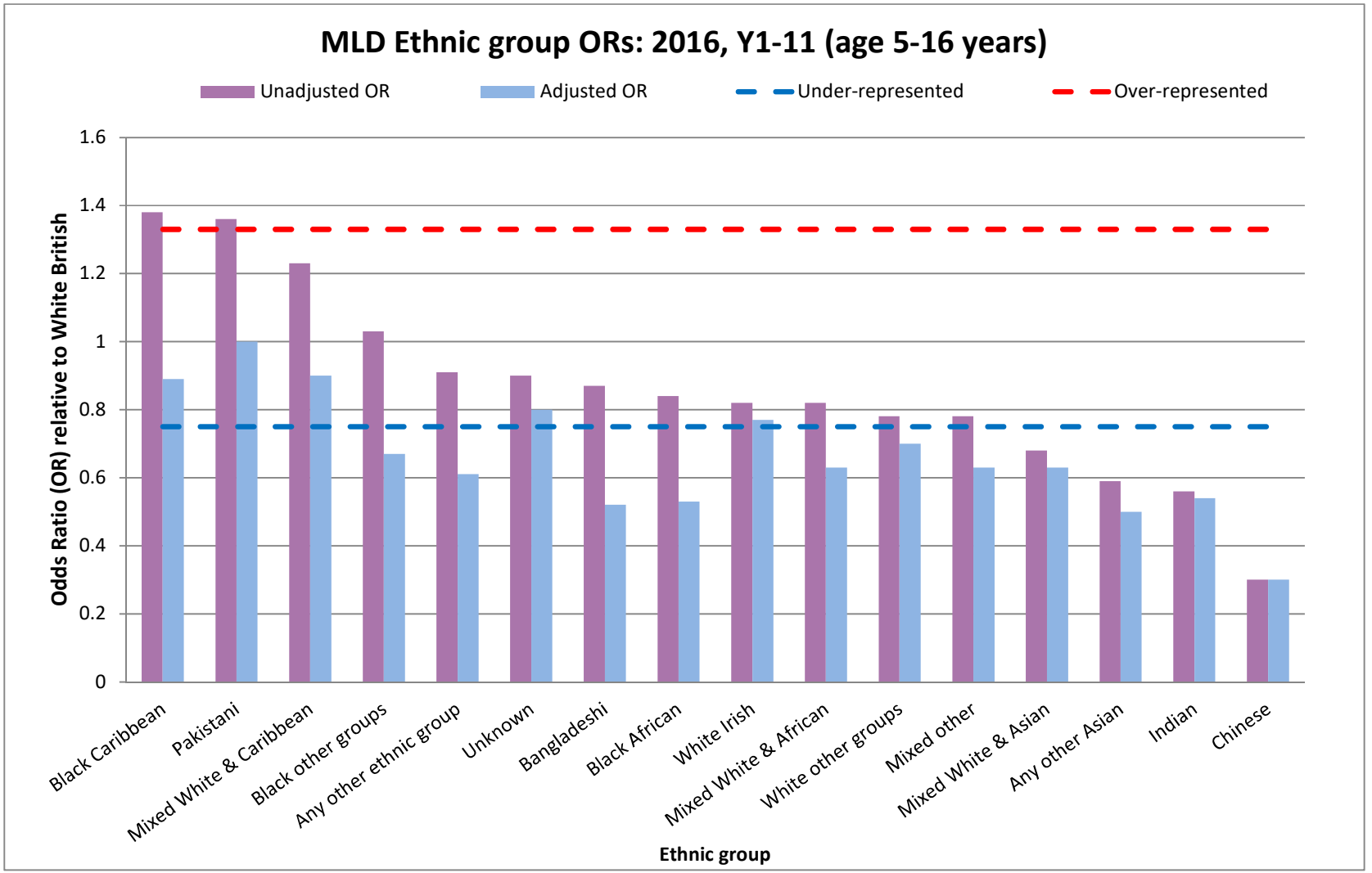
*Ethnic group, no additional pupil predictors/controls.

Table 1.9: Adjusted Odds Ratios by ethnic minority group, 2016 (Y1-11)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other				
	SpLD		MLD		SLD		PMLD		SEMH		SLCN		ASD		HI		VI		MSI		PD		Other		NSA		N
White Irish	1.12	*	0.77	*	0.86		1.29		0.85	*	0.94		1.01		1.21		0.74		1.42		0.95		0.99		1.18		19044
Traveller Irish	***2.05***	*	***2.45***	*	1.01		1.21		***1.53***	*	1.47	*	0.31	*	1.29		1.49		1.13		0.94		***2.02***	*	***3.16***	*	4869
Traveller Gypsy/Roma	1.50	*	***2.70***	*	***1.61***	*	0.91		1.17	*	***1.78***	*	0.33	*	***2.76***	*	1.36	*	0.29		0.85		1.41	*	***2.51***	*	21735
White other groups	0.56	*	0.70	*	0.80	*	1.10		0.53	*	1.08	*	0.61	*	0.75	*	0.61	*	0.66	*	0.54	*	0.83	*	1.01		367017
Mixed White & African	0.73	*	0.63	*	0.84	*	1.18		0.92	*	0.90	*	0.86	*	0.60	*	0.41	*	0.65		0.62	*	0.77	*	0.80	*	45042
Mixed White & Caribbean	0.96	*	0.90	*	0.72	*	1.01		1.38	*	0.90	*	0.97		0.65	*	0.84	*	0.65	*	0.76	*	0.95		1.13	*	96033
Mixed White & Asian	0.54	*	0.63	*	0.83	*	0.96		0.67	*	0.82	*	0.81	*	0.73	*	0.68	*	0.51	*	0.68	*	0.67	*	0.72	*	78940
Any other mixed	0.67	*	0.63	*	1.00		1.38	*	0.88	*	0.90	*	0.99		0.76	*	0.69	*	0.82		0.68	*	0.86	*	0.85	*	122534
Indian	0.28	*	0.54	*	0.87	*	1.10		0.23	*	0.80	*	0.47	*	0.79	*	0.73	*	0.42	*	0.58	*	0.59	*	0.57	*	179111
Pakistani	0.35	*	1.00		1.30	*	***2.44***	*	0.36	*	1.07	*	0.48	*	***1.89***	*	***1.87***	*	0.70	*	1.02		0.80	*	1.18	*	275269
Bangladeshi	0.31	*	0.52	*	0.91		***1.72***	*	0.26	*	1.09	*	0.65	*	1.21	*	0.77	*	0.49	*	0.53	*	0.59	*	0.89	*	108478
Any other Asian	0.31	*	0.50	*	1.06		1.48	*	0.27	*	0.97		0.60	*	0.89	*	0.59	*	0.34	*	0.59	*	0.61	*	0.71	*	110319
Black African	0.40	*	0.53	*	1.09	*	1.35	*	0.52	*	1.11	*	0.97		0.55	*	0.60	*	0.53	*	0.53	*	0.66	*	0.85	*	235333
Black Caribbean	0.89	*	0.89	*	0.92		1.21		1.43	*	1.34	*	1.12	*	0.73	*	0.70	*	0.65		0.73	*	1.01		1.19	*	79909
Black other groups	0.55	*	0.67	*	1.27	*	1.40	*	0.84	*	1.16	*	1.13	*	0.58	*	0.54	*	0.76		0.64	*	0.82	*	0.98		46924
Chinese	0.25	*	0.30	*	0.80		0.87		0.21	*	1.09	*	0.96		0.85		0.47	*	0.64		0.28	*	0.46	*	0.59	*	25993
Any other ethnic group	0.44	*	0.61	*	0.81	*	1.41	*	0.40	*	1.06	*	0.61	*	0.95		0.62	*	0.59	*	0.54	*	0.76	*	1.10	*	111023
Unknown	0.86	*	0.80	*	1.19	*	1.19		0.95	*	1.01		1.13	*	0.84	*	0.97		1.16		0.86	*	1.02		1.30	*	60484
N	142397		260295		24389		7871		180463		176809		86018		17932		9983		1944		27750		50271		33793		6490615

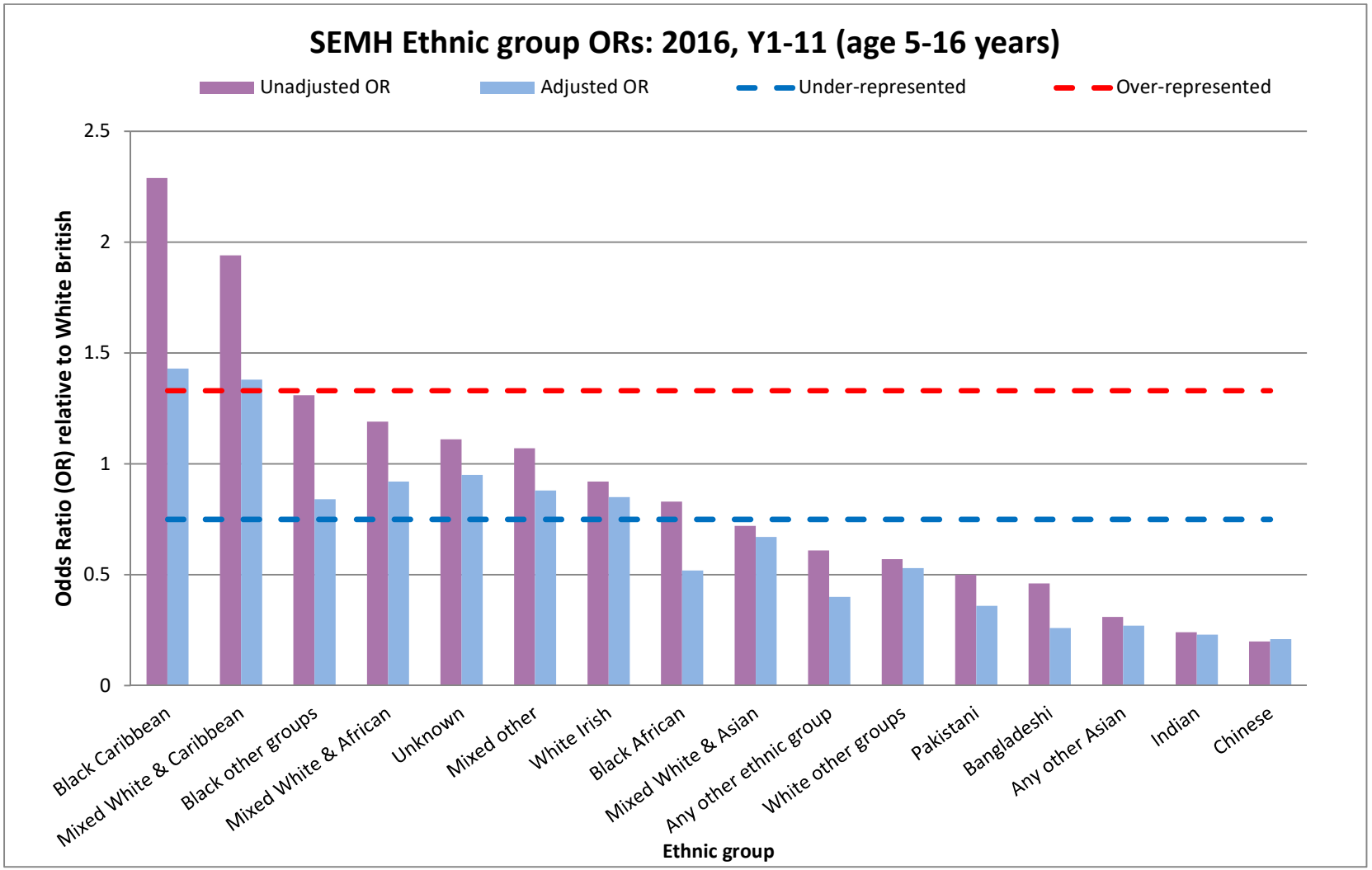
* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared = 0.096

*After adjusting for: Normalised IDACI, DSM eligibility, gender, birth season, Year Group,



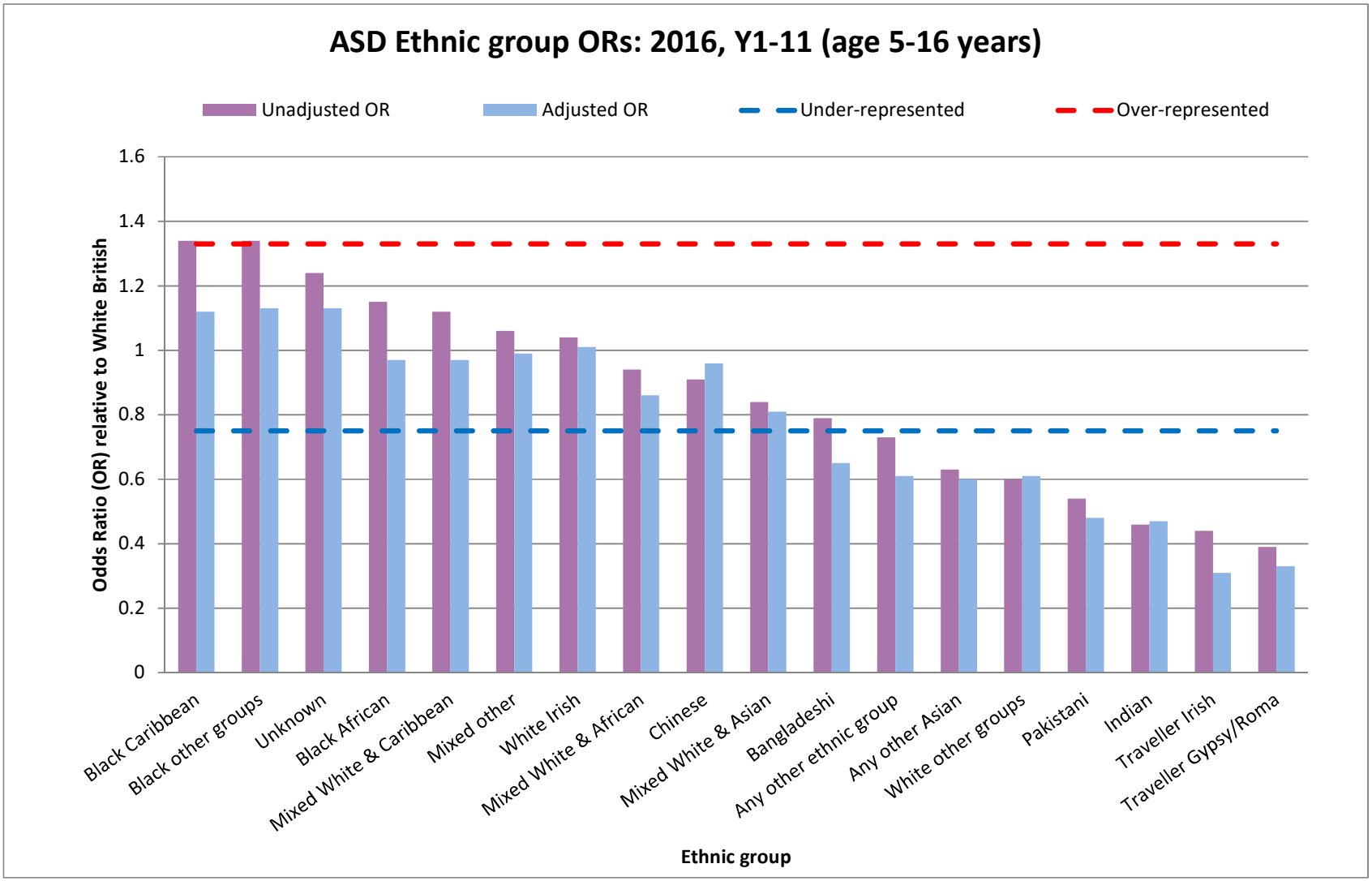
Note: Traveller groups have been omitted from the figure above as their inclusion altered the scale, complicating interpretation; the relevant ORs were included in the unadjusted and adjusted OR tables, however.

Figure 1-2A: Ethnic group adjusted and unadjusted ORs (2016): Bar chart, MLD



Note: Traveller groups have been omitted from the figure above as their inclusion altered the scale, complicating interpretation; the relevant ORs were included in the unadjusted and adjusted OR tables, however.

Figure 1-2B: Ethnic group adjusted and unadjusted ORs (2016): Bar chart, SEMH



Note: Traveller groups have been omitted from the figure above as their inclusion altered the scale, complicating interpretation; the relevant ORs were included in the unadjusted and adjusted OR tables, however.

Figure 1-2C: Ethnic group adjusted and unadjusted ORs (2016): Bar chart, ASD

Association of additional pupil background characteristics with type of SEN

It is useful to examine the effects of each of the additional pupil background variables directly. Table 1.10 shows the exponentiated coefficients (ORs) for each of the pupil background variables (FSM eligibility, gender, birth season, year group, and normalised IDACI score) on each of the different types of SEN. The ethnic group coefficients are included but not re-reported, as they were presented in the previous section. Overall, the additional pupil controls explained a considerable amount of the variance in SEN. For example in a model of predicting any SEN, the model with all pupil background characteristics explained 9.6% of the variance in the outcome, compared to just 1.1% of the variance in the model with ethnic group as the only independent variable.

Socio-economic disadvantage: The FSM coefficient represents the contrast between a pupil on FSM and one not on FSM, and the IDACI score represents a 2SD change, or moving from 1SD below the average deprivation to 1SD above the average deprivation. We also report a 'combined deprivation' effect for both socio-economic variables – this is not an additional variable in the multinomial logistic regression model, rather it is the combined effects of IDACI and FSM calculated post-hoc.[8] This is strongly associated with only some types of SEN, particularly MLD, SEMH, SLCN, and the Other/Unspecified types. FSM, on the other hand, appears to have a considerably substantial effect on the odds of identification for all SEN types; this is, however, a very coarse measure as FSM eligibility only distinguishes between a small proportion of very economically disadvantaged pupils and all others. According to the 'combined deprivation' effect sizes, socio-economic disadvantage/deprivation is most strongly associated with MLD, SLD and SEMH identification, and least strongly associated with SpLD, HI, VI and MSI (notably, most of the Sensory and Physical types of primary need).

Sex: Boys were over-represented for every SEN type except HI, with effect sizes only over the less-conservative threshold of 1.33 for PMLD and VI. This gender effect was most pronounced for ASD (with boys having over 5 times the odds of being identified), SEMH (with boys having over 3 times the odds of being identified), and SLCN (with boys having over 2.5 the odds of being identified).[9]

[8] *Normalised IDACI 2SD effect was calculated as e^(2*IDACI logit), and the combined deprivation effect was calculated as e^(IDACI logit+FSM logit).*

[9] *These results may appear slightly different from those reported by Strand and Lindsay (2009), who found that the odds of identification for Sensory and Physical types of need were not substantially different for boys and girls; however, such differences are the result of different analytical choices including the use of less stringent cutoffs for assessing ORs (0.75 and 1.33 versus 0.67 and 1.50) and different predictors being included (here, birth season).*

Table 1.10: ORs for additional pupil control variables, 2016

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/Other			
Pupil variables	SpLD		MLD		SLD		PMLD		SEMH		SLCN		ASD		HI		VI		MSI		PD		Other		NSA	
FSM																										
Eligible	***1.68***	*	***2.41***	*	***3.51***	*	***2.75***	*	***3.08***	*	***2.09***	*	***2.31***	*	***1.80***	*	***1.99***	*	***1.74***	*	***2.29***	*	***2.04***	*	***1.97***	*
Gender																										
Boy	***1.78***	*	***1.72***	*	***2.06***	*	1.46	*	***3.17***	*	***2.53***	*	***5.37***	*	1.16	*	1.38	*	***1.98***	*	***1.55***	*	***1.73***	*	***1.70***	*
Birth Season																										
Summer	***1.52***	*	***1.83***	*	1.26	*	1.10	*	1.16	*	***1.64***	*	1.09	*	1.12	*	1.07	*	1.27	*	1.18	*	1.43	*	***1.71***	*
Spring	1.24	*	1.35	*	1.09	*	1.05		1.07	*	1.29	*	1.04	*	1.07	*	1.00		1.06		1.09	*	1.19	*	1.27	*
Year Group																										
Y2	***1.97***	*	***1.65***	*	1.14	*	0.95		1.31	*	0.90	*	1.09	*	1.18	*	1.17	*	1.09		1.12	*	1.33	*	1.36	*
Y3	***2.80***	*	***1.99***	*	1.16	*	0.93		1.48	*	0.72	*	1.15	*	1.23	*	1.19	*	1.07		1.13	*	1.40	*	1.31	*
Y4	***3.74***	*	***2.29***	*	1.20	*	0.98		***1.70***	*	***0.62***	*	1.23	*	1.28	*	1.29	*	0.95		1.12	*	***1.55***	*	1.28	*
Y5	***4.53***	*	***2.49***	*	1.28	*	0.90	*	***1.86***	*	***0.53***	*	1.33	*	1.26	*	1.34	*	0.97		1.11	*	***1.65***	*	1.17	*
Y6	***5.25***	*	***2.71***	*	1.44	*	0.88	*	***1.98***	*	***0.49***	*	1.42	*	1.38	*	1.37	*	***0.63***	*	1.13	*	***1.72***	*	1.20	*
Y7	***5.41***	*	***2.15***	*	1.31	*	0.86	*	***1.78***	*	***0.36***	*	***1.56***	*	***1.52***	*	1.42	*	0.72	*	1.13	*	***2.15***	*	1.02	
Y8	***5.41***	*	***2.05***	*	1.29	*	0.76	*	***1.80***	*	***0.33***	*	***1.54***	*	***1.56***	*	***1.59***	*	***0.57***	*	1.05		***1.99***	*	0.84	*
Y9	***5.30***	*	***1.87***	*	1.34	*	0.75	*	***1.91***	*	***0.29***	*	***1.54***	*	***1.55***	*	***1.57***	*	***0.54***	*	1.04		***1.82***	*	***0.59***	*
Y10	***5.40***	*	***1.75***	*	1.35	*	0.74	*	***2.06***	*	***0.27***	*	1.47	*	***1.61***	*	***1.61***	*	***0.60***	*	1.02		***1.80***	*	***0.53***	*
Y11	***5.43***	*	***1.72***	*	1.39	*	0.70	*	***2.35***	*	***0.24***	*	1.49	*	***1.54***	*	***1.58***	*	***0.53***	*	1.07	*	***1.91***	*	***0.62***	*
Deprivation																										
Normalised IDACI (2SD)	1.12	*	***1.91***	*	1.18	*	1.00		***1.92***	*	***1.72***	*	1.17	*	1.19	*	1.25	*	1.06		1.14	*	1.42	*	***1.66***	*
Combined deprivation (IDACI+FSM)	***1.78***		***3.33***		***3.82***		***2.75***		***4.27***		***2.75***		***2.50***		***1.96***		***2.22***		***1.79***		***2.44***		***2.43***		***2.53***	

* indicates significance at the p<0.05 level; not marked for combined deprivation as the individual components are listed (with appropriate significance indicators) elsewhere in the above table.

Birth season: This was associated with some SEN types but not others. Summer-born pupils had substantially higher odds of identification than Autumn-born pupils for SpLD, MLD, SLCN, and the Other/Unspecified categories, and for MLD, Spring-born pupils also had somewhat higher odds of identification than Autumn-born pupils. It is important not to make strong inferences of cause and effect here. However, there is a possibility that this may be a result of teachers or schools not making appropriate allowances for pupils' development according to their age.

Year group: This appeared to be at least somewhat associated with identification for many SEN types, though not for PD. Pupils in the later years of secondary school (Y9-Y11) had lower odds of PMLD identification and SEN support with no specialist assessment (relative to Y1), but no substantial differences in the odds of identification for other year groups. For SLCN there was a general pattern of lower odds of identification for pupils in higher year groups. For MLD the odds declined somewhat in Y9-Y11 while the odds for SEMH increased, perhaps reflecting changes in the definition of the primary need of some pupils between the two types. The odds of ASD identification were higher for later primary year groups (Y5 and Y6) and then remained relatively consistent for secondary year groups. Odds of identification for HI and VI were somewhat higher for older pupils, and much higher for older pupils for SpLD. Year group was not strongly associated with SLD identification.

Because including these additional pupil characteristics made some substantial differences to the ethnic group ORs, the additional variables explained considerably more variance than did ethnic group alone, and each additional pupil control variable had significant and substantial effects for at least some SEN types, these were again included in subsequent multilevel models (the results of which are given in Part 2).

English as an Additional Language (EAL)

We did not include EAL along with the other pupil controls for the analysis presented here. This is because the EAL measure available in the School Census is not in fact a measure of pupils' English fluency, rather, it indicates whether or not a child was exposed to a first language other than English at home during early development. It is also highly confounded with ethnicity. The effects of such a measure, therefore, become potentially problematic to interpret or to inform inferences. For a further discussion see Strand et. al. (2015) and Strand & Hessell (2017). Appendix B presents the adjusted ORs after including EAL in addition to the other predictors already discussed in this section. Importantly, including EAL as a predictor did not explain additional variance in the multinomial regression (Nagelkerke's Pseudo-R^2= 0.096), further justifying the exclusion of this variable in subsequent analyses.

Considerations of school phase

Appendix C presents analogous results (unadjusted and adjusted ORs) separated by phase of schooling (primary and secondary). While ORs for individual ethnic groups

do vary across phases, most of the key patterns of under- and over-representation hold across phases (e.g. Black Caribbean and Mixed White and Caribbean over-representation for SEMH); the exception to this is that Asian groups are less consistently under-represented for ASD in the primary year groups (Y1-6).

Analysis by level of need

For all of the results presented above, we considered SEN type for any level of need. However, SEN identification at the level of SEN support and at the level of a statement or EHC plan are quite different processes and may exhibit different patterns of disproportionality as a result. Appendices D & E present analogous results for only pupils with SEN support as their level of need, and for statement/EHC plan as their level of need, respectively. There are differences in ethnic group ORs evident in these results, but our main analyses proceed with a focus on all those identified with a focal type of SEN (not separated by level of need) as the statement/EHC incidence for any given SEN type is very low in the overall population, which is prone to causing complications (unstable estimates) for more complex statistical analyses.

Interactions between ethnic group, FSM and gender

Because FSM and gender had substantial ORs for MLD, SEMH and ASD, we further investigated possible interactions between these and ethnic group to understand whether the likelihood of being identified was affected differently within minority ethnic groups for girls and boys, and for pupils eligible for FSM (i.e. experiencing socioeconomic deprivation) and those who were not.

*Ethnic group * FSM*

Table 1.11 gives the ORs for pupils with and without FSM from a model including ethnic group, additional pupil controls as described in the section above, and an interaction between ethnic group and FSM. The ORs in the table are evaluated at the mean (or reference value) of all other independent variables (e.g. IDACI).

Most of the substantive interpretations associated with ethnic group ORs do not change greatly with the inclusion of either an ethnic group by FSM interaction or an ethnic group by gender interaction. There are some exceptions, however:

- For MLD, the largest differences in the ORs for pupils entitled to FSM compared to those not entitled to FSM are for the Irish Traveller and Gypsy/Roma groups, but these are very small groups, which generates much volatility. In terms of substantive differences, the largest are for Black Caribbean, Mixed White and Black Caribbean and Pakistani pupils. We noted that these three groups were over-represented for MLD in the raw results but not over-represented after adjusting for pupil background. The current results indicate that these three ethnic groups are even under-represented for MLD among those entitled to FSM. However, given around three-quarters of pupils in these ethnic groups are not

entitled to FSM, the substantive conclusion remains the same, that these groups are not over-represented for MLD.

- For SEMH, once an ethnic group and FSM interaction has been included, the big difference is that White British pupils on FSM are at higher risk of SEMH identification than is apparent in the model without the interaction terms. Ethnic groups that are under-represented relative to White British pupils amongst pupils without FSM entitlement are more pronouncedly under-represented amongst pupils with FSM (e.g. Black African, OR=0.70 for pupils with no FSM, OR=0.32 for pupils with FSM). Further, some groups that are not under-represented amongst pupils with no FSM (e.g. Mixed White & African, OR=1.07; Black Other, OR=1.09) are under-represented amongst pupils with FSM (Mixed White & African, OR=0.70; Black Other, OR=0.55). Perhaps more strikingly, the groups that are over-represented amongst pupils with no FSM (Mixed White & Black Caribbean, OR=1.61; Black Caribbean, OR=1.80) are not over-represented relative to the White British group amongst pupils with FSM (Mixed White & Black Caribbean, OR=1.10; Black Caribbean, OR=1.00).

- For ASD, after including the ethnic group and FSM interaction, differences between ORs across pupils with and without FSM entitlement are mostly minor in terms of their substantive interpretations. A few groups that are not under-represented relative to the White British group amongst pupils with no FSM (Mixed White & African, OR=0.92; Mixed White & Asian, OR=0.87; Bangladeshi, OR=0.78) do appear under-represented amongst pupils with FSM (Mixed White & African, OR=0.71; Mixed White & Asian, OR=0.66; Bangladeshi, OR=0.43). Most groups that are under-represented amongst pupils with no FSM are more pronouncedly so amongst pupils with FSM entitlement.

Figures 1-3A, 1-3B and 1-3C show the effect of the ethnic group by FSM interactions on the predicted probabilities for MLD, SEMH and ASD respectively, including only those interactions that were statistically significant at the level of $p<0.01$[10]. For MLD, the probability of identification is higher for pupils entitled to FSM than those not entitled to FSM for all ethnic groups, but the increase in risk is markedly higher for White British pupils. As a result, the under-representation relative to the White British majority is particularly evident among pupils on FSM. For SEMH, Figure 1-3B tells a similar story with regard to under-represented groups for SEMH; however, over-represented groups (specifically Black Caribbean and Mixed White and Black Caribbean pupils) are most pronouncedly so amongst pupils not eligible for FSM, the difference being small or non-existent among those entitled to FSM. For ASD

[10] We use $p<0.01$, a more stringent cutoff than considered elsewhere, in the interest of parsimony and to avoid making too much of interactions that might be borderline at a more relaxed significance threshold.

identification, under-representation is generally greater among pupils entitled to FSM, but it tends to be the same ethnic groups that are under-represented.

Table 1.11: Odds ratios after including ethnic group by FSM interactions (2016, Y1-11)

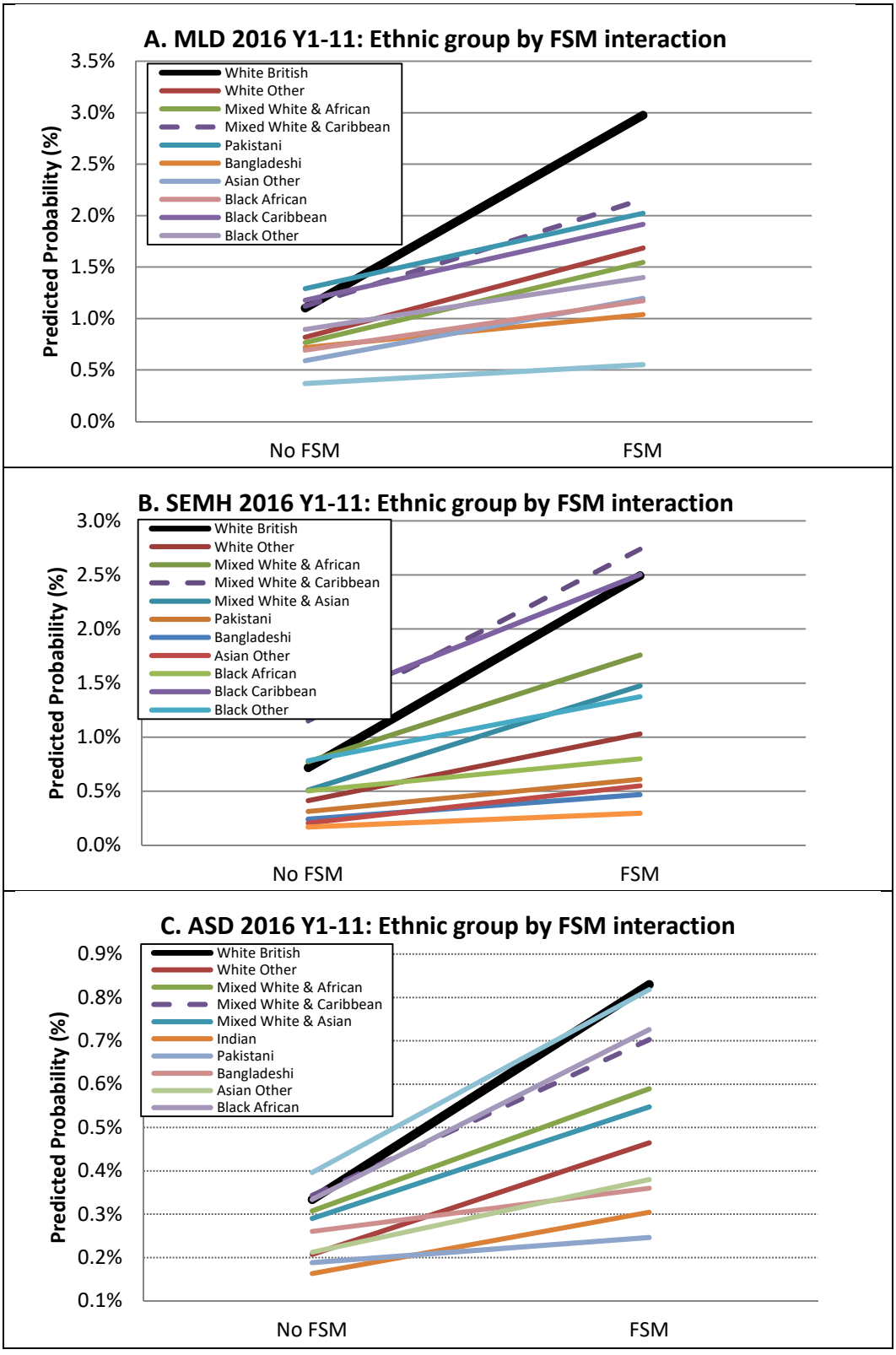
Ethnic group	MLD			SEMH			ASD		
	No FSM	FSM[a]		No FSM	FSM[a]		No FSM	FSM[a]	
White Irish	0.67	0.98	*	0.79	0.98		1.02	1.00	
Traveller Irish	***4.07***	***1.83***	*	***2.04***	1.27	*	***0.29***	***0.29***	
Traveller Roma	***3.12***	***2.10***	*	1.26	1.01	*	***0.32***	***0.33***	
White Other	0.74	***0.56***	*	***0.57***	***0.41***	*	***0.62***	***0.56***	
Mixed White & African	0.69	***0.51***	*	1.07	0.70	*	0.92	0.71	*
Mixed White & Caribbean	1.02	0.72	*	***1.61***	1.10	*	1.03	0.85	*
Mixed White & Asian	***0.62***	***0.63***		0.71	***0.59***	*	0.87	***0.66***	*
Mixed Other	0.67	***0.55***	*	0.96	0.74	*	0.99	0.96	
Indian	***0.56***	***0.52***		***0.24***	***0.22***		***0.49***	***0.36***	*
Pakistani	1.17	0.67	*	***0.44***	***0.24***	*	***0.56***	***0.29***	*
Bangladeshi	***0.65***	***0.34***	*	***0.33***	***0.18***	*	0.78	***0.43***	*
Asian Other	***0.53***	***0.39***	*	***0.29***	***0.22***	*	***0.63***	***0.46***	*
Black African	***0.63***	***0.39***	*	0.70	***0.32***	*	1.00	0.87	*
Black Caribbean	1.07	***0.64***	*	***1.80***	1.00	*	1.13	1.02	
Black Other	0.81	***0.46***	*	1.09	***0.55***	*	1.19	0.98	*
Chinese	***0.33***	***0.18***	*	***0.23***	***0.12***	*	0.96	0.96	
Any Other	0.73	***0.43***	*	***0.49***	***0.29***	*	0.72	***0.42***	*
Unknown	0.85	0.69	*	1.01	0.83	*	1.12	1.12	

Note: '*' here denotes interactions that are significant at the p<0.01 level.

ORs in the table are given after holding all other independent variables at their mean/reference values.

[a] The base for ORs for minority groups *with FSM* is the White British FSM group, while the ethnic minority group *no FSM* base for ORs is the White British no FSM group (for the sake of comparing like with like).

Figure 1.3:Predicted probabilities for ethnic by FSM interactions: MLD, SEMH & ASD



Note: Some very small groups were omitted from the above where their inclusion would have substantially changed the scale. These groups are still included in the OR tables (Table 1.11 and 1.12).

*Ethnic group * Sex*

Table 1.12 shows the ORs for boys and girls from a model including ethnic group, additional pupil background, and an interaction between ethnic group and sex. Again, ORs in the table are evaluated at the mean (or reference value) of all other independent variables like IDACI. Figures 1-4A, 1-4B and 1-4C show predicted probabilities for the ethnic group by sex interactions, again displaying (for each SEN type) only those interactions found to be statistically significant ($p<0.01$).

- For MLD, the differences in ORs for boys and girls in each ethnic group are minimal.

- For SEMH, the differences in ORs for boys and girls in each ethnic group are again small. The over-representation of Mixed White & Black Caribbean pupils, and the under-representation of Indian and Bangladeshi pupils, are slightly greater among boys than among girls.

- For ASD, the differences in ORs for boys and girls in each ethnic group are again minimal. The only statistically significant result was for the Pakistani group, and the substantive difference was small (Pakistani Girls OR=0.54; Pakistani Boys OR=0.46).

Summary

These findings are important because they alert us to nuances in the interrelationship between ethnicity, class and gender. The 'effect' of one dimension (e.g. social class) may not be consistent across all levels of another (e.g. ethnicity). They also mirror similar interactions between ethnicity, SES and gender seen for educational attainment at age 5 (Strand, 1999), age 11 (Strand 2014b) and at age 16 (Strand, 2014a).

Nevertheless, the Nagelkerke's Pseudo R squared values associated with these models (9.6% for the multinomial model with no interactions, 9.6% for the model with an ethnic group by gender interaction, and 9.7% for the model with an ethnic group by FSM interaction) suggests that including the interaction terms does not substantially improve model fit (i.e. the power of the model to accurately predict pupil's classification in terms of SEN identification by type of need) but does increase processing times given the size of the NPD datasets. Therefore, in subsequent analyses in parts 2 & 3 we do not include these interaction terms.

Table 1.12: Odds ratios after including ethnic group by sex interactions (2016, Y1-11)

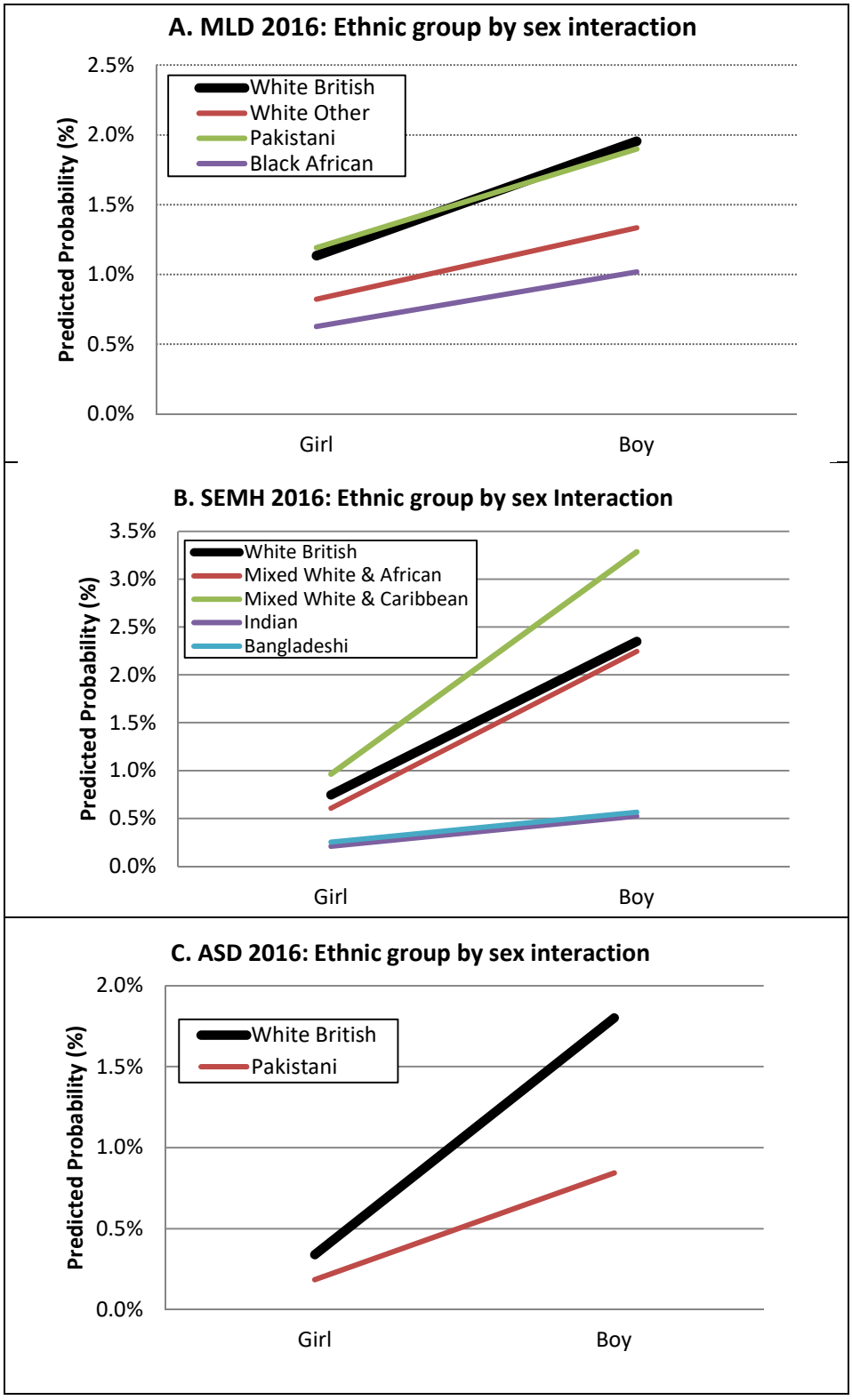
Ethnic group	MLD			SEMH			ASD		
	Girl	Boy[a]		Girl	Boy[a]		Girl	Boy[a]	
White Irish	0.80	0.74		0.84	0.85		0.97	1.01	
Traveller Irish	***2.32***	***2.54***		1.20	***1.66***		***0.29***	***0.31***	
Traveller Roma	***3.03***	***2.45***	*	1.15	1.15		***0.44***	***0.30***	
White Other	0.72	0.68	*	***0.53***	***0.53***		***0.59***	***0.61***	
Mixed White & African	***0.64***	***0.62***		0.81	0.96	*	0.79	0.87	
Mixed White & Caribbean	0.87	0.92		1.29	1.41	*	0.95	0.98	
Mixed White & Asian	***0.60***	***0.65***		0.68	***0.67***		0.81	0.81	
Mixed Other	***0.63***	***0.64***		0.91	0.87		0.99	0.99	
Indian	***0.54***	***0.54***		***0.28***	***0.22***	*	***0.43***	***0.48***	
Pakistani	1.05	0.97	*	***0.37***	***0.35***		***0.54***	***0.46***	*
Bangladeshi	***0.54***	***0.51***		***0.34***	***0.24***	*	0.69	***0.64***	
Asian Other	***0.51***	***0.49***		***0.28***	***0.26***		***0.58***	***0.60***	
Black African	***0.55***	***0.52***	*	***0.51***	***0.52***		0.96	0.97	
Black Caribbean	0.88	0.91		1.37	1.46		1.03	1.14	
Black Other	***0.65***	0.68		0.83	0.84		1.27	1.10	
Chinese	***0.27***	***0.32***		***0.24***	***0.20***		1.04	0.95	
Any Other	***0.67***	***0.57***	*	***0.43***	***0.39***		***0.54***	***0.62***	
Unknown	0.76	0.83		1.04	0.92		1.12	1.13	

Note: '*' denotes interactions that are significant at the p<0.01 level.

ORs in the table are given after holding all other independent variables at their mean/reference values.

[a] The base for ORs for minority group Boy is White British Boys, while the base for ethnic minority group Girl is White British Girls (for the sake of comparing like with like).

Figure 1.4: Predicted probabilities for ethnic by sex interactions: MLD, SEMH & ASD



Note: Some very small groups were omitted from the above where their inclusion would have substantially changed the scale. These groups are still included in OR tables, however (Table 1.11 and 1.12).

Trends over time in unadjusted and adjusted ORs for ethnic groups

Tables 1.13 presents the unadjusted ethnic ORs from analogous models from 2005 to 2016.

Tables 1.14 presents the adjusted ethnic ORs from analogous models from 2005 to 2016.

Ethnic disproportionality seems an extraordinarily stable and consistent feature of the data over this 12 year period. This is true despite the changes in prevalence rates we saw in Table 1.7. Figure 1.4 to 1.6 displays the unadjusted ORs, showing only those ethnic groups who were under-represented or over-represented in at least one year during the time series.

MLD

The over-representation of Black Caribbean pupils is consistent across all years, and Pakistani pupils have been consistently either over-represented or close to the threshold. Chinese, Indian, Other Asian, and Mixed White & Asian have been consistently under-represented. White Other groups have been at or just above the threshold for under-representation in all years.

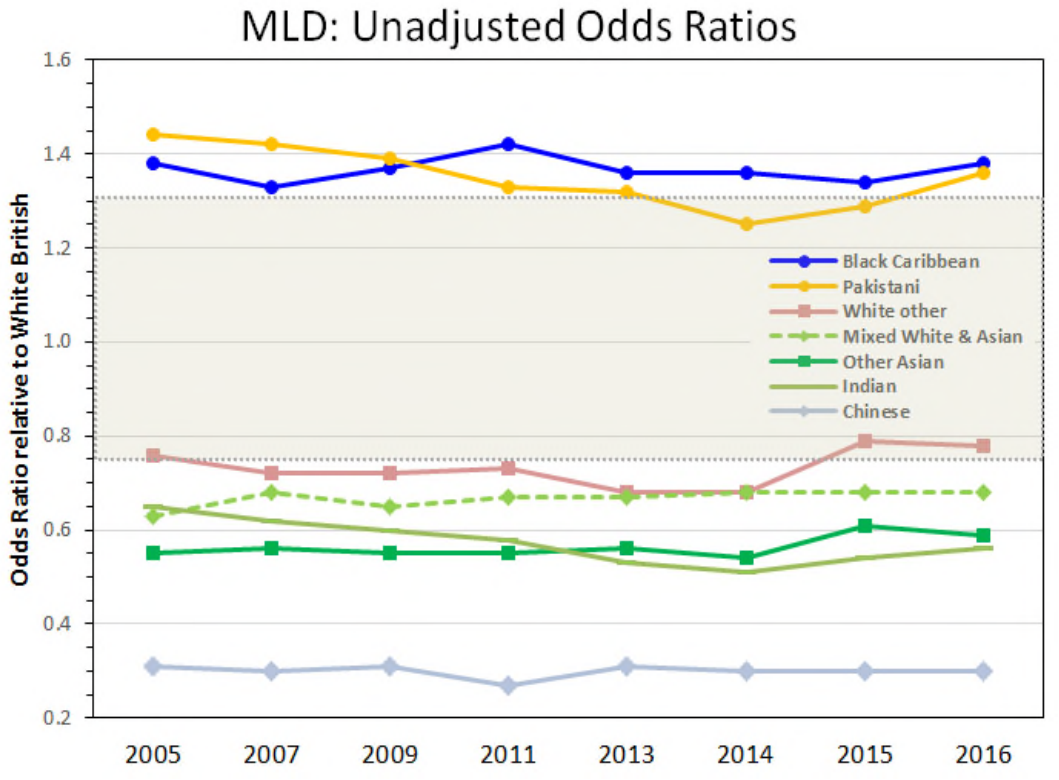


Figure 1.4: Unadjusted ethnic group Odds Ratios for MLD 2005-2016

Note: The figure excludes ethnic groups that were never over or under-represented in any year, and also excludes the small traveller groups. The shaded area represents the region OR> 0.75 and OR <1.33 so not substantially different from White British. The 2015 data is based on all pupils with a recorded primary need.

Table 1.13: Unadjusted OR time trends by ethnic group for focal SEN types (MLD, SEMH, ASD)

	MLD									BESD/SEMH									ASD								
Ethnic Group	*2005*	*2007*	*2009*	*2011*	*2013*	*2014*	*2015 (L.B.) **	*2015 (U.B.) **	*2016*	*2005*	*2007*	*2009*	*2011*	*2013*	*2014*	*2015 (L.B.) **	*2015 (U.B.) **	*2016*	*2005*	*2007*	*2009*	*2011*	*2013*	*2014*	*2015 (L.B.) **	*2015 (U.B.) **	*2016*
White Irish	1.01	0.99	0.98	0.83	0.88	0.82	0.93	0.83	0.82	1.08	1.11	1.13	0.99	0.93	0.93	0.87	0.85	0.92	1.22	1.15	1.01	1.11	1.16	1.11	1.16	1.13	1.04
Traveller Irish	***6.73***	***6.33***	***6.59***	***6.37***	***5.19***	***4.35***	***4.59***	***4.77***	***4.35***	***3.55***	***3.36***	***3.29***	***3.05***	***3.01***	***2.81***	***2.76***	***2.89***	***2.87***	***0.26***	***0.35***	***0.30***	***0.43***	***0.45***	***0.57***	***0.55***	***0.56***	***0.44***
Traveller Roma	***5.69***	***5.20***	***5.04***	***3.88***	***3.17***	***3.00***	***3.40***	***3.38***	***3.78***	***2.72***	***2.83***	***2.43***	***1.92***	***1.69***	***1.71***	***1.68***	***1.75***	***1.64***	***0.50***	***0.42***	***0.44***	***0.39***	***0.36***	***0.36***	***0.34***	***0.33***	***0.39***
White other	0.76	0.72	0.72	0.73	0.68	0.68	0.78	0.79	0.78	0.71	0.69	0.68	0.67	***0.59***	***0.56***	***0.56***	***0.58***	***0.57***	0.96	0.92	0.83	0.70	***0.64***	***0.64***	***0.62***	***0.60***	***0.60***
MWBA	0.77	0.85	0.90	0.87	0.84	0.81	0.83	0.85	0.82	1.44	1.36	1.24	1.23	1.18	1.20	1.22	1.21	1.19	1.14	0.97	0.97	1.00	1.00	0.97	1.03	1.00	0.94
MWBC	1.00	1.11	1.16	1.21	1.22	1.24	1.24	1.21	1.23	***2.03***	***2.04***	***1.99***	***1.95***	***1.93***	***1.97***	***1.92***	***1.92***	***1.94***	1.03	0.93	1.03	1.08	1.08	1.09	1.12	1.11	1.12
MWAS	***0.63***	0.68	***0.65***	0.67	0.67	0.68	0.67	0.68	0.68	0.68	0.69	0.67	0.73	0.73	0.69	0.69	0.69	0.72	1.08	1.08	0.99	0.95	0.86	0.86	0.87	0.84	0.84
Mixed other	0.79	0.80	0.80	0.80	0.81	0.80	0.83	0.78	0.78	1.22	1.20	1.16	1.09	1.06	1.10	1.09	1.07	1.07	1.26	1.24	1.20	1.20	1.18	1.13	1.10	1.06	1.06
Indian	***0.65***	***0.62***	***0.60***	***0.58***	***0.53***	***0.51***	***0.55***	***0.54***	***0.56***	***0.22***	***0.22***	***0.23***	***0.24***	***0.24***	***0.21***	***0.22***	***0.23***	***0.24***	***0.42***	***0.44***	***0.46***	***0.44***	***0.46***	***0.47***	***0.48***	***0.44***	***0.46***
Pakistani	1.44	1.42	1.39	1.33	1.32	1.25	1.28	1.29	1.36	***0.45***	***0.43***	***0.46***	***0.44***	***0.40***	***0.40***	***0.42***	***0.45***	***0.50***	***0.46***	***0.47***	***0.47***	***0.49***	***0.49***	***0.52***	***0.55***	***0.52***	***0.54***
Bangladeshi	0.97	0.99	0.95	0.86	0.80	0.80	0.85	0.84	0.87	***0.35***	***0.35***	***0.40***	***0.42***	***0.35***	***0.35***	***0.41***	***0.44***	***0.46***	***0.38***	***0.40***	***0.49***	***0.54***	***0.64***	0.74	0.79	0.73	0.79
Asian other	***0.55***	***0.56***	***0.55***	***0.55***	***0.56***	***0.54***	***0.62***	***0.61***	***0.59***	***0.30***	***0.30***	***0.35***	***0.33***	***0.33***	***0.32***	***0.31***	***0.33***	***0.31***	***0.59***	***0.61***	***0.55***	***0.60***	***0.64***	***0.65***	***0.64***	***0.59***	***0.63***
Black African	0.85	0.93	1.00	0.94	0.86	0.83	0.86	0.87	0.84	1.05	1.01	0.99	0.90	0.82	0.78	0.75	0.83	0.83	1.09	1.07	1.19	1.20	1.22	1.25	1.23	1.15	1.15
Black Caribbean	1.38	1.33	1.37	1.42	1.36	1.36	1.34	1.34	1.38	***2.32***	***2.34***	***2.34***	***2.30***	***2.21***	***2.21***	***2.14***	***2.27***	***2.29***	1.12	1.21	1.32	1.36	1.41	1.39	1.39	1.35	1.34
Other Black	0.98	1.06	1.13	1.13	1.09	1.07	1.08	1.05	1.03	***1.88***	***1.79***	***1.80***	***1.51***	1.32	1.38	1.31	1.34	1.31	1.38	1.42	***1.60***	***1.56***	1.49	1.50	1.49	1.41	1.34
Chinese	***0.31***	***0.30***	***0.31***	***0.27***	***0.31***	***0.30***	***0.31***	***0.30***	***0.30***	***0.17***	***0.22***	***0.24***	***0.20***	***0.19***	***0.17***	***0.22***	***0.22***	***0.20***	0.95	0.86	0.78	0.76	0.89	0.93	0.91	0.85	0.91
Any other	0.79	0.83	0.90	0.92	0.91	0.89	0.95	0.93	0.91	***0.62***	***0.57***	***0.64***	0.67	***0.62***	***0.60***	***0.60***	***0.62***	***0.61***	***0.53***	***0.52***	***0.60***	***0.63***	0.69	0.71	0.74	0.70	0.73
Unknown	1.10	1.10	0.98	0.97	0.99	0.97	1.04	0.95	0.90	1.41	1.36	1.24	1.15	1.09	1.14	1.18	1.16	1.11	0.97	1.07	1.12	1.26	1.25	1.18	1.16	1.12	1.24

**Note: 'L.B.' ('lower bound') indicates results calculated after discounting SEN types in 2015 for those pupils recorded as having 'School Action' in 2014; was done to account for the fact that in previous years, the School Census requested SEN types for only those pupils with needs identified as 'School Action Plus' or above. 'U.B.' ('upper bound') indicates results calculated using all available SEN type records from the January 2015 School Census.*

MWBA=Mixed White and Black African; MWBC=Mixed White and Black Caribbean; MWAS=Mixed White and Asian (abbreviated for the sake of formatting).

Table 1.14: Adjusted OR time trends by ethnic group for focal SEN types (MLD, SEMH, ASD), 2005-16

Ethnic Group	MLD									BESD									ASD								
	2005	2007	2009	2011	2013	2014	2015 (LB.) *	2015 (UB) *	2016	2005	2007	2009	2011	2013	2014	2015 (LB.) *	2015 (UB) *	2016	2005	2007	2009	2011	2013	2014	2015 (LB.) *	2015 (UB.) *	2016
White Irish	0.86	0.85	0.84	0.74	0.80	0.76	0.87	0.77	0.77	0.91	0.96	0.96	0.88	0.83	0.84	0.80	0.78	0.85	1.24	1.15	0.99	1.09	1.13	1.09	1.14	1.10	1.01
Traveller Irish	***3.75***	***3.68***	***3.64***	***3.59***	***2.81***	***2.32***	***2.48***	***2.67***	***2.45***	***2.20***	***2.18***	***2.02***	***1.86***	***1.67***	***1.53***	1.49	***1.57***	***1.53***	***0.24***	***0.32***	***0.27***	***0.37***	***0.37***	***0.43***	***0.40***	***0.41***	***0.31***
Traveller Roma	***3.84***	***3.71***	***3.55***	***2.76***	***1.98***	***1.85***	***2.25***	***2.27***	***2.70***	***1.97***	***2.15***	***1.84***	1.47	1.07	1.07	1.11	1.16	1.17	***0.47***	***0.40***	***0.42***	***0.37***	***0.31***	***0.30***	***0.28***	***0.28***	***0.33***
White Other	***0.63***	***0.61***	***0.61***	***0.63***	***0.60***	***0.60***	0.71	0.71	0.70	***0.59***	***0.60***	***0.59***	***0.60***	***0.54***	***0.51***	***0.53***	***0.54***	***0.53***	0.93	0.90	0.82	0.69	***0.64***	***0.64***	***0.63***	***0.61***	***0.61***
MWBA	***0.56***	***0.63***	***0.65***	***0.65***	***0.64***	***0.62***	***0.64***	***0.65***	***0.63***	1.09	1.05	0.94	0.95	0.92	0.94	0.96	0.94	0.92	1.10	0.94	0.94	0.96	0.95	0.91	0.95	0.92	0.86
MWBC	0.69	0.77	0.79	0.84	0.86	0.89	0.89	0.88	0.90	1.44	1.46	1.39	1.39	1.38	1.42	1.38	1.37	1.38	1.01	0.90	0.98	1.01	0.98	0.98	0.99	0.97	0.97
MWAS	***0.58***	***0.62***	***0.58***	***0.62***	***0.62***	***0.64***	***0.63***	***0.63***	***0.63***	***0.65***	***0.65***	***0.62***	0.69	0.70	***0.66***	***0.66***	***0.65***	0.67	1.03	1.04	0.97	0.93	0.85	0.85	0.85	0.82	0.81
Mixed other	***0.63***	***0.64***	***0.63***	***0.64***	***0.66***	***0.65***	0.68	***0.64***	***0.63***	1.00	0.99	0.94	0.90	0.88	0.92	0.91	0.88	0.88	1.20	1.19	1.17	1.17	1.13	1.08	1.04	1.00	0.99
Indian	***0.62***	***0.59***	***0.54***	***0.55***	***0.52***	***0.51***	***0.55***	***0.52***	***0.54***	***0.21***	***0.21***	***0.21***	***0.23***	***0.23***	***0.22***	***0.23***	***0.23***	***0.23***	***0.41***	***0.43***	***0.45***	***0.44***	***0.46***	***0.48***	***0.49***	***0.45***	***0.47***
Pakistani	0.94	0.94	0.88	0.91	0.93	0.90	0.95	0.95	1.00	***0.29***	***0.29***	***0.29***	***0.31***	***0.29***	***0.29***	***0.31***	***0.32***	***0.36***	***0.44***	***0.44***	***0.43***	***0.45***	***0.45***	***0.47***	***0.49***	***0.46***	***0.48***
Bangladeshi	***0.48***	***0.50***	***0.45***	***0.46***	***0.44***	***0.46***	***0.52***	***0.51***	***0.52***	***0.17***	***0.18***	***0.19***	***0.22***	***0.19***	***0.20***	***0.24***	***0.26***	***0.26***	***0.36***	***0.36***	***0.44***	***0.47***	***0.55***	***0.62***	0.67	***0.61***	***0.65***
Asian other	***0.44***	***0.46***	***0.45***	***0.47***	***0.48***	***0.47***	***0.54***	***0.53***	***0.50***	***0.23***	***0.25***	***0.28***	***0.28***	***0.29***	***0.28***	***0.28***	***0.28***	***0.27***	***0.57***	***0.59***	***0.53***	***0.58***	***0.62***	***0.63***	***0.62***	***0.57***	***0.60***
Black African	***0.45***	***0.50***	***0.52***	***0.51***	***0.49***	***0.50***	***0.55***	***0.54***	***0.53***	***0.57***	***0.56***	***0.52***	***0.50***	***0.47***	***0.48***	***0.48***	***0.51***	***0.52***	1.06	1.00	1.08	1.06	1.05	1.07	1.06	0.98	0.97
Black Caribbean	0.88	0.86	0.85	0.88	0.84	0.85	0.87	0.87	0.89	1.49	***1.54***	1.46	1.44	1.38	1.39	1.35	1.42	1.43	1.17	1.21	1.27	1.25	1.24	1.21	1.19	1.14	1.12
Other Black	***0.58***	***0.64***	0.67	0.69	0.67	0.67	0.70	0.67	0.67	1.08	1.09	1.08	0.94	0.83	0.88	0.84	0.85	0.84	1.37	1.36	1.49	1.42	1.31	1.29	1.27	1.19	1.13
Chinese	***0.29***	***0.29***	***0.29***	***0.28***	***0.32***	***0.31***	***0.32***	***0.31***	***0.30***	***0.16***	***0.20***	***0.22***	***0.20***	***0.20***	***0.18***	***0.23***	***0.23***	***0.21***	0.99	0.88	0.79	0.78	0.93	0.99	0.97	0.90	0.96
Any other	***0.47***	***0.49***	***0.52***	***0.57***	***0.58***	***0.58***	***0.65***	***0.62***	***0.61***	***0.36***	***0.34***	***0.37***	***0.42***	***0.40***	***0.39***	***0.40***	***0.40***	***0.40***	***0.49***	***0.47***	***0.55***	***0.56***	***0.60***	***0.61***	***0.64***	***0.59***	***0.61***
Unknown	1.03	1.00	0.89	0.87	0.87	0.84	0.91	0.84	0.80	1.25	1.16	1.04	0.97	0.92	0.95	1.01	0.98	0.95	1.00	1.06	1.08	1.20	1.17	1.09	1.07	1.03	1.13

**Note: 'LB.' ('lower bound') and UB. ('Upper bound') defined as in Table 1.13.*

ORs adjusted for: Sex, FSM, IDACI (normalised), birth season, year group

MWBA=Mixed White and Black African; MWBC=Mixed White and Black Caribbean; MWAS=Mixed White and Asian (abbreviated for the sake of formatting).

SEMH/BESD

The prevalence rate for BESD/SEMH increased from 1.9% in 2005 to 2.8% in 2016, partly reflecting the increase since 2015 in the number of pupils for whom data on type of need is requested. Additionally, there was a change of category in 2015 removing BESD and introducing SEMH as the nearest analogous category, although it is not considered a direct replacement. Despite these changes Black Caribbean and Mixed White & Black Caribbean over-representation remains unchanged 2014 to 2016. Perhaps this is not surprising since displaying "challenging, disruptive or disturbing behaviour" remains central to the definition of SEMH/BESD, whatever the putative drivers of such behaviour. The only trend seems to be that Mixed White and Black African pupils have not been over-represented since 2007, the over-representation of Black Other pupils is borderline in the most recent 2013-2016 data, and Gypsy/Roma pupils are no longer over-represented in the adjusted ORs after 2011.

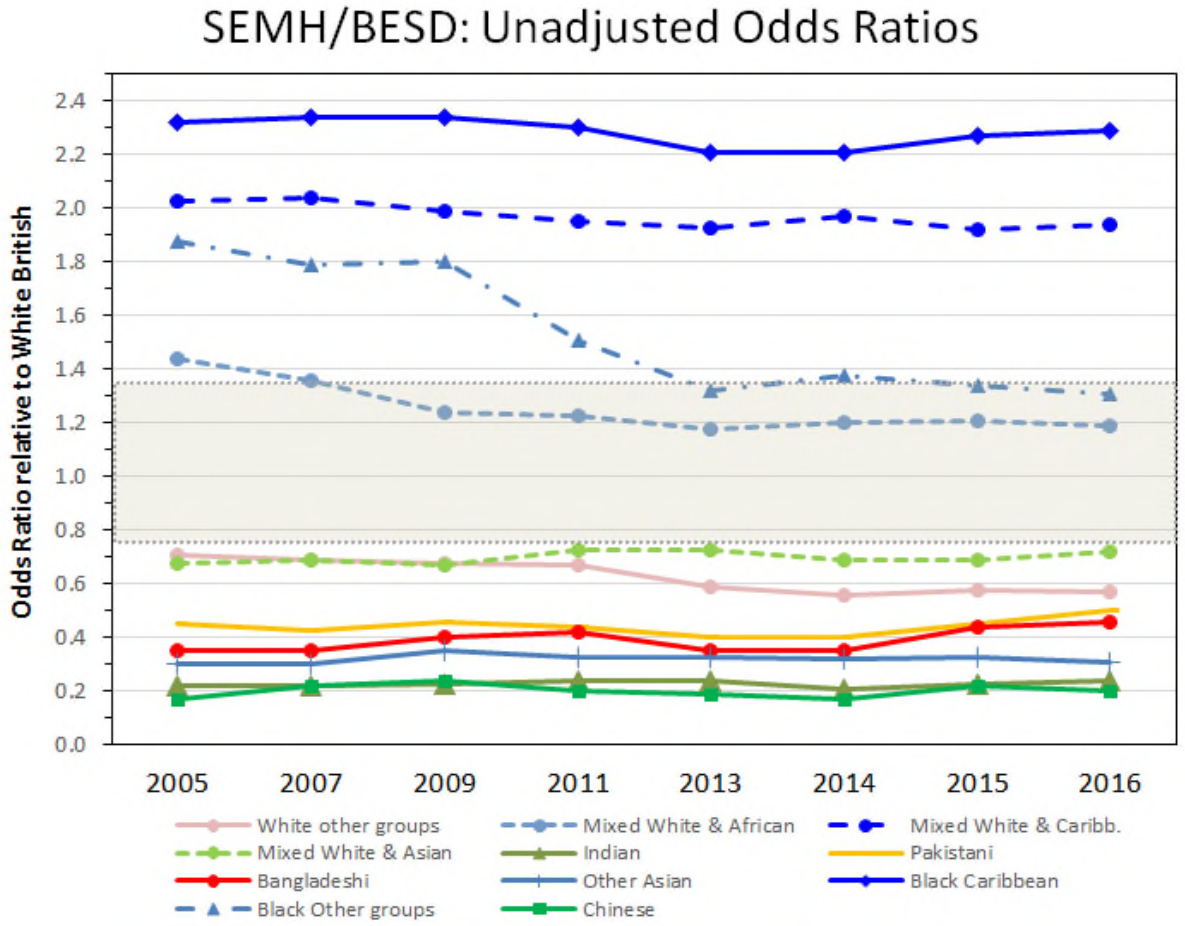


Figure 1.5: Unadjusted ethnic group Odds Ratios for SEMH/BESD 2005-2016

Note: The figure excludes ethnic groups that were never over or under-represented in any year, and also excludes the small traveller groups. The shaded area represents the region OR> 0.75 and OR <1.33 so not substantially different from White British. The 2015 data is based on all pupils with a recorded primary need.

ASD

It is apparent that Black Other pupils have been consistently over-represented, and Indian, Pakistani and Other Asian pupils have been consistently under-represented, relative to White British pupils in all 12 years.

There was more variation in ethnic disproportionality over time for ASD than was the case for the other focal types of SEN. There are three ethnic groups with noticeable changes/trends over time:

- Black Caribbean pupils were not over-represented 2005-2009 but have been consistently over-represented since 2011 (OR= 1.12 in 2005 to OR= 1.34 by 2016).
- White Other groups were not under-represented 2005-2009 but have been consistently under-represented since 2011 (OR= 0.96 in 2005 to OR= 0.60 by 2016).
- The under-representation of Bangladeshi pupils has steadily declined and in 2016 they were no longer under-represented (OR= 0.38 in 2005 to OR= 0.79 by 2016).

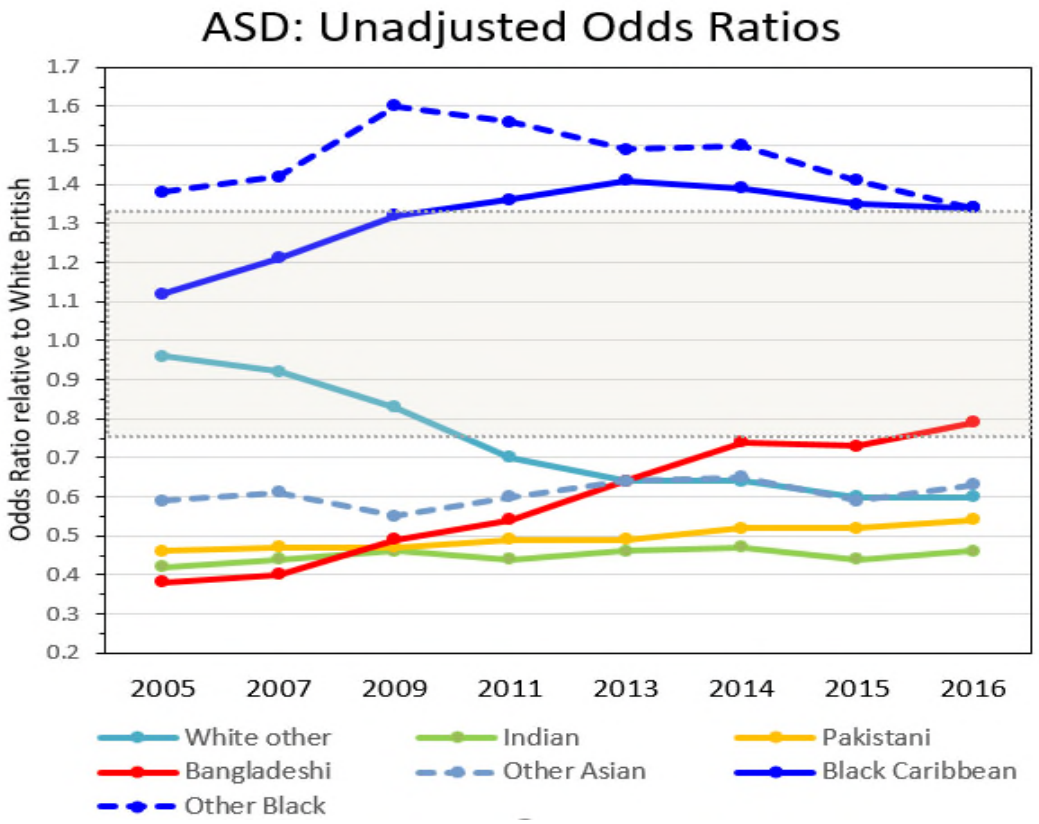


Figure 1.6: Unadjusted ethnic group Odds Ratios for ASD 2005-2016

Note: The figure excludes ethnic groups that were never over or under-represented in any year, and also excludes the small traveller groups. The shaded area represents the region OR> 0.75 and OR <1.33 so not substantially different from White British. The 2015 data is based on all pupils with a recorded primary need.

What lies behind these changes is not clear. Given there has been no change in the over-representation of Black Caribbean pupils recorded as BESD/SEMH, it appears to be unlikely to be ‘symptom substitution’ of ASD for BESD. It is possible that awareness of ASD has grown within the Bangladeshi community reducing their under-representation, but this raises the question as to why there has been no similar change among Indian, Pakistani and Other Asian communities. Further research to address these trends is needed.

There were only two trends in ASD adjusted ORs: White Other pupils have become under-represented in the adjusted ORs as well (OR=0.93 in 2005 to OR= 0.61 in 2016) and Black Other pupils are no longer over-represented after adjusting for pupil background (OR= 1.37 in 2005 to OR=1.13 in 2016).

Part 2: Multi-level results: LA and school effects

The main aim of the analysis reported in this chapter was to investigate whether ethnic disproportionality varied by school and by Local Authority (LA), the extent to which this was the case, and if so whether such variation could be explained by known school and LA characteristics. We focused on maintained mainstreamed schools (unlike in other sections of the report, where we include pupils in all schools - except independent schools - including special schools), and we looked separately at Primary and secondary schools, to ensure that we were making reasonable comparisons.

The first section of the chapter describes some of the school level data from the 2016 School Census, and explains the approach to analysis taken.

The second section reports findings regarding the extent of school and LA variation in identification for our focal SEN types (MLD, SEMH and ASD), as well as findings regarding which school variables were associated with the likelihood of being identified with these three types of SEN, and the nature of these associations. The third section reports findings regarding interactions between school- and pupil-level factors.

Key findings:

- Accounting for the fact that pupils are clustered in schools, and schools are clustered in LAs, made little overall difference to the estimates of disproportionality (relative to the White British majority) for each ethnic group. This was true across both Primary and secondary phases, with a notable exception: for SEMH in the secondary phase, Black Caribbean and Mixed White and Black Caribbean pupils appear far less over-represented after accounting for school and LA clustering, suggesting that there are differences across schools in the extent to which these groups of pupils are over-identified.
- Being in a school with a higher proportion of pupils entitled to FSM is strongly associated with higher odds of MLD identification, as well as higher odds of SEMH identification, across both Primary and secondary phases.
- Being in a school with a higher proportion of Black Caribbean and Mixed White and Black Caribbean pupils was associated with slightly higher odds of SEMH identification, particularly in the secondary phase.
- Being in the smallest schools (in terms of enrolment) was associated with somewhat higher odds of MLD, SEMH and ASD identification, particularly in the Primary phase.
- Being in a Grammar school was, somewhat unsurprisingly, associated with substantially lower odds of identification for all three focal types of primary SEN, likely as a consequence of academic selectivity.
- There were significant interactions between school % FSM and individual pupil FSM for MLD and SEMH; Generally increasing school deprivation raised the risk of MLD/SEMH identification for the non-FSM pupils, while pupils entitled to FSM tended to have the same raised risk regardless of the level of school-wide deprivation.

What we did

Filtering

The exclusion criteria used for the purpose of multilevel modelling were necessarily different from the exclusion criteria used for the single-level multinomial regression analysis presented in the previous section. As before, duplicate records and records with missing IDACI scores were excluded, and we restricted our analysis to only records for pupils in Y1 through to Y11. However, in addition, we also excluded special schools as these specifically cater for pupils with SEN, and therefore have high levels of SEN identification which cannot be compared directly to mainstream schools. We also excluded schools/settings with fewer than 10 pupils on roll according to the school-level census for 2016. This was to avoid having schools with extremely (or implausibly) low enrolment bias calculations of school-level proportions (e.g. FSM eligibility) and inter-school variation for our focal outcomes (as before, MLD, SEMH, and ASD). A further difference from the single-level analysis presented in the previous section was that for multilevel analysis, we separated the data into primary (Years 1 to 6) and secondary (Years 7 to 11) datasets. These were analysed separately to allow for the possibility that some variables might have different distributions and different effects across phases (as found by Strand & Lindsay, 2012). After filtering, the primary dataset contained records for 3,666,196 pupils in 16,730 schools (within 152 LAs), and the secondary dataset contained records for 2,662,921 pupils in 3,353 schools (within 151 LAs[11]).

To summarise, the estimates of ethnic disproportionality presented in Part 1 are the most appropriate for assessing population wide outcomes. However, the results presented in Part 2 are aimed at identifying a slightly different question: What role, if any, do factors in mainstream schools play in the identification of SEN?

Approach to analysis

We used binary logistic regression for each of our three focal types of SEN (MLD, SEMH and ASD) instead of multinomial logistic multi-level regression models for practical reasons, as the latter led to extremely long computation times.[12] We began with 'empty' models (models with only the outcome variable included and no predictor variables) for each focal outcome (MLD, SEMH, ASD). This is standard practice in multilevel modelling, and provides a baseline for the proportions of variance at the school and LA levels.

We then ran models including all pupil predictors as in the single-level models in the previous section. In order to ensure that we were comparing like with like, we ran single level logistic regression models (separately for primary and secondary) with

---

[11] The City of London LA, with a very small number of total pupils, had no records for pupils in Y7-11.

[12] The results from binary and multinomial logistic regression models are slightly different, as for binary models the comparison is between outcomes of 'identified with specific SEN type' and 'No SEN or any other SEN type', while for multinomial models the comparison is between 'identified with specific SEN type' and 'No SEN' only.

binary (0=not identified, 1=identified) outcome variables for MLD, SEMH, and ASD using the exclusion criteria described above so as to compare ORs in the single- and multi-level models with identical predictors, filtering, and outcomes (all pupil-level predictors included FSM eligibility, gender, birth season, Year group within phase, and normalised pupil IDACI score). Where results from single-level models appear different to those in the previous section, differences can be attributed to the fact that we have filtered out special schools, so that results in this section must be interpreted only with regard to pupils in mainstream schools.

Next, we controlled for school-level variables in models for MLD, SEMH, and ASD to investigate how these affected the ethnic group odds ratios and other pupil-level fixed effects, and to assess how much variation existed at the school and LA levels for each focal SEN type. School-level variables included:

- School proportion FSM (an underlying continuous variable coded into quintiles)
- School proportion White British (an underlying continuous variable coded into quintiles)
- School type (a 9-category variable with 'Community School' as the reference category)
- School size (quintiles specific to primary and secondary phases)

Finally, we investigated cross-level interactions to allow for the possibility of differential effects of school composition variables for pupils with different characteristics (e.g. FSM eligibility or ethnic group).

Approach to interpretation

Results of multilevel logistic regression models are somewhat more complicated to interpret than those of multilevel linear regression models. To facilitate the reader's understanding, we present here a brief explanation of the statistics we report and what they mean, as these are related to but somewhat different from the statistics reported for single-level models and those reported for multilevel linear regression models. To assess variation at each level of a given model (beginning with 'empty' models with no predictors included), we report Variance Partition Coefficients (VPC; a percentage of the total variation that lies at a particular level, e.g. between schools or between LAs). To assess the association of particular predictors with MLD/SEMH/ASD identification, we report *Odds Ratios (OR)* as in Part 1. For higher-level (in particular, school-level) predictors, the ORs are average values over all schools, so we also report statistics that provide some information about the distribution across schools of a particular school variable's association with the odds of identification. These include *Median Odds Ratios (MOR),* essentially the median of a distribution of comparisons between schools on an OR scale; *Interval Odds Ratios (IOR)*, a range in which the middle 80% of ORs would fall if computed based on all possible comparisons between pupils with similar predictor values across

schools with different values for a given school variable (e.g. different school types, or different school % FSM quintiles); and *Proportion of Opposed Odds Ratios (POOR)*, a measure of – for each school-level predictor – the proportion of schools for which the association would be in the opposite direction to the overall average association. Simply put, the MOR gives a measure of the scale of overall school and LA heterogeneity, and the IOR and POOR give some sense of the extent to which the association between an aspect of school context/composition and focal SEN type identification varies across schools.

Appendix F provides more detailed information on the calculations and interpretations of these additional statistics that we report for multilevel logistic regression models.

General effects of school and LA clustering: Variance and heterogeneity

For linear multilevel regression models with a continuous, normally-distributed outcome variable, it is usual to calculate the proportion of variance attributable to each level (e.g. pupil, school, LA) and to use the variance components at each level to calculate the variance explained as models are built up to include new predictors.

For logistic multilevel regression models with a binary outcome variable, comparing variance across models with different predictors is more complicated. First, the use of a link function inherent to logistic regression means that there is a rescaling of variances for each model run. When using a logit link, pupil-level variance is always rescaled to the variance of the logistic distribution, $\frac{\pi^2}{3}$. As a consequence, variances at higher levels are also rescaled, so that school and LA variance cannot necessarily be directly compared from one model to the next. This is particularly problematic when entering pupil-level variables, as explaining pupil-level variation leads to a different rescaling of variances at each level. However, higher-level (school or LA) variables cannot explain pupil-level variation, so when such variables are included, rescaling is not an issue. Because of this, and because one of the main aims of multilevel analysis was to investigate the extent to which school characteristics/context may affect the odds of identification with the focal SEN types (MLD, SEMH, ASD), we assess variance explained for the steps at which higher-level predictors are included (with no changes to pupil predictors from preceding models).

What we found

Mainstream and non-mainstream schools

Table 2.1 shows that the proportion of pupils enrolled in non-mainstream schools/settings (special schools, alternative provision and pupil referral units) is quite low overall (1.3%) but, unsurprisingly, much higher for pupils with Statements or EHC plans (44.9%). This proportion varies quite widely across specific SEN types, and is highest for PMLD (79.2%) and SLD (76.2%). Amongst our three focal types of

primary need, ASD has the highest proportion (26.0%) enrolled in non-mainstream settings overall compared to MLD (5.5%) and SEMH (12.2%). However, amongst only those with Statements or EHC Plans, MLD has the highest proportion of pupils (53.0%) enrolled in non-mainstream settings, though only by a small margin, compared to SEMH (49.8%) and ASD (45.2%).

This means that roughly half of the pupils with Statements or EHC plans who are identified with MLD, SEMH, or ASD are excluded from our multilevel analysis, which underscores the importance of careful interpretation of results (i.e. with regard only to mainstream schools and their pupils).

Table 2.1: Proportion of pupils with each SEN type by SEN level in special and mainstream schools

	SEN support				EHC Plan (or Statement)				Total (any identification)			
	Special sch.		*Mainstream sch.*		*Special sch.*		*Mainstream sch.*		*Special sch.*		*Mainstream sch.*	
Primary SEN type	***N***	***%***	***N***	***%***	***N***	***%***	***N***	***%***	***N***	***%***	***N***	***%***
SpLD	290	0.2	134255	99.8	1186	14.4	7066	85.6	1476	1.0	141321	99.0
MLD	516	0.2	234339	99.8	13923	53.0	12350	47.0	14439	5.5	246689	94.5
SLD	122	4.8	2411	95.2	18531	84.4	3431	15.6	18653	76.2	5842	23.8
PMLD	61	11.4	476	88.6	6203	84.2	1166	15.8	6264	79.2	1642	20.8
SEMH	8564	5.6	145350	94.4	13628	49.8	13737	50.2	22192	12.2	159087	87.8
SLCN	181	0.1	149046	99.9	5296	18.9	22784	81.1	5477	3.1	171830	96.9
ASD	230	0.6	36986	99.4	22186	45.2	26847	54.8	22416	26.0	63833	74.0
HI	32	0.2	12837	99.8	1048	20.5	4070	79.5	1080	6.0	16907	94.0
VI	17	0.2	7175	99.8	540	19.2	2279	80.8	557	5.6	9454	94.4
MSI	2	0.1	1433	99.9	168	32.4	351	67.6	170	8.7	1784	91.3
PD	36	0.2	16917	99.8	2727	25.1	8153	74.9	2763	9.9	25070	90.1
Other	285	0.6	45480	99.4	983	21.2	3662	78.8	1268	2.5	49142	97.5
NSA	595	1.8	33059	98.2	85	36.0	151	64.0	680	2.0	33210	98.0
Any type (Total)	10931	1.3	819764	98.7	86504	44.9	106047	55.1	97435	9.5	925811	90.5

Notes: Focal SEN types highlighted in grey.

Appendix G also shows the numbers and percentage of pupils in mainstream and non-mainstream (Special/PRU/AP) school settings by ethnic group for each focal type of SEN.

Descriptive information: School variables

Tables 2.2 (for primary, Y1-6) and 2.3 (for secondary, Y7-11) provide descriptive information about the school variables included in models used to assess the associations of school composition/context with odds of identification and impact on disproportionality estimates for minority ethnic groups. Appendix H provides analogous school descriptive information after additional filtering of the analytic sample (i.e. excluding records for schools with <2 pupils in the combined ethnic group of interest).

Table 2.2: Primary schools (2016, Y1-6) descriptive information

PRIMARY		*N*	*% of schools*	*M*	*SD*	*Min*	*Max*
School type	Foundation	698	4.2	--	--	--	--
	Academy - Converter	2008	12.0	--	--	--	--
	Academy - Sponsored	961	5.7	--	--	--	--
	Church	5197	31.1	--	--	--	--
	Other (Free/CTC/UTC)	126	0.8	--	--	--	--
	Community	7740	46.3	--	--	--	--
School FSM	Highest	3338	20.0	31.3	7.7	22.0	78.1
	Average-High	3351	20.0	17.2	2.5	13.4	21.9
	Average	3330	19.9	10.5	1.5	8.2	13.3
	Low-Average	3325	19.9	6.2	1.1	4.5	8.1
	Lowest	3386	20.2	2.5	1.3	0	4
	OVERALL	16730	100.0	13.5	10.8	0.0	78.1
School % Asian	Highest	3337	19.9	20.7	15.4	8.9	100.0
(excl. Pakistani;	Average-High	3356	20.1	6.1	1.4	4.1	8.8
for MLD model)	Average	3379	20.2	2.9	0.6	2.1	4.0
	Low-Average	3335	19.9	1.4	0.4	0.8	2.0
	Lowest	3323	19.9	0.2	0.3	0	1
	OVERALL	16730	100.0	6.2	10.2	0.0	100.0
School % Black Car.	Highest	3365	20.1	7.9	5.6	3.1	47.6
/Mixed Wh. & Car.	Average-High	3219	19.2	2.1	0.5	1.4	3.0
	Average	3440	20.6	0.9	0.2	0.6	1.3
	Low-Average	1256	7.5	0.4	0.1	0.1	0.5
	Lowest	5450	32.6	0.0	0.0	0.0	0.0
	OVERALL	16730	100.0	2.2	3.9	0.0	47.6
School % Asian	Highest	3352	20.0	30.7	23.3	9.2	100.0
(Indian/Pakistani/	Average-High	3319	19.8	5.3	1.8	2.9	9.1
Bangladeshi/Other)	Average	3343	20.0	1.8	0.5	1.0	2.8
for ASD model)	Low-Average	2071	12.4	0.6	0.2	0.2	0.9
	Lowest	4645	27.8	0.0	0.0	0	0
	OVERALL	16730	100.0	7.6	15.7	0.0	100.0
School size (roll)	Smallest	3315	19.8	71.1	25.2	10	110
	Small-Average	3183	19.0	146.4	19.2	115	170
	Average	3500	20.9	188.1	14.8	175	225
	Average-Large	3266	19.5	281.2	35.3	230	340
	Largest	3467	20.7	440.8	176.7	345	2750
	OVERALL	16730	100.0	227.5	152.9	10	2750

Table 2.3: Secondary schools (2016, Y7-11) descriptive information

SECONDARY		N	% of schools	M	SD	Min	Max
School type	Foundation	262	7.8	--	--	--	--
	Academy - Converter	1253	37.4	--	--	--	--
	Academy - Sponsored	581	17.3	--	--	--	--
	Church	319	9.5	--	--	--	--
	Grammar	163	4.9	--	--	--	--
	Other (Free/CTC/UTC)	191	5.7	--	--	--	--
	Community	584	17.4	--	--	--	--
School FSM	Highest	673	20.1	30.1	7.6	21.7	62.7
	Average-High	668	19.9	17.5	2.2	14.2	21.6
	Average	667	19.9	11.7	1.4	9.5	14.1
	Low-Average	672	20.0	7.5	1.1	5.7	9.4
	Lowest	673	20.1	3.7	1.3	0	6
	OVERALL	3353	100.0	14.1	10.0	0.3	62.7
School % Asian (excl. Pakistani; for MLD model)	Highest	671	20.0	23.8	17.0	10.1	97.2
	Average-High	678	20.2	6.9	1.6	4.6	10.0
	Average	660	19.7	3.3	0.7	2.3	4.5
	Low-Average	652	19.4	1.7	0.3	1.2	2.2
	Lowest	692	20.6	0.7	0.3	0	1
	OVERALL	3353	100.0	7.2	11.5	0.0	97.2
School % Black Car. /Mixed Wh. & Car.	Highest	676	20.2	9.5	6.1	3.8	48.8
	Average-High	643	19.2	2.4	0.6	1.6	3.7
	Average	714	21.3	1.1	0.2	0.8	1.5
	Low-Average	614	18.3	0.5	0.1	0.4	0.7
	Lowest	706	21.1	0.1	0.1	0.0	0.3
	OVERALL	3353	100.0	2.7	4.4	0.0	48.8
School % Asian (Indian/Pakistani/ Bangladeshi/Other) for ASD model)	Highest	667	19.9	37.2	23.1	13.3	98.7
	Average-High	678	20.2	7.8	2.6	4.2	13.2
	Average	649	19.4	2.7	0.8	1.6	4.1
	Low-Average	687	20.5	1.0	0.3	0.6	1.5
	Lowest	672	20.0	0.2	0.2	0	1
	OVERALL	3353	100.0	9.7	17.4	0.0	98.7
School size (roll)	Smallest	674	20.1	374.5	157.4	13	585
	Small-Average	656	19.6	708.7	66.4	590	815
	Average	687	20.5	925.4	64.1	820	1030
	Average-Large	663	19.8	1154.6	71.9	1035	1285
	Largest	673	20.1	1534.5	218.7	1290	2750
	OVERALL	3353	100.0	939.9	414.9	13	2750

MLD in the primary and secondary phases

General contextual (school/LA) effects

School and LA variance and heterogeneity

One of the key reasons for using multilevel analysis was to assess whether there was substantial variation between schools and between LAs in terms of SEN identification, here for MLD.

Table 2.4 shows that for MLD across both school phases, around one-quarter of the variance is at the school level (22.4% for primary, 25.5% for secondary), with a much lower proportions of variance at the LA level (5.4% for primary, 5.7% for secondary). We compared 2- and 3-level models (pupils nested within schools, and then schools additionally nested within LAs); we deemed LA variance to be sufficient to retain the 3-level model so as not to overestimate variation at another level of the model (i.e. schools). High school MOR values (close to 3) indicate substantial heterogeneity across schools in the odds of identification for MLD, while LA MORs are lower but still substantial; these exceed the more-stringent threshold used previously to assess ORs (OR>1.50)[13], suggesting that there is heterogeneity in the odds of MLD identification across LAs.

The values presented in Table 2.4 are used as a baseline for later comparisons.

Table 2.4: MLD empty models: Baseline variance and heterogeneity, primary (Y1-Y6) and secondary (Y7-Y11), 2016

MLD VPCs		*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Primary (Y1-6)	2-level	--	1.210	--	0.269	--	2.86
	3-level	0.248	1.020	0.054	0.224	1.61	2.93
Secondary (Y7-11)	2-level	--	1.478	--	0.310	--	3.19
	3-level	0.271	1.217	0.057	0.255	1.64	3.20

OR comparison: Effects of clustering on disproportionality

A second reason for using multilevel analysis was to assess whether – and to what extent – accounting for school and LA clustering affected the ORs for pupil ethnic groups and other pupil characteristics; that is, whether and how much ethnic disproportionality (and/or the effects of other pupil characteristics) varied across schools and LAs. The coefficients for the explanatory variables in single level model reflects all sources of variability, both pupil and school level. The multilevel model removes all effects of schools so the resulting coefficients represent the average 'within school' effect. The extent to which the ethnic coefficients change between

[13] *For the empty models, VPC is the most straightforward statistic to interpret; however, MORs are reported here as a baseline for comparison with subsequent explanatory models with pupil and school predictors.*

single level and multilevel models gives us an indication of the importance of school factors in accounting for identification.

Table 2.5 presents the ORs for ethnic groups and other pupil predictors, before and after accounting for clustering (single- and multi-level models, respectively), for MLD identification in the primary and secondary phases. We refrain from making inferences about changes in the ORs for ethnic groups that constitute very small proportions of the total population (e.g. White Irish and Traveller groups), as results for these groups are more volatile. Similarly, we refrain from making inferences about categories that lack a substantive group definition (i.e. the 'Unknown', 'Any other', and 'Any other mixed' categories).

Primary (Y1-6)

For MLD in the primary phase, for the most part, ethnic group ORs do not change substantially after accounting for school and LA clustering. There are some increases in the ORs for a few under-represented groups, but these tend to be small increases with little change to substantive interpretation. Most Asian pupils remain substantially under-represented after accounting for clustering, with the exception of the Pakistani group (OR=1.08 in the single-level model, OR=0.90 in the multilevel model). With regard to other pupil variables, changes to ORs after accounting for clustering are small and do not change substantive interpretations.

Secondary (Y7-11)

For MLD in the secondary phase, changes in ethnic group ORs after accounting for clustering are also mostly minor with regard to substantive meaning, except for some very small groups. The Pakistani group OR changes direction after accounting for clustering i.e. this group is slightly but not substantially over-represented in the single-level model (OR= 1.11) and slightly but not substantially under-represented in the multi-level model (OR= 0.88); other Asian groups remain under-represented. The ORs associated with other pupil background characteristics change little after accounting for clustering, none so much as to change substantive interpretations substantially.

On the whole, accounting for clustering had minimal impact on disproportionality estimates for MLD, suggesting little variation in disproportionality across schools (though we note this is distinct from variation in the overall likelihood of being identified).

Table 2.5: MLD primary (Y1-6) and secondary (Y7-11) 2016 OR comparisons in adjusted single- and multi-level models

MLD		Primary					Secondary				
		Single-level		*Multi-level*			*Single-level*		*Multi-level*		
		Exp(B)		*Exp(B)*			*Exp(B)*		*Exp(B)*		
Pupil ethnic group	White Irish	0.84	*	1.02			0.69	*	0.87	*	
	Traveller Irish	***2.36***	*	***2.77***	*		***2.03***	*	***2.32***	*	
	Traveller Gypsy/Roma	***2.58***	*	***2.61***	*		***2.63***	*	***2.04***	*	
	White other groups	0.73	*	0.81	*		0.77	*	0.79	*	
	Mixed White & African	0.68	*	0.74	*		***0.60***	*	0.69	*	
	Mixed White & Caribbean	0.91	*	0.92	*		0.88	*	0.90	*	
	Mixed White & Asian	0.67	*	0.67	*		0.67	*	0.69	*	
	Any other mixed	0.67	*	0.75	*		***0.62***	*	0.71	*	
	Indian	***0.57***	*	***0.51***	*		***0.62***	*	***0.57***	*	
	Pakistani	1.08	*	0.90	*		1.11	*	0.88	*	
	Bangladeshi	***0.55***	*	***0.66***	*		***0.61***	*	***0.62***	*	
	Any other Asian	***0.53***	*	***0.54***	*		***0.53***	*	***0.57***	*	
	Black African	***0.55***	*	***0.62***	*		***0.60***	*	*0.67*	*	
	Black Caribbean	0.87	*	0.96			0.90	*	0.96		
	Black other groups	0.72	*	0.80	*		***0.66***	*	0.73	*	
	Chinese	***0.32***	*	***0.35***	*		***0.31***	*	***0.39***	*	
	Any other	***0.65***	*	0.70	*		0.69	*	0.74	*	
	Unknown	0.84	*	0.94			0.74	*	0.87	*	
Pupil FSM	Eligible	***2.06***	*	***2.04***	*		***2.01***	*	***1.93***	*	
Pupil gender	Boy	***1.57***	*	***1.61***	*		*1.43*	*	*1.46*	*	
Birth season	Summer	***1.99***	*	***2.06***	*		***1.51***	*	***1.53***	*	
	Spring	1.41	*	1.43	*		1.25	*	1.25	*	
Pupil year group	Primary: Y6	***2.62***	*	***2.82***	*	Y11	0.77	*	0.75	*	
	Y5	***2.43***	*	***2.61***	*	Y10	0.79	*	0.78	*	
	Y4	***2.24***	*	***2.39***	*	Y9	0.86	*	0.86	*	
	Y3	***1.98***	*	***2.09***	*	Y8	0.95	*	0.95	*	
	Y2	***1.64***	*	***1.67***	*						
Pupil IDACI	(Normalised, 2SD)	***1.74***	*	***1.54***	*		***1.99***	*	***1.67***	*	
	Combined deprivation (FSM+IDACI)	***2.72***	*	***2.53***	*		***2.84***	*	***2.50***	*	

*=significant at the level p<0.05;highlighting = ***OR<0.67***; OR<0.75; OR>1.33; ***OR>1.50***

Associations between school characteristics and MLD identification

Table 2.6 gives the OR (and IOR and POOR values) for each of the school-level variables included in the MLD models for both primary and secondary school phases, and also provides the school and LA variances, residual VPCs and MORs for each phase. Table 2.7 facilitates discussion of changes to variance and heterogeneity measures by comparing the variance, VPC and MOR to empty models, models with only pupil-level predictors, and models with pupil and school-level predictors for each phase.

- For the primary phase, the proportion of pupils in a school who were entitled to FSM was quite strongly associated with the odds of MLD identification (OR= 1.61, 1.48, 1.28 and 1.16 from highest to second-lowest quintile, with the lowest quintile as the reference category). That is, pupils in schools with higher proportions of FSM entitlement had, on average, higher odds of identification. Being in a smaller school was also associated with higher odds of identification (OR= 1.48 for the smallest quintile of schools). Pupils in 'Other' (e.g. Free) schools had somewhat lower odds of identification (OR= 0.70); however, this school type makes up a very small proportion of schools overall, so we avoid drawing strong inferences from this result. There was no significant or substantial association between the school percentage of pupils in Asian groups (defined here as Asian except for Pakistani, based on patterns of MLD under-representation noted in Part 1 of this report) and MLD identification for primary pupils.

- For the secondary phase, there was also a strong association between high school proportions of FSM entitlement and the odds of MLD identification (OR= 1.76, 1.48, 1.35, and 1.12, from highest to second-lowest quintile, respectively). There was some association between being in a smaller school and the odds of MLD identification, although this was less pronounced than in primary schools (for secondary, OR= 1.24 for the smallest quintile). Pupils in Grammar schools had incredibly low odds of being identified (OR= 0.05), which is not surprising given the definition of MLD and the academic selectivity of Grammar schools. Schools with the highest proportions of Asian pupils (defined as noted above) were associated with slightly lower odds of MLD identification (OR=0.82 and 0.80 for the highest two quintiles), but this effect was less substantial than the others noted above.

IORs and POORs across both phases show large amounts of heterogeneity across schools for most school variables, though less so for the strongest associations which had lower POOR values and smaller IOR ranges (specifically, school percent FSM and school size across both primary and secondary phases, and particularly Grammar school in the secondary phase).

Table 2.6: MLD primary (Y1-6) and secondary (Y7-11): Specific school context/composition variables

MLD School contextual effects		PRIMARY				SECONDARY			
		Exp(B)		*IOR*	*POOR*	*Exp(B)*		*IOR*	*POOR*
School type	Foundation	1.08		(0.16,7.43)	0.480	1.07		(0.16,7.02)	0.482
	Academy - Converter	0.91	*	(0.13,6.29)	0.476	0.92		(0.14,6.07)	0.479
	Academy - Sponsored	0.85	*	(0.12,5.87)	0.458	0.91		(0.14,6.01)	0.476
	Church	0.94	*	(0.14,6.44)	0.482	0.89		(0.13,5.82)	0.467
	Grammar	--		--	--	***0.05***		(0.01,0.36)	0.024
	Other (Free/CTC/UTC)	0.70	*	(0.1,4.79)	0.405	0.84		(0.13,5.55)	0.454
School FSM	Highest	***1.61***	*	(0.23,11.09)	0.376	***1.76***	*	(0.27,11.54)	0.351
	Average-High	1.48	*	(0.21,10.17)	0.398	1.48	*	(0.23,9.73)	0.395
	Average	1.28	*	(0.19,8.79)	0.436	1.35	*	(0.21,8.87)	0.419
	Low-Average	1.16	*	(0.17,8.02)	0.460	1.12		(0.17,7.38)	0.468
School ethnic group %	Highest	0.99		(0.14,6.84)	0.498	0.82	*	(0.12,5.37)	0.446
	Average-High	0.98		(0.14,6.73)	0.494	0.80	*	(0.12,5.29)	0.441
	Average	0.97		(0.14,6.68)	0.492	0.89		(0.14,5.84)	0.468
	Low-Average	1.01		(0.15,6.94)	0.498	0.93		(0.14,6.14)	0.482
School size (roll)	Smallest	1.48	*	(0.21,10.16)	0.398	1.24	*	(0.19,8.15)	0.442
	Small-Average	1.17	*	(0.17,8.07)	0.458	1.12	*	(0.17,7.37)	0.469
	Average	1.12	*	(0.16,7.74)	0.469	1.12	*	(0.17,7.33)	0.47
	Average-Large	1.10	*	(0.16,7.6)	0.474	1.04		(0.16,6.82)	0.49
Variance/heterogeneity	LA Variance	0.205	*			0.219	*		
	LA (residual) VPC	0.046				0.050			
	LA MOR	1.54				1.56			
	School Variance	0.928	*			0.859	*		
	School (residual) VPC	0.210				0.197			
	School MOR	2.76				2.69			

Notes: prop.=proportion; VPC=Variance Partition Coefficient; MOR=Median Odds Ratio; * indicates significance at p<0.05
Pupil level variables (not including EAL) are also controlled for in these models but coefficients for these are not reported here.

Combined ethnic group of interest for school composition: Asian (excluding Pakistani) groups.

Table 2.7: MLD primary and secondary 2016: Sequential model step variance/heterogeneity

MLD Primary (Y1-6)	*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Empty	0.248	1.020	0.05	0.220	1.61	2.93
Pupil pred.	0.218	0.961	0.05	0.220	1.56	2.82
Pupil and school pred.	0.205	0.928	0.05	0.210	1.54	2.76
MLD Secondary (Y7-11)	*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Empty	0.271	1.217	0.057	0.255	1.64	3.20
Pupil pred.	0.240	1.068	0.052	0.232	1.59	2.98
Pupil and school pred.	0.219	0.859	0.050	0.197	1.56	2.69

Note: 'var.'=variance; 'pred.'=predictors; 'Empty' refers to a model with no predictors.

Table 2.7 shows that including these school variables explained roughly 6.0% of the LA variance and 3.4% of the school variance in the primary phase, roughly 8.8% of the LA variance and 19.6% of the school variance in the secondary phase (compared to the corresponding models with only pupil-level predictors). LA MORs in both primary and secondary phases were only slightly reduced compared to the models with only pupil-level predictors, as was the school MOR for the primary phase, but the secondary school MOR was more substantially reduced. Together, these results indicate that our school composition variables explain more of the school variation in MLD identification for secondary schools than they do for primary schools.

As a robustness check, Appendix I includes results from models run on a filtered sample that excludes any schools with fewer than two pupils in the combined group of interest for ethnic composition for MLD (Asian, excluding Pakistani). This made a negligible difference to ethnic group ORs and to general and specific school effects in both primary and secondary phases.

SEMH in the primary and secondary phases

General contextual (school/LA) effects

School and LA variance and heterogeneity

Table 2.8 shows that for SEMH across both school phases, the highest proportions of variance is at the school level (13.1% for primary, 15.5% for secondary), much lower proportions of variance exists across LAs SEMH (1.4% for primary, 2.0% for secondary). School MOR values (above 2 for primary and secondary) indicate substantial heterogeneity across schools in the odds of identification, while both primary and secondary MOR values at the LA level for SEMH do not exceed even the less-stringent threshold (OR>1.33) used in previous sections to assess ORs, indicating that there is relatively little heterogeneity in the odds of identification across LAs even before including any predictor variables in the model. For this focal type of SEN, then, the analysis proceeded with only 2-level models given the extremely small amount of LA variance and heterogeneity in SEMH identification.

Table 2.8: SEMH empty models: Baseline variance and heterogeneity, primary and secondary, 2016

SEMH VPCs		*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Primary (Y1-6)	2-level	--	0.558	--	0.145	--	2.04
	3-level	0.055	0.506	0.014	0.131	1.25	2.04
Secondary (Y7-11)	2-level	--	0.698	--	0.175	--	2.22
	3-level	0.081	0.618	0.020	0.155	1.31	2.22

OR comparison: Effects of clustering on disproportionality

Table 2.9 includes ORs from both single level and multilevel models for comparison, for both primary and secondary phases.

Primary (Y1-6)

For SEMH in the primary phase, most ethnic group ORs change only minimally after accounting for clustering, and the substantive meanings of the ethnic group ORs remain largely unchanged. There are no substantial changes in the ORs for other pupil variables.

Secondary (Y7-11)

For the most part, ethnic group ORs in the secondary phase also change negligibly after accounting for school clustering, with one exception. There were substantial reductions in the ORs for Mixed White and Black Caribbean and Black Caribbean pupils (both OR=1.47 in the single-level model, and OR=1.29 and 1.14 respectively in the multilevel model respectively). To check whether this finding was a consequence of many schools having no pupils from the relevant ethnic groups, we re-ran the same models filtering out schools with no Black Caribbean or Mixed White and Black Caribbean pupils, but the results did not change (see Appendix I). These results, seem to suggest that the over-representation of these ethnic groups for

SEMH apparent in single-level analyses is a phenomenon that varies in strength between schools.

Table 2.9: SEMH primary (Y1-6) and secondary (Y7-11) 2016 OR comparisons in adjusted single- and multi-level models

SEMH		Primary					Secondary				
		Single-level		*Multi-level*			*Single-level*		*Multi-level*		
		Exp(B)		*Exp(B)*			*Exp(B)*		*Exp(B)*		
Pupil ethnic group	White Irish	0.84	*	0.83	*		0.84	*	0.83	*	
	Traveller Irish	1.03		0.94			1.33	*	1.28		
	Traveller Gypsy/Roma	0.80	*	0.70	*		1.16	*	0.81	*	
	White other groups	***0.58***	*	***0.52***	*		***0.57***	*	***0.44***	*	
	Mixed White & African	0.98		0.93	*		1.02		0.91	*	
	Mixed White & Caribbean	1.35	*	1.34	*		1.47	*	1.29	*	
	Mixed White & Asian	0.69	*	0.67	*		0.80	*	0.72	*	
	Any other mixed	0.93	*	0.89	*		0.90	*	0.77	*	
	Indian	***0.27***	*	***0.24***	*		***0.29***	*	***0.23***	*	
	Pakistani	***0.38***	*	***0.33***	*		***0.43***	*	***0.33***	*	
	Bangladeshi	***0.31***	*	***0.24***	*		***0.32***	*	***0.23***	*	
	Any other Asian	***0.34***	*	***0.29***	*		***0.28***	*	***0.22***	*	
	Black African	***0.66***	*	***0.59***	*		***0.56***	*	***0.44***	*	
	Black Caribbean	1.46	*	1.40	*		1.47	*	1.14	*	
	Black other groups	0.91	*	0.84	*		0.91	*	0.75	*	
	Chinese	***0.25***	*	***0.23***	*		***0.24***	*	***0.24***	*	
	Any other	***0.46***	*	***0.40***	*		***0.45***	*	***0.32***	*	
	Unknown	0.93		0.90	*		0.92	*	0.89	*	
Pupil FSM	Eligible	***2.47***	*	***2.40***	*		***2.53***	*	***2.46***	*	
Pupil gender	Boy	***3.14***	*	***3.19***	*		***2.23***	*	***2.29***	*	
Birth season	Summer	1.10	*	1.10	*		1.08	*	1.07	*	
	Spring	1.05	*	1.05	*		1.04	*	1.04	*	
Pupil year group	Primary: Y6	***1.75***	*	***1.80***	*	Y11	1.04	*	1.02		
	Y5	***1.68***	*	***1.73***	*	Y10	1.03	*	1.03		
	Y4	***1.57***	*	***1.60***	*	Y9	1.01		1.01		
	Y3	1.40	*	1.43	*	Y8	0.97	*	0.97		
	Y2	1.26	*	1.27	*						
Pupil IDACI	(Normalised, 2SD)	***1.61***	*	1.47	*		***1.83***	*	***1.56***	*	
	Combined deprivation (FSM+IDACI)	***3.13***	*	***2.91***	*		***3.43***	*	***3.08***	*	

*=significant at the level p<0.05;highlighting = ***OR<0.67***; OR<0.75; OR>1.33; ***OR>1.50***

Associations between school characteristics and SEMH Identification

Table 2.10 gives the OR (plus IOR and POOR values) for each of the school-level variables included in the SEMH models for both primary and secondary school phases. It also provides the school variances, residual VPCs and MORs for each school phase. Table 2.11 facilitates discussion of changes to variance and heterogeneity measures by comparing the variance, VPC, and MOR values for

schools across empty models, models with only pupil-level predictors and models with pupil and school predictors.

- For primary schools, there were again significant associations with school characteristics in the primary phase. As was the case for MLD, the odds of identification were substantially higher for schools with higher proportions of pupils entitled to FSM (OR= 1.54, 1.49, 1.32 and 1.12 from highest to second-lowest quintile), and being in a smaller school was somewhat associated with higher odds of SEMH identification (OR= 1.28 for the smallest quintile). School type and the proportion of pupils in a school who were Black Caribbean or Mixed White & Black Caribbean did not have substantial ORs.

- For secondary schools, a higher school proportion of pupils entitled to FSM was again associated with higher odds of SEMH identification (OR= 1.56, 1.37, 1.32, and 1.10, from highest to second-lowest quintile). Being in a Grammar school was again associated with substantially lower odds of SEMH identification (OR= 0.42), and being in the smallest quintile of schools in terms of enrolment was associated with slightly higher odds of identification (OR= 1.20). Unlike in the primary phase, being in a school with a higher proportion of Black Caribbean and Mixed White & Black Caribbean pupils was somewhat associated with higher odds of SEMH identification in the secondary phase (OR= 1.28, 1.19, and 1.19 for the highest three quintiles); however, including this composition variable did not substantially change the individual ethnic group ORs for these groups of pupils.

Table 2.11 shows that including these school variables explained 4.9% of the school variance in the primary model, compared to the corresponding model with only pupil characteristics. In the secondary model, the same variables explained 12.2% of the school variance. School MORs in both primary and secondary models are only slightly reduced after adding school variables. IORs and POORs again indicate substantial heterogeneity across schools in the specific effects of the above school composition and contextual factors, although slightly less so for the substantial and significant effects noted above. Here again, the effects of school composition /context appear to be greater for the secondary phase for SEMH, but perhaps to less of an extent than was true for MLD.

Appendix I includes results from the models including school variables, run on a filtered sample that excludes any schools with fewer than two pupils in the combined group of interest for ethnic composition for SEMH (Black Caribbean and Mixed White & Black Caribbean). This was done as a robustness check for the composition effect. This alternative filtering made a negligible difference to ethnic group ORs and very minor differences to general and specific school effects in both primary and secondary phases, except for a noticeable reduction in the effect of being in the smallest quintile of schools.

Table 2.10: SEMH primary (Y1-6) and secondary (Y7-11): Specific school context/composition variables

SEMH School contextual effects		PRIMARY				SECONDARY			
		Exp(B)		*IOR*	*POOR*	*Exp(B)*		*IOR*	*POOR*
School type	Foundation	0.98		(0.28,3.38)	0.492	0.96		(0.26,3.52)	0.485
	Academy - Converter	0.90	*	(0.26,3.10)	0.457	0.91	*	(0.25,3.34)	0.464
	Academy - Sponsored	0.98		(0.29,3.39)	0.493	0.91	*	(0.25,3.31)	0.461
	Church	0.93	*	(0.27,3.22)	0.471	0.84	*	(0.23,3.09)	0.434
	Grammar	--		--	--	***0.42***	*	(0.12,1.55)	0.198
	Other (Free/CTC/UTC)	0.97		(0.28,3.35)	0.488	1.04		(0.29,3.82)	0.483
School FSM	Highest	***1.54***	*	(0.45,5.30)	0.328	***1.56***	*	(0.43,5.70)	0.331
	Average-High	1.49	*	(0.43,5.15)	0.339	1.37	*	(0.37,5.00)	0.379
	Average	1.32	*	(0.38,4.54)	0.388	1.32	*	(0.36,4.85)	0.391
	Low-Average	1.12	*	(0.33,3.86)	0.453	1.10	*	(0.30,4.02)	0.463
School ethnic group %	Highest	1.07	*	(0.31,3.67)	0.474	1.28	*	(0.35,4.67)	0.405
	Average-High	1.07	*	(0.31,3.67)	0.474	1.19	*	(0.32,4.35)	0.433
	Average	1.06	*	(0.31,3.64)	0.478	1.19	*	(0.33,4.36)	0.432
	Low-Average	0.97		(0.28,3.33)	0.486	1.04		(0.28,3.81)	0.484
School size (roll)	Smallest	1.28	*	(0.37,4.42)	0.398	1.20	*	(0.33,4.40)	0.428
	Small-Average	1.13	*	(0.33,3.88)	0.451	0.97		(0.27,3.56)	0.489
	Average	1.02		(0.29,3.50)	0.493	1.02		(0.28,3.74)	0.492
	Average-Large	1.04		(0.30,3.57)	0.486	1.01		(0.27,3.68)	0.498
Variance/heterogeneity	School Variance	0.466				0.513			
	School (residual) VPC	0.124				0.135			
	School MOR	1.92				1.98			

*Notes: prop.= proportion; VPC=Variance Partition Coefficient; MOR=Median Odds Ratio; * indicates significance at p<0.05. Pupil level variables are also controlled for in these models but coefficients for these are not reported here. Combined ethnic group of interest for school composition: Black Caribbean and Mixed White & Black Caribbean.*

Table 2.11: SEMH primary and secondary: Sequential model step variance/heterogeneity

SEMH primary (Y1-6)	*School var.*	*School VPC*	*School MOR*
Empty	0.558	0.145	2.04
Pupil pred.	0.490	0.130	1.95
Pupil and school pred.	0.466	0.124	1.92
SEMH Secondary (Y7-11)	*School var.*	*School VPC*	*School MOR*
Empty	0.698	0.175	2.22
Pupil pred.	0.584	0.151	2.07
Pupil and school pred.	0.513	0.135	1.98

Note: 'var.'=variance; 'pred.'=predictors; 'Empty' refers to a model with no predictors. LA level excluded as LA VPCs and MOR indicate minimal variation between LAs.

ASD in the primary and secondary phases

General contextual (school/LA) effects

School and LA variance and heterogeneity

Table 2.12 shows that ASD has a lower proportion of variance at the school level than MLD or SEMH, but this variation is still non-negligible (11.3% for primary, 12.2% for secondary). At the LA level, as for the other two types of SEN, proportions of variance at the LA level were lower (4.6% for primary, 3.7% for secondary). School MOR values (above 2 for primary and secondary phases) indicate substantial heterogeneity across schools in the odds of identification as was true for the other two focal SEN types, and the LA-level MOR values are somewhat borderline according to the same thresholds previously used to assess OR size, indicating that there is some heterogeneity in the odds of ASD identification across LAs before including any predictor variables in the model (this supports the decision to proceed with 3-level models).

Table 2.12: ASD empty models: Baseline variance and heterogeneity, primary and secondary, 2016

ASD VPCs		*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Primary (Y1-6)	2-level[14]	--		--		--	
	3-level	0.178	0.442	0.046	0.113	1.50	2.12
Secondary (Y7-11)	2-level	--	0.598	--	0.154	--	2.09
	3-level	0.144	0.476	0.037	0.122	1.44	2.12

OR comparison: Effects of clustering on disproportionality

Table 2.13 provides ORs for comparison across single- and multi-level models with the same sets of pupil-level predictors, for both primary and secondary phases.

Primary (Y1-6)

For ASD in the primary phase, ethnic group ORs mostly do not change in substantive meaning, although the magnitudes of some of the differences in ORs between single- and multi-level models are somewhat larger than those for MLD and SEMH identification. Most ethnic group ORs decreased after accounting for clustering, but not to the extent that this changed substantive interpretations (e.g. Black groups went from being very slightly over-represented to no longer being so after accounting for clustering, and some under-represented groups appeared slightly more under-represented after accounting for clustering). Other pupil variable ORs changed only negligibly between the single- and multi-level models for the primary phase.

[14] *The 2-level model for the Primary phase had convergence issues in SPSS; because we have encountered this as a software-specific issue in other use of this software, and because the 3-level models showed sufficient variance at the LA level to proceed with a 3-level model, we did not pursue the 2-level model here.*

Secondary (Y7-11)

For ASD in the secondary phase, accounting for clustering also made little difference to most of the ethnic group ORs. Under-represented groups remained under-represented, and although some groups (e.g. Black Caribbean and White Other) had ORs that decreased enough to cross the less stringent OR<0.75 threshold, the absolute changes were small (e.g. OR=0.76 to 0.71 from single- to multi-level models for the White Other group). The effects of other pupil variables (in terms of their ORs) in the secondary phase vary negligibly before and after accounting for clustering. In short, accounting for clustering made little difference to substantive interpretations with regard to disproportionality in ASD identification.

Table 2.13: ASD primary and secondary OR comparisons in adjusted single- and multi-level models

ASD		Primary					Secondary				
		Single-level Exp(B)		*Multi-level Exp(B)*			*Single-level Exp(B)*		*Multi-level Exp(B)*		
Ethnic group	White Irish	1.09		0.98			1.03		0.96		
	Traveller Irish	***0.26***	*	***0.25***	*		***0.22***	*	***0.21***	*	
	Traveller Gypsy/Roma	***0.25***	*	***0.24***	*		***0.14***	*	***0.13***	*	
	White other groups	0.74	*	***0.66***	*		***0.47***	*	***0.45***	*	
	Mixed White & African	0.88		0.78	*		0.72	*	***0.65***	*	
	Mixed White & Caribbean	1.03		0.89	*		0.98		0.85	*	
	Mixed White & Asian	0.93		0.88	*		0.76	*	0.71	*	
	Any other mixed	1.06		0.93			0.90	*	0.82	*	
	Indian	***0.62***	*	***0.61***	*		***0.27***	*	***0.27***	*	
	Pakistani	***0.52***	*	***0.54***	*		***0.30***	*	***0.28***	*	
	Bangladeshi	0.87	*	0.75	*		***0.32***	*	***0.30***	*	
	Any other Asian	0.71	*	***0.64***	*		***0.31***	*	***0.27***	*	
	Black African	1.18	*	0.97			***0.49***	*	***0.43***	*	
	Black Caribbean	1.15	*	0.90	*		0.86	*	0.70	*	
	Black other groups	1.17	*	0.91			***0.62***	*	***0.53***	*	
	Chinese	1.18	*	1.03			***0.60***	*	***0.56***	*	
	Any other	0.75	*	***0.65***	*		***0.36***	*	***0.34***	*	
	Unknown	1.17	*	1.06			0.92		0.85	*	
Pupil FSM	Eligible	***1.66***	*	***1.65***	*		***1.76***	*	***1.75***	*	
Pupil gender	Boy	***4.94***	*	***4.95***	*		***4.65***	*	***4.69***	*	
Birth season	Summer	0.96	*	0.96	*		1.07	*	1.06	*	
	Spring	0.99		0.99			1.01		1.01		
Pupil year group	Primary: Y6	1.36	*	1.38	*	Y11	0.96	*	0.95	*	
	Y5	1.29	*	1.30	*	Y10	0.95	*	0.94	*	
	Y4	1.20	*	1.21	*	Y9	0.99		0.98		
	Y3	1.10	*	1.11	*	Y8	0.98		0.98		
	Y2	1.04		1.04							
Pupil IDACI	(Normalised, 2SD)	1.06	*	1.00			1.04	*	1.03		
	Combined deprivation (FSM+IDACI)	***1.71***	*	***1.65***	*		***1.79***	*	***1.77***	*	

*=significant at the level p<0.05;highlighting = ***OR<0.67***; OR<0.75; OR>1.33; ***OR>1.50***

Associations between school characteristics and ASD identification

Table 2.14 gives the OR (plus IOR and POOR values) for each of the school-level variables included in the ASD models for both school phases. These also provide the school and LA variances, residual VPCs and MORs for each phase. Table 2.15 facilitates discussion of changes to variance and heterogeneity measures by comparing the variance, VPC, and MOR values for LAs and schools across empty models, models with only pupil-level predictors and models with pupil and school predictors.

Overall, there were fewer significant associations with school variables for ASD than for the other two focal SEN types based on the overall ORs.

- For primary schools, the proportion of pupils entitled to FSM in a school was significantly but not linearly associated with the odds of ASD identification (OR=1.17, 1.25, 1.23, and 1.11, from highest to second-lowest quintiles); there was no substantial association between ASD identification and the proportion of pupils in a school who were Asian (defined here as Indian, Bangladeshi, Pakistani, or other Asian, according to patterns of ASD under-representation). As was true for MLD and SEMH, being in the smallest quintile of schools was associated with higher odds of identification for ASD in the primary phase (OR=1.39).

- For secondary schools, there was also no substantial association between the school proportion of pupils entitled to FSM and individual odds of ASD identification, nor any substantial association between the school percent of pupils who were Asian (as defined above) and individual odds of ASD identification. Pupils in Grammar schools had lower odds of identification (OR=0.61), and those in the smallest schools had slightly higher odds of identification (OR=1.19 for the smallest quintile).

The IORs for all of these variables included 1, and the POORs were all quite close to 0.5 except for Grammar schools in secondary (POOR=0.312) and the smallest quintile of schools in primary (POOR=0.383). To an even greater extent than was true in relation to the other two focal SEN types, despite the associations apparent in the overall average ORs for some school variables as noted above, the effects (and directions of the effects) of these school variables are actually very heterogeneous across different schools.

Additionally, including these variables in the primary and secondary models for ASD identification did not explain LA-level variance, and the LA MORs changed negligibly as compared to previous empty and pupil-variable-only models for ASD. Including school variables explained only 3.5% of the school variance in the primary model and 4.8% of the school variance in the secondary model as compared to the models with only pupil predictors; correspondingly, the school MOR was negligibly reduced in the primary model (from MOR=2.12 to 2.10) and in the secondary model (MOR=2.01 to 1.99) as compared to the corresponding pupil-variable-only models. Appendix I includes robustness checks for ethnic composition (%Asian) as described before, which made a negligible difference to ethnic group ORs. Together, these results indicate that the school variables discussed above do little to explain the school and LA general contextual effects, which were quite small even before including pupil or school predictors.

Table 2.14: ASD primary (Y1-6) and secondary (Y7-11): Specific school context/composition variables

ASD School contextual effects		PRIMARY				SECONDARY			
		Exp(B)		*IOR*	*POOR*	*Exp(B)*		*IOR*	*POOR*
School type	Foundation	0.97		(0.24,3.99)	0.490	1.06		(0.29,3.90)	0.478
	Academy - Converter	0.93	*	(0.23,3.83)	0.475	0.97		(0.26,3.56)	0.486
	Academy - Sponsored	0.97		(0.24,3.97)	0.487	0.93		(0.25,3.44)	0.473
	Church	0.89	*	(0.22,3.66)	0.459	0.97		(0.26,3.56)	0.487
	Grammar	--		--	--	***0.61***	*	(0.16,2.24)	0.312
	Other (Free/CTC/UTC)	0.92		(0.22,3.78)	0.47	1.21	*	(0.33,4.47)	0.425
School FSM	Highest	1.17	*	(0.29,4.82)	0.442	0.96		(0.26,3.52)	0.482
	Average-High	1.25	*	(0.30,5.12)	0.421	1.02		(0.28,3.76)	0.492
	Average	1.23	*	(0.30,5.04)	0.426	1.11	*	(0.30,4.10)	0.458
	Low-Average	1.11	*	(0.27,4.55)	0.463	1.08	*	(0.29,3.99)	0.469
School ethnic group %	Highest	1.01		(0.25,4.16)	0.495	1.08		(0.29,3.97)	0.471
	Average-High	1.11	*	(0.27,4.56)	0.462	1.11	*	(0.30,4.10)	0.459
	Average	1.05		(0.26,4.30)	0.483	1.02		(0.28,3.74)	0.494
	Low-Average	1.07	*	(0.26,4.38)	0.477	1.02		(0.28,3.77)	0.491
School size (roll)	Smallest	1.39	*	(0.34,5.70)	0.383	1.19	*	(0.32,4.38)	0.432
	Small-Average	1.08	*	(0.26,4.43)	0.473	1.12	*	(0.30,4.13)	0.455
	Average	1.11	*	(0.27,4.57)	0.462	1.05		(0.29,3.88)	0.480
	Average-Large	0.99		(0.24,4.06)	0.496	1.01		(0.27,3.72)	0.496
Variance/heterogeneity	LA Variance	0.196				0.149			
	LA (residual) VPC	0.050				0.039			
	LA MOR	1.53				1.45			
	School Variance	0.412				0.369			
	School (residual) VPC	0.106				0.097			
	School MOR	2.10				1.99			

*Notes: prop.= proportion; VPC=Variance Partition Coefficient; MOR=Median Odds Ratio; * indicates significance at p<0.05. Pupil level variables are also controlled for in these models but coefficients for these are not reported here. Combined ethnic group of interest for school composition= Asian (Indian, Pakistani, Bangladeshi and Asian Other).*

Table 2.15: ASD primary and secondary: Sequential model step variance/heterogeneity

ASD Primary (Y1-6)	*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Empty	0.178	0.442	0.05	0.11	1.50	2.12
Pupil pred.	0.193	0.427	0.05	0.11	1.52	2.12
Pupil and school pred.	0.196	0.412	0.05	0.11	1.52	2.10
ASD Secondary (Y7-11)	*LA var.*	*School var.*	*LA VPC*	*School VPC*	*LA MOR*	*School MOR*
Empty	0.144	0.476	0.037	0.122	1.44	2.12
Pupil pred.	0.149	0.388	0.039	0.101	1.44	2.01
Pupil and school pred.	0.149	0.369	0.039	0.097	1.45	1.99

Note: 'var.'=variance; 'pred.'=predictors; 'Empty' refers to a model with no predictors.

Cross level interactions

In order to further investigate some of the school effects noted above, we further investigated cross-level interactions between school composition factors and pupil characteristics.

MLD cross-level interactions

Because of the strong association between higher proportions of pupils entitled to FSM in a school and the odds of identification, a cross-level interaction between school proportion entitled to FSM and individual pupil FSM was tested and found to be significant in both primary and secondary phases.

Figures 2-1A/B display these cross-level interactions for primary and secondary phases in terms of predicted probabilities of MLD identification. Overall, in both primary and secondary, gaps between pupils with and without FSM entitlement are narrower in schools with higher proportions of pupils entitled to FSM. In particular, school-level deprivation makes more of a difference to pupils *not* entitled to FSM (i.e. they are more likely to be identified with MLD in schools with higher proportions of FSM-entitled pupils) than it does to pupils with FSM entitlement.

Including these interactions made a negligible difference to school and LA variance component estimates in either the primary or the secondary phase.

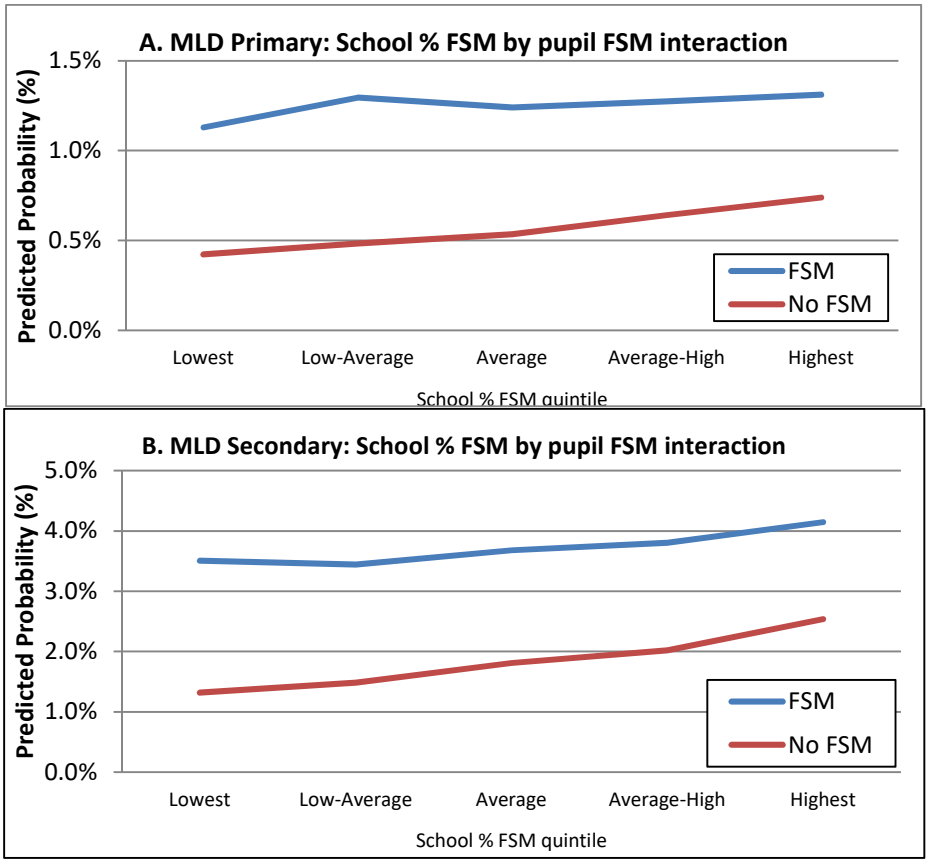


Figure 2-1: MLD: School %FSM by pupil FSM interaction (predicted probabilities)

SEMH cross-level interactions

We tested interactions between the school percent of pupils that were Black Caribbean/Mixed White & Black Caribbean, and individual student ethnicity (specifically Black Caribbean and Mixed White and Black Caribbean). These interactions were not significant, even after filtering out schools with fewer than two pupils in the relevant ethnic groups (to avoid skewing the results based on schools without any pupils in these ethnic groups).

However, there was a significant interaction between school percent FSM and pupil FSM across both primary and secondary phases. Figures 2-2A and 2-2B show these interactions in terms of predicted probabilities. Higher levels of school deprivation (%FSM) were specifically associated with increased probability of SEMH identification among non-FSM pupils. The probability of being identified with SEMH was consistently high for pupils entitled to FSM whatever the overall level of deprivation in the school. Similar interaction effects between %FSM and FSM have been reported in relation to educational attainment (see Strand, 2014b). This may be an artefact because FSM as a simple binary indicator is not able to differentiate levels of economic deprivation within the non-FSM group, and it is likely that non-FSM pupil's level socio-economic disadvantage is greater in the more deprived schools. Including this interaction did not, however, reduce the school-level variance components in either the primary or the secondary models.

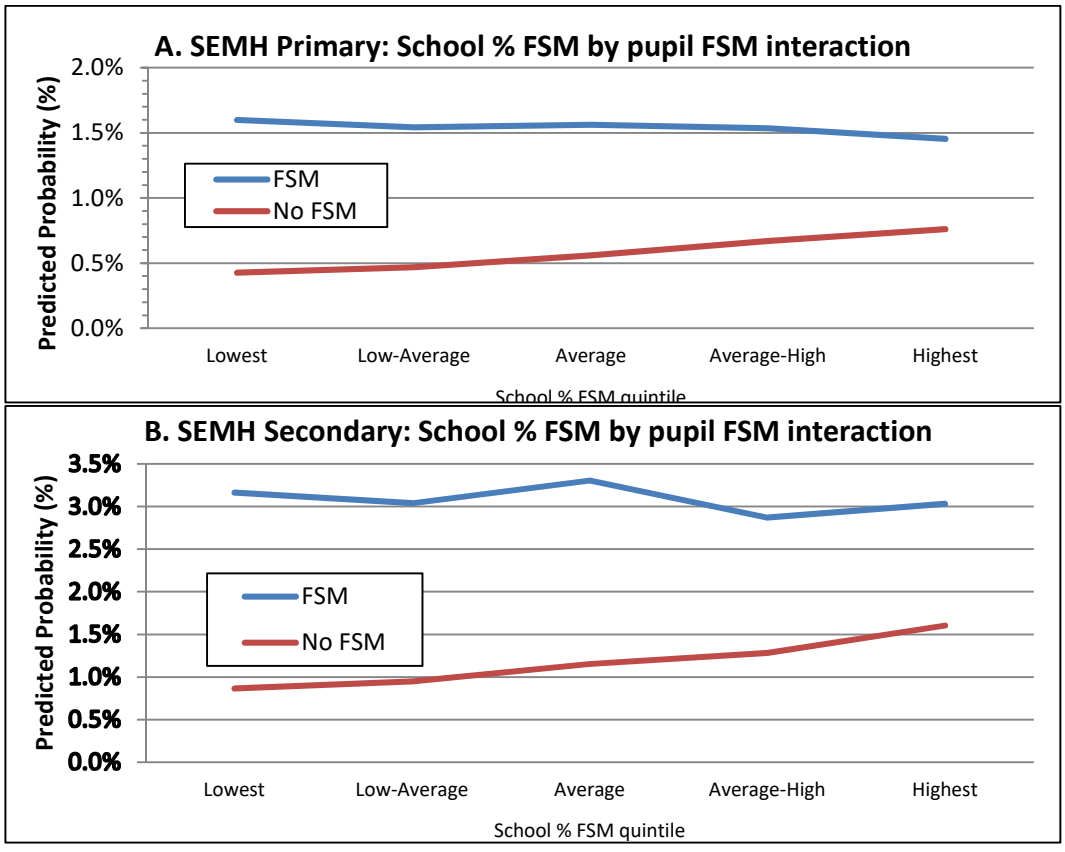


Figure 2-2: SEMH: School %FSM by pupil FSM interaction (predicted probabilities)

ASD cross-level interactions

Because of the under-representation of Asian groups for ASD across both primary and secondary phases, we additionally tested an interaction between the school proportion of pupils in Asian groups (specifically Indian, Pakistani, Bangladeshi, and Asian Other), and individual pupil ethnic group (specifically those same four Asian groups). These interactions were only significant between school percent Asian and individual Pakistani group membership in both primary and secondary phases, but this was not robust to filtering (in particular, excluding schools with fewer than two Asian pupils by the above definition led to non-statistically significant interactions between Pakistani individual ethnic group and school proportion Asian, and to a borderline-significant ($p<0.05$) interaction for only some quintiles in the secondary phase between Bangladeshi individual ethnic group and school proportion Asian.

We further explored whether there was an interaction between individual Pakistani classification and school proportion of pupils in the Pakistani group; this was only significant for the highest two quintiles of school proportion Pakistani, only in the primary phase, and only after filtering out schools with fewer than two Pakistani pupils. Figures 2-3 visualises the interactions of Pakistani pupil ethnic group with school proportion Pakistani in terms of predicted probabilities.

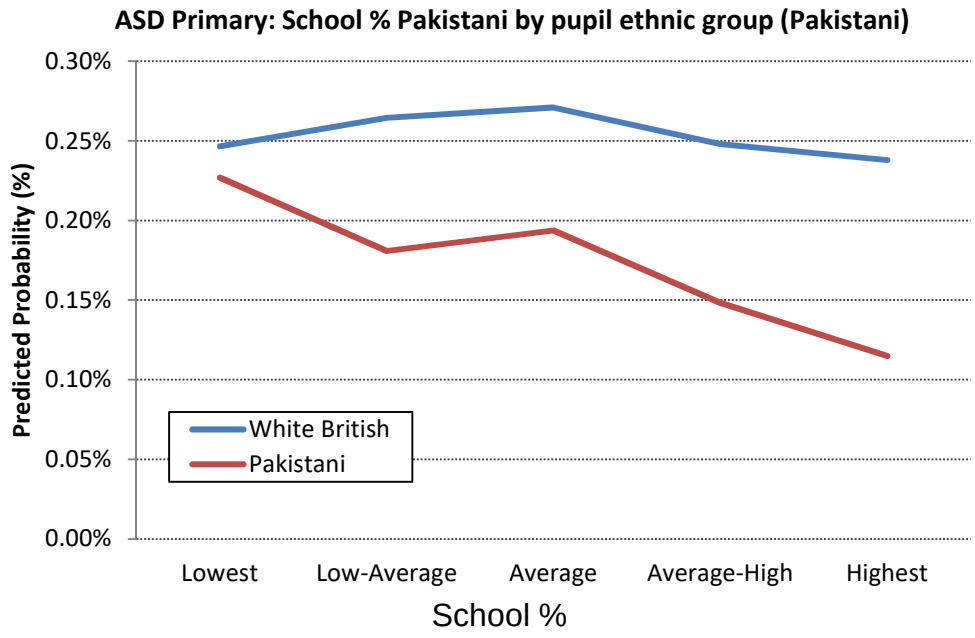


Figure 2-3: Primary ASD: %Asian by Pakistani interaction (predicted probabilities)

Note: The above results come from a model with schools filtered out if they have <2 Pakistani pupils, with variables including pupil ethnic group and all additional pupil background factors, but only school % Pakistani and the cross-level Pakistani by School %Pakistani interaction (N=1,587,025 pupils, N=5,317 schools).

Part 3: SEN identification over time – Longitudinal analyses of the NPD

Summary

The aim of this chapter is to investigate ethnic disproportionality in SEN identification over time as children and young people's progress through their primary and secondary school careers.

The first part of the chapter describes matched cohort data from the National Pupil Database (NPD) used for analysis and the general analytic approach taken.
The second part of the chapter presents results of a longitudinal analysis of the risk of identification - specifically, MLD, SEMH/BESD, and ASD identification - over the course of children's time in primary school for the primary cohort (pupils who were in Reception in January 2009, tracked through to Y6 in 2015), accounting for a range of pupil variables including prior attainment.

The third part of the chapter presents results from an analysis of the odds of ever being identified with MLD, SEMH/BESD or ASD over the course of secondary schooling for young people in the secondary cohort (pupils who were in Y7 in 2011, tracked through Y11 in 2015, with baseline information from Y6 in 2010), accounting for a range of pupil variables including prior attainment.

Key findings:

- Disproportionate risk of identification accumulates over time; for example, over-represented groups (such as Black Caribbean and Mixed White and Black Caribbean pupils for SEMH/BESD identification) experience a greater cumulative risk of identification compared to their White British majority peers over the course of their experience in Primary school than is visible from a cross-sectional analysis alone.
- Socioeconomic deprivation (SED) has strong associations with SEN identification, particularly for more judgmental types of SEN (i.e. much more for MLD and BESD than for ASD). SED can account for over-representation of some ethnic groups, but does not account for the over-representation of Black Caribbean and Mixed White & Black Caribbean pupils for SEMH/BESD, nor the under-representation of some Asian groups for MLD and ASD.
- Attainment and development at the start of school (end of reception year) is associated with SEN identification over the course of a pupil's primary school career. For MLD, literacy and numeracy attainment was the strongest predictor; for SEMH/BESD and ASD, personal, social and emotional development (PSED) was the strongest predictor. These measures do not, however, fully account for the identified ethnic disproportionality, such as Asian under-representation for ASD and MLD.

Key findings (continued):

- Prior attainment can have a strong association with the odds that a pupil will ever be identified with SEN in secondary school, but the strength of this relationship varies widely by type of SEN and is much stronger for MLD and ASD than for SEMH/BESD. These measures do not, however, fully account for the identified ethnic disproportionality, such as Asian under-representation for ASD and MLD.

- Although school context and composition are not the strongest predictors of SEN identification, they do appear to play a role, and which school characteristics matter (and how much) varies by type of SEN.

For MLD, schools serving more deprived communities (higher school % FSM) are associated with higher risk of identification.

  - For SEMH/BESD, smaller school size, higher school % FSM, and higher proportions of over-represented groups (Black Caribbean and Mixed White & Black Caribbean) are all associated with higher risk of identification.

  - For ASD, schools serving more deprived populations (higher % FSM) are associated with a slightly lower risk of identification, and schools with larger proportions of under-identified (Asian) groups are associated with higher risk of identification.

What we did

Data source

As in Parts 1 and 2, School Census data from the National Pupil Database (NPD) obtained from the DFE were used for the analysis discussed here. Data were first received in a longitudinally matched format, with one dataset for a primary cohort and one for a secondary cohort. For each of these datasets, the data was linked by the NPD data team before being sent to the research team. The cohort began with all pupils enrolled in the relevant baseline year group at the time of the January School Census (Reception in January 2009 for the primary cohort, Y6 in January 2010 for the secondary cohort) and matching these to the records for those same pupils in subsequent years up to and including the January 2015 School Census. As would be reasonably expected, some pupils left the cohort and some re-joined from year to year. Table 3.1 shows the numbers of pupils joining and leaving each year; this does not include any pupils joining the cohort who were not included in the baseline year as subsequent matching was based on those baseline pupil records.

Table 3.1: Primary and secondary cohorts: Number of pupils by year, joiners and leavers

Academic Year	Year Group	# of nonduplicate pupil records*	Joined since previous January Census	Left since previous January Census
Primary				
2008-09	YR	562274	--	--
2009-10	Y1	551820	--	10454
2010-11	Y2	546867	1656	6609
2011-12	Y3	540680	1695	7882
2012-13	Y4	536389	2389	6680
2013-14	Y5	532788	2170	5771
2014-15	Y6	530003	2033	4818
Secondary				
2010-11	Y7	544541	--	--
2011-12	Y8	539072	--	5469
2012-13	Y9	534087	1170	6155
2013-14	Y10	531931	2638	4794
2014-15	Y11	526164	1273	7040

*on the January School Census date of the relevant year.

For survival analysis, individuals were not tracked beyond their first instance of leaving their cohort; that is, records for pupils who were enrolled in earlier years, then missing in a later year, and then enrolled again in year after that, were treated as having left the cohort (i.e. if such individuals had not already been identified with some SEN before the first instance of leaving the cohort, any identification after this

was not accounted for in the analysis, for reasons addressed below under 'approach to analysis'). This applies only to the primary cohort as it was not possible to conduct survival analysis for the secondary cohort, the reason for which is also addressed below in 'approach to analysis'. Table 3.2 shows the numbers of pupils in each year for the primary cohort after accounting only for leavers and not for re-joiners; this is still a fairly small proportion of the initial total number of pupils (7.2%). It should be noted that this is not selective attrition from a sample, but reflects the genuine level of mobility within the population.

Table 3.2: Primary and secondary cohorts: Leavers (excluding re-joiners) by year

Academic Year	Year Group	# of nonduplicate pupil records*	# of leavers since previous January Census	Leavers as cumulative % of initial total pupils
Primary				
2008-09	YR	562274	--	--
2009-10	Y1	551820	10454	1.9
2010-11	Y2	545211	6609	3.0
2011-12	Y3	537470	7741	4.4
2012-13	Y4	531552	5918	5.5
2013-14	Y5	526081	5471	6.4
2014-15	Y6	521641	4440	7.2
Secondary				
2010-11	Y7	544541	--	--
2011-12	Y8	539072	5469	1.0
2012-13	Y9	534087	4985	1.9
2013-14	Y10	531931	2156	2.3
2014-15	Y11	526164	5767	3.4

Filtering

Duplicate records (307 in the primary cohort, eight in the Secondary cohort, both negligible percentages of the total available records) and records missing values on any of the variables to be used in analyses were excluded, as was done for the analyses presented in previous sections (i.e. listwise deletion was used). The only pupil-level variables missing values on any records were IDACI score (0.4% missing in each dataset, primary and Secondary), components of the Early Years Foundation Stage Profile (EYFSP) used as a measure of prior attainment in the primary cohort dataset, and Key Stage 2 results used as a measure of prior attainment in the Secondary cohort dataset. In the primary dataset, approximately 1.2% of pupil records were missing EYFSP scores. In the Secondary dataset, a more substantial proportion were missing KS2 scores (approximately 27.3% in English, 27.0% in Mathematics), mostly due to a partial school boycott of national tests in 2010.

However, comparing the distribution at *pupil level* between the sample completing the tests and the whole Y6 population reveals negligible differences; this is addressed in more detail below under 'Measures'.

We did not filter out special schools for the analysis reported in this section, which is important to take into consideration in drawing comparisons to the results presented in the previous section (which were based only on maintained mainstream schools).

Measures

The outcome of interest was SEN identification as in the previous sections, and again we focus particularly on our three focal primary types of SEN: MLD, SEMH (or BESD, prior to 2015), and ASD.

Prior attainment

Early Years Foundation Stage Profile (EYFSP)

Scores on several components of the EYFSP were used as measures of Reception attainment/development for the primary cohort. Appendix J provides information on the scores used and the EYFSP in more detail. Scores were standardised (Z-scored) to aid interpretation of coefficients. These were not strictly measures of prior attainment since they were completed in May while the Reception school census data was gathered in January, but the time gap is small and there are no national measures of attainment prior to the EYFSP.

End of Key Stage 2 (KS2) national tests

We used the end of KS2 (age 11) English (reading & writing) and mathematics test scores completed at the end of Y6 as measures of prior attainment for the secondary cohort. We used the 'finely-graded' levels from the NPD; these use the marks awarded in the relevant tests to calculate a fine-grade (decimalised) National Curriculum test level which can run from a low of 2.5 through to a maximum of 6.5. Descriptive statistics are presented below in Table 3.3.

Table 3.3: Descriptive statistics: 2010 KS2 English and Maths finely-graded levels

	N	Min.	Max.	Mean	SD
KS2 English finely-graded level	409241	2.50	5.97	4.50	0.78
KS2 mathematics finely-graded level	411010	2.50	6.50	4.58	0.84
Valid listwise	406813				

Students who did not score not enough marks to be awarded a level (N), and those who were working towards the level of the test (B), are typically included in the measures with the floor score of 2.5. The distributions therefore have a small peak at the lowest score (see Appendix K) but this does allow the maximum number of students to be included. A very small number (0.5%) of pupils with extreme SEN who are either Disapplied from the National Curriculum (D) or Working Below Level 1 (W)

are necessarily excluded. Pupils who were absent from school on the day of the test (A) have no score, but these students are randomly distributed, as are a very small number whose results were lost or invalid.

A more substantial challenge was offered by the fact that in 2010 there had been a partial boycott of national testing arrangements by some primary schools in England. In 2010, 15,518 maintained primary schools were expected to administer KS2 tests, but 4,005 (26 per cent) of these schools did not administer them. The decision to boycott the tests (or not) was made at the school (head teacher or senior leadership level), and therefore no pupils at boycott schools will have taken the tests. However comparing the distribution at *pupil level* between the sample completing the tests and the whole Y6 population reveals very few differences. For our key variable of ethnicity, 77.8% of the tested sample were White British compared to 76.3% of the population. Apart from some small under-representation of Pakistani (2.9% of those tested vs 3.7% of population) and Bangladeshi (1.2% vs. 1.5%) pupils, the proportion of ethnic minorities in the tested sample closely parallels the whole population (see Appendix K). The same result was found for primary SEN type: 88.2% of the tested sample had no identified SEN compared to 88.8% of the Y6 population. For the individual SEN types the difference between the proportion in the tested sample and the population never differed by more than 0.1%.

These results agree with those of a DFE commissioned report into dealing with the missing data (Saunders et al., 2016) which concluded there were few pupil-level predictors of pupils having missing KS2 results. They report "*The 'missingness mechanism' for KS2 test results occurred at the school level rather than at the individual pupil level, because it was the schools – or rather their head teachers – who chose to boycott the tests or not*" and "*the results of the analyses undertaken for this project suggest that complete-cases analyses using only pupil-level data that include a random effect for secondary school should be unbiased" (p7).*

There is no reason therefore to consider our sample biased, and while the number of students is somewhat reduced we are still tracking an extremely large sample of approximately 400,000 students. We also take our school level variables (where we test them) from the secondary schools the students moved to in Y7 and not from the primary schools.

Like the EYPFS scores for the primary cohort, the finely-graded levels in English and maths for the secondary cohort were normal score transformed to Mean=0 and SD=1 in order to facilitate easier substantive interpretation (i.e. regression coefficients reflect the effect of a 1SD difference).

Other pupil level explanatory variables

Pupil ethnic group and background variables were identical to those described in previous sections. For the primary cohort, we used information from the 2009

January Census, when pupils were in Reception. For the secondary cohort, we used information from the 2010 January Census, when pupils were in Y6.

Attendance: It was not possible to account for attendance in the primary cohort, as the available data included very few non-missing attendance records (<2000); this may be due to the fact that children are not all of statutory school age (the term after the child's fifth birthday) by the time of the January School Census in a given year, so that attendance recording is consistent only by the January School Census in Y1. In the secondary cohort, attendance was included in the analysis in terms of a binary persistent absence indicator (defined as missing more than 63 days of Y6 in 2010).

Descriptive statistics for pupil level variables by ethnic group

Tables 3.4 and 3.5 present descriptive information for the above listed explanatory variables broken down by ethnic group in the primary and secondary cohorts, respectively.

Looking first at the primary cohort (see Table 3.4), it is apparent that while some of these variables are fairly consistent on average across ethnic groups (e.g. proportion by gender), others vary substantially (e.g. FSM entitlement, which is 34.6% for the Black African group and only 5.7% in the Indian group).

Similarly, in the secondary cohort (see Table 3.5), gender and birth season are largely consistent across ethnic groups while FSM entitlement, IDACI, attendance and KS2 finely-graded English and maths levels have more substantial differences across ethnic groups.

This underscores the importance of accounting for those pupil background variables for which group differences are observable, given the emphasis on ethnic group as the focal variable of our analyses. Meanwhile, it is still important to control for the other pupil background characteristics (birth season, gender), as these may be linked to individual differences in the likelihood that a pupil will be identified with some SEN.

Table 3.4: Primary cohort descriptive statistics for explanatory variables by ethnic group

PRIMARY COHORT	*Total*	*FSM entitlement*		*Gender*		*Birth season*						*IDACI score**		*EYFSP*					
		Entitled		*Boy*		*Autumn*		*Spring*		*Summer*				*CLL*		*PSRN*		*PSE*	
Ethnic group		*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
White British	352175	49036	13.9%	180992	51.4%	119115	33.8%	115930	32.9%	117130	33.3%	20.0	16.8	25.8	6.3	20.4	4.3	20.8	4.0
White Irish	1327	236	17.8%	659	49.7%	461	34.7%	419	31.6%	447	33.7%	23.2	18.9	26.3	6.1	20.8	4.2	21.4	3.8
Irish Traveller	436	224	51.4%	210	48.2%	152	34.9%	130	29.8%	154	35.3%	30.0	19.6	18.0	7.3	14.6	5.9	17.4	4.7
Gypsy/Roma	836	289	34.6%	432	51.7%	249	29.8%	294	35.2%	293	35.0%	26.9	18.0	18.5	7.2	15.1	5.7	17.9	4.6
White Other	20209	2397	11.9%	10319	51.1%	6777	33.5%	6682	33.1%	6750	33.4%	26.0	19.2	24.1	7.0	19.1	5.0	20.2	4.3
Mixed White & African	2741	607	22.1%	1362	49.7%	874	31.9%	889	32.4%	978	35.7%	31.0	20.4	25.4	6.4	20.1	4.5	20.6	4.1
Mixed White & Caribbean	6115	1967	32.2%	3068	50.2%	2060	33.7%	1969	32.2%	2086	34.1%	31.7	19.5	25.0	6.2	19.8	4.3	20.4	4.0
Mixed White & Asian	4940	776	15.7%	2480	50.2%	1664	33.7%	1612	32.6%	1664	33.7%	22.0	17.6	26.3	6.5	20.6	4.4	21.2	4.0
Mixed Other	8249	1788	21.7%	4277	51.8%	2646	32.1%	2823	34.2%	2780	33.7%	28.6	20.4	25.4	6.5	20.0	4.6	20.6	4.2
Indian	12114	694	5.7%	6230	51.4%	3978	32.8%	3999	33.0%	4137	34.2%	27.4	17.7	25.9	6.4	20.2	4.4	20.8	4.0
Pakistani	17449	3024	17.3%	8911	51.1%	6027	34.5%	5756	33.0%	5666	32.5%	38.0	17.1	23.0	7.0	18.1	5.1	19.3	4.4
Bangladeshi	6200	1312	21.2%	3164	51.0%	2104	33.9%	2093	33.8%	2003	32.3%	45.1	20.0	22.9	6.9	18.0	5.1	19.3	4.3
Asian Other	7297	858	11.8%	3669	50.3%	2523	34.6%	2387	32.7%	2387	32.7%	29.8	17.8	24.6	6.8	19.4	4.8	20.1	4.3
Black African	14444	5004	34.6%	7283	50.4%	4841	33.5%	4608	31.9%	4995	34.6%	43.1	19.3	24.0	6.9	18.9	4.9	19.6	4.4
Black Caribbean	5757	1732	30.1%	2954	51.3%	2026	35.2%	1878	32.6%	1853	32.2%	41.2	18.0	24.1	6.5	19.1	4.6	19.6	4.3
Black Other	3170	1002	31.6%	1652	52.1%	1107	34.9%	1016	32.1%	1047	33.0%	42.0	20.0	23.8	6.9	18.8	4.9	19.6	4.4
Chinese	1591	144	9.1%	755	47.5%	517	32.5%	524	32.9%	550	34.6%	25.9	20.3	25.3	6.8	20.3	4.7	20.6	4.2
Any Other	6674	1645	24.6%	3491	52.3%	2198	32.9%	2214	33.2%	2262	33.9%	37.6	21.6	23.4	7.0	18.5	5.0	19.6	4.4
Unknown	90550	16881	18.6%	46488	51.3%	30037	33.2%	29601	32.7%	30912	34.1%	27.7	20.8	25.2	6.6	19.9	4.7	20.5	4.2
Total (Overall)	562274	89616	15.9%	288396	51.3%	189356	33.7%	184824	32.9%	188094	33.5%	24.1	19.1	25.4	6.5	20.0	4.5	20.6	4.1

Note: EYFSP = Early Years Foundation Stage Profile; CLL=Communication, Language, and Literacy; PSRN=Problem Solving, Reasoning and Numeracy; PSE= Personal, Emotional and Social development.
**The total pupil N for IDACI (559753) is slightly lower than those for FSM and gender due to a small proportion of missing values for IDACI. The total pupil Ns are again lower for EYFSP (555792 for CLL, 555762 for PSRN, 555809 for PSE) scores because of missing values on these variables.*

Table 3.5: Secondary cohort descriptive statistics for explanatory variables by ethnic group

SECONDARY COHORT	*Total*	*FSM entitlement*		*Gender*		*Birth season*						*IDACI score**		*Attendance**		*Key Stage 2 attainment (finely-graded levels)**			
		Entitled		*Boy*		*Autumn*		*Spring*		*Summer*				*Persistently absent*		*English*		*Maths*	
Ethnic group		*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*M*	*SD*	*N*	*%*	*M*	*SD*	*M*	*SD*
White British	429484	65614	15.3%	219963	51.2%	144481	33.6%	138648	32.3%	146355	34.1%	19.4	16.4	5552.0	1.3%	4.5	0.8	4.6	0.8
White Irish	1917	345	18.0%	941	49.1%	639	33.3%	622	32.4%	656	34.2%	24.6	19.2	42.0	2.2%	4.7	0.8	4.7	0.8
Traveller Irish	448	286	63.8%	249	55.6%	144	32.1%	133	29.7%	171	38.2%	33.2	20.9	158.0	35.3%	3.5	0.9	3.6	0.9
Traveller Gypsy/Roma	1228	460	37.5%	621	50.6%	391	31.8%	427	34.8%	410	33.4%	29.5	18.5	290.0	23.7%	3.3	0.9	3.5	0.9
White Other	19403	2984	15.4%	10023	51.7%	6353	32.7%	6272	32.3%	6778	34.9%	28.4	20.3	364.0	1.9%	4.3	0.9	4.5	0.9
Mixed White & African	2371	657	27.7%	1175	49.6%	829	35.0%	771	32.5%	771	32.5%	31.1	20.4	35.0	1.5%	4.6	0.7	4.6	0.8
Mixed White & Carib.	7323	2309	31.5%	3699	50.5%	2580	35.2%	2299	31.4%	2444	33.4%	31.5	20.1	156.0	2.1%	4.5	0.8	4.5	0.8
Mixed White & Asian	4613	889	19.3%	2384	51.7%	1517	32.9%	1490	32.3%	1606	34.8%	22.7	18.4	66.0	1.4%	4.7	0.7	4.8	0.8
Mixed Other	7799	1829	23.5%	4007	51.4%	2685	34.4%	2430	31.2%	2684	34.4%	27.8	20.3	109.0	1.4%	4.6	0.8	4.6	0.8
Indian	13031	1233	9.5%	6801	52.2%	4382	33.6%	4189	32.1%	4460	34.2%	27.3	17.3	74.0	0.6%	4.7	0.7	4.8	0.8
Pakistani	20651	5558	26.9%	10603	51.3%	6722	32.6%	6875	33.3%	7054	34.2%	39.2	17.7	311.0	1.5%	4.3	0.8	4.4	0.9
Bangladeshi	8301	2969	35.8%	4283	51.6%	2773	33.4%	2809	33.8%	2719	32.8%	48.4	20.8	86.0	1.0%	4.4	0.8	4.5	0.9
Asian Other	6803	943	13.9%	3517	51.7%	2269	33.4%	2222	32.7%	2312	34.0%	29.2	18.3	61.0	0.9%	4.5	0.8	4.7	0.9
Black African	16024	6614	41.3%	8071	50.4%	5320	33.2%	5215	32.5%	5489	34.3%	43.7	19.7	81.0	0.5%	4.4	0.8	4.4	0.9
Black Caribbean	7893	2314	29.3%	4047	51.3%	2761	35.0%	2537	32.1%	2595	32.9%	41.1	19.1	83.0	1.1%	4.4	0.8	4.3	0.8
Black Other	2857	965	33.8%	1425	49.9%	995	34.8%	887	31.0%	975	34.1%	40.1	20.1	35.0	1.2%	4.3	0.8	4.3	0.9
Chinese	1846	192	10.4%	914	49.5%	647	35.0%	568	30.8%	631	34.2%	24.7	20.4	3.0	0.2%	4.7	0.8	5.1	0.7
Any Other	7436	2407	32.4%	3862	51.9%	2369	31.9%	2550	34.3%	2517	33.8%	37.8	21.5	85.0	1.1%	4.3	0.9	4.5	0.9
Unknown	3711	684	18.4%	1921	51.8%	1232	33.2%	1203	32.4%	1276	34.4%	22.1	17.9	70.0	1.9%	4.5	0.8	4.5	0.8
Total (Overall)	563139	99252	17.6%	288506	51.2%	189089	33.6%	182147	32.3%	191903	34.1%	23.0	18.6	7661.0	1.4	4.5	0.8	4.6	0.8

**The total pupil Ns for IDACI (561156) and attendance () are slightly lower than those for FSM and gender due to a small proportion of missing values for each of these variables. The total pupil Ns are again lower for KS2 scores (409236 for English, 411005 for Maths) scores because of missing values on these variables.*

School level variables

School variables were also similar to those described in Part 2. We again used information from the 2009 School Census for the primary cohort (i.e. information about pupils' schools from their Reception year); however, for the secondary cohort we used information from the 2011 School Census (when pupils were in Y7) to ensure that we were using information on pupils' secondary schools. The variables used included:

- School type: a categorical variable, with Community schools as the reference category
- School % FSM eligibility: coded into quintiles over all schools for which there were pupil records in the relevant sample for analysis, with the lowest quintile as the reference category
- School % ethnic group composition: coded into quintiles over all schools for which pupil records were included in the sample for analysis, with the groups defined according to those previously identified as persistently under- or over-represented in previous analyses, with the lowest quintile as the reference category[15]. Where we have combined groups for the purpose of exploring school composition effects, we do not intend to imply shared identity or culture; rather, these groupings are empirically driven based on combining groups with similar profiles for a given primary type of SEN.
- School size: determined based on the number of pupils on roll from the January school level census, coded into quintiles with 'Largest' as the reference category.

Descriptive statistics for school level variables by phase

Tables 3.6 and 3.7 provide descriptive information at the school level for the above school variables by category for the primary and secondary cohort, respectively. Descriptive statistics are reported over all schools available in the main analysis as well as in samples filtered specifically to exclude schools with fewer than two pupils across the ethnic groups relevant to school ethnic composition variables for each SEN type (MLD, SEMH/BESD and ASD), as quintiles were recalculated for these different samples.

[15]. For MLD, the combined ethnic group of interest for school composition was Asian (all groups excluding Pakistani pupils, as Pakistani pupils have an identification profile distinct from the other Asian groups); for ASD, the combined group of interest was again Asian (but defined as Indian, Pakistani, Bangladeshi, and Asian Other); and for SEMH/BESD, the combined group of interest was Black Caribbean and Mixed White & Black Caribbean.

Table 3.6: Primary cohort schools descriptive information

PRIMARY - Schools		All available						Filtered for SEMH-BESD model*					
		N	%	Min	Max	Mean	SD	N	%	Min	Max	Mean	SD
School type	Foundation	326	2.0	--	--	--	--	155	2.3	--	--	--	--
	Academy	12	0.1	--	--	--	--	11	0.2	--	--	--	--
	Church	5946	36.9	--	--	--	--	2010	30.0	--	--	--	--
	Special/PRU/AP	457	2.8	--	--	--	--	153	2.3	--	--	--	--
	Community	9364	58.1	--	--	--	--	4364	65.2	--	--	--	--
School % FSM	Highest	3230	20.1	26.4	79.8	37.8	9.3	1334	19.9	31.9	79.8	42.4	8.5
	Average-High	3205	19.9	14.3	26.3	19.7	3.5	1344	20.1	19.9	31.8	25.4	3.4
	Average	3237	20.1	7.6	14.2	10.5	1.9	1337	20.0	10.8	19.8	15.1	2.6
	Low-Average	3175	19.7	3.7	7.5	5.5	1.1	1339	20.0	5.3	10.7	7.9	1.6
	Lowest	3258	20.2	0.0	3.6	1.8	1.2	1339	20.0	0.0	5.2	2.9	1.4
School % Asian	Highest	3236	20.1	6.5	100.0	17.6	15.5	--	--	--	--	--	--
(except Pakistani)	Average-High	3134	19.5	2.9	6.4	4.3	1.0	--	--	--	--	--	--
	Average	3245	20.1	1.3	2.8	2.0	0.4	--	--	--	--	--	--
	Low-Average	2626	16.3	0.2	1.2	0.8	0.3	--	--	--	--	--	--
	Lowest	3864	24.0	0.0	0.0	0.0	0.0						
School % Black Caribbean	Highest	3253	20.2	2.8	55.9	8.7	7.5	1327	19.8	7.1	55.9	15.1	8.2
/ MWBC	Average-High	3128	19.4	1.1	2.7	1.7	0.5	1335	19.9	3.5	7.0	4.9	1.0
	Average	2798	17.4	0.1	1.0	0.6	0.2	1334	19.9	2.1	3.4	2.7	0.4
	Low-Average	6926	43.0	0.0	0.0	0.0	0.0	1291	19.3	1.3	2.0	1.6	0.2
	Lowest	--	--	--	--	--	--	1406	21.0	0.2	1.2	0.9	0.2
School % Asian	Highest	3219	20.0	6.9	100.0	28.0	24.4	--	--	--	--	--	--
(Indian/Pakistani/	Average-High	3186	19.8	2.0	6.8	3.8	1.3	--	--	--	--	--	--
Bangladeshi/Asian other)	Average	3176	19.7	0.6	1.9	1.2	0.4	--	--	--	--	--	--
	Low-Average[a]	591	3.7	0.1	0.5	0.4	0.1	--	--	--	--	--	--
	Lowest	5933	36.8	0.0	0.0	0.0	0.0						
School Size (Roll)	Smallest	3245	20.1	0	87	57.4	20.5	1330	19.9	20	147	106.8	28.1
	Small-Average	3195	19.8	88	141	112.9	15.4	1352	20.2	148	177	165.6	8.3
	Average	3157	19.6	142	177	162.8	10.2	1328	19.8	178	242	202.5	21.3
	Average-Large	3303	20.5	178	264	211.3	27.4	1339	20.0	243	334	287.8	27.4
	Largest	3205	19.9	265	1636	350.3	85.2	1344	20.1	335	1636	398.5	100.5
Total schools		16105	100.0					6693	100.0				

**In the BESD model with school variables, overall school quintile calculations were complicated by the large number of schools with 0% pupils in the groups of interest Black Caribbean/Mixed White & Black Caribbean (MWBC). The model was therefore run (and results presented) for a sample with schools with <2 pupils in the groups of interest filtered out.*

[a] *The apparent imbalance in quintile frequencies for School % Asian for the ASD model is a consequence of having a large number of primary schools with no pupils belonging to the relevant ethnic groups clustered in the lowest quintile. This is resolved in the alternatively filtered version for which quintiles were recomputed below.*

PRIMARY – Schools		Filtered for MLD model*						Filtered for ASD model*					
		N	%	Min	Max	Mean	SD	N	%	Min	Max	Mean	SD
School type	Foundation	231	2.3	--	--	--	--	181	2.2	--	--	--	--
	Academy	11	0.1	--	--	--	--	9	0.1	--	--	--	--
	Church	3273	32.2	--	--	--	--	2496	29.8	--	--	--	--
	Special/PRU/AP	278	2.7	--	--	--	--	273	3.3	--	--	--	--
	Community	6370	62.7	--	--	--	--	5419	64.7	--	--	--	--
School % FSM	Highest	2033	20.0	28.6	79.8	39.7	8.9	1670	19.9	30.4	79.8	41.2	8.7
	Average-High	2035	20.0	16.7	28.5	22.3	3.4	1687	20.1	19.0	30.3	24.3	3.3
	Average	2043	20.1	8.9	16.6	12.4	2.3	1672	20.0	10.3	18.9	14.3	2.5
	Low-Average	2027	19.9	4.3	8.8	6.5	1.3	1679	20.0	4.9	10.2	7.4	1.5
	Lowest	2025	19.9	0.0	4.2	2.3	1.2	1670	19.9	0.0	4.8	2.7	1.3
School % Asian	Highest	2024	19.9	10	100	23.4	17.1						
(except Pakistani)	Average-High	2023	19.9	5.1	9.9	7.1	1.4						
	Average	2117	20.8	3	5	3.9	0.6						
	Low-Average	2103	20.7	1.8	2.9	2.3	0.4						
	Lowest	1896	18.7	0.2	1.7	1.2	0.3						
School % Asian	Highest							1674	20.0	16.2	100	44.0	24.63
(Indian/Pakistani/	Average-High							1685	20.1	6.4	16.1	10.3	2.77
Bangladeshi/Asian other)	Average							1654	19.7	3.3	6.3	4.6	0.87
	Low-Average							1653	19.7	1.8	3.2	2.4	0.44
	Lowest							1712	20.4	0.4	1.7	1.2	0.33
School Size (Roll)	Smallest	2036	20.0	17	130	95.4	25.2	1668	19.9	6	141	104.0	26.3
	Small-Average	2045	20.1	131	172	156.1	11.8	1676	20.0	142	175	162.4	9.5
	Average	2015	19.8	173	215	186.1	12.2	1691	20.2	176	232	194.6	17.7
	Average-Large	2035	20.0	216	315	261.8	27.7	1664	19.9	233	326	276.4	27.5
	Largest	2032	20.0	316	1636	382.6	90.4	1679	20.0	327	1636	391.1	94.3
Total schools		10163	100.0					8378	100.0				

**The alternative sample used to check ethnic group composition effect robustness for MLD filtered out schools with <2 Asian (excluding Pakistani) pupils; the sample used for this purpose for ASD filtered out schools with <2 Asian (Indian/Pakistani/Bangladeshi/Asian Other pupils).*

Table 3.7: Secondary cohort schools descriptive information

SECONDARY - Schools		All available						Filtered for SEMH-BESD model*					
		N	%	Min	Max	Mean	SD	N	%	Min	Max	Mean	SD
School type	Foundation	706	17.7	--	--	--	--	639	21.2	--	--	--	--
	Academy-Converter	70	1.8	--	--	--	--	61	2.0	--	--	--	--
	Academy-Sponsored	265	6.7	--	--	--	--	236	7.8	--	--	--	--
	Church	573	14.4	--	--	--	--	487	16.1	--	--	--	--
	Selective/Grammar	161	4.0	--	--	--	--	139	4.6	--	--	--	--
	Special/PRU/AP	785	19.7	--	--	--	--	317	10.5	--	--	--	--
	Community	1421	35.7	--	--	--	--	1141	37.8	--	--	--	--
School % FSM	Highest	795	20.0	31.9	100.0	44.8	11.8	602	19.9	29.4	100.0	42.7	11.9
	Average-High	799	20.1	19.8	31.8	25.3	3.5	604	20.0	17.4	29.3	23.0	3.5
	Average	788	19.8	11.5	19.7	15.3	2.4	602	19.9	10.5	17.3	13.4	2.0
	Low-Average	795	20.0	6.5	11.4	8.9	1.5	598	19.8	6.1	10.4	8.1	1.3
	Lowest	804	20.2	0.0	6.4	3.9	1.7	614	20.3	0.0	6.0	3.8	1.5
School % Asian (except Pakistani)	Highest	791	19.9	7.7	95.1	19.5	15.0						
	Average-High	810	20.3	3.4	7.6	5.1	1.2						
	Average	799	20.1	1.7	3.3	2.4	0.5						
	Low-Average	805	20.2	0.8	1.6	1.2	0.3						
	Lowest	776	19.5	0.0	0.7	0.3	0.3						
School % Black Caribbean /Mixed White &Caribbean	Highest	798	20.0	3.8	43.5	9.9	6.6	598	19.8	5.3	43.5	11.7	6.7
	Average-High	789	19.8	1.4	3.7	2.3	0.7	620	20.5	2.1	5.2	3.3	0.9
	Average	738	18.5	0.7	1.3	1.0	0.2	604	20.0	1.0	2.0	1.4	0.3
	Low-Average	937	23.5	0.2	0.6	0.4	0.1	513	17.0	0.6	0.9	0.7	0.1
	Lowest	719	18.1	0.0	0.1	0.0	0.0	685	22.7	0.1	0.5	0.4	0.1
School % Asian (Indian/Pakistani/ Bangladeshi/Asian other)	Highest	795	20.0	10.2	99.5	30.0	21.2						
	Average-High	800	20.1	3.2	10.1	5.8	2.0						
	Average	768	19.3	1.2	3.1	2.0	0.6						
	Low-Average	768	19.3	0.4	1.1	0.7	0.2						
	Lowest	850	21.4	0.0	0.3	0.1	0.1						
School Size (Roll)	Smallest	795	20.0	11	205	93.3	41.9	604	20.0	11	590	268.3	188.5
	Small-Average	798	20.0	206	713	516.0	142.0	604	20.0	591	855	733.8	75.7
	Average	795	20.0	714	960	837.5	68.4	605	20.0	857	1059	953.6	61.1
	Average-Large	797	20.0	961	1229	1091.2	78.1	604	20.0	1060	1309	1175.9	71.5
	Largest	796	20.0	1230	2577	1494.3	221.0	603	20.0	1310	2577	1550.6	212.3
Total schools		3981	100.0					3020	100.0				

**The sample used for checking the robustness of an ethnic group composition effect for SEMH/BESD filtered out schools with <1 Black Caribbean and Mixed White & Caribbean pupils.*

SECONDARY – Schools (continued)		Filtered for MLD model*						Filtered for ASD model*					
		N	%	Min	Max	Mean	SD	N	%	Min	Max	Mean	SD
School type	Foundation	689	19.7	--	--	--	--	636	19.8	--	--	--	--
	Academy-Converter	69	2.0	--	--	--	--	61	1.9	--	--	--	--
	Academy-Sponsored	255	7.3	--	--	--	--	249	7.7	--	--	--	--
	Church	547	15.7	--	--	--	--	525	16.3	--	--	--	--
	Selective/Grammar	161	4.6	--	--	--	--	161	5.0	--	--	--	--
	Special/PRU/AP	434	12.4	--	--	--	--	399	12.4	--	--	--	--
	Community	1340	38.3	--	--	--	--	1182	36.8	--	--	--	--
School % FSM	Highest	698	20.0	29.0	100.0	40.8	10.3	643	20.0	29.8	100.0	41.6	10.2
	Average-High	698	20.0	17.8	28.9	22.9	3.2	642	20.0	18.2	29.7	23.6	3.3
	Average	700	20.0	10.7	17.7	13.7	2.0	644	20.0	10.8	18.1	14.0	2.1
	Low-Average	704	20.1	6.1	10.6	8.2	1.3	639	19.9	6.1	10.7	8.3	1.4
	Lowest	695	19.9	0.0	6.0	3.7	1.6	645	20.1	0.0	6.0	3.7	1.6
School % Asian (except Pakistani)	Highest	697	19.9	8.8	95.1	21.0	15.4						
	Average-High	713	20.4	4.0	8.7	5.9	1.3						
	Average	682	19.5	2.2	3.9	2.9	0.5						
	Low-Average	725	20.7	1.1	2.1	1.5	0.3						
	Lowest	678	19.4	0.1	1.0	0.7	0.2						
School % Asian (Indian/Pakistani/ Bangladeshi/Asian other)	Highest							645	20.1	13.0	99.5	34.3	21.3
	Average-High							635	19.8	4.9	12.9	8.1	2.3
	Average							641	20.0	2.0	4.8	3.2	0.8
	Low-Average							646	20.1	0.8	1.9	1.3	0.3
	Lowest							646	20.1	0.1	0.7	0.4	0.2
School Size (Roll)	Smallest	698	20.0	22	494	209.0	142.8	642	20.0	22	532	220.7	157.0
	Small-Average	699	20.0	495	797	666.8	85.1	642	20.0	533	826	699.1	81.7
	Average	702	20.1	798	1019	905.7	62.8	644	20.0	827	1038	929.6	61.9
	Average-Large	699	20.0	1020	1277	1139.3	73.1	644	20.0	1039	1291	1157.5	71.9
	Largest	697	19.9	1278	2577	1525.7	214.8	641	20.0	1292	2577	1539.2	213.3
Total schools		3495	100.0					3213	100.0				

**The alternative sample used to check ethnic group composition effect robustness for MLD filtered out schools with <2 Asian (excluding Pakistani) pupils; the sample used for this purpose for ASD filtered out schools with <2 Asian (Indian/Pakistani/Bangladeshi/Asian Other pupils).*

Approach to analysis

A combination of Cox's regression and logistic regression was used to analyse the longitudinal data. Approaches to analysis differed for the primary and secondary cohorts because of structural differences in the data and implications of analytical choices for each phase.

Cox's proportional hazards regression, a form of survival analysis, facilitates an investigation of the 'risk' of an event happening over a specified period of time. For our purposes, the event is *first-time*[16] SEN identification, though we do not attribute positive or negative value to the 'risk' of this event as such. As noted previously, identification may be associated with stigma or with benefits due to additional support, and these issues go beyond the scope of our analysis. Logistic regression does not inherently incorporate the time element per se in the way that Cox regression does, but where there are structural issues with the data (as for our secondary cohort, discussed below), it does allow for the investigation of the odds of an event (SEN identification) ever occurring over a specified period of time e.g. the whole five years of a student's secondary schooling. This is substantively different from, and can be compared with, the cross-sectional results from the analysis of data from all year groups in a particular year (e.g. 2016 as in Part 1 and Part 2 of this report).

Regardless of the specific type of regression model, however, the modelling approach remained similar. Blocks of individual pupil-level and finally school-level variables were entered sequentially (i.e. variables from each stage carry over to the next stage), and these blocks were entered in the following sequence:

Model 1: Pupil ethnic group (from the baseline year) only.

Model 2: Additional variables included: Pupil gender, birth season, FSM eligibility, and normalised pupil IDACI score (from the baseline year).

Model 3: Additional variables included:

Primary -Three EYFSP scores from Reception (Communication, language and learning; problem solving, reasoning and numeracy; personal, social, and emotional development).

Secondary – Attendance, then KS2 finely grained levels in English and maths.

Model 4: Additional school level variables included: School type, school % FSM, school % ethnic group composition, school size, sourced in the baseline year.

[16] We focus on a pupil's first instance of identification for simplicity here, although further research could investigate further changes in identification over time via, for example, a multi-state, competing risks model that would account for changes in the type of need or dropping SEN identification entirely.

For the primary cohort, in which pupils began in Reception in the 2008/09 school year and completed Y6 in the 2014/15 school year, Cox Proportional Hazards (P-H) regression formed the main emphasis of the analytical approach, with time to *first* SEN identification as the focus of our analysis. Logistic regression was used as a robustness check for the Cox analysis for each focal outcome (with *whether or not a pupil was ever identified* with a particular primary type of SEN - MLD, BESD/SEMH, or ASD - as the dependent variable for logistic regression models).

For the Secondary cohort, because there was significant left-censoring[17] due to a substantial number of pupils at secondary school entry (Y7) being already identified with a primary type of SEN, the main analytical approach was logistic regression (with dependent variables defined as described above for the primary cohort). We did not consider left-censoring an issue for the primary cohort in the same way, instead taking identifications in Y1 as new (on the premise that the statutory school age range in England is 5-16, and our analysis focuses on identification in the context of compulsory schooling).

As noted above in the description of data filtering, we treated pupils who left their cohorts without having yet been identified with SEN as having left permanently (i.e. they were not included in the analysis past the initial instance of a missing record for a given year). This was to avoid problems with interval censoring in the survival analysis[18]; logistic analysis did not have the same inherent issue, and used all available information on all pupils in the dataset whether or not they left and/or re-joined their cohorts.

Interpretation

For logistic and multinomial regression results, ORs are reported and interpreted as in Part 1, and we adopt the same threshold values to assess under- and over-representation.

Cox proportional hazards regression results are reported in terms of Hazard Ratios (HR). These have an interpretation somewhat similar to ORs, but with a time aspect. An HR gives the 'risk' of identification *per unit time* (in the context of this research, one year) for a given condition (e.g. a particular ethnic group) relative to the reference group (e.g. White British pupils). So, for example, an HR of 2.0 for a minority ethnic group would indicate that this group has twice the risk of first-time identification in any particular year, relative to the White British majority reference group. These risks are compounded year on year, indicating differential rates of growth in the risk of identification for the minority relative to the majority group. This

---

[17] *In survival analysis, left-censoring occurs when the event of interest has already occurred for an individual at the earliest available time point. This is a non-trivial problem that potentially affects results and their substantive interpretation.*

[18]*. Likewise, interval censoring in survival analysis occurs when the event of interest occurs (or may occur) at an unspecified time in between time points with known information (e.g. an individual leaves the cohort without any SEN identification, then returns several years later with some SEN identification which may have occurred in any interim year for which information on the individual was missing).*

is probably best illustrated graphically, and we will see examples in the presentation of the results later.

What we found

Incidence of SEN identification over time

Simple rates of identification across years within each cohort give an initial indication of patterns over time. It is important to keep in mind that these are within-cohort patterns; that is, the trends in the rates of SEN identification are for a particular group of pupils as they progress through primary or secondary school, and as such these trends may differ from the overall patterns for all pupils in England (irrespective of Year Group) over the same period of time.

Any identified SEN over time

Tables 3.8 and 3.9 show the rates of identification with any type of SEN across primary (Reception to Y6, 2009 to 2015) and Secondary (Y7 to Y11, 2011 to 2015) cohorts. In the primary years, there is a steady increase in the rate of ascription of a particular type of SEN, with a more marked increase from 2014 to 2015; although some ethnic groups had more stable rates up to 2014, the increase in 2015 was apparent across ethnic groups. In the secondary cohort, however, the rate of ascription of a particular type of SEN was relatively stable up until 2014 with a substantial increase in 2015; this was relatively consistent across ethnic groups, although the decrease over time up to and including 2014 was more marked for some. On the other hand, trends over time in SEN identification at any level of need (including School Action, for which no primary type of SEN was reported on January School Census returns) are quite different from the trends for only School Action Plus and above (and including SEN support in 2015). In the primary cohort, there is an increase in the rates of identification at any level of need from Reception (2009) to Y3 (or Y2 for some ethnic groups), followed by annual decreases up to and including 2015. In the secondary cohort, there is an annual decrease in the rates of identification at any level of need, with a sharper decline from 2014 to 2015, and this is fairly consistent across ethnic groups.

This suggests an important consideration to inform the interpretations of subsequent findings relevant to particular types of SEN, because even though fewer pupils overall were identified as having SEN in 2015, more were identified in such a way that they would have had a type of primary need reported for the January 2015 School Census, not necessarily because more pupils had more severe needs, but because of policy changes that led to different reporting categories and requirements.

Table 3.8: Primary cohort, rates of identification (any SEN)

Primary cohort	N of pupils	*% identified (any SEN, School Action Plus and above*)*							*% identified (any SEN, any level including School Action*)*						
	R 2009	R 2009	Y1 2010	Y2 2011	Y3 2012	Y4 2013	Y5 2014	Y6 2015	R 2009	Y1 2010	Y2 2011	Y3 2012	Y4 2013	Y5 2014	Y6 2015
White Irish	1327	3.5	5.6	7.6	9.2	9.1	10.0	13.4	8.2	15.7	20.4	22.3	22.0	21.0	18.0
Traveller Irish	436	8.9	15.0	19.2	21.9	22.3	20.2	36.1	17.4	39.7	52.6	55.6	49.4	49.6	42.1
Traveller Gypsy/Roma	836	6.8	12.6	18.6	20.8	21.6	23.5	33.6	17.6	38.7	51.7	52.2	50.3	48.9	42.3
White other groups	20209	4.2	6.8	8.3	9.2	9.5	9.5	14.5	9.0	18.5	22.0	22.1	21.3	20.4	17.5
Mixed White & African	2741	4.8	7.2	8.7	9.4	10.1	10.7	15.0	10.1	19.3	22.8	22.1	22.0	20.6	17.9
Mixed White & Caribbean	6115	5.2	7.8	10.7	11.9	12.6	13.6	20.1	10.8	21.1	26.0	26.5	26.8	26.8	24.4
Mixed White & Asian	4940	3.8	5.6	6.0	6.3	6.7	6.9	11.2	7.5	14.8	16.2	16.6	16.1	15.7	13.7
Any other mixed	8249	5.1	7.7	9.3	9.8	10.3	10.7	15.7	9.6	18.1	21.9	21.9	21.4	21.3	19.0
Indian	12114	3.3	4.5	5.1	5.2	5.1	5.0	8.0	7.0	12.7	14.5	13.5	12.8	12.0	10.4
Pakistani	17449	5.7	8.1	9.1	9.6	9.6	9.7	15.1	13.4	22.8	25.9	24.6	23.3	22.6	19.5
Bangladeshi	6200	4.4	6.6	8.0	7.9	7.7	7.8	11.9	10.4	20.0	22.8	20.9	19.5	18.5	15.4
Any other Asian	7297	4.1	6.0	6.7	6.9	6.5	6.4	9.8	8.8	16.1	18.1	16.5	15.5	14.4	12.1
Black African	14444	6.9	9.8	10.7	10.8	10.6	10.6	14.7	13.5	22.7	24.8	23.6	22.5	21.7	18.3
Black Caribbean	5757	6.7	11.0	13.0	14.2	15.4	15.8	23.1	12.9	25.3	30.5	31.3	31.4	31.5	28.4
Black other groups	3170	6.6	9.8	11.9	12.5	13.4	13.1	19.0	13.6	23.9	28.6	28.7	27.8	27.6	24.1
Chinese	1591	3.8	5.8	6.6	6.0	6.3	5.0	7.5	8.9	13.2	14.3	13.4	13.2	11.1	9.4
Any other group	6674	4.3	7.1	9.3	9.6	9.6	9.4	14.3	9.9	19.7	23.3	22.7	21.5	20.9	17.4
Unknown	90550	5.4	8.1	8.8	9.5	10.0	10.3	15.9	10.2	19.0	22.0	22.6	22.0	21.2	19.4
White British	352175	4.7	6.9	8.5	9.5	10.2	10.6	16.0	9.2	17.5	21.2	22.2	22.1	21.8	19.4
Total	562274	4.9	7.2	8.7	9.4	10.0	10.4	15.7	9.7	18.2	21.6	22.3	22.0	21.5	19.1

**Up to and including the 2014 January School Census, pupils were recorded as having levels of SEN categorised as School Action, School Action Plus, or Statement; only those with School Action Plus or Statements had a type of primary need reported. After 2014, there was a transition to a new set of categories for levels of need (with SEN support replacing School Action and School Action Plus, Education Health and Care Plans replacing Statements, and all pupils with any level of need having a type of primary need recorded), but this transition was still in progress by January 2015 so that a combination of old and new categories were reported.*

Table 3.9: Secondary cohort, rates of identification (any SEN)

Secondary cohort	N of pupils	*% identified (any SEN, School Action Plus and above*)*					*% identified (any SEN, any level including School Action*)*				
	Y7 2010	Y7 2011	Y8 2012	Y9 2013	Y10 2014	Y11 2015	Y7 2011	Y8 2012	Y9 2013	Y10 2014	Y11 2015
White Irish	1793	12.0	11.9	11.5	11.1	15.2	24.7	24.5	23.2	21.8	18.5
Traveller Irish	275	41.1	42.2	44.4	49.3	56.9	65.8	68.6	64.5	65.3	57.8
Traveller Gypsy/Roma	872	26.6	26.1	25.0	27.0	34.5	56.4	55.6	51.6	47.1	39.2
White other groups	18089	9.2	9.1	8.2	8.2	12.7	26.2	24.9	21.7	20.2	15.7
Mixed White & African	2273	11.5	11.7	11.5	11.3	15.5	25.9	24.8	23.2	21.9	18.3
Mixed White & Caribbean	7167	14.1	14.1	13.5	14.3	19.3	29.9	28.4	25.9	25.5	22.3
Mixed White & Asian	4350	8.9	9.2	9.0	8.8	12.3	20.0	19.1	17.3	16.7	14.2
Any other mixed	7359	11.1	10.9	10.4	11.0	15.2	25.7	24.0	22.1	21.3	17.5
Indian	12277	5.7	5.3	5.1	4.8	7.3	16.9	15.7	13.8	11.8	9.3
Pakistani	19396	10.5	10.1	9.6	9.0	12.3	29.1	27.2	24.4	21.6	16.0
Bangladeshi	7702	8.9	8.2	7.4	7.7	11.7	24.6	23.4	20.4	18.0	13.9
Any other Asian	6356	6.1	5.6	5.5	5.0	8.1	19.3	17.9	15.9	13.6	10.0
Black African	15103	11.0	10.8	9.9	9.7	14.5	28.8	27.6	24.8	22.4	17.4
Black Caribbean	7743	15.1	16.4	15.1	15.3	21.0	33.4	32.9	30.4	28.6	24.3
Black other groups	2760	13.6	13.4	12.8	13.0	17.7	32.3	31.3	28.8	26.2	20.9
Chinese	1683	5.6	5.2	4.7	4.5	6.2	14.6	12.0	10.0	9.4	7.6
Any other group	6924	10.0	9.5	9.1	8.7	13.2	27.3	26.0	23.8	21.1	16.3
Unknown	3557	11.7	11.6	11.0	10.6	15.6	27.4	25.8	23.6	22.3	18.6
White British	418862	11.1	10.7	10.3	10.2	14.9	25.4	23.8	21.9	20.6	17.5
Total	544541	10.9	10.6	10.1	10.0	14.5	25.6	24.1	22.0	20.6	17.2

**Up to and including the 2014 January School Census, pupils were recorded as having levels of SEN categorised as School Action, School Action Plus, or Statement; only those with School Action Plus or Statements had a type of primary need reported. After 2014, there was a transition to a new set of categories for levels of need (with SEN support replacing School Action and School Action Plus, Education Health and Care Plans replacing Statements, and all pupils with any level of need having a type of primary need recorded), but this transition was still in progress by January 2015 so that a combination of old and new categories were reported.*

Type of SEN over time

Tables 3.10 and 3.11 show the rates of identification by ethnic group for each of the focal primary types of SEN (MLD, BESD/SEMH and ASD), across the primary and secondary cohorts.

ASD

Across the primary years, there is an apparent but fairly slight increase in ASD incidence over time, which appears largely consistent across ethnic groups. In the secondary cohort, the rates of ASD identification are relatively stable or very slightly increasing over time across ethnic groups.

BESD/SEMH

There is a small increase in the rates of identification over time in the primary cohort, more so for some groups than others; for example, the rate of identification rises considerably more over the period considered amongst Black Caribbean pupils than amongst White British pupils. The rates of BESD/SEMH identification are higher and also increasing in the secondary cohort over time, and this is consistent across ethnic groups with few exceptions (the Indian group, however, has a fairly stable and low rate of identification for this type of need).

MLD

There appears to be a fairly steady increase in the rates of identification over primary years, followed by a more dramatic increase in 2015 (even after considering a 'lower bound' for counting identification that does not take into account any identification at the level of SEN support in 2015 if a pupil had no identification in 2014). Across secondary years, there is a similarly steady decrease in MLD identification rates followed by a sharp increase in 2015 that is again fairly consistent across ethnic groups. The jump in the apparent incidence of MLD identification in 2015 is largely consistent across ethnic groups (with the exception of Indian and Chinese groups in the secondary cohort, which have fairly stable MLD identification rates from 2014 to 2015).

Considered alongside the overall decrease in the incidence of any SEN identification at any level, it seems that more of those pupils who would previously have been identified as having SEN with School Action were subsequently identified as having MLD than other types of primary need, when transitioned to the new SEN support category in 2015.

Table 3.10: Primary cohort, rates of identification (focal types of SEN)

Primary cohort	N of pupils	*MLD - % identified*								*BESD - % identified*								*ASD - % identified*							
	R 2009	R 2009	Y1 2010	Y2 2011	Y3 2012	Y4 2013	Y5 2014	Y6 2015 l.b.*	2015 u.b.*	R 2009	Y1 2010	Y2 2011	Y3 2012	Y4 2013	Y5 2014	Y6 2015 l.b.*	2015 u.b.*	R 2009	Y1 2010	Y2 2011	Y3 2012	Y4 2013	Y5 2014	Y6 2015 l.b.*	2015 u.b.*
White Irish	1327	0.1	0.6	1.4	2.2	1.9	2.2	3.2	3.4	0.6	0.8	1.5	2.1	1.9	1.8	2.5	3.0	0.5	0.6	0.6	0.7	0.8	0.9	1.0	1.0
Traveller Irish	436	2.3	5.6	9.2	10.3	8.5	8.6	16.1	17.3	3.4	2.1	2.7	3.7	5.9	4.9	5.8	7.0	0.0	0.5	0.3	0.6	0.6	0.3	0.6	0.6
Traveller Gypsy/Roma	836	2.0	3.5	7.0	8.4	8.3	9.1	13.2	13.9	1.3	2.0	2.7	3.1	3.0	3.5	3.1	3.6	0.2	0.5	0.9	0.9	0.8	0.9	1.3	1.3
White other groups	20209	0.3	0.7	1.4	1.9	2.0	1.9	3.7	4.0	0.6	1.2	1.6	1.8	2.0	1.9	2.0	2.3	0.5	0.7	0.7	0.9	0.9	1.0	1.2	1.2
Mixed White & African	2741	0.2	0.7	1.2	1.6	1.8	1.9	3.0	3.3	0.8	1.3	2.2	2.2	2.7	2.9	3.4	4.0	0.8	0.9	0.9	1.0	1.0	1.2	1.4	1.4
Mixed White & Caribbean	6115	0.3	0.9	2.1	2.7	3.0	3.3	5.1	5.6	1.2	2.0	3.0	3.6	3.9	4.4	5.2	5.6	0.5	0.7	0.8	0.9	1.0	1.2	1.4	1.5
Mixed White & Asian	4940	0.4	0.6	1.1	1.0	1.3	1.4	2.6	2.9	0.4	0.8	0.9	1.2	1.3	1.5	1.8	2.1	0.5	0.6	0.6	0.6	0.7	0.7	0.9	0.9
Any other mixed	8249	0.3	0.8	1.6	1.9	2.1	2.3	3.5	3.9	0.7	1.4	2.0	2.0	2.3	2.4	2.7	3.3	0.7	0.9	1.0	1.1	1.1	1.2	1.5	1.5
Indian	12114	0.3	0.7	1.1	1.2	1.2	1.2	2.3	2.6	0.3	0.4	0.5	0.5	0.5	0.5	0.5	0.6	0.3	0.3	0.4	0.5	0.5	0.5	0.5	0.5
Pakistani	17449	0.8	1.7	2.3	2.6	2.7	2.8	5.4	5.9	0.4	0.7	0.8	0.8	0.8	0.8	1.0	1.1	0.3	0.3	0.4	0.5	0.5	0.6	0.6	0.6
Bangladeshi	6200	0.4	1.0	1.5	1.8	1.9	1.9	3.4	3.7	0.3	0.5	0.5	0.4	0.4	0.5	0.9	1.1	0.6	0.6	0.6	0.7	0.8	0.8	0.8	0.8
Any other Asian	7297	0.3	0.8	1.2	1.3	1.3	1.3	2.7	2.9	0.3	0.5	0.6	0.8	0.6	0.6	0.8	1.0	0.6	0.7	0.7	0.7	0.8	0.9	0.9	0.9
Black African	14444	0.5	1.0	1.5	1.8	1.9	2.0	3.2	3.6	0.9	1.5	1.8	2.0	2.0	2.1	2.2	2.5	1.0	1.1	1.2	1.3	1.3	1.4	1.5	1.5
Black Caribbean	5757	0.5	1.5	2.3	2.6	3.0	3.0	5.2	5.8	1.3	2.9	3.9	4.3	5.1	5.4	6.2	6.8	0.9	1.0	1.2	1.3	1.3	1.5	1.7	1.7
Black other groups	3170	0.6	1.1	2.0	2.2	2.6	2.3	4.5	5.0	0.9	1.6	2.2	2.2	2.7	3.1	3.4	3.9	1.1	1.3	1.2	1.3	1.4	1.6	1.7	1.7
Chinese	1591	0.4	0.6	0.5	0.5	0.7	0.4	0.7	0.9	0.3	0.3	0.3	0.3	0.2	0.3	0.7	0.8	0.3	0.6	0.7	0.8	0.8	0.8	0.9	0.9
Any other group	6674	0.5	1.1	2.1	2.3	2.4	2.2	3.7	4.2	0.5	0.9	1.2	1.3	1.4	1.5	1.8	2.0	0.4	0.5	0.7	0.8	0.8	0.8	1.0	1.0
Unknown	90550	0.8	1.6	1.9	2.3	2.4	2.5	4.2	4.7	0.6	1.2	1.5	1.7	1.9	2.1	2.4	2.8	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.2
White British	352175	0.4	0.9	1.7	2.2	2.5	2.6	4.3	4.7	0.7	1.2	1.6	1.9	2.1	2.3	2.5	2.9	0.4	0.5	0.6	0.8	0.9	1.0	1.3	1.3
Total	562274	0.5	1.0	1.7	2.2	2.4	2.5	4.2	4.6	0.7	1.2	1.6	1.8	2.0	2.1	2.4	2.8	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.3

**l.b. denotes a 'lower bound' for 2015 identification with the relevant primary type of SEN, for which pupils with no SEN in 2014 but SEN support in 2015 were recoded as having no SEN in 2015. u.b. denotes an 'upper bound', which includes all SEN identifications with the relevant type as recorded in the pupil-level NPD data.*

Table 3.11: Secondary cohort, rates of identification over time (focal types of SEN)

Secondary cohort	N of pupils	*MLD*						*BESD/SEMH*						*ASD*					
	Y7 2010	Y7 2011	Y8 2012	Y9 2013	Y10 2014	Y11 2015 l.b.	Y11 2015 u.b.	Y7 2011	Y8 2012	Y9 2013	Y10 2014	Y11 2015 l.b.	Y11 2015 u.b.	Y7 2011	Y8 2012	Y9 2013	Y10 2014	Y11 2015 l.b.	Y11 2015 u.b.
White Irish	1793	2.4	2.3	2.3	1.8	2.4	2.6	2.5	2.6	2.8	3.3	2.7	3.5	1.8	1.8	1.7	1.6	1.8	2.0
Traveller Irish	275	17.8	16.2	17.3	13.7	11.1	11.1	11.6	15.3	15.1	23.5	25.4	28.6	0.4	0.4	0.6	1.3	2.4	2.4
Traveller Gypsy/Roma	872	11.4	10.4	10.4	8.3	9.7	10.4	6.5	7.9	7.0	10.1	11.1	11.7	0.2	0.1	0.3	0.2	0.2	0.2
White other groups	18089	2.2	2.0	1.6	1.5	2.9	3.1	1.8	2.2	2.1	2.4	2.5	2.9	0.9	0.9	0.9	0.9	1.0	1.0
Mixed White & African	2273	2.3	1.7	1.8	1.6	2.4	2.7	3.3	3.9	3.9	4.3	4.5	5.2	1.2	1.3	1.4	1.4	1.5	1.6
Mixed White & Caribbean	7167	3.3	2.9	2.4	2.0	2.8	3.2	5.1	5.7	5.6	6.5	7.1	8.0	1.3	1.4	1.4	1.6	1.7	1.7
Mixed White & Asian	4350	2.1	1.9	1.7	1.4	2.2	2.4	2.2	2.7	2.5	2.7	3.0	3.6	1.0	1.1	1.2	1.2	1.3	1.3
Any other mixed background	7359	2.3	2.0	1.8	1.9	2.9	3.0	3.2	3.6	3.6	4.2	4.3	5.0	1.5	1.5	1.4	1.5	1.6	1.7
Indian	12277	2.1	1.7	1.6	1.4	2.0	2.2	0.6	0.6	0.7	0.7	0.8	1.0	0.5	0.5	0.5	0.5	0.6	0.6
Pakistani	19396	4.3	3.7	3.4	2.7	4.0	4.3	1.1	1.4	1.4	1.6	1.8	2.1	0.5	0.5	0.5	0.5	0.5	0.6
Bangladeshi	7702	3.2	2.6	2.3	2.1	3.3	3.5	0.9	1.1	1.1	1.6	2.0	2.5	0.4	0.4	0.4	0.4	0.4	0.4
Any other Asian	6356	2.1	1.7	1.4	1.1	2.0	2.2	0.8	0.9	1.0	1.0	1.2	1.5	0.6	0.5	0.6	0.6	0.6	0.6
Black African	15103	2.7	2.4	2.0	1.9	3.5	3.9	2.5	2.7	2.5	2.7	3.0	3.5	1.1	1.0	1.0	1.0	1.1	1.1
Black Caribbean	7743	3.6	3.5	3.0	2.4	3.7	4.1	5.1	6.3	6.0	6.8	6.6	7.5	1.6	1.7	1.7	1.8	1.9	2.0
Black other groups	2760	3.8	3.4	3.0	2.4	3.8	4.3	3.3	4.0	3.8	4.5	4.3	4.8	1.4	1.4	1.3	1.3	1.3	1.3
Chinese	1683	1.2	0.7	0.6	0.7	0.8	0.8	0.5	0.6	0.5	0.4	0.5	0.6	0.7	0.7	0.8	0.8	0.8	0.8
Any other ethnic group	6924	3.0	2.4	2.2	1.8	3.2	3.6	1.7	2.1	2.3	2.6	2.8	3.1	0.5	0.6	0.6	0.6	0.6	0.6
Unknown	3557	3.1	2.9	2.4	2.0	3.1	3.2	2.6	2.9	3.0	3.3	3.7	4.1	1.0	1.1	1.1	1.1	1.3	1.4
White British	418862	3.0	2.6	2.3	2.0	3.1	3.2	2.5	2.7	2.9	3.2	3.4	3.9	1.1	1.1	1.2	1.2	1.4	1.4
Total	544541	3.0	2.6	2.3	2.0	3.1	3.3	2.4	2.7	2.8	3.1	3.3	3.8	1.0	1.1	1.1	1.2	1.3	1.3

**l.b. denotes a 'lower bound' for 2015 identification with the relevant primary type of SEN, for which pupils with no SEN in 2014 but SEN support in 2015 were recoded as having no SEN in 2015. u.b. denotes an 'upper bound', which includes all SEN identifications with the relevant type as recorded in the pupil-level NPD data.*

Secondary cohort 'Ever identified' outcomes

As noted in the Approach to Analysis above, for the secondary cohort we investigated outcomes in terms of whether or not a pupil was ever identified (during the secondary school phase) as having MLD, having BESD/SEMH, or having ASD. The incidences are higher when considering identification over time in this way, i.e. many more pupils are identified *at some point* during secondary school than are identified within a particular year examined cross-sectionally.

The rates of ever being identified with MLD, BESD/SEMH, ASD, or with any primary SEN type (SAP or above prior to 2015, and including SEN support in 2015) are given in Table 3.12.

Table 3.12: 'Ever identified' (MLD, BESD/SEMH, ASD and Any SEN) incidence by ethnic group in the secondary cohort

Ethnic group	*N*	*Ever MLD*	*Ever BESD /SEMH*	*Ever ASD*	*Ever SEN*
White Irish	1793	4.2%	6.1%	2.3%	20.1%
Traveller Irish	275	20.0%	25.1%	1.1%	52.4%
Traveller Roma	872	17.7%	14.7%	0.5%	41.5%
White other	18089	4.6%	5.1%	1.3%	17.2%
Mixed White & Black African	2273	4.3%	8.3%	1.8%	21.0%
Mixed White & Black Caribbean	7167	5.6%	12.3%	2.0%	25.5%
Mixed White & Asian	4350	3.8%	5.8%	1.5%	16.2%
Mixed other	7359	4.4%	8.1%	2.0%	20.1%
Indian	12277	3.4%	1.8%	0.7%	9.6%
Pakistani	19396	7.2%	3.7%	0.7%	17.3%
Bangladeshi	7702	5.4%	3.8%	0.5%	15.6%
Asian other	6356	3.7%	2.6%	0.7%	11.1%
Black African	15103	5.7%	6.5%	1.4%	19.7%
Black Caribbean	7743	6.9%	12.9%	2.2%	28.5%
Black other	2760	6.8%	9.0%	1.7%	23.9%
Chinese	1683	1.7%	1.3%	0.9%	8.7%
Any other	6924	5.5%	5.2%	0.7%	18.1%
Unknown	3557	5.6%	6.6%	1.6%	20.2%
White British	418862	5.2%	6.3%	1.7%	19.5%
Total	544541	5.2%	6.2%	1.6%	19.2%

Note: Counts include pupils with records in Y7 from the 2011 census; duplicate records excluded.

Consideration of changes in policy from 2014 to 2015

In order to understand and appropriately treat identification over time, it was important to consider how rates of identification may have changed from the previous (2014 and earlier) set of categories for level of SEN (School Action, School Action Plus, Statemented) to the new set of categories (from 2015 on). The shift was

not immediate; rather, in 2015 there were pupils recorded as having levels of need based on the previous set of categories, and there were others recorded as having SEN support or Education, Health and Care (EHC) Plans according to the policy change. Our approach to dealing with this in our data was to repeat analyses two ways. First, we included all SEN identification as given in the NPD for 2015, and called this an 'upper bound'. Second, we did not include as identified with a given type (MLD, BESD/SEMH, ASD) any pupil recorded as having SEN support in 2015 but who had no SEN identification in 2014, and called this a 'lower bound', on the premise that these pupils were likely to be School Action equivalent, it being relatively unlikely to have transitioned from no identification to SAP equivalent within a single year.

The lower bound analysis served as a robustness check for our results. We ultimately report our 'upper bound' analysis, on the basis that this takes into account maximum available information about pupil SEN identification. It is therefore important to take this into account in interpreting any changes (e.g. in incidence rates) from 2014 to 2015.

Table 3.13 gives the numbers and percentages of pupils in each cohort who were identified as having SEN with either the old or new (beginning in 2015) levels of need, to give the reader a sense of how much shift occurred (to the new way of defining levels of need) in the first year of the change in policy. These figures demonstrate that while only a small percentage of pupils at the highest level of need were recorded as having EHC Plans by January 2015, and the overall proportions with Statements or EHC Plans in 2015 were identical to the proportions with Statements in 2014, there was a substantial change from the previous (SA and SAP) categories for recording levels of need to the new 'SEN support' category across both primary and secondary phases.

Table 3.13: Numbers and % of pupils in each cohort recorded with old/new levels of need 2014-2015

	2014		2015	
Census Record	N (pupils)	%	N (pupils)	%
PRIMARY				
No SEN	418354	78.5	428845	80.9
School Action/SAP	99454	18.7	30261	5.7
SEN support	--	--	54153	10.2
Statemented	14980	2.8	14708	2.8
EHC	--	--	2036	0.4
SECONDARY				
No SEN	424750	79.4	438349	82.8
School Action/SAP	89470	16.7	22092	4.2
SEN support	--	--	48415	9.1
Statemented	20909	3.9	18343	3.5
EHC	--	--	2305	0.4

**Note: SAP='School Action Plus', EHC = 'Education, Health, and Care Plan'.*

Stability of SEN identification: Descriptive information

Stability of any SEN identification

Taking into consideration only identification at the level of School Action Plus (SAP) or above (Statement, EHC Plan), whether or not a pupil is identified with any SEN at any level is quite stable. In the primary cohort, roughly three-quarters (73.0%) of those with any type of SEN identified in Y2 were still identified with some type of SEN in Y6 (from the 2011 to 2015 January School Censuses). In the secondary cohort, similarly, 72.1% of those identified with any type of SEN in Y7 (2011) were still identified in Y11 (2015).

These figures drop to 60.1% for the primary cohort and 53.7% for the secondary cohort if we take into account SEN with School Action (SA) and no specific type of need recorded; this shows that identification at SA level appears to be more volatile as well as somewhat less clearly defined as we cannot ascertain from NPD data what the type of need was for a particular pupil at SA for 2014 and earlier. The figures may also be less stable than in previous years because of the sizeable decrease in the proportion of all pupils recorded as having any SEN from 22% in 2010 to 14% in 2016 (DFE, 2017), which partly reflect an increasing number of students being moved from SA to No SEN. For these reasons we focus in this chapter on SEN with School Action Plus or above.

Stability of level of identification

As shown in Table 3.14, the stability of SEN identification varies somewhat for different levels of need (None, School Action, School Action Plus, SEN support – from 2015 on, and Statement or EHC Plan).

In the primary cohort, about half (51.5%) of those with School Action in Y2 [19] were without any SEN identification in Y6, while only 11.2% were still identified at the level of School Action by Y6; 5.6% had moved to School Action Plus, 29.8% had the new 'SEN support' classification which does not show whether the actual level of need changed, and only a few had been given statements or EHC plans by Y6 (1.9%). Of those with School Action Plus in Y2, just over a quarter (27.7%) had no identified SEN by Y6, 13.3% remained with School Action plus, 38.1% had been reclassified under the new 'SEN support' category that had absorbed School Action and School Action plus, and 5.8% had been shifted to only having School Action, while 15.1% had by Y6 received statements or EHC plans. Statements and EHC plans were more stable than other levels of identification, somewhat unsurprisingly. Of those with statements in Y2, 97.0% had statements or EHC plans in Y6, and only 1.3% no longer had any SEN identified, while very small percentages had been moved out of

[19]. *Note that we use Y2 as a base here due to lower rates of identification in Reception and Y1; see Tables 3.8 and 3.10, for rates of identification in each year of Primary school for the 2009-15 (Reception to Y6) cohort.*

this higher level of need to the new 'SEN support' category (1.2%), School Action Plus (0.3%) or School Action (0.2%).

In the secondary cohort, a large percentage of those with School Action in Y7 are not identified with any SEN by Y11 (62.8%), while only 10.2% were still recorded as having SA, 2.3% had been moved up to SAP, about a quarter (24.0%) had been moved to the new 'SEN support' level classification, and very few (0.7%) had received statements or EHC plans. Of those with SAP in Y7, 33.9% no longer had any SEN recorded in Y11, 10.7% still had the same level of need recorded, 43.1% may have retained the same level of need but were re-classified with 'SEN support', only 6.8% had shifted down to SA, and 5.5% had received a statement or EHC plan. As in the primary cohort, the statement or EHC plan category was more stable; of those with a statement or EHC plan in Y7, only 1% had no SEN identified in Y11; 96.8% still had a statement or EHC plan, and very small numbers had moved to 'SEN support (1.6%), SAP (0.4%), or SA (0.2%).

Table 3.14: Primary and secondary cohorts, cross-tabulations of level of need categories, 2011 by 2015 (Y2 to Y6 and Y7 to Y11), percentages

Primary cohort		Y6					
		No SEN	School Action	School Action Plus	SEN support	State-ment / EHC plan	Total N
Y2	No SEN	92.3	1.9	0.8	4.7	0.2	412488
	School Action	51.5	11.2	5.6	29.8	1.9	68999
	School Action Plus	27.7	5.8	13.3	38.1	15.1	36051
	Statement	1.3	0.2	0.3	1.2	97.0	9336
	Total	--	--	--	--	--	526874
Secondary cohort		Y11					
		No SEN	School Action	School Action Plus	SEN support	State-ment / EHC plan	Total N
Y7	No SEN	95.0	0.9	0.6	3.4	0.1	394310
	School Action	62.8	10.2	2.3	24.0	0.7	76731
	School Action Plus	33.9	6.8	10.7	43.1	5.5	37292
	Statement	1.0	0.2	0.4	1.6	96.8	17831
	Total	--	--	--	--	--	526164

Stability of primary type of SEN

The percentages given in Table 3.15 show the stability of each type of primary SEN (including no SEN) between Y2 (2011) and Y6 (2015) for the cohort that was in Reception as of the 2009 School Census. Although the vast majority of pupils without an identified type of primary need remained without SEN identification by the end of primary school (89.8%), stability varied by the type of SEN for those who had

been identified in Y2. We concentrate here on our focal types of SEN: MLD, BESD/SEMH, and ASD.

- MLD, 27.9% of those identified as of Y2 were no longer identified with any type of SEN in Y6, and slightly less than half (47.3%) were still identified with MLD in Y6. Relatively few had switched to a different type of primary need by Y6, but of those, the most common changes were to SpLD (6.7%), BESD/SEMH (5.3%), and SLCN (4.6%).

- BESD/SEMH exhibited a pattern of stability fairly similar to that for MLD described above. Of those identified with BESD in Y2 in 2011, 28.2% were not identified with any type of SEN in Y6 in 2015, and again slightly less than half (44.7%) were still identified with SEMH in 2015 after the change in policy that eliminated BESD and introduced this new label. Of those who changed types between Y2 and Y6, the most common switches were to ASD (9.6%), MLD (6.7%), SLCN (3.7%) and SpLD (3.6%).

- ASD appears to have been a much more stably-identified type of primary need across the primary school years, as compared to MLD and BESD/SEMH. Of those identified with ASD in Y2, very few (5.4%) were without any type of SEN by Y6, while the vast majority (85.0%) were still identified with ASD in Y6. Very few pupils changed from being identified with ASD in Y2 to a different type of need in Y6; of those who did, the most common changes were to SLD (3.2%), MLD (2.2%), SLCN (1.7%) and BESD/SEMH (1.6%).

The percentages given in Tables 3.16 show the stability of each type of SEN (including no SEN) between Y7 (2011) and Y11 (2015) for the secondary cohort. As in the primary cohort results given above, the vast majority of those without SEN type identification in Y7 were still without any type of primary SEN in Y11 (92.2%), and the stability of individual types of need largely followed patterns similar to those in the primary cohort.

- For MLD, about a third (31.9%) of those identified in Y7 had no type of SEN recorded in Y11, while slightly less than half (47.5%) were still identified with MLD in Y11. Of those who changed to a different type, the most common switches were to BESD/SEMH (6.3%), SpLD (5.9%), and SLCN (2.2%).

- For BESD, 34.4% of those identified in Y7 had no SEN type recorded in Y11, while roughly half (49.2%) were still recorded as having SEMH in Y11. Of those who switched primary types of need between Y7 and Y11, the most common changes were to ASD (4.5%), MLD (4.0%) and SpLD (3.1%).

- ASD, just as in the primary cohort, was a more stably-identified type of need than MLD or BESD/SEMH. Of those identified with ASD in Y7, 7.9% had no

type of need recorded in Y11, and 82.2% were still identified with ASD as their primary type of SEN in Y11. Of the few who changed to another type of need between Y7 and Y11, the most common switches were to BESD (2.5%), SLCN (2.3%) and MLD (2.0%).

Throughout the rest of this section of the report, we focus on the instance of first identification when conducting Cox proportional hazards regression, and on whether or not a pupil was ever identified with each focal type when conducting logistic regression analyses. That is, we do not focus on changes between types of SEN or whether a pupil was identified and subsequently moved out of being identified. Future work might extend this research to account for such changes in individual pupils' identification over time, but in this report we are primarily concerned with ethnic disproportionality with regard to first identification and/or whether or not a pupil was ever identified with MLD, BESD/SEMH, or ASD.

Table 3.15: Primary cohort, cross-tabulation of primary SEN Y2 to Y6 (row percentages)

		Y6 No SEN	SpLD	MLD	SLD	PMLD	BESD/ SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA	Total N
Y2	No SEN	89.8	2.2	3.6	0.0	0.0	2.0	0.9	0.3	0.1	0.1	0.0	0.1	0.5	0.4	481488
	SpLD	29.6	42.0	9.7	1.7	0.2	4.4	4.9	3.3	0.4	0.1	0.1	0.8	1.3	1.4	2802
	MLD	27.9	6.7	47.3	2.1	0.1	5.3	4.6	2.8	0.2	0.1	0.0	0.6	1.3	1.2	9081
	SLD	2.9	2.0	7.4	70.4	3.7	1.6	2.7	7.0	0.1	0.2	0.0	1.0	0.8	0.1	1351
	PMLD	1.2	1.0	3.9	10.5	76.3	0.3	1.3	1.0	0.0	0.0	0.2	3.3	1.0	0.0	608
	BESD/SEMH	28.2	3.6	6.7	0.6	0.1	44.7	3.7	9.6	0.1	0.1	0.1	0.4	1.4	0.9	8277
	SLCN	35.3	4.0	8.9	1.3	0.1	3.9	39.2	4.7	0.4	0.1	0.0	0.4	1.0	0.6	14317
	ASD	5.4	0.4	2.2	3.2	0.2	1.6	1.7	85.0	0.0	0.0	0.0	0.1	0.3	0.0	3484
	HI	16.3	1.4	1.7	0.9	0.2	1.0	1.0	0.7	75.8	0.2	0.1	0.6	0.1	0.0	999
	VI	19.1	1.0	3.1	1.2	0.8	1.5	0.5	1.2	0.3	70.2	0.2	0.8	0.0	0.2	608
	MSI	13.8	6.3	6.3	2.5	2.5	3.8	1.3	3.8	1.3	1.3	51.3	5.0	1.3	0.0	80
	PD	18.8	1.5	4.0	2.6	1.1	1.9	1.6	1.5	0.4	0.2	0.1	64.1	2.1	0.1	2038
	Other	33.2	6.1	8.8	2.8	0.5	4.5	4.8	4.2	0.2	0.3	0.3	2.7	30.7	0.9	1741
	Total	--	--	--	--	--	--	--	--	--	--	--	--	--	--	526874

Table 3.16: Secondary cohort, cross-tabulation of primary SEN Y7 to Y11 (row percentages)

		Y11 No SEN	SpLD	MLD	SLD	PMLD	BESD/ SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA	Total N
Y7	No SEN	92.2	1.6	1.8	0.0	0.0	2.6	0.3	0.3	0.1	0.1	0.0	0.1	0.7	0.2	471044
	SpLD	33.7	50.7	4.7	0.4	0.0	4.5	1.9	1.7	0.2	0.1	0.0	0.3	1.3	0.6	8329
	MLD	31.9	5.9	47.5	1.2	0.0	6.3	2.2	1.9	0.3	0.2	0.0	0.4	1.5	0.7	15216
	SLD	3.4	1.9	7.2	76.3	2.3	1.3	1.7	4.2	0.1	0.1	0.1	1.1	0.4	0.1	1989
	PMLD	0.2	0.2	3.1	6.7	84.4	0.8	1.0	1.5	0.0	0.0	0.2	1.2	0.6	0.0	481
	BESD/SEMH	34.4	3.1	4.0	0.2	0.0	49.2	1.6	4.5	0.1	0.1	0.0	0.2	1.9	0.7	11491
	SLCN	25.4	3.6	5.6	1.1	0.0	3.4	54.2	4.4	0.4	0.1	0.0	0.4	0.9	0.4	5835
	ASD	7.9	0.8	2.0	1.7	0.1	2.5	2.3	82.2	0.0	0.0	0.0	0.2	0.3	0.1	5267
	HI	17.9	0.7	0.9	0.0	0.0	2.0	0.5	0.6	74.0	0.3	1.0	1.1	1.0	0.0	1306
	VI	14.8	1.1	0.9	0.5	0.0	1.7	0.9	0.6	0.6	76.6	0.6	0.9	0.5	0.3	650
	MSI	20.0	3.6	3.6	3.6	3.6	5.5	0.0	3.6	1.8	0.0	52.7	1.8	0.0	0.0	55
	PD	14.6	1.3	3.8	1.9	0.8	1.3	1.0	1.7	0.5	0.6	0.3	70.9	0.9	0.3	1901
	Other	45.4	5.2	5.6	0.6	0.1	8.2	1.7	2.5	0.2	0.3	0.0	2.0	27.8	0.5	2600
	Total	--	--	--	--	--	--	--	--	--	--	--	--	--	--	526164

SEN Identification over time in the primary cohort: What we found

In this section we present results from the Cox proportional hazards regression models, first without adjusting for any predictors other than ethnic group, then after controlling for additional pupil background characteristics, next after controlling for Early Years Foundation Stage Profile scores in Communication, Language and Literacy (CLL), Problem Solving, Reasoning and Numeracy (PSRN), and Personal, Social and Emotional development (PSE), and finally after controlling for school factors. Because the higher-level effects were demonstrated to be small in Part 2, we do not extend to multilevel models in this section but we do test analogous school predictors in single-level models to identify any possible school effects in the longitudinal analysis.

As in previous sections, we refrain from emphasizing inferences about some groups (e.g. Traveller groups, Unknown and Any other ethnic group) - although we still include the relevant results in tables - because these groups are either not specifically-enough defined to have substantive collective meaning or too small to inform reliable conclusions.

Appendix L provides analogous results from logistic regression analyses with the same sets of predictors as a basis for comparison, both to the results of Cox regression analyses and to the secondary cohort results presented subsequently. Appendix M further provides analogous results from models run with 'Any SEN type' as the outcome of interest, as a basis for comparison to individual focal types of primary need.

MLD in the Primary Cohort

Table 3.17 shows the HRs for each ethnic minority group (and for other predictors) in models with first MLD identification as the outcome of interest, proceeding through the various stages of hierarchical entry of predictors described above. Figures 3-1A to 3-1D provide plots of the cumulative hazards for each ethnic group (holding all other variables at reference/average values); this provides a visual display of the ways in which relative over- and under-representation change as different variables are accounted for as well as how the relative risks accumulate over time for ethnic minority groups compared to the White British reference group.

Model 1: Unadjusted results (Ethnic group only)

With only ethnic group as an explanatory variable, several Asian groups were substantially under-represented, specifically Indian (HR=0.55), Chinese (HR=0.28), Mixed White & Asian (HR=0.62), and Other Asian (HR=0.63).

Black Caribbean pupils were somewhat over-represented (HR=1.37), and Traveller pupils were substantially over-represented (HR=4.69 and 3.58 for Irish Traveller and Roma pupils, respectively).

The reduction in deviance associated with including ethnic group as a predictor is (Δ-2LL=807.84).

Model 2: Pupil background

After accounting for other pupil background characteristics, all Asian groups appeared substantially under-represented except for Pakistani pupils (HR=0.95). Additionally, Black African pupils (HR=0.51) and to a lesser extent Black Other (HR=0.73) were also under-represented at this stage.

Black Caribbean over-representation appeared to be largely attributable to other background characteristics, based on a new HR=0.91; Traveller over-representation (HR=3.05 and 2.69 for Irish Traveller and Gypsy/Roma pupils, respectively) was somewhat lessened after controlling for additional pupil background characteristics.

Each of the pupil background variables also had a significant effect based on the relevant HRs. Being born later in the year (HR=1.33 for Spring, HR=1.84 for Summer), being a boy (HR=1.66), being entitled to FSM (HR=2.05), or living in a more deprived area (HR=1.81 for a 2SD difference in neighbourhood IDACI), were all associated with a higher hazard of MLD identification for a child in any given year of primary schooling.

This model improves upon the unadjusted model, as demonstrated by a large difference in the deviance statistic (Δ-2LL=10545.69).

Model 3: Reception attainment and development

After accounting for Reception attainment and development by controlling for pupils' EYFSP scores in Communication, Language and Literacy (CLL); Problem Solving, Reasoning and Numeracy (PSRN); and Personal, Social & Emotional Development (PSED) (Model 3), no groups appeared over-represented.

All Asian groups were substantially under-represented at this stage, including the Pakistani group (HR=0.60), and the White Other group was now additionally substantially under-represented (HR=0.49). No other ethnic coefficients changes markedly after the inclusion of prior attainment.

The effects of the other pupil background variables were considerably reduced at this stage, suggesting that EYFSP score mediated much of the effect of socio-economic deprivation. Gender (HR=1.23 for being a boy) and birth season (HR=0.93 for Spring, HR=0.90 for Summer) effects also appeared smaller after controlling for Reception-year attainment and development, as the EYFSP scores were also reflecting gender and birth season differences.

The independent effects of 1SD differences in scores for CLL (HR=0.40) and PSRN (HR=0.67) were substantial and significant; higher scores were associated with considerably lower risk of being identified with MLD in any year of primary school, which is unsurprising given the way in which MLD is defined in the SEN Code of

Practice. The effect of a 1SD difference in PSED score (HR=1.14), however, was not substantial, underscoring the particular relationship of MLD to academic or cognitive aspects rather than behaviour and social-emotional development.

Accounting for Reception attainment and development via EYFSP scores substantially improved upon the previous model, with a large reduction in the deviance statistic of Δ-2LL =40344.28.

Model 4: School composition and context

Accounting for school composition and context variables (including school type, school % FSM entitlement, school size, and school % Asian) made a negligible difference to ethnic group HRs. This was also true for the pupil background and Reception attainment/development HRs.

Furthermore, most of the school variables had minimal effects of their own based on their HRs. The exceptions to this were: First, Special schools (HR=0.51) were associated with a lower risk of MLD identification, which appears counter-intuitive but may arise because special schools were more likely to serve other types of SEN (e.g. Profound and Multiple or Severe Learning Difficulties). Second, higher school proportions of pupils entitled to FSM were associated with higher risk of MLD identification, although these were not large effects (HR=1.17 for the highest quintile of schools).

Including these school variables did not lead to much improvement in model fit, with only a relatively small reduction in the deviance statistic (Δ-2LL =195.68).

Overall, the cumulative hazard plots in Figures 3-1A to 3-1D illustrate the importance of accounting for pupil background in order to understand ethnic disproportionality with regard to MLD identification; they also illustrate that although prior attainment/development had substantive importance based on their HRs, in fact accounting for this made little difference to the ethnic group cumulative hazards, which revealed substantial under-representation for most ethnic minority groups.

Table 3.17: Primary cohort (2009-2015, age 4-11) MLD: Hazard ratios

		Model 1		Model 2		Model 3		Model 4	
		Exp(B)		Exp(B)		Exp(B)		Exp(B)	
Ethnic group	White Irish	0.77	*	0.72	*	0.75	*	0.75	*
	Traveller Irish	***4.69***	*	***3.05***	*	0.97		0.95	
	Traveller Gypsy/Roma	***3.58***	*	***2.69***	*	0.78	*	0.77	*
	White other groups	0.87	*	0.79	*	***0.49***	*	***0.48***	*
	Mixed White & African	0.77	*	***0.61***	*	***0.65***	*	***0.65***	*
	Mixed White & Caribbean	1.22	*	0.89	*	0.98		0.97	
	Mixed White & Asian	***0.62***	*	***0.58***	*	***0.58***	*	***0.57***	*
	Any other mixed	0.85	*	0.68	*	0.69	*	0.69	*
	Indian	***0.55***	*	***0.52***	*	***0.48***	*	***0.48***	*
	Pakistani	1.26	*	0.95		***0.60***	*	***0.59***	*
	Bangladeshi	0.83	*	***0.55***	*	***0.36***	*	***0.36***	*
	Any other Asian	***0.63***	*	***0.55***	*	***0.42***	*	***0.41***	*
	Black African	0.84	*	***0.51***	*	***0.46***	*	***0.45***	*
	Black Caribbean	1.37	*	0.91		0.93		0.91	
	Black other groups	1.16		0.73	*	***0.63***	*	***0.62***	*
	Chinese	***0.28***	*	***0.26***	*	***0.21***	*	***0.21***	*
	Any other ethnic group	0.96		***0.66***	*	***0.45***	*	***0.44***	*
	Unknown	1.11	*	0.94	*	0.89	*	0.89	*
FSM	Entitled to FSM			***2.05***	*	1.39	*	1.36	*
Gender	Boy			***1.66***	*	1.23	*	1.23	*
Birth Season	Spring			1.33	*	0.93	*	0.93	*
	Summer			***1.84***	*	0.90	*	0.89	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.81***	*	1.23	*	1.15	*
Combined Deprivation	(IDACI 1SD+FSM)			***2.75***	*	***1.54***	*	1.46	*
EYFSP	CLL 1SD					***0.40***	*	***0.41***	*
	PSRN 1SD					0.67	*	***0.67***	*
	PSE 1SD					1.14	*	1.13	*
School Type	Foundation							1.16	*
	Academy							1.04	
	Church							0.99	
	Special							***0.51***	*
School % FSM	Highest							1.17	*
	Average-High							1.09	*
	Average							1.06	*
	Low-Average							1.01	
School % Asian	Highest							1.02	
(except Pakistani)	Average-High							1.04	*
	Average							1.01	
	Low-Average							0.95	*
School Size	Smallest							1.10	*
	Small-Average							1.11	*
	Average							1.10	*
	Average-Large							1.04	*
Deviance (-2LL)	Initial (null): 787760.63	786952.79		776407.10		736062.82		735867.14	
Δ-2LL from empty model		807.84		10545.69		40344.28		195.68	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI; Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL – Communication, Learning and Literacy; PSRN – Problem Solving, Reasoning and Numeracy; and PSE – Personal, Social and Emotional Development); and Model 4 additionally includes school variables. N=553264 pupils are included in all four models. Deviance=-2*Log-Likelihood (-2LL) and deviance change from previous model (Δ-2LL) are used to assess model fit. *=significant at the level of p<0.05.

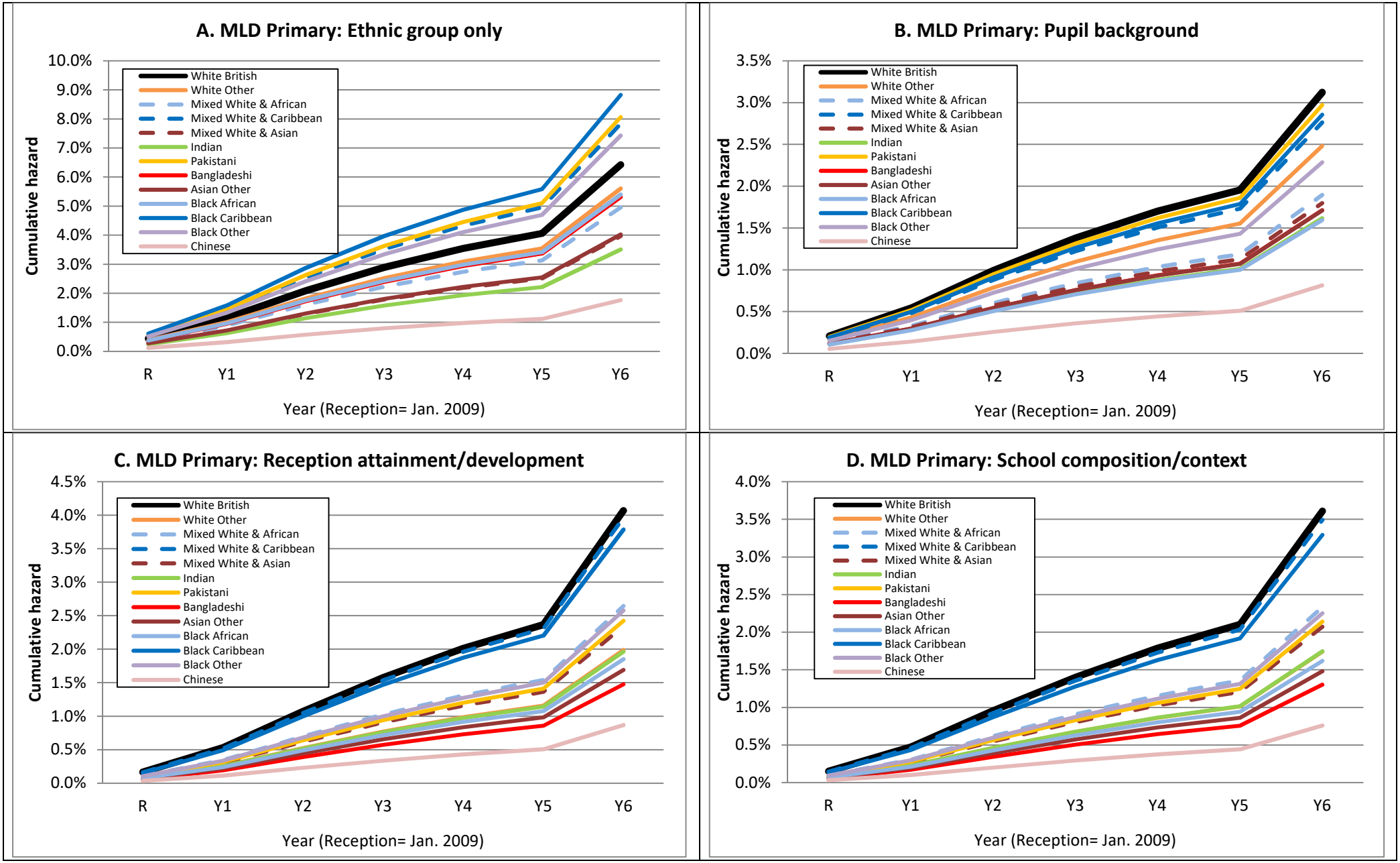


Figure 3-1: Primary cohort (2009-2015, age 4-11) MLD: Cumulative hazard plots

SEMH/BESD in the Primary Cohort

Table 3.18 shows the HRs for each ethnic minority group (and for other predictors) in models with first SEMH/BESD identification as the outcome of interest, while Figures 3-2 A to 3-2 D provide plots of the cumulative hazards for each ethnic group (holding all other variables at reference/average values).

Model 1: Unadjusted results (Ethnic group only)

With only ethnic group as an explanatory variable, Asian groups were generally substantially under-represented, while Black Caribbean (HR=2.31) and Mixed White & Black Caribbean (HR=1.86) were substantially over-represented. Black Other (HR=1.54) and Traveller (HR=2.64 and HR=1.78 for Irish Traveller and Roma respectively) groups were also substantially over-represented, and to a lesser extent so were Mixed White and Black African (MWBA) pupils (HR=1.34).

The reduction in deviance associated with including ethnic group as a predictor was Δ-2LL=1477.92.

Model 2: Pupil background

After accounting for other pupil background characteristics, Asian groups were still generally under-represented for SEMH/BESD identification. Black Caribbean (HR=1.55) and Mixed White and Black Caribbean (HR=1.35) pupils were still over-represented, but less so, suggesting that the initially observed over-representation was partly attributable to other pupil background characteristics such as socioeconomic deprivation. Other groups which were over-represented in the unadjusted model showed diminished or considerably reduced HRs indicating that much or all of the over-representation for those groups was attributable to other pupil background factors.

Most of the included pupil background characteristics variables were also strongly associated with higher risk of SEMH/BESD identification in any given year of primary schooling, including FSM entitlement (HR=2.21), being a boy (HR=3.33), and neighbourhood deprivation (HR=1.73 for a 2SD increase in IDACI score).

This model improves upon the unadjusted model, as demonstrated by a large difference in the deviance statistic (Δ-2LL=11991.98).

Model 3: Reception attainment and development

After accounting for Reception age attainment and development by controlling for pupils' EYFSP scores in Communication, Language and Literacy (CLL); Problem Solving, Reasoning and Numeracy (PSRN); and Personal, Social and Emotional Development (PSED), Asian under-representation remained largely unchanged according to the ethnic group HRs. Black Caribbean (HR=1.42) and Mixed White and Black Caribbean (HR=1.46) remained over-represented, indicating that this over-representation was not attributable merely to differences in early attainment and

development indicators. The White Other group (HR=0.74) also appeared somewhat under-represented for SEMH/BESD after accounting for Reception attainment and development.

The effects of some of the pupil background variables were somewhat reduced at this stage, including FSM (HR=1.80), gender (HR=2.46 associated with being a boy), and neighbourhood deprivation (HR=1.39 for a 2SD difference in IDACI score).

Literacy and language attainment in Reception was not strongly associated with a relative risk of SEMH/BESD identification (HR=0.96 for a 1SD difference in CLL score), while stronger numeracy and reasoning skills were somewhat counter-intuitively but not strongly associated with a higher risk of this type of identification (HR=1.23 for a 1SD difference in PSRN score). The effect of a 1SD difference in PSE score (HR=0.39), however, was substantial, with higher scores in this area of the EYFSP associated with much lower risk of SEMH/BESD identification in any given year of primary school. This is intuitively reasonable, given the closely-related definition of SEMH/BESD as a type of SEN to behaviour, social and emotional development.

Accounting for Reception attainment and development via EYFSP scores substantially improved upon the previous model, with a large reduction in the deviance statistic of Δ-2LL =15711.03.

Model 4: School composition and context

Accounting for school composition and context variables (including school type, school % FSM entitlement, school size, and school % Black Caribbean/MWBC) made a negligible difference to most ethnic group HRs. Exceptions to this were the Black Caribbean (HR=1.30) and Mixed White and Black Caribbean (HR=1.38) groups, for whom over-representation was somewhat lessened after accounting for school context/composition.

Except for neighbourhood IDACI, for which the effect was less substantial after accounting for school context/composition (HR=1.23 for a 2SD difference), the HRs for other pupil background and EYFSP variables did not change substantially.

In contrast to the results for MLD given above, more of the school variables were associated with differential risk of identification with SEMH/BESD. Being in an Academy (HR=1.86) was associated with a substantially higher risk of identification, although it is important to note that there were at most only 12 primary school academies in 2009 when the data was sourced so the result must be interpreted cautiously. Special schools were again associated with lower risk of identification (HR=0.05); however, this too must be interpreted with caution as it is likely due to the emphasis on other types of SEN in many Special schools or alternative settings. Higher school proportions of pupils eligible for FSM was associated with a higher risk of identification (HR=1.26 for the highest quintile, HR=1.21 for the second-highest),

and although these effects were not very large, they were significant and consistent across quintiles. Similarly, smaller school size was associated with a higher risk of SEMH/BESD identification (HR=1.23 for the smallest quintile of schools, HR=1.17 for the second smallest), but these were not large effects.

Schools with the highest proportions of Black Caribbean and Mixed White and Black Caribbean pupils (HR=1.18) were associated with somewhat higher odds of SEMH/BESD identification, however, this was substantially reduced when the same model was run on a sample filtering out those schools with fewer than two pupils in the Black Caribbean/MWBC group although other school composition and contextual effects remained consistent before and after this filtering (see Appendix N). Interactions between pupil ethnic group (specifically, the groups of interest for this composition effect) and school proportion Black Caribbean or Mixed White and Black Caribbean, were tested but found not to be significant (this was true for the full sample and the sample filtered for schools with <2 pupils in the groups of interest).

Including these school variables did not lead to a large improvement in model fit, with only a relatively small reduction in the deviance statistic (Δ-2LL =635.70).

On the whole, the cumulative hazard plots illustrate how the over-representation of Black Caribbean and Mixed White and Black Caribbean pupils for SEMH/BESD builds up over time (i.e. the hazard for this group has a higher slope as well as higher instantaneous risk) compared to the White British reference group, while the reverse is true for under-represented groups for whom the low risk of identification increases little over time.

Table 3.18: Primary cohort (2009-2015, age 4-11) SEMH/BESD: Hazard ratios

		Model 1		Model 2		Model 3		Model 4	
		Exp(B)		Exp(B)		Exp(B)		Exp(B)	
Ethnic group	White Irish	0.86	*	0.81	*	0.93		0.92	
	Traveller Irish	***2.64***	*	***1.70***	*	1.22		1.15	
	Traveller Gypsy/Roma	***1.78***	*	1.34	*	0.90		0.86	
	White other groups	0.88	*	0.81	*	0.74	*	0.70	*
	Mixed White & African	1.35	*	1.10		1.17		1.13	
	Mixed White & Caribbean	***1.86***	*	1.35	*	1.46	*	1.38	*
	Mixed White & Asian	***0.64***	*	***0.60***	*	***0.65***	*	***0.64***	*
	Any other mixed	1.22	*	0.99		1.05		1.01	
	Indian	***0.26***	*	***0.25***	*	***0.26***	*	***0.25***	*
	Pakistani	***0.49***	*	***0.37***	*	***0.32***	*	***0.32***	*
	Bangladeshi	***0.36***	*	***0.24***	*	***0.21***	*	***0.21***	*
	Any other Asian	***0.38***	*	***0.34***	*	***0.30***	*	***0.29***	*
	Black African	1.12	*	0.69	*	***0.66***	*	***0.61***	*
	Black Caribbean	***2.31***	*	***1.55***	*	1.42	*	1.30	*
	Black other groups	***1.54***	*	0.97		0.91		0.85	*
	Chinese	***0.32***	*	***0.32***	*	***0.31***	*	***0.29***	*
	Any other ethnic group	0.73	*	***0.50***	*	***0.45***	*	***0.42***	*
	Unknown	0.97		0.82	*	0.84	*	0.81	*
FSM	Entitled to FSM			***2.21***	*	***1.80***	*	***1.73***	*
Gender	Boy			***3.33***	*	***2.46***	*	***2.47***	*
Birth Season	Spring			1.04	*	0.90	*	0.89	*
	Summer			1.08	*	0.78	*	0.77	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.73***	*	1.39	*	1.23	*
Combined Deprivation	(IDACI 1SD+FSM)			***2.91***	*	***2.12***	*	***1.91***	*
EYFSP	CLL 1SD					0.96	*	0.97	
	PSRN 1SD					1.23	*	1.21	*
	PSE 1SD					***0.39***	*	***0.39***	*
School Type	Foundation							1.04	
	Academy							***1.86***	*
	Church							0.93	*
	Special							***0.05***	*
School % FSM	Highest							1.26	*
	Average-High							1.21	*
	Average							1.11	*
	Low-Average							1.06	*
School % Black Caribbean / MWBC	Highest							1.18	*
	Average-High							1.05	*
	Average							1.05	*
	Lowest[a]							-	
School Size	Smallest							1.23	*
	Small-Average							1.17	*
	Average							1.13	*
	Average-Large							1.04	*
Deviance (-2LL)	Initial (null): 568253.94	566776.02		554784.04		539073.01		538437.31	
Δ-2LL from empty model		1477.92		11991.98		15711.03		635.70	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI; Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL – Communication, Learning and Literacy; PSRN – Problem Solving, Reasoning and Numeracy; and PSE – Personal, Social and Emotional Development); and Model 4 additionally includes school variables. N=553264 pupils are included in all four models. Deviance=-2*Log-Likelihood (-2LL) and deviance change from previous model (Δ-2LL) are used to assess model fit. *=significant at the level of p<0.05.

[a]Note: Only 3 categories are given for School % Black Caribbean/Mixed White & Black Caribbean due to the large number of schools with 0%.

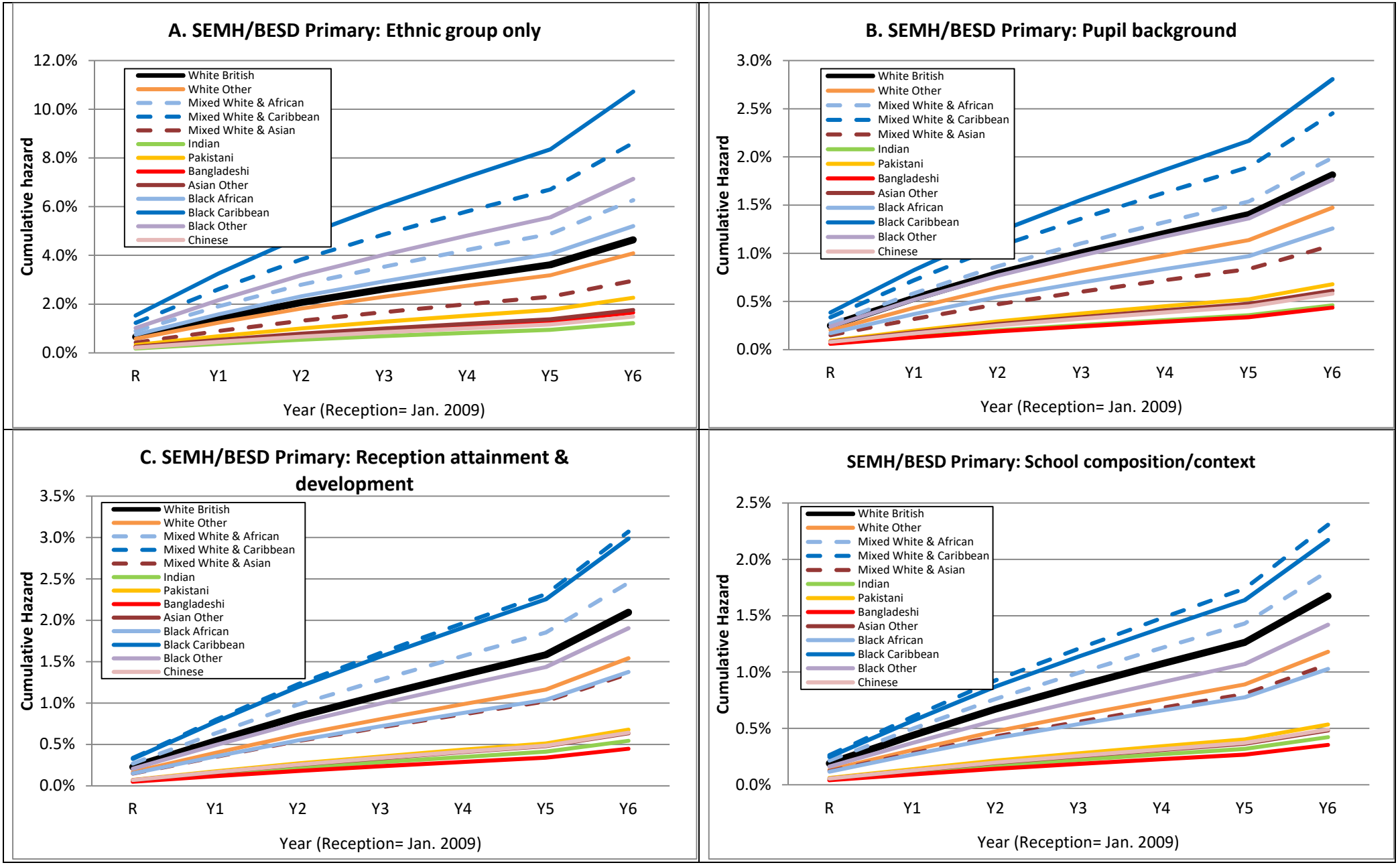


Figure 3-2: Primary cohort (2009-2015, age 4-11) SEMH/BESD: Cumulative hazard plots

ASD in the Primary Cohort

Table 3.19 shows the HRs for each ethnic minority group (and for other predictors) in models with first ASD identification as the outcome of interest, while Figures 3-3 A to 3-3 D provide plots of the cumulative hazards for each ethnic group (holding all other variables at reference/average values).

Model 1: Unadjusted results (Ethnic group only)

With only ethnic group as an explanatory variable (Model 1), Black groups appeared to be generally somewhat over-represented while some Asian groups were under-represented (specifically, Indian and Pakistani pupils, with HR=0.44 and HR=0.43, respectively). Traveller groups were also under-represented for ASD identification in the primary cohort at this stage of analysis.

There was a small reduction in deviance associated with including ethnic group as a predictor (Δ-2LL=144.99).

Model 2: Pupil background

After accounting for other pupil background characteristics (Model 2), the HRs for each of the over-represented Black groups were reduced, but only slightly so. Black Caribbean (HR=1.35) and Black Other (HR=1.77) pupils were still over-represented, while Black African pupils (HR=1.31) still had an HR that was borderline according to our thresholds, suggesting that apparent over-representation of these groups was not merely a consequence of the other included pupil background factors. Indian (HR=0.43), Pakistani (HR=0.41) and Traveller groups remained under-represented with negligible affected HRs, indicating that this under-representation was not attributable to other pupil background characteristics such as deprivation and gender.

In fact, most of the included pupil background characteristics had little association with the risk of ASD identification in any given year of primary schooling, except for gender; being a boy was associated with much higher risk of ASD identification in any year of a pupil's primary school education (HR=5.13).

This model improves upon the unadjusted model, as demonstrated by a difference in the deviance statistic of Δ-2LL=2070.27. Notably, however, pupil background variables were of far less consequence for model fit for ASD than for the other focal outcomes discussed above.

Model 3: Reception attainment and development

After accounting for Reception age attainment and development by controlling for pupils' EYFSP scores in Communication, Language and Literacy (CLL); Problem Solving, Reasoning and Numeracy (PSRN); and Personal, Social and Emotional Development (PSE) (Model 3), additional Asian groups were under-represented (including Chinese, HR=0.70; Bangladeshi, HR=0.69; Asian Other, HR=0.66), while

the previously under-represented groups remained so. Meanwhile, much of the Black over-representation appeared to be accounted for once Reception attainment and development were included as controls, as the HRs for the relevant groups decreased (including for the Black African group, HR=1.06; Black Caribbean, HR=1.09; and Black Other, still over-represented but considerably reduced to HR=1.46).

The HRs of pupil background characteristics were somewhat reduced after accounting for Reception attainment and development, to the point that some appeared counter-intuitive; for example, Summer-born children (HR=0.57) and combined deprivation (calculated as a 1SD change in neighbourhood IDACI plus FSM, HR=0.68) are pupil background factors known to be associated with higher odds of SEN identification based on previous findings, but here look as though they are working in the opposite direction. This suggests that interactions may exist between these pupil background variables and Reception attainment/development variables; however, as our main interest here is in the effect of controlling for these additional pupil variables to obtain improved model fit and identify any resulting change in ethnic group coefficients, in the interest of parsimony we leave this as a possibility for future further investigation.

Of the three areas of the EYFSP included in the analysis, only PSE score had a significant and large association with the risk of ASD identification, with a 1SD difference in PSE score (HR=0.31) associated with much lower risk of identification. The HR for CLL (HR=0.94) was not even significant at the $p<0.05$ level, and an increase of 1SD in PSRN score was significant and associated with only very slightly higher odds of identification (HR=1.10).

Accounting for Reception attainment and development via EYFSP scores substantially improved upon the previous model, with a large reduction in the deviance statistic of Δ-2LL =9122.18.

Model 4: School composition and context

Accounting for school composition and context variables (including school type, school % FSM entitlement, school size, and school % Asian) made a negligible difference to the ethnic group, pupil background characteristic and Reception attainment and development HRs, with the exception of a slightly raised value for neighbourhood deprivation (HR=0.92 for a 2SD difference in IDACI score).

Most of the school variables did not have significant effects at the $p<0.05$ level; this is likely a consequence of the low incidence of ASD identification as compared to the other focal types of primary SEN. Of those that were significant, Special schools were associated with a higher risk of ASD identification (HR=2.47), which essentially flags up that pupils identified with ASD are more likely to be placed in this type of setting as compared to the other two focal outcomes.

The highest two quintiles of school % FSM entitlement were associated with somewhat lower risk of identification (HR=0.75 for the highest and second-highest quintile). This indicates that when holding pupil level risk factors constant (PSED scores and individual socio-economic deprivation), there is a lower likelihood of ASD identification in more deprived schools/neighbourhoods. This is an interesting finding which supports some hypotheses in the academic literature (e.g. Durkin et al, 2010). We will return to this further in Part 4 where we have the opportunity to work with family socio-economic status variables like parental occupation, education and income.

There was also an association between higher proportions of Asian students in a school (here defined based on Indian, Pakistani, Bangladeshi and Asian Other groups according to previously established patterns of ASD under-identification for these four groups) and higher likelihood of ASD identification, which was not quite linear and may appear counter-intuitive. The highest two quintiles of school % Asian (HR=1.19 and HR=1.24, respectively) in particular were associated with higher levels of ASD identification than the lowest quintile, after control for all other pupil and school risk factors. As a robustness check, the same model was run with a sample excluding schools with the smallest numbers of pupils in the groups of interest (specifically, <2); this showed that the highest quintile was no longer associated with higher risk of identification, and indeed had reversed the direction of this association (HR=0.87), while the second highest (HR=1.23) and middle (HR=1.16) retained fairly consistent associations with somewhat and slightly higher risks of identification. As a further check for this variable, for which the effect was not straightforwardly interpreted, models were run on the full and filtered (no fewer than two Asian pupils) samples in which school % Asian was the only school variable; this showed similar patterns to the models run with the other school variables included, so we can be reasonably confident that this compositional effect was not an artefact of an excluded interaction with another school variable. That said, we must interpret these effects with caution, as the sensitivity of this effect to different filtering highlights a challenge of investigating such a low-incidence outcome parcelled out across a large number of other categories (ethnic group, etc.) even in national population data (e.g. building in an additional interaction term would be likely to lead to very small counts across cross-tabulated categories of the relevant variables, which in turn would risk results being more statistical artefact than substantive finding).

Ultimately, including these school variables led to a relatively small improvement in model fit, with only a relatively small reduction in the deviance statistic (Δ-2LL =256.14).

In contrast to the other two focal outcomes, the cumulative hazard plots for ASD show less substantial differences in the slopes of cumulative hazards for the various ethnic groups. Visually, it is apparent that accounting for Reception attainment / development makes the greatest difference in the cumulative hazard rates and the ordering of ethnic groups at any particular time).

Table 3.19: Primary cohort (age 5-11) ASD: Hazard ratios

		Model 1		Model 2		Model 3		Model 4	
		Exp(B)		Exp(B)		Exp(B)		Exp(B)	
Ethnic group	White Irish	1.15		1.16		1.48		***1.52***	
	Traveller Irish	***0.40***		***0.39***		***0.14***		***0.18***	
	Traveller Gypsy/Roma	***0.18***		***0.18***		***0.08***	*	***0.09***	*
	White other groups	0.99		0.98		0.78	*	0.79	*
	Mixed White & African	1.20		1.19		1.12		1.09	
	Mixed White & Caribbean	1.02		0.99		1.06		1.01	
	Mixed White & Asian	0.93		0.93		1.00		0.98	
	Any other mixed	1.34	*	1.30	*	1.25	*	1.23	
	Indian	***0.44***	*	***0.43***	*	***0.41***	*	***0.41***	*
	Pakistani	***0.43***	*	***0.41***	*	***0.29***	*	***0.29***	*
	Bangladeshi	0.95		0.89		0.69	*	0.71	*
	Any other Asian	0.89		0.87		***0.66***	*	***0.65***	*
	Black African	1.39	*	1.31	*	1.06		1.08	
	Black Caribbean	1.43	*	1.35	*	1.09		1.12	
	Black other groups	***1.90***	*	***1.77***	*	1.46	*	1.42	*
	Chinese	0.77		0.81		0.70		0.71	
	Any other ethnic group	0.71	*	***0.66***	*	***0.53***	*	***0.55***	*
	Unknown	1.05		1.02		0.97		0.97	
FSM	Entitled to FSM			1.06		0.76	*	0.81	*
Gender	Boy			***5.13***	*	***3.16***	*	***3.18***	*
Birth Season	Spring			0.99		0.79	*	0.82	*
	Summer			0.89	*	***0.57***	*	***0.61***	*
Neighbourhood Depr.	Normalised IDACI 2SD			1.12	*	0.78	*	0.92	
Combined Deprivation	(IDACI 1SD+FSM)			1.12	*	0.68	*	0.78	*
EYFSP	CLL 1SD					0.94		0.91	*
	PSRN 1SD					1.10	*	1.19	*
	PSE 1SD					***0.31***	*	***0.33***	*
School Type	Foundation							1.08	
	Academy							1.01	
	Church							0.98	
	Special							***2.47***	*
School % FSM	Highest							0.75	*
	Average-High							0.75	*
	Average							0.98	
	Low-Average							1.05	
School % Asian	Highest							1.19	*
(Indian/Pakistani/	Average-High							1.24	*
Bangladeshi/Asian other)	Average							1.12	*
	Low-Average							0.94	
School Size	Smallest							1.01	
	Small-Average							1.08	
	Average							1.06	
	Average-Large							1.02	
Deviance (-2LL)	Initial (null): 109150.04	109005.05		106934.78		97812.60		97556.46	
Δ-2LL from empty model		144.99		2070.27		9122.18		256.14	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI; Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL – Communication, Learning and Literacy; PSRN – Problem Solving, Reasoning and Numeracy; and PSED – Personal, Social and Emotional Development); and Model 4 additionally includes school variables. N=553264 pupils are included in all four models. Deviance=-2*Log-Likelihood (-2LL) and deviance change from previous model (Δ-2LL) are used to assess model fit. *=significant at the level of p<0.05.

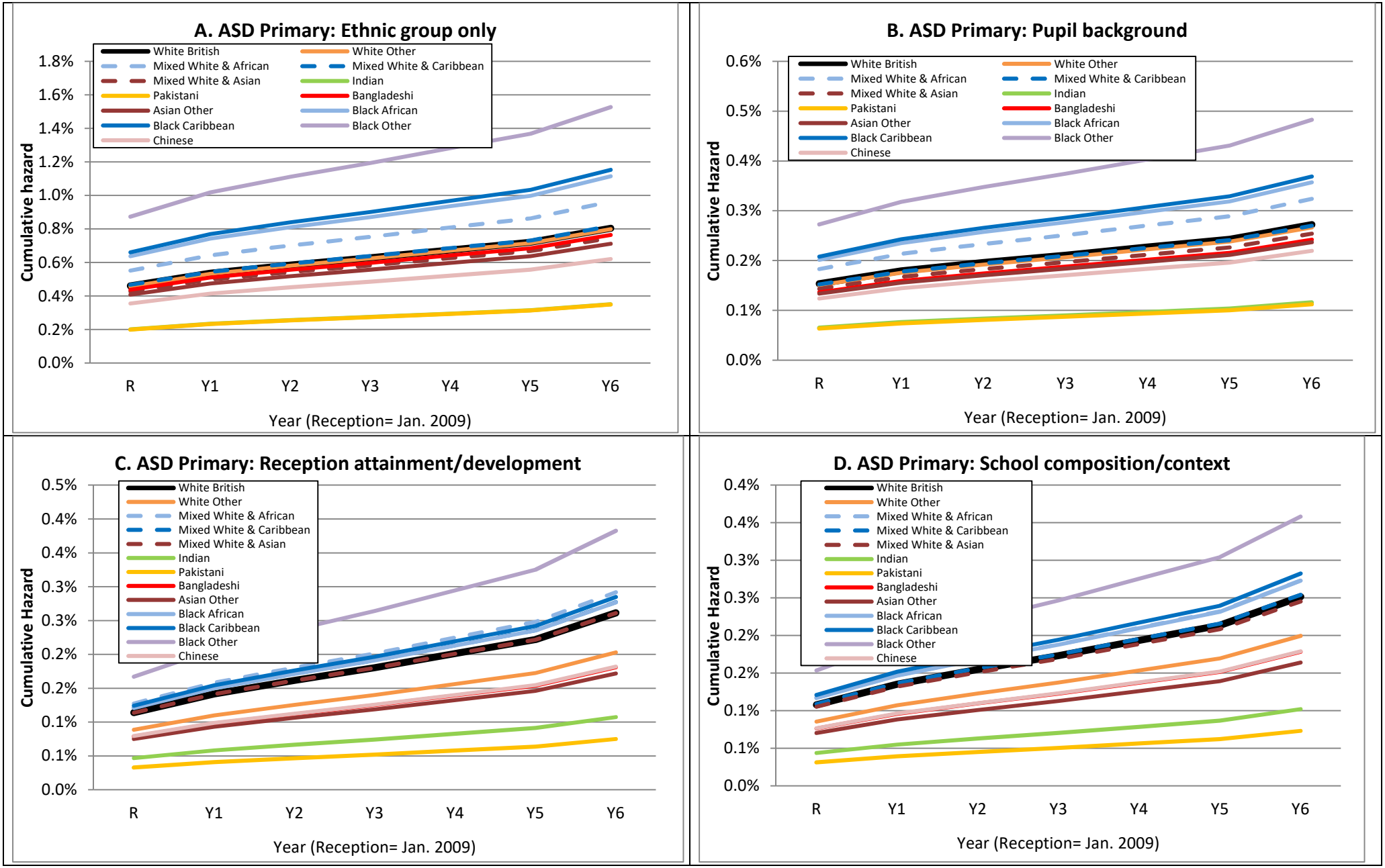


Figure 3-3: Primary cohort (2009-2015, age 4-11) ASD: Cumulative hazard plots

SEN Identification over time in the secondary cohort: What we found

In this section we present results from logistic regression models with whether or not a pupil was ever identified with MLD, SEMH/BESD or ASD as the outcomes of interest, first without adjusting for any predictors other than ethnic group, then after controlling for additional pupil background characteristics, next after controlling for a measure of attendance (persistent absence), after that controlling for prior attainment at KS2, and finally after controlling for school factors. To maintain some degree of consistency throughout this chapter, we do not extend to multilevel models in this section, but we do test analogous school predictors in single-level models to identify any possible school effects that might serve as a robustness check for the multilevel analysis.

Appendix M includes results from comparable models run with 'Any SEN type' as the outcome of interest, as a basis for comparison to individual types of primary need.

MLD in the Secondary Cohort

Table 3.20 shows the ORs for each ethnic minority group (and for other predictors) in models with MLD identification as the outcome of interest, proceeding through the various stages of hierarchical entry of predictors described above.

Model 1: Unadjusted results (Ethnic group only)

With only ethnic group as an explanatory variable, several Asian groups were under-represented, specifically Indian (OR=0.67), Chinese (OR=0.30), Mixed White & Asian (OR=0.75), and Other Asian (OR=0.72).

Black Caribbean (OR=1.38) and Black Other (OR=1.37) pupils were somewhat over-represented, and Traveller pupils were substantially over-represented (OR=5.09 and 4.04 for Irish Traveller and Roma pupils, respectively). Pakistani pupils (OR=1.42) were considerably more over-represented than was the case in the primary cohort for a comparable logistic regression model (Appendix L).

Model 2: Pupil background

After accounting for other pupil background characteristics, all Asian groups appeared substantially under-represented except for Pakistani pupils (OR=0.94). Additionally, Black African pupils (OR=0.58) and Mixed White & African pupils (OR=0.55) were also under-represented at this stage of analysis.

Black Caribbean over-representation appeared to be largely attributable to other background characteristics, based on a new OR=0.84, as was the case in the primary cohort; Traveller over-representation (OR=2.48 and 2.74 for Irish Traveller and Roma pupils respectively) was somewhat lessened after controlling for additional pupil background characteristics as well.

Each of the included pupil background characteristics variables also had a significant effect based on the relevant ORs. Being younger in the year group (OR=1.22 for spring born, OR=1.44 for summer born), being a boy (OR=1.47), being entitled to FSM (OR=2.21), or living in a more deprived area (OR=2.02 for a 2SD difference in neighbourhood IDACI), were all associated with a higher odds of MLD identification over the course of secondary school. Pupils who had been persistent absentees during Y6 of primary school (Model 2b), also had a substantial association, doubling the odds of MLD identification relative to pupils who were not persistent absentees (OR=2.11).

This model improves considerably upon the unadjusted model, as demonstrated by the difference between the Nagelkerke's Pseudo R squared values (0.4% in Model 1; 5.5% in Model 2; and 5.7% in Model 2b).

Model 3: Prior attainment

After accounting for KS2 prior attainment in English and maths, no groups remained over-represented. There were some changes to ethnic group ORs after accounting for KS2 attainment. Indian (OR=0.83) and Mixed White and Asian (OR=0.86) under-representation was somewhat mitigated, while other Asian groups (except Pakistani, with OR=0.82) as well as Black African (OR=0.64) and Mixed White and African (OR=0.67) groups remained more substantially under-represented. Traveller groups were no longer over-represented, indicating that the previously observed over-representation could be attributed to prior attainment.

The effects of all of the pupil background variables were considerably reduced at this stage, suggesting that prior attainment might be accounting for pupil socioeconomic deprivation in a more fine-grained way than, say, a binary FSM indicator (OR=1.30 compared to 2.18 in the previous model). Gender (OR=1.04 for being a boy) and birth season (OR=1.04 for Spring, OR=1.10 for Summer) associations with odds of being identified with MLD over the course of secondary school also appeared less strong after controlling for prior attainment, as prior attainment may already have reflected gender and birth season differences.

The independent effects of 1SD differences in scores for KS2 English (OR=0.39) and Maths (OR=0.64) were substantial and significant; higher scores were associated with considerably lower odds of ever being identified with MLD in secondary school (which is analogous to the results found for the primary cohort based on Reception attainment).

Accounting for prior attainment substantially improved upon the previous model, with a Nagelkerke's Pseudo R squared of 32.4% suggesting that these prior attainment

measures were powerful predictors of whether or not a pupil was ever identified with MLD in secondary school.[20]

Model 4: School composition and context

Accounting for school composition and context variables (including school type, school % FSM entitlement, school size, and school % Asian) made little difference to most ethnic group ORs (although Pakistani pupils appeared somewhat under-represented at this stage of analysis, OR=0.75). This was largely also true for the pupil background and prior attainment ORs, although the effect of neighbourhood deprivation was somewhat reduced (OR=1.23 for a 2SD difference in IDACI).

Some school variables had significant associations with the odds of ever being identified with MLD. Being in a Converter Academy was associated with slightly lower odds of identification (OR=0.75), while being in a Grammar school was associated with substantially lower odds of identification (OR=0.44), which is consistent with an intuitive understanding of selective admissions policies as well as with results from Part 2 of this report. Special schools (OR=0.50) were associated with lower odds of MLD identification, quite possibly because this SEN type is less likely to warrant special school placement than is a more "severe" learning difficulty. As was the case for the primary cohort, higher percentages of FSM in a school were associated with higher odds of identification (OR=1.23 and OR=1.19 for the highest two quintiles), although this was not a very large effect it appeared relatively linear and consistent across quintiles. There was no marked effect of ethnic group composition, with the highest percentage Asian (excluding Pakistani) quintiles associated with only slightly higher odds of identification (OR=1.09 and OR=1.08 for the highest two quintiles). Finally, the smallest schools (OR=0.78) were associated with slightly lower odds of ever being identified with MLD in secondary school.

Including these school variables led to only a small improvement in predictive power, with a Nagelkerke's Pseudo R squared of 33.1%.

20. An interaction term was subsequently included between the persistent absence variable and KS2 English and Maths attainment (Model 3b), because including the prior attainment controls led to a counter-intuitive OR=0.85 for persistent absence; interaction terms were significant (OR=1.20 for English by persistent absence, OR=1.21 for Maths by persistent absence) but including them in the model did not lead to a substantially higher Nagelkerke's Pseudo R squared value (32.5%). This did aid in substantive interpretation of the absence effect, however. These interactions are visualised in Appendix O.

Table 3.20: Secondary Cohort (age 11-16) MLD: Odds ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 2b Exp(B)		Model 3 Exp(B)		Model 3b Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	0.76	*	0.67	*	***0.67***	*	0.92		0.93		0.93	
	Traveller Irish	***5.09***	*	***2.48***	*	***2.10***	*	1.05		1.09		0.98	
	Traveller Gypsy/Roma	***4.04***	*	***2.74***	*	***2.44***	*	***0.63***	*	***0.65***	*	***0.54***	*
	White Other	0.89	*	0.75	*	0.75	*	***0.51***	*	***0.51***	*	***0.46***	*
	Mixed White & African	0.75	*	***0.55***	*	***0.55***	*	0.67	*	0.67	*	***0.63***	*
	Mixed White & Caribbean	1.14	*	0.80	*	0.80	*	0.94		0.94		0.91	
	Mixed White & Asian	0.75	*	0.68	*	0.67	*	0.86		0.86		0.85	
	Mixed Other	0.85	*	***0.67***	*	***0.67***	*	0.77	*	0.77	*	0.75	*
	Indian	***0.67***	*	***0.62***	*	***0.63***	*	0.83	*	0.83	*	0.78	*
	Pakistani	1.42	*	0.94		0.95		0.82	*	0.82	*	0.75	*
	Bangladeshi	1.05		***0.54***	*	***0.55***	*	***0.66***	*	***0.66***	*	***0.59***	*
	Asian Other	0.72	*	***0.61***	*	***0.62***	*	***0.62***	*	***0.62***	*	***0.56***	*
	Black African	1.10	*	***0.58***	*	***0.60***	*	***0.64***	*	***0.64***	*	***0.59***	*
	Black Caribbean	1.38	*	0.84	*	0.85	*	0.90		0.90		0.83	*
	Black Other	1.37	*	0.82	*	0.83	*	0.76	*	0.76	*	0.72	*
	Chinese	***0.30***	*	***0.28***	*	***0.29***	*	***0.39***	*	***0.39***	*	***0.40***	*
	Any other group	1.04		***0.64***	*	***0.64***	*	***0.55***	*	***0.55***	*	***0.48***	*
	Unknown	1.19	*	1.09		1.09		1.06		1.07		1.06	
FSM	Entitled to FSM			***2.21***	*	***2.18***	*	1.30	*	1.30	*	1.27	*
Gender	Boy			1.47	*	1.47	*	1.04	*	1.04	*	1.07	*
Birth season	Spring			1.22	*	1.22	*	1.04	*	1.04	*	1.03	
	Summer			1.44	*	1.44	*	1.10	*	1.10	*	1.07	*
Neighbourhood Depr.	Normalised IDACI 2SD			***2.02***	*	***2.01***	*	1.38	*	1.38	*	1.23	*
Combined Depr.	(IDACI 1SD + FSM)			***3.15***	*	***3.08***	*	***1.53***	*	***1.53***	*	1.41	*
Attendance	Persistent absence (>63 days)					***2.11***	*	0.85	*	***1.69***	*	***1.59***	*
Prior attainment	KS2 English Finely Graded Level							***0.39***	*	***0.39***	*	***0.37***	*
	KS2 Maths Finely Graded Level							***0.64***	*	***0.64***	*	***0.60***	*
Attend.* Prior attainment	KS2 English by Persistent abs.									1.20	*	1.18	*
	KS2 Maths by Persistent abs.									1.21	*	1.16	*
School type	Foundation											0.99	
	Academy-Converter											0.75	*
	Academy-Sponsored											0.96	
	Church											0.98	
	Selective/Grammar											***0.44***	*
	Special/PRU/AP											***0.50***	*
School % FSM	Highest											1.23	*
	Average-High											1.19	*
	Average											0.99	
	Low-Average											0.96	
School % Asian (except Pakistani)	Highest											1.09	*
	Average-High											1.08	*
	Average											1.05	
	Low-Average											1.01	
School Size	Smallest											0.78	*
	Small-Average											0.97	
	Average											1.03	
	Average-Large											1.01	
	Nagelkerke Pseudo R Squared	0.004		0.055		0.057		0.324		0.325		0.331	
	Model-specific N (# of pupils)	392708		392708		392465		392465		392465		392230	

*Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 2b includes a persistent absence indicator (absent>63 days over a year); Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 3b additionally includes an interaction between KS2 scores and attendance, and Model 4 additionally includes school variables. Note: *=Significant at the level of p<0.05. "Depr."=Deprivation; "Attend."=Attendance.*

SEMH/BESD in the Secondary Cohort

Table 3.21 shows the ORs for each ethnic minority group (and for other predictors) in models with SEMH/BESD identification as the outcome of interest, proceeding through the various stages of hierarchical entry of predictors described above.

Model 1: Unadjusted results (ethnic group only)

With only ethnic group as an explanatory variable, most Asian groups were under-represented while Black Caribbean (OR=2.20) and Mixed White and Black Caribbean (OR=2.08) pupils were over-represented, as were Traveller and to a lesser extent Black Other (OR=1.41) groups. Overall patterns were not dissimilar to the analogous logistic model for the primary cohort for which results are given in Appendix L.

Model 2: Pupil background

After accounting for other pupil background characteristics, Asian groups remained under-represented, joined now also by the White Other (OR=0.71) and Black African (OR=0.52) groups. Over-representation of Black Other pupils was been accounted for, and Traveller groups were still over-represented but substantially less so. Black Caribbean and Mixed White & Black Caribbean over-representation (OR=1.32 and OR=1.45) were substantially lessened, although not fully accounted for by other pupil background factors. Interestingly, socioeconomic deprivation and other pupil background appeared to account for more of the over-representation of these two groups than was true in the primary cohort; this is also slightly different to previous cross-sectional results in which other pupil background factors accounted for less of the Black Caribbean over-representation.

Most of the included pupil background characteristics variables also had a significant association with identification based on the relevant ORs: Being a boy (OR=1.98), being entitled to FSM (OR=2.80) and living in a more deprived area (OR=1.99 for a 2SD difference in neighbourhood IDACI) were strongly associated with higher odds of ever being identified with SEMH/BESD in secondary school. Birth season was not associated with SEMH/BESD. Persistent absence in Y6 (model 2b) was strongly associated with subsequent SEMH/BESD (OR=2.85).

This model improves considerably upon the unadjusted model, as demonstrated by the difference between the Nagelkerke's Pseudo R squared values (1.0% in Model 1; 8.8% in Model 2; 9.2% in Model 2b).

Model 3: Prior attainment

After accounting for KS2 prior attainment in English and maths, many ethnic group ORs were little changed, but there were differences for some over-represented groups. Black Caribbean (OR=1.37) and Mixed White & Black Caribbean (OR=1.53) showed exacerbated over-representation (although only on a small scale for the former). Traveller groups, on the other hand, appeared less over-represented.

The associations of most pupil background variables were somewhat accounted for by controlling for prior attainment, for example, the effects of FSM (OR=2.27), being a boy (OR=1.83), neighbourhood deprivation (OR=1.74 for a 2SD difference in IDACI), and attendance (OR=2.12 for persistent absence).

The independent effects of 1 SD change in scores for KS2 English (OR=0.80) and Maths (OR=0.79) were significant, with higher scores associated with slightly lower odds of ever being identified with SEMH/BESD in secondary school. However this effect was small relative to the prior attainment association with MLD in the secondary cohort. This makes some intuitive sense given the nature of SEMH/BESD identification and its definitional relationship to externalising and internalising disorders rather than cognitive/academic performance or ability.

Accounting for prior attainment improved upon the previous model but not by a large amount, with a Nagelkerke's Pseudo R squared of 12.2% suggesting that these prior attainment measures were less powerful predictors of whether or not a pupil was ever identified with SEMH/BESD in secondary school than was the case for MLD identification.

Model 4: School composition and context

Accounting for school composition and context variables (including school type, school % FSM entitlement, school size, and school % Black Caribbean and Mixed White and Black Caribbean) made little difference to most ethnic group ORs. Notably, however, including school variables seemed here to account for a good deal of the Black Caribbean (OR=1.16) and Mixed White & Black Caribbean (OR=1.37) over-representation observed in preceding models, suggesting (in line with what was found with regard to the effect of accounting for school clustering in Part 2) that at least some of this over-representation may be attributed to differences between schools, an effect that is particular to the secondary phase as over-representation of these groups was not accounted for in analogous primary cohort logistic regression models for which results are given in Appendix L.

There was little change in the pupil background and prior attainment ORs after accounting for school composition and context, although the effect of neighbourhood deprivation was somewhat reduced (OR=1.47 for a 2SD difference in IDACI).

Some school variables had small to moderate associations with the odds of ever being identified with SEMH/BESD.

- Being in a selective Grammar school was associated with lower odds of identification (OR=0.67).

- As was the case for the primary cohort, higher percentages of FSM in a school were associated with higher odds of SEMH/BESD identification (OR=1.39, OR=1.29 and OR=1.18 for the top three quintiles). This was a

larger effect than for MLD identification, and it appeared relatively linear and consistent across quintiles.

- There was an effect of ethnic group composition; schools with the highest percentage in the combined Black Caribbean/MWBC quintiles were associated with somewhat higher odds of identification (OR=1.24 and OR=1.14 for the highest two quintiles); this was robust to filtering dropping schools with fewer than two pupils in the groups of interest, and in fact this increased the effect size (see Appendix N). Interactions between pupil ethnic group (specifically, the groups of interest for this composition effect) and school proportion Black Caribbean/MWBC, were tested but found not to be significant (this was true for the full sample and the sample filtered for schools with <2 pupils in the groups of interest).

- The smallest schools appeared to be associated with somewhat higher odds of ever being identified with SEMH/BESD (OR=1.45), and special schools with lower odds of identification (OR=0.38). However these findings can be disregarded because of substantial collinearity. The smallest 20% of schools by roll size were attended by just 9,200 pupils or 1.7% of the cohort, and Special schools were attended by just over 8,100 pupils (1.5% of the cohort). There was substantial overlap between these variables with 93% of the special school pupils (n=7525) attending the smallest schools. Dropping the school size and special school variables made no difference to the pupil level coefficients and further strengthened the ethnic composition and %FSM results.

Including these school variables led to a very small improvement in predictive power, with a Nagelkerke's Pseudo R squared of 12.7%.

Table 3.21: Secondary cohort (age 11-16) SEMH/BESD: Odds ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 2b Exp(B)		Model 3 Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	0.90		0.79		0.77	*	0.86		0.84	
	Traveller Irish	***5.27***	*	***2.28***	*	***1.80***	*	1.38		1.31	
	Traveller Gypsy/Roma	***2.97***	*	***1.89***	*	***1.59***	*	1.00		0.92	
	White Other	0.84	*	0.71	*	0.71	*	0.68	*	***0.62***	*
	Mixed White & African	1.27	*	0.91		0.92		1.01		0.93	
	Mixed White & Caribbean	***2.08***	*	1.45	*	1.44	*	***1.53***	*	1.37	*
	Mixed White & Asian	0.89		0.80	*	0.80	*	0.89		0.85	*
	Mixed Other	1.30	*	1.01		1.01		1.09		1.02	
	Indian	***0.27***	*	***0.25***	*	***0.25***	*	***0.28***	*	***0.26***	*
	Pakistani	***0.66***	*	***0.42***	*	***0.42***	*	***0.41***	*	***0.37***	*
	Bangladeshi	***0.52***	*	***0.25***	*	***0.26***	*	***0.28***	*	***0.25***	*
	Asian Other	***0.43***	*	***0.36***	*	***0.37***	*	***0.39***	*	***0.35***	*
	Black African	1.04		***0.52***	*	***0.53***	*	***0.57***	*	***0.50***	*
	Black Caribbean	***2.20***	*	1.32	*	1.35	*	1.37	*	1.16	*
	Black Other	1.41	*	0.81	*	0.83	*	0.81	*	0.72	*
	Chinese	***0.20***	*	***0.19***	*	***0.19***	*	***0.24***	*	***0.23***	*
	Any other group	0.82	*	***0.47***	*	***0.48***	*	***0.50***	*	***0.43***	*
	Unknown	1.16		1.06		1.06		1.05		1.02	
FSM	Entitled to FSM			***2.80***	*	***2.73***	*	***2.27***	*	***2.19***	*
Gender	Boy			***1.98***	*	***1.97***	*	***1.83***	*	***1.85***	*
Birth season	Spring			0.99		0.98		0.92	*	0.92	*
	Summer			1.00		1.00		0.90	*	0.89	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.99***	*	***1.97***	*	***1.74***	*	1.47	*
Combined Depr.	(IDACI 1SD + FSM)			***3.96***	*	***3.84***	*	***2.99***	*	***2.66***	*
Attendance	Persistent absence (>63 days/yr)					***2.85***	*	***2.12***	*	***2.15***	*
Prior attainment	KS2 English Finely Graded Level							0.80	*	0.79	*
	KS2 Maths Finely Graded Level							0.79	*	0.78	*
School type	Foundation									0.99	
	Academy-Converter									0.94	
	Academy-Sponsored									0.96	
	Church									0.86	*
	Selective/Grammar									0.67	*
	Special/PRU/AP									***0.38***	*
School % FSM	Highest									1.39	*
	Average-High									1.29	*
	Average									1.18	*
	Low-Average									1.03	
School % Black Caribbean & MWBC	Highest									1.24	*
	Average-High									1.14	*
	Average									1.03	
	Low-Average									0.97	
School Size	Smallest									1.45	*
	Small-Average									0.97	
	Average									1.02	
	Average-Large									1.05	*
	Nagelkerke Pseudo R Squared	0.010		0.088		0.092		0.122		0.127	
	Model-specific N (# of pupils)	392708		392708		392465		392465		392230	

*Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI) and Model 2b includes a persistent absence indicator (absent>63 days over a year); Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 4 additionally includes school variables. Note: *=Significant at the level of p<0.05.*

ASD in the Secondary Cohort

Table 3.22 shows the ORs for each ethnic minority group (and for other predictors) in models with ASD identification as the outcome of interest, proceeding through the various stages of hierarchical entry of predictors described above.

Model 1: Unadjusted results (ethnic group only)

With only ethnic group as an explanatory variable, most Asian groups were under-represented (except for Mixed White & Asian, OR=0.96), as were Traveller groups. Black Caribbean (OR=1.51) pupils were over-represented. Overall patterns were largely not dissimilar to those in the analogous logistic primary cohort model for which results are given in Appendix L, but over- and under-representation appear to be somewhat more pronounced for ethnic groups in the secondary cohort.

Model 2: Pupil background

After accounting for other pupil background characteristics, Asian groups (except Mixed White & Asian, OR=0.93) remained under-represented, as did the Traveller groups. Black Caribbean over-representation (OR=1.33) was reduced, although not fully accounted for by other pupil background factors.

The associations of the included pupil background characteristics were less strongly associated with whether or not a pupil was ever identified with ASD in secondary school than was the case for the other focal types of SEN; this was consistent with findings from analogous primary cohort results (see Appendix L). Only being a boy was substantially (OR=4.81) associated with higher odds of ASD identification in secondary school, with much smaller effects of socioeconomic deprivation (OR=1.22 for a 2SD difference in IDACI; OR=1.18 for FSM). Persistent absence in Y6 (Model 2b) had a substantial association with subsequent ASD identification (OR= 2.27) although we need to be appropriately cautious about drawing cause and effect conclusions.

This model nonetheless improves considerably upon the unadjusted model, as demonstrated by the difference between the Nagelkerke's Pseudo R^2 values (0.5% in Model 1; 5.6% in Model 2; 5.7% in Model 2b).

Model 3: Prior attainment

After accounting for KS2 prior attainment in English and maths, most ethnic group ORs remained unchanged in substantive interpretation (with the exception of the very small White Irish and Chinese groups, about which we avoid over-interpreting these changes as they may be an artefact of crossing a relatively low-frequency group with a very low-incidence outcome).

Measures of socioeconomic deprivation had considerably reduced ORs after accounting for prior attainment, including FSM (OR=0.75) and IDACI (OR=0.86 for a 2SD difference). This may appear counter-intuitive, however, it might indicate that –

for pupils with the same prior attainment – a lack of access to resources and services on the part of the economically disadvantaged is linked to lower odds of ASD identification, as we discussed with regard to the primary cohort results.

The independent effects of 1SD differences in scores for KS2 English (OR=0.63) and maths (OR=0.63) were substantial; higher scores were associated with considerably lower odds of ever being identified with ASD in secondary school.

Accounting for prior attainment substantially improved upon the previous model, with a Nagelkerke's Pseudo R squared of 15.6% suggesting that these prior attainment measures were moderately powerful predictors of whether or not a pupil was ever identified with ASD in secondary school.

Model 4: School composition and context

Accounting for school composition and context variables, including school type, school % FSM entitlement, school size, and school % Asian (measured as combined Indian, Pakistani, Bangladeshi and Asian Other according to those groups persistently identified as under-represented for ASD in previous findings) made little difference to most ethnic group ORs or pupil background ORs.

There were three school variables with low to moderate associations with the odds of a pupil ever being identified with ASD in secondary school:

- Attending a special schools were associated with much higher odds of identification (OR=6.34), possibly because ASD – in comparison to the other focal outcomes – may be perceived as a type of need more frequently appropriate for a specialised or alternative setting.

- The highest quintile in terms of school % FSM was associated with very slightly lower odds of ever being identified. The fact that this was also true in the primary cohort helps to suggest that this is a robust if small effect, possibly due to the aforementioned possibility that ASD identification may be linked to access to resources and services that may be less available in deprived areas. This is a finding that warrants further investigation in future research to better understand the underlying mechanism(s) at play.

- Schools with the highest percentages of Asian pupils were associated with raised odds of identification (OR=1.32 and OR=1.24 for the highest two quintiles); this was moderately robust to different filtering dropping schools with fewer than two pupils in the groups of interest, although this decreased the effect size and consistency across quintiles somewhat (see Appendix N). This was further checked against models in which the ethnic composition variable was the only school variable included, and the effect of school % Asian remained consistent, suggesting that the ethnic composition effect is not the consequence of an interaction with another school variable.

Table 3.22: Secondary cohort (age 11-16) ASD: Odds ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 2b Exp(B)		Model 3 Exp(B)		Model 3b Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	1.27		1.26		1.27		***1.63***	*	***1.63***	*	1.48	*
	Traveller Irish	***0.60***		***0.46***		***0.36***		***0.19***	*	***0.21***	*	***0.27***	
	Traveller Gypsy/Roma	***0.26***	*	***0.23***	*	***0.13***	*	***0.05***	*	***0.05***	*	***0.07***	*
	White Other	0.77	*	0.73	*	0.72	*	***0.61***	*	***0.61***	*	***0.65***	*
	Mixed White & African	1.12		1.07		1.07		1.23		1.23		1.22	
	Mixed White & Caribbean	1.16		1.08		1.08		1.20		1.19		1.14	
	Mixed White & Asian	0.96		0.93		0.94		1.18		1.18		1.03	
	Mixed Other	1.05		0.99		0.99		1.11		1.11		1.00	
	Indian	***0.41***	*	***0.39***	*	***0.39***	*	***0.49***	*	***0.49***	*	***0.43***	*
	Pakistani	***0.45***	*	***0.40***	*	***0.40***	*	***0.36***	*	***0.36***	*	***0.34***	*
	Bangladeshi	***0.35***	*	***0.30***	*	***0.29***	*	***0.33***	*	***0.33***	*	***0.32***	*
	Asian Other	***0.41***	*	***0.38***	*	***0.39***	*	***0.42***	*	***0.42***	*	***0.41***	*
	Black African	0.84	*	0.72	*	0.74	*	0.79	*	0.79	*	0.75	*
	Black Caribbean	***1.51***	*	1.33	*	1.35	*	1.35	*	1.34	*	1.32	*
	Black Other	1.23		1.10		1.10		1.01		1.01		0.96	
	Chinese	***0.55***	*	***0.55***	*	***0.56***	*	0.76		0.76		***0.62***	
	Any other group	***0.42***	*	***0.36***	*	***0.37***	*	***0.35***	*	***0.35***	*	***0.37***	*
	Unknown	0.98		0.94		0.94		0.89		0.90		0.87	
FSM	Entitled to FSM			1.18	*	1.15	*	0.75	*	0.75	*	0.74	*
Gender	Boy			***4.81***	*	***4.79***	*	***4.07***	*	***4.07***	*	***3.95***	*
Birth season	Spring			1.04		1.04		0.92	*	0.92	*	0.96	
	Summer			1.08	*	1.08	*	0.87	*	0.87	*	0.93	*
Neighbourhood Depr.	Normalised IDACI 2SD			1.22	*	1.21	*	0.86	*	0.86	*	0.94	
Combined Depr.	(IDACI 1SD + FSM)			1.30	*	1.26	*	0.69	*	0.69	*	0.72	*
Attendance	Persistent absence (>63 days)					***2.27***	*	1.16		***2.27***	*	***2.40***	*
Prior attainment	KS2 English Finely Graded Level							***0.63***	*	***0.63***	*	0.71	*
	KS2 Maths Finely Graded Level							***0.63***	*	***0.62***	*	0.73	*
	KS2 Maths by Persistent abs.									***1.54***	*	***1.71***	*
School type	Foundation											1.02	
	Academy-Converter											1.21	*
	Academy-Sponsored											1.11	
	Church											1.19	*
	Selective/Grammar											1.11	
	Special/PRU/AP											***6.34***	*
School % FSM	Highest											0.82	*
	Average-High											0.96	
	Average											1.01	
	Low-Average											1.07	
School % Asian (Indian/Pakistani/ Bangladeshi/Asian Other)	Highest											1.32	*
	Average-High											1.24	*
	Average											1.18	*
	Low-Average											1.13	*
School Size	Smallest											1.02	
	Small-Average											1.07	
	Average											1.01	
	Average-Large											1.06	
	Nagelkerke Pseudo R Squared	0.005		0.056		0.057		0.156		0.157		0.184	
	Model-specific N (# of pupils)	392708		392708		392465		392465		392465		392230	

*Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 2b includes a persistent absence indicator (absent>63 days over a year); Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 3b additionally includes an interaction between KS2 maths scores and attendance, and Model 4 additionally includes school variables. Note: *=Significant at the level of p<0.05.*

- Thus while Asian groups are substantially under-represented relative to White British pupils, this is mitigated somewhat in schools with large Asian populations.

Including these school variables led to only a small improvement in predictive power, with a Nagelkerke's Pseudo R squared of 18.4%.

Limitations to the analysis

There are limitations to the analytical approach taken here. Where we investigate a particular type of SEN (MLD, BESD/SEMH or ASD) using Cox regression, we treat the first instance of identification as the outcome, and so do not take into account possible later transitions to identification with a different type of need or to no SEN identification in a later year. Additionally, both Cox and logistic regression models account for whether or not a pupil is identified with a particular type of need using binary indicators that treat other types of SEN and no SEN as the collective alternative to being identified specifically with MLD, BESD/SEMH or ASD, rather than treating different types of primary need as competing 'risks'.

Further research as an extension of this work might extend to more complex models to investigate multi-state transitions (in and out of SEN identification and between SEN types) as well as accounting for different types of SEN as simultaneous competing risks. Alternatively, another extension to this work might be to run ordinal regression analyses with outcomes defined as whether or the course of the relevant school phase a pupil was: (1) never identified with a given type; (2) identified at least once with a given type, or (3) identified for the majority of the given period with the given type of primary need.

Part 4: Analysis of the Second Longitudinal Study of Young People in England (LSYPE2)

Summary

The SES measures available in the NPD are limited to entitlement to a FSM and a measure of economic deprivation at the neighbourhood level through the IDACI. The question remains as to whether more differentiated measures of Socio-economic Status (SES), based on parent's occupational classification, parent's educational qualifications and family income, might account for the Black Caribbean/Mixed White & Black Caribbean over-representation for BESD/SEMH.

This chapter use the Second Longitudinal Study of Young People in England (LSYPE2) to investigate this question. LSYPE2 is a nationally representative sample of some 12,000 students, drawn from the same population as our secondary longitudinal cohort, who were studied intensively, including interviews with the young people and their parents and linking to data from the NPD. Three SEN outcomes are evaluated: whether a student was ever identified in any January School Census between Y7 and Y11 with SEMH/BESD, with ASD or with MLD.

The main findings were:

For SEMH/BESD, the over-representation of Black Caribbean and Mixed White & Black Caribbean students is not accounted for using detailed measures of SES such as parental occupation, parental education and family income. Indeed the NPD measures of FSM and IDACI account for a greater proportion of the Black Caribbean over-representation than measures of parental occupation, education and income. Combined with our finding regarding early attainment also not accounting for the over-representation (see part 3), this is evidence to suggest social processes may be involved in the over-representation of Black Caribbean and MWBC students among those identified with SEMH/BESD.

For ASD, gender was the strongest predictor of ASD identification (boys were six times more likely to be identified than girls), but ethnicity was the second strongest predictor, higher than any of the SES variables. After holding economic factors such as family social class and entitlement to FSM constant, there were raised odds of identification in the most highly educated families (one or more parents holds a degree). The combined Pakistani/Bangladeshi ethnic group were significantly and consistently under-represented in identification for ASD even after controls.

For MLD, entitlement to FSM was the strongest predictor, followed by sex, IDACI, Family SEC and Parent Education. Ethnicity was barely significant after the inclusion of SES variables.

Our two major conclusions are:

First, the over-representation of Black Caribbean and Mixed White & Black Caribbean students for BESD/SEMH is not accounted for using detailed measures of SES such as parental occupation, parental education and family income.

Second, the fact that we often did not find statistically significant relationships with ethnicity in the LSYPE2 sample, where we know these relationships do exist in the full cohort, indicates that recent studies from the US based upon longitudinal surveys need to be interpreted with substantial caution. Further studies such as ours, based on national population data, are urgently needed.

Introduction

Purpose of this chapter

We have seen from the NPD analysis in Parts 1 to 3 that SES variables have a strong relationship with SEN identification, and that they were able to account for the over-representation of Black Caribbean, MWBC and Pakistani students among those identified with MLD. However, they were not able to account for the over-representation of BCRB/MWBC students among those identified with BESD/SEMH. The SES measures available in the NPD are limited to entitlement to a FSM and a measure of economic deprivation (the number of children aged 0-16 in families entitled to benefits) at neighbourhood level (IDACI). The question remains as to whether more differentiated measures of SES, based on parent's occupational status, educational qualifications and income, might account for the Black Caribbean/MWBC over-representation for SEMH. We considered two longitudinal studies as potential sources of rich background data that could also be matched to the NPD to pick up reliable data on SEN identification.

Millennium Cohort Study (MCS): The MCS was explored as an additional source of data to explore the risk factors for SEN identification emerging during the early years. For example MCS Sweeps 2 & 3 collects parents' rating of child behaviour through the Strengths and Difficulties Questionnaire (SDQ) and achievement scores at age 3 & 5. Matching Sweep 4 (age 7, 2008) and Sweep 5 (age 11, 2012) against the relevant January School Census to add SEN type would allow the completion of a similar analysis to that completed by Hibel et. al. (2010) using the US ELCS-K. However, there were substantial challenges in sample size. For example we could only use the England sample, and by Sweep 5 at age 11 this is reduced to just over 8,000 children. For low incidence outcomes such as the intersection of ethnic minorities and SEN identification the sample was simply too small to offer appropriate statistical power.

Second Longitudinal Study of Young People in England 2 (LSYPE2): Focussing on the secondary school phase, the LSYPE2 proved a much more robust data source. This study contains rich data on family SES collected in detailed interviews with the young person, the main parent and a second parent if there was one in the household. It is recent and representative, based on a large sample (n=13,100) of students aged

13/14 in 2012/13. Indeed this sample was drawn from the same cohort we employed in our secondary longitudinal analysis. It has also been matched to the NPD, and although the standard datasets lodged with the UKDS do not include type of SEN, the DFE LSYPE team agreed to match in and supply this data to us.

The research question addressed by this section is: *Do traditional measures of family SES (parent's occupational status, educational qualifications and income), account for the over-representation of BCRB/MWBC students for SEMH/BESD during the secondary school phase?* Since we have seen low SES tends to be associated with higher levels of SEN identification, and that ethnic minorities are general more economically deprived, it is unlikely that SES will explain the under-representation of ethnic minority groups for SEN. Nevertheless we run the same models for ASD and MLD for completeness.

Method

The LSYPE2 dataset

The primary sample frame for LSYPE2 was the English School Census, which was used to identify sample members in state-funded education. This provides access to pupil-level characteristics information about these young people, which was used to stratify the sample. The stratification has been designed to maintain minimum numbers in certain subgroups of interest right through to the planned end of the survey, to ensure robust analyses of these groups can continue. These subgroups include those with free school meals (FSM), those with special educational needs (SEN), and certain ethnic groups. The sample also included pupils from independent schools and pupil referral units (PRUs), these schools/settings were sampled first and then asked to supply contact details for pupils. Interviews took place with both the young person and at least one parent in the first three waves (i.e. until the young person is aged 15/16). In Wave 1 the interviews took place over a five month period, starting in early April 2013 and finishing in early September 2013. In Wave 1 LSYPE2 achieved a response rate of 71 per cent, representing an achieved sample of 13,100.

The analytic sample

As stated above, there were 13,100 responding young people in Wave 1 of LSYPE2. Of these, 12,177 gave permission for linkage to the NPD. Some of those giving permission were in independent schools (n=460) which do not complete the school census, leaving 11,710 available records with NPD linkage. The DFE supplied the primary and secondary type of SEN variables from each January school census between 2010 and 2015 for all young people that had consented to NPD linkage. This was matched into the LSYPE2 records. For 40 pupils in Alternative Provision (AP) and 20 in Pupil Referral Units (PRU) their SEN records were manually matched in by the author. In subsequent analyses a small number of cases (n=59) that were missing the continuous variable IDACI were excluded, giving an analytic sample of 11,651 records.

Variables

The outcome variables

Interviews for LSYPE took place between 08/04/13 and 13/09/13. We initially drew SEN type from the first census following interviews, i.e. January 2014. The advantage of this particular census was that, being so close to the interviews, there was little time for attrition so we could be very confident that the absence of a type of SEN genuinely meant no SEN identification. A breakdown by type is given below.

Table 4.1: Type of SEN School Census January 2014: LSYPE2 Sample

SEN type	N	%	Cum. %
None recorded	10386	88.7	88.7
ASD	133	1.1	89.8
BESD	429	3.7	93.5
HI	31	0.3	93.8
MLD	277	2.4	96.1
OTH	70	0.6	96.7
PD	44	0.4	97.1
PMLD	2	0.0	97.1
SLCN	125	1.1	98.2
SLD	35	0.3	98.5
SPLD	157	1.3	99.8
VI	21	0.2	100.0
Total	11710	100.0	

We see for example that 429 pupils or 3.7% had BESD as their primary need in 2014. However this may be a conservative sample for three reasons. First, this records the primary type of need at the time of the census. However many children have multiple needs, so while one need may be primary at the time of the census another need may still have been identified. Indeed a further 124 pupils had BESD recorded as their secondary type of SEN in 2014. Second, and more importantly, the primary need may vary over time. For example BESD may not have been the primary need in 2014, but it may have been the primary need in an earlier or later year of secondary school. Third, prior to 2015 identification of the type of need was only required for pupils with higher levels of need (SAP & Statemented). However from 2015 the distinction between School Action and School Action Plus was removed and the type of need was requested for pupils in the new combined 'SEN Support' category. Thus while the total size of the group with some form of SEN did not change, the number of pupils for whom the type of need was requested and reported increased quite sharply. Indeed, 183 pupils were identified with SEMH in 2015 who had not previously been identified with BESD. Since recording the type of need for all pupils receiving SEN Support is the system for recording SEN going forward, we did not want to exclude this group.

For all the above reasons we chose a more inclusive definition of ever identified in Y7-Y11, including students who at any time during the secondary school phase had

BESD/SEMH identified as their primary need (which we denote Ever BESD/SEMH). We followed the same process to create Ever ASD and Ever MLD measures. In addition we calculated an Ever SEN variable for students who had been identified in any census with any form of SEN (at School Action, SAP, SEN Support of statement/EHC). A breakdown for our focus SEN types and Ever SEN by ethnic group is given below.

Table 4.2: Identified with SEN in any census Y7-Y11 by ethnic group: LSYPE2

Ethnic group	N	Ever SEMH / BESD	Ever ASD	Ever MLD	Ever SEN
White Irish	36	8.3%		8.3%	36.1%
GRT	14	14.3%		35.7%	85.7%
White Other	338	3.6%	0.6%	6.2%	33.7%
MWBC	214	10.3%	2.8%	6.5%	42.5%
MWBA	75	12.0%	1.3%	2.7%	42.7%
Mixed White & Asian	114	7.9%	1.8%	4.4%	25.4%
Mixed Other groups	107	6.5%		8.4%	37.4%
Indian	252	2.4%	0.4%	3.2%	18.3%
Pakistani	394	3.3%	0.8%	7.9%	40.6%
Bangladeshi	277	3.2%		8.3%	34.3%
Chinese	26		3.8%	3.8%	15.4%
Asian Other	158	2.5%	0.6%	3.8%	28.5%
Black African	543	5.5%	0.7%	5.5%	36.8%
Black Caribbean	382	16.0%	1.3%	9.2%	49.7%
Black Other	39	10.3%		2.6%	28.2%
Any other ethnic group	87	8.0%	1.1%	9.2%	49.4%
White British	8595	7.3%	1.7%	6.7%	35.3%
Total	11651	7.1%	1.5%	6.7%	35.7%

Note: Ever SEN incudes pupils at any stage of SEN (School Action, School Action Plus, SEN Support, Statement/EHC). Ever SEMH/BESD, Ever ASD and Ever MLD includes pupil identified with that as their primary type of need. The Ever SEMH, Ever ASD and Ever MLD categories are not mutually exclusive: 69 pupils had both BESD/SEMH and MLD identified as a primary need at some time during secondary school, 19 had BESD/SEMH and ASD, and 15 both MLD and ASD. Three pupils had all three identified as their primary need at different times during the secondary school phase.

There are three points to make from Table 4.2.

- First, the approach increases the proportion of pupils identified in our focus SEN type. The proportion with SEMH/BESD increases from 3.7% to 7.1% for the EVER SEMH/BESD and for MLD from 2.4% to 6.7% for Ever MLD. However, the proportion of pupils with ASD remains relatively low, increasing only from 1.1% to 1.5% for the Ever ASD measure.

- Second, a snapshot of recorded SEN at any single point in time gives a very different picture of the incidence of SEN compared to the longitudinal data. For example in the school census for 2015, 17.9% of pupils of secondary school age (Y7-Y11) were identified with some form of SEN (DFE, 2015)[21]. However, the cumulative picture for pupils tracked longitudinally between Y7 in 2011 and Y11 in 2015 indicates that more than double the number of pupils had an SEN need

21. The comparable figure in 2011 was 21.4% (DFE, 2011).

identified at some point during secondary school. Cumulatively over one-third (35.7%) of pupils had a SEN identified at some point during their secondary school career.

- Third, there is evidence of ethnic disproportionality in the data. For example, while overall around 35% of White British pupils were Ever SEN the figure was even higher for some ethnic minority groups, for example 42% of MWBC and 50% of Black Caribbean pupils. These two groups were also notably over-represented for the BESD/SEMH measure, at 10.3% and 16.0% respectively compared to 7.3% for White British. Issues of under-representation are also apparent, for example for Asian ethnic groups for SEMH/BESD and ASD, though not for Pakistani or Bangladeshi groups for MLD. In short the patterns are similar to those we identified in the secondary national cohort, but not identical.

Explanatory variables

Ethnicity

We used the ethnic group coding collected as part of the LSYPE2 interview with each young person. These uses the same 18 ethnic group categories we have discussed throughout the report, but is drawn directly from the young person. For a small number of 'don't know' responses (n=25) we used the code from the student's record in 2013 school census.

FSM

We used the entitlement to FSM variable from the 2013 school census, as this was the variable we had used in our NPD analysis. We also had the EVER6 measure - i.e. whether a young person had been entitled to FSM at any census time-point over the last 6 years - available for sensitivity analyses.

Income Deprivation Affecting Children Index (IDACI)

IDACI is produced by the Department for Communities and Local Government (DCLG). The index is based on 32,482 Super Output Areas (SOAs) in England, which are geographical regions of around 1,500 residents, designed to include those of similar social backgrounds. The IDACI score is the percentage of under-16s in the SOA living in income deprived households (primarily defined by being in receipt of certain benefits). This variable is highly skewed and so for the purpose of the current analysis the measure was normal score transformed to give a variable with a mean of 0 and SD=1. A score above 0 indicate greater than average deprivation, and score below 0 indicate less than average deprivation, relative to the average for the LSYPE2 sample. Both 2001 and 2007 IDACI measures were included in the LSYPE2 file. The means of the two were nearly identical (24.7% and 25.7%) and they correlated r=0.97, so the more recent 2007 values were used. Further information about IDACI can be found at: https://www.gov.uk/government/publications/english-indices-of-deprivation-2010.

Family Socio-economic Classification (SEC)

We utilised the ONS eight category Socio-Economic Classification (SEC). A Family SEC variable is included in LSYPE2 based upon the Household Reference Person

(HRP), but in a large number of cases the HRP was not interviewed (n=487) or the individual was not classifiable (n=121). We therefore created our own Family SEC measure. First we took the SEC for the main parent, which had fewer missing or unclassifiable instances (n=116). Second, to create a family measure, we substituted the SEC of the second parent (if present) if it was higher than for main parent. As a robustness check we completed the same process taking the highest of the mother's or father's SEC. This measure was very highly correlated (r=0.996) with the MP/SP version, but the MP/SP version had fewer missing cases (n=116 as opposed to n=502) so was preferred.

Table 4.3: ONS Socio-economic classification (SEC) categories: LSYPE2 Sample

Code	SEC 8 category	SEC 3 category
8	Higher managerial & professional	Professional
7	Lower professional & higher technical	
6	Intermediate occupations	Intermediate
5	Small employers & own account workers	
4	Lower supervisory & technical	
3	Semi-routine occupations	Low
2	Routine occupations	
1	Never worked or long-term unemployed	

Educational Qualifications

We took the highest educational qualification of the main parent, substituting the highest qualification of the second parent (where present) if it was higher. A small number of cases (n=37) which were coded as 'entry level qualifications' were combined with 'Other qualifications'. This created a 7 point scale ranging from 'No educational qualifications' through to 'Degree or equivalent'.

Family Income

Household income is based on a survey response, with respondents picking a band from a list to represent the annual household income from all sources. The results have been edited to take account of implausible responses, primarily through the use of self-reported earnings data. Earnings data was generally more credible, not least because parents reported their own earnings, over the time period of their choice, rather than having to combine sources and annualise the results. This data has also been edited where implausible, such as where what looked like an annual salary for the stated occupation was reported as being paid weekly. Where the plausible earnings of a household were greater than the annual income selected, the earnings have been used instead. This is likely to underestimate the true income, as it excludes other sources such as benefits, but should still represent an improvement on the self-reported estimate.

Table 4.4: Fifteen income bands and distribution of responses: LSYPE2 Sample

Code	Income Band	N	%	Valid %	Cum. %	Band mid-point
1	Under £2,600	135	4.4	1.3	1.3	1,300
2	£ 2,600 - £ 5,199	308	2.6	3.0	4.3	3,900
3	£ 5,200 - £10,399	990	8.5	9.7	14.0	7,800
4	£10,400 - £15,599	1484	12.7	14.5	28.4	13,000
5	£15,600 - £20,799	1349	11.6	13.2	41.6	18,200
6	£20,800 - £25,999	1065	9.1	10.4	52.0	23,400
7	£26,000 - £31,199	933	8.0	9.1	61.1	28,600
8	£31,200 - £36,399	598	5.1	5.8	66.9	33,800
9	£36,400 - £39,999	505	4.3	4.9	71.8	38,200
10	£40,000 - £44,999	524	4.5	5.1	76.9	42,500
11	£45,000 - £49,999	495	4.2	4.8	81.8	47,500
12	£50,000 - £59,999	547	4.7	5.3	87.1	55,000
13	£60,000 - £74,999	597	5.1	5.8	92.9	67,500
14	£75,000 - £99,999	386	3.3	3.8	96.7	87,500
15	£100,000 or more	340	2.9	3.3	100.0	100,000
	Total Valid	10256	88.0	100.0		
	Missing	1395	12.0			
	Grand total	11651	100.0			

The data were collected in 15 bands allowing a high degree of differentiation. For descriptive purposes we used the mid-point of the ranges as the data value rather than the band number to give a mean income in pounds per annum. It should be noted that income data is notoriously difficult to collect accurately via household surveys, and LSYPE2 is no exception. There was a high level of non-response (12%) as indicated in Table 4.4. Notwithstanding the editing and banding of this data, which is intended to mitigate its limitations, all analyses based on this should be treated with caution.

Missing data treatment

To minimise data loss, and to model non-linear relationships between predictors and the outcome, we treated SEC, Parent Education and family income as categorical variables and explicitly included missing values as discrete categories for each variable.

We accounted for clustering at the school level by using multilevel logistic regression models. For software we predominantly employed MLWin, with a logistic link and estimation by IGLS (PQL2). For some analyses we used the SPSS GENLIN MIXED command.

We did not use sample weights for two reasons. First, we were primarily interested in the relationship between variables, not in simply recapturing descriptive statistics for the relevant population. In these cases the use of weights can be problematic (Solon, Haider & Woodridge, 2015). Second, and more importantly the Primary Sampling Unit identifiers (PSU) and stratum variable (FinalStratum) we would need had not been

included in the public release of LSYPE2, so it is not possible to use features of STATA or SPSS that would allow the calculation of robust standard errors. The 'LSYPE2 User Guide to the Datasets: Wave 1' (TNS-BMRB, 2015, p12) states:

"In order to use the complex samples options, you would need to specify a file plan, which tells SPSS what the PSUs are, what the strata are and what the selection weights are. We have taken the decision not to release the PSUs and strata data for all cases, to avoid identifying the schools attended by those not consenting to NPD linkage. As such, the approach described above will not be possible with the available data".

We were therefore not able to use the modules for robust SE's. In terms of interpretation the LSYPE guide suggests the following:

"If you are running cross tabulations and your conclusions are highly significant, e.g. p < 0.005, it is almost certain that the conclusion will be significant at the 95% level and you don't need to worry. However, if your conclusion is only just significant at the 95% level, it is advisable to be cautious in your interpretation".

Results

Descriptive statistics

Table 4.5 presents descriptive statistics for the sample, broken down by ethnicity. For simplicity we show the mean and SD of each variable, e.g. the mean SEC category (1-8). We present further more detailed analysis of the three SES measures from the LSYPE2 in subsequent tables.

Table 4.5: Mean and SD of predictors by ethnic group: LSYPE2 Sample

	IDACI normal		Entitled FSM		sex		Parent SEC		Parent Educ.		Income	
Ethnic group	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
White Irish	0.10	*1.07*	0.28	*0.45*	0.50	*0.51*	5.2	*2.2*	4.1	*2.2*	32,959	26822
GRT	0.29	*0.90*	0.64	*0.50*	0.43	*0.51*	3.4	*1.9*	2.3	*1.4*	21,763	14811
White Other	0.41	*0.88*	0.23	*0.42*	0.50	*0.50*	4.3	*2.1*	4.5	*2.2*	26,110	20053
MWBC	0.32	*0.92*	0.35	*0.48*	0.50	*0.50*	5.0	*2.2*	4.4	*1.9*	28,260	23473
MWBA	0.50	*0.87*	0.39	*0.49*	0.49	*0.50*	4.7	*2.3*	4.7	*2.1*	26,449	21312
Mixed White & Asian	0.02	*0.96*	0.32	*0.47*	0.46	*0.50*	5.4	*2.3*	5.1	*2.0*	37,202	31288
Mixed Other groups	0.52	*1.12*	0.36	*0.48*	0.50	*0.50*	4.9	*2.0*	4.7	*1.9*	23,888	19959
Indian	0.20	*0.86*	0.14	*0.35*	0.47	*0.50*	5.5	*2.0*	5.0	*2.0*	31,829	21943
Pakistani	0.61	*0.75*	0.37	*0.48*	0.47	*0.50*	4.1	*2.2*	4.1	*2.2*	21,683	21767
Bangladeshi	1.15	*0.91*	0.51	*0.50*	0.47	*0.50*	3.8	*1.9*	3.3	*2.1*	16,647	13471
Chinese	-0.34	*1.17*	0.12	*0.33*	0.31	*0.47*	4.7	*2.3*	5.0	*2.3*	27,330	23579
Asian Other	0.40	*0.93*	0.30	*0.46*	0.55	*0.50*	4.8	*2.3*	5.2	*2.1*	25,591	17696
Black African	0.86	*0.86*	0.41	*0.49*	0.49	*0.50*	4.5	*2.3*	5.0	*2.1*	23,917	18885
Black Caribbean	0.83	*0.82*	0.41	*0.49*	0.49	*0.50*	4.8	*2.1*	4.6	*1.9*	22,647	17721
Black Other	0.77	*0.86*	0.38	*0.49*	0.69	*0.47*	5.1	*2.2*	5.6	*1.6*	21,854	14898
Any other ethnic group	0.83	*1.08*	0.55	*0.50*	0.46	*0.50*	4.8	*2.7*	4.9	*2.2*	28,356	25099
White British	-0.22	*0.93*	0.22	*0.42*	0.49	*0.50*	5.4	*2.0*	4.7	*1.8*	34,189	24312
Total	0.00	*1.00*	0.26	*0.44*	0.49	*0.50*	5.2	*2.1*	4.6	*1.9*	31,980	23850
Vaid observations	11651		11651		11651		11535		11597		10256	

Notes: IDACI is 2007 version, high scores indicate greater deprivation. GRT= Gypsy Roma Traveller; MWBC= Mixed White & Black Caribbean, MWBA= Mixed White & Black African.

Family SEC

Table 4.6 and Figure 4.1 present the results for Family SEC. For simplicity we have presented the three-category SEC classification. For White British pupils just over one-quarter (27%) of pupil's are from low SEC homes, while around 43% are from managerial and professional households. Indian, Mixed White & Asian, White British, White Irish and Any other group, also have over 40% of pupil's come from managerial and professional homes. In contrast the proportion of pupils from low SEC home is around 40% or over for Black Caribbean, Chinese, Black African, White Other, Pakistani and Bangladeshi groups.

Parental educational qualifications

Table 4.7 and Figure 4.2 present parental educational qualifications. Black Caribbean and White British groups show very similar levels of parental educational qualification, with around 35% of parents with higher education (HE) Qualifications of some kind and around 10% with no educational qualifications. There are substantially higher proportions of parents with low educational qualifications among the Mixed White & Black Caribbean, Bangladeshi, Pakistani and Gypsy/Roma Traveller (GRT) students.

Family income

Table 4.8 and Figure 4.3 present the mean income data by ethnic group. The average income for White British families was £34,189. No ethnic minority group had mean income significantly higher than the White British mean, though the average for White Irish, Indian, Mixed White & Asian, and any other group did not differ significantly from White British. All other groups had significantly lower average income than White British. Bangladeshi had the lowest mean income of £16,647, less than half the White British mean.

Table 4.6: Ethnic group by socio-economic classification (SEC) of the home: LSYPE2 Sample

		Socio-economic classification			
Ethnic group	N	Low	Intermediate	Professional	Missing
White Irish	36	33.3%	19.4%	47.2%	0.0%
GRT	14	64.3%	28.6%	7.1%	0.0%
White Other	338	45.4%	32.9%	21.6%	3.0%
MWBC	214	37.3%	25.5%	37.3%	0.9%
MWBA	75	45.9%	16.2%	37.8%	1.3%
Mixed White & Asian	114	29.2%	23.9%	46.9%	0.9%
Mixed Other groups	107	33.7%	35.6%	30.8%	2.8%
Indian	252	24.4%	28.4%	47.2%	0.8%
Pakistani	394	47.9%	33.4%	18.7%	2.0%
Bangladeshi	277	45.8%	43.2%	11.1%	2.2%
Chinese	26	42.3%	23.1%	34.6%	0.0%
Asian Other	158	37.2%	25.6%	37.2%	1.3%
Black African	543	47.9%	19.8%	32.3%	1.5%
Black Caribbean	382	38.9%	31.1%	30.0%	2.4%
Black Other	39	34.2%	26.3%	39.5%	2.6%
Any other ethnic group	87	39.3%	20.2%	40.5%	3.4%
White British	8595	27.1%	30.3%	42.5%	0.7%
Total	11651	30.8%	29.9%	39.3%	1.0%

Notes: SEC are given as percentage of valid values, missing is given as percentage of all cases. Low= Routine, semi-routine & Long term unemployed. MWBC= Mixed White & Black Caribbean; MWBA = Mixed White & Black African; GRT= Gypsy & Roma Traveller.

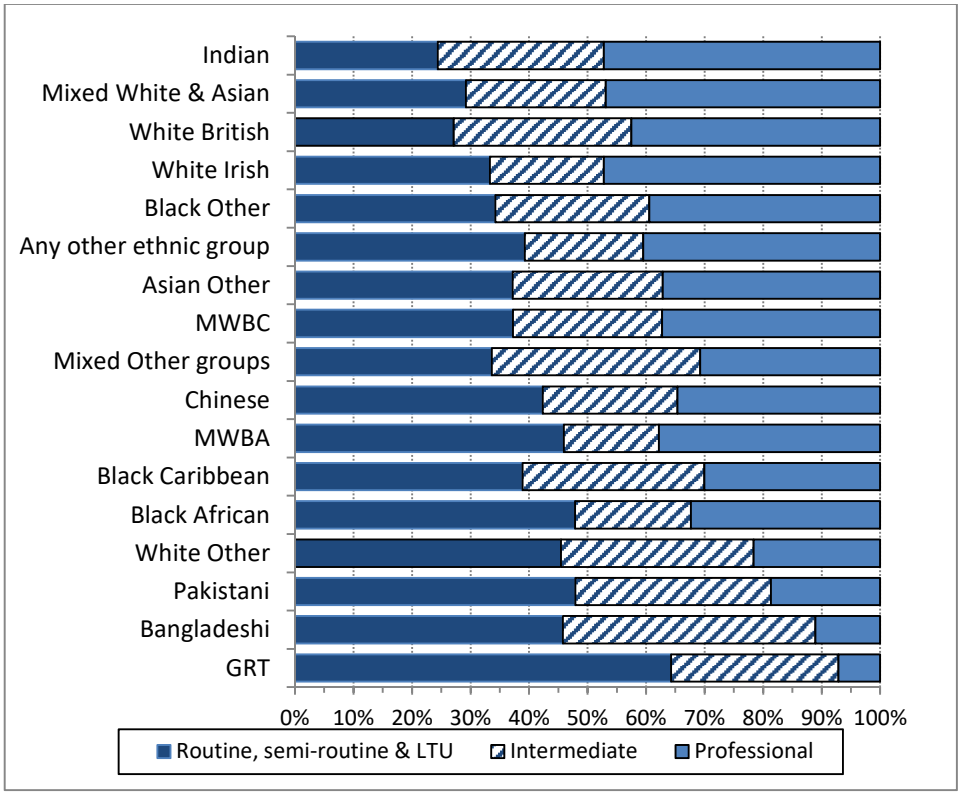


Figure 4-1: Socio-economic classification (SEC) of the home by ethnic group: LSYPE2 Sample

Table 4.7: Ethnic group by parent's highest educational qualification: LSYPE2 Sample

Ethnic group	N	No Qual-ifications	Other quals	Some GCSE or equiv.	5+ GCSEs A*-C or equiv.	A/AS levels or equiv.	HE below degree (e.g. HND)	Degree (e.g. BA, BSc, MA)	Missing
White Irish	36	22.2%	0.0%	22.2%	11.1%	11.1%	13.9%	19.4%	0.0%
GRT	14	42.9%	14.3%	21.4%	14.3%	7.1%	0.0%	0.0%	0.0%
White Other	338	16.9%	5.4%	11.2%	10.0%	15.1%	12.1%	29.3%	2.1%
MWBC	214	12.7%	0.9%	22.5%	13.1%	16.0%	16.9%	17.8%	0.5%
MWBA	75	13.5%	0.0%	24.3%	5.4%	13.5%	12.2%	31.1%	1.3%
Mixed White & Asian	114	9.6%	2.6%	10.5%	11.4%	14.9%	12.3%	38.6%	0.0%
Mixed Other groups	107	10.3%	3.7%	13.1%	13.1%	19.6%	16.8%	23.4%	0.0%
Indian	252	11.2%	0.4%	13.2%	10.8%	13.2%	15.2%	36.0%	0.8%
Pakistani	394	21.6%	3.1%	17.0%	13.7%	13.7%	7.7%	23.2%	1.5%
Bangladeshi	277	34.9%	3.6%	19.3%	14.9%	10.5%	3.6%	13.1%	0.7%
Chinese	26	15.4%	3.8%	7.7%	7.7%	11.5%	7.7%	46.2%	0.0%
Asian Other	158	10.2%	4.5%	6.4%	9.6%	13.4%	10.8%	45.2%	0.6%
Black African	543	13.1%	3.0%	12.4%	7.1%	11.2%	17.6%	35.6%	1.7%
Black Caribbean	382	10.1%	2.4%	17.0%	17.5%	16.2%	17.5%	19.4%	1.3%
Black Other	39	2.6%	0.0%	12.8%	12.8%	7.7%	23.1%	41.0%	0.0%
Any other ethnic group	87	13.8%	2.3%	14.9%	9.2%	10.3%	3.4%	46.0%	0.0%
White British	8595	7.7%	1.2%	18.4%	20.5%	14.5%	16.3%	21.4%	0.2%
Total	11651	9.8%	1.6%	17.5%	18.2%	14.2%	15.4%	23.2%	0.5%

Notes: Qualifications are given as percentage of valid values. Missing is given as a percentage of all cases. Routine= Routine, semi-routine & Long term unemployed. MWBC= Mixed White & Black Caribbean; MWBA = Mixed White & Black African; GRT= Gypsy & Roma Traveller.

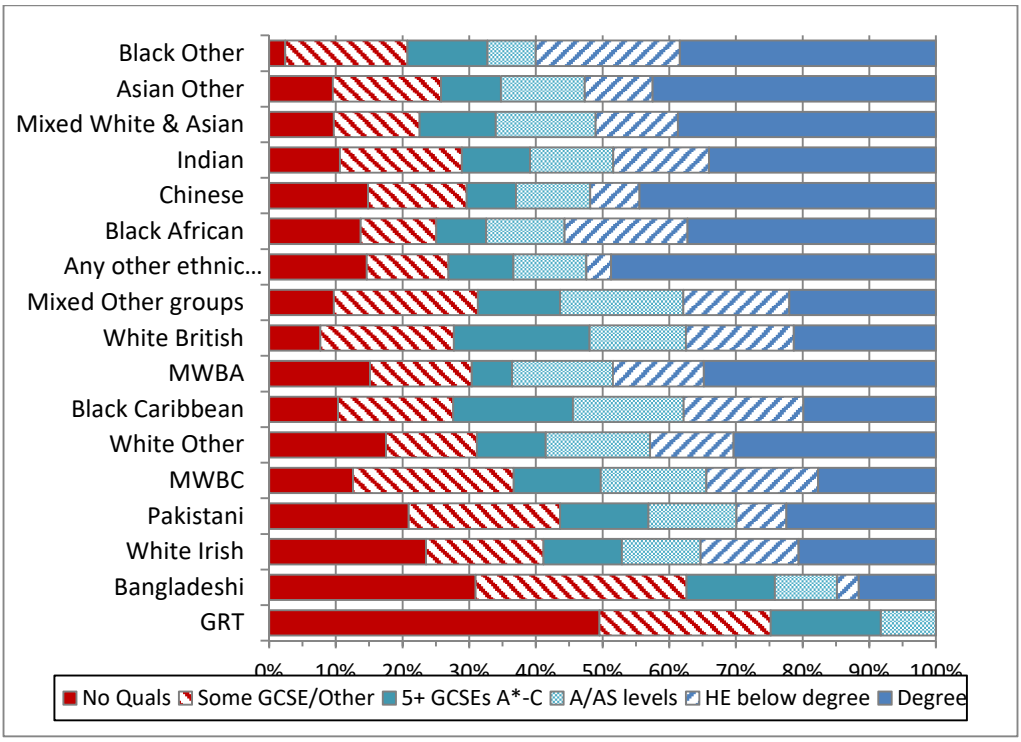


Figure 4-2: Parents' highest educational qualification by ethnic group: LSYPE2

Table 4.8: Mean income by ethnic group: LSYPE2 Sample

Ethnic Group	N	<£15,600	£15,600-25,999	£26,000-£44,999	£45,000+	Missing	Mean Band
White Irish	36	31.3%	21.9%	25.0%	21.9%	11.1%	32,959
GRT	14	50.0%	12.5%	25.0%	12.5%	42.9%	21,763
White Other	338	32.3%	31.0%	24.5%	12.2%	13.0%	26,110
MWBC	214	37.4%	24.6%	17.3%	20.7%	16.4%	28,260
MWBA	75	39.3%	23.0%	19.7%	18.0%	18.7%	26,449
Mixed White & Asian	114	28.1%	26.0%	17.7%	28.1%	15.8%	37,202
Mixed Other groups	107	42.2%	31.3%	14.5%	12.0%	22.4%	23,888
Indian	252	25.5%	23.1%	29.2%	22.2%	15.9%	31,829
Pakistani	394	54.4%	21.5%	13.8%	10.4%	24.4%	21,683
Bangladeshi	277	59.3%	26.8%	11.3%	2.6%	30.0%	16,647
Chinese	26	34.8%	30.4%	17.4%	17.4%	11.5%	27,330
Asian Other	158	36.4%	23.3%	24.8%	15.5%	18.4%	25,591
Black African	543	38.4%	28.5%	23.0%	10.1%	16.0%	23,917
Black Caribbean	382	41.7%	26.9%	21.8%	9.6%	18.3%	22,647
Black Other	39	34.3%	37.1%	22.9%	5.7%	10.3%	21,854
Any other group	87	42.5%	23.3%	13.7%	20.5%	16.1%	28,356
White British	8595	24.7%	22.6%	26.4%	26.2%	9.6%	34,189
Total	11651	28.4%	23.5%	25.0%	23.1%	12.0%	31,980

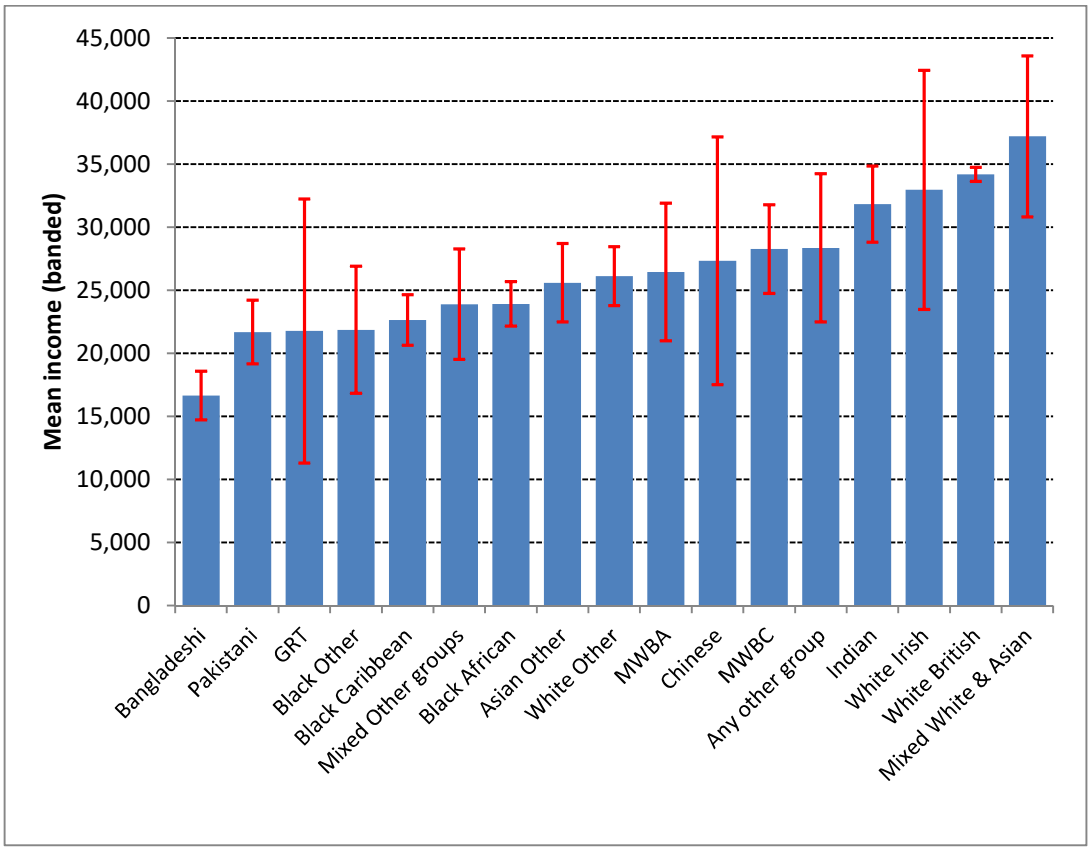


Figure 4-3: Mean income by ethnic group: LSYPE2 Sample

SEMH/BESD identification

Approach to analysis

There were 11,651 pupils with complete data attending 736 secondary schools and settings. A null model, including just intercept terms at level 1 (student) and level 2 (school) indicted that 17.0% of the variance was at the school level. In our experience longitudinal samples tend to inflate estimates of the variance at school level because of the inclusion of some schools with very small samples. We repeated the analysis only including schools with at least 10 observations. This removed 195 pupils (1.7% of the sample) and 28 schools, leaving 11,456 pupils in 708 schools. The resulting school variance estimate was reduced to 13.2%[22].

We ran five models.

- Null model: to establish the school variance in a base model.
- Model 1: Including just ethnicity to establish the simple relationship between ethnicity and SEN identification, controlling for clustering.
- Model 2: Including the SES measures from the NPD, i.e. sex, entitlement to FSM and IDACI normal score.
- Model 3: Including the SES measures from the LSYPE2, i.e. sex, parent's socio-economic classification (SEC), Parental educational qualifications and family income.
- Model 4: Including the SES measures from both the NPD and the LSYPE2, i.e. sex, entitlement to FSM, IDACI normal score, parent's socio-economic classification (SEC), Parental Education and family income.

The models failed to converge when the very small group of Chinese pupils were included, so the Chinese (n=26) and GRT (n=14) were collapsed into the Any Other ethnic group. Table 4.9 present the results of the regression analyses.

Results: Multilevel regression analysis

Model 1: Ethnicity only

The association with ethnicity were strong with Black Caribbean students being significantly over-represented (OR=2.20), and Indian, Pakistani, Bangladeshi, Asian other and White other groups all significantly under-represented (OR ranged from 0.27 to 0.39) relative to White British students. Black African students were also significantly under-represented (OR= 0.61). Interestingly Mixed White & Black Caribbean and Mixed White & Black African were both over-represented (OR=1.34 and 1.53 respectively) but the results were not statistically significant.

---

[22]. SPSS GENLIN MIXED gave a slightly lower estimate for school level variation with the full sample (13.3%). However when the reduced dataset was used the school level variance dropped to 0.0% and GENLIN MIXED gave the following warning: glmm: The final Hessian matrix is not positive definite although all convergence criteria are satisfied. The procedure continues despite this warning. Subsequent results produced are based on the last iteration. Validity of the model fit is uncertain. For these reasons MLWin was preferred as the software for analysis for BESD.

Table 4.9: Regression Models for Ever BESD: LSYPE2 Sample

EVER BESD	Ethnic				NPD				LSYPE Only				NPD+LSYPE			
	Coeff	SE	OR		Coeff	SE	OR		Coeff	SE	OR		Coeff	SE	OR	
Intercept	-2.76	0.058			-3.53	0.09			-4.91	0.47			-4.72	0.48		
White Irish	-0.11	0.67	0.90		-0.19	0.68	0.82		-0.20	0.68	0.82		-0.24	0.68	0.79	
GRT	0.51	0.84	1.67		0.11	0.84	1.12		0.07	0.85	1.07		-0.04	0.85	0.96	
White Other	-0.95	0.33	0.39	**	-1.08	0.33	0.34	**	-1.20	0.33	0.30	***	-1.19	0.34	0.30	***
MWBC	0.29	0.25	1.34		0.08	0.26	1.08		0.12	0.26	1.13		0.05	0.26	1.05	
MWBA	0.42	0.40	1.53		0.16	0.40	1.17		0.28	0.40	1.32		0.15	0.40	1.16	
Mixed White & Asian	-0.09	0.39	0.91		-0.25	0.39	0.78		-0.12	0.39	0.88		-0.20	0.39	0.82	
Mixed Other groups	-0.23	0.44	0.80		-0.60	0.44	0.55		-0.43	0.44	0.65		-0.60	0.44	0.55	
Indian	-1.30	0.47	0.27	**	-1.30	0.47	0.27	**	-1.36	0.47	0.26	**	-1.33	0.47	0.26	**
Pakistani	-1.07	0.33	0.34	**	-1.39	0.33	0.25	***	-1.44	0.33	0.24	***	-1.47	0.33	0.23	***
Bangladeshi	-1.11	0.40	0.33	**	-1.73	0.40	0.18	***	-1.59	0.39	0.20	***	-1.80	0.40	0.17	***
Asian Other	-1.46	0.57	0.23	*	-1.69	0.56	0.18	**	-1.60	0.56	0.20	**	-1.70	0.56	0.18	**
Black African	-0.50	0.22	0.61	*	-0.95	0.23	0.39	***	-0.73	0.22	0.48	**	-0.92	0.23	0.40	***
Black Caribbean	0.78	0.17	2.19	***	0.43	0.18	1.54	*	0.63	0.17	1.87	***	0.43	0.18	1.54	*
Black Other	0.08	0.59	1.09		-0.12	0.59	0.89		0.22	0.59	1.24		-0.07	0.60	0.93	
Any other ethnic group	-0.32	0.45	0.73		-0.90	0.46	0.41	*	-0.41	0.46	0.66		-0.71	0.46	0.49	
Entitled FSM (vs. not)					1.09	0.09	2.97	***					0.84	0.10	2.32	***
IDACI (normal score 1SD)					0.31	0.05	1.37	***					0.22	0.05	1.24	***
Boy (vs. girl)					0.75	0.09	2.12	***	0.76	0.09	2.14	***	0.76	0.09	2.14	***
SEC:Missing									0.97	0.40	2.62	*	0.70	0.41	2.01	
Never worked or LT unemployed									0.76	0.29	2.13	*	0.24	0.30	1.27	
Routine occupations									0.94	0.25	2.56	***	0.61	0.25	1.83	*
Semi-routine occupations									0.72	0.24	2.05	**	0.45	0.24	1.56	
Lower supervisory & technical									0.73	0.26	2.08	**	0.60	0.27	1.82	*
Small employers & own account									0.55	0.25	1.73	*	0.48	0.25	1.61	
Intermediate occupations									0.48	0.24	1.62	*	0.36	0.24	1.44	
Lower prof. & higher technical									0.45	0.22	1.57	*	0.37	0.22	1.45	
Parent Education: Missing									-0.26	0.68	0.77		-0.39	0.68	0.68	
No qualifications									0.67	0.18	1.94	***	0.47	0.18	1.60	**
Other qualifications									0.46	0.30	1.58		0.32	0.30	1.38	
Some GCSE passes or equivalent									0.20	0.16	1.22		0.10	0.17	1.10	
5+ GCSEs at A*-C or equiv									0.10	0.16	1.11		0.06	0.16	1.07	
A/AS levels or equivalent									0.20	0.17	1.22		0.17	0.17	1.18	
HE below degree level (e.g. HND)									0.24	0.16	1.28		0.25	0.16	1.28	
Income: Missing									1.21	0.48	3.36	*	0.91	0.49	2.49	
Under £2,600									1.83	0.54	6.22	**	1.38	0.55	3.99	*
£ 2,600 - £ 5,199									1.58	0.51	4.87	**	1.13	0.51	3.09	*
£ 5,200 - £10,399									1.40	0.48	4.06	**	0.93	0.49	2.54	
£10,400 - £15,599									1.17	0.48	3.23	*	0.77	0.49	2.15	
£15,600 - £20,799									1.14	0.48	3.11	*	0.80	0.49	2.23	
£20,800 - £25,999									1.08	0.49	2.94	*	0.87	0.49	2.39	
£26,000 - £31,199									0.52	0.50	1.68		0.37	0.50	1.45	
£31,200 - £36,399									0.74	0.51	2.09		0.65	0.51	1.91	
£36,400 - £39,999									1.02	0.51	2.78	*	0.93	0.51	2.52	
£40,000 - £44,999									0.42	0.53	1.52		0.41	0.53	1.50	
£45,000 - £49,999									0.41	0.54	1.51		0.38	0.54	1.45	
£50,000 - £59,999									0.14	0.55	1.15		0.12	0.55	1.13	
£60,000 - £74,999									0.44	0.52	1.54		0.36	0.53	1.43	
£75,000 - £99,999									0.28	0.58	1.33		0.27	0.58	1.30	
Random Part																
School Variance (SE)	0.718	0.102			0.594	0.094			0.593	0.09			0.587	0.09		
School Variance %	17.9%				15.3%				15.3%				15.1%			
Akaike IC (Corrected)	64,105				62,666				62,668				62,577			

Notes: Computed with MLWin estimation by IGLS (PQL2). Akaike from SPSS GenLin Mixed command. All models based on 11,651 pupils in 736 educational establishments. SEC contrast against higher professional & technical; Parent education contrast against degree (BA, BSc, MA etc.); Income contrast against £100,000 or more.

Model 2: SES controls using NPD variables (FSM & IDACI)

Next we introduced the controls for gender, entitlement to FSM and IDACI. All were significantly associated with the identification of SEN, as we saw in Part 3. In relation to Black Caribbean over-identification, the results mirrored the secondary national longitudinal sample; accounting for gender and NPD SES controls reduced but did not eliminate the over-representation (OR=1.54).

Model 3: SES controls using Parent SEC, education and family income

In this model we used the SES measures available from LSYPE. Compared to the highest SEC group the ORs were significantly higher for all other groups, especially for the four lowest categories (pupils from families in the lower supervisory through to long-term unemployed groups) where the odds of being identified were at least twice as high (OR range= 2.1 to 2.6). Income was also strongly related to SEMH identification. Relative to the highest income group (£100,000+) the odds were significantly raised for all families with income <£26,000 by at least OR=3.0, up to OR=6.2 for families with income <£2,600. Parental education had a smaller association with identification, given SEC and income were already in the model, the only significant result being that the odds of identification for pupils from families were no parent held any educational qualifications were nearly twice as high (OR=1.94) as for pupils from families where one or more parents had a degree level qualification.

The impact upon Black-Caribbean over-representation was to reduce the OR to 1.87. The effect was more limited than for the NPD controls, which reduced the OR to 1.54. So rather than being more likely to explain Black Caribbean over-representation, traditional SES measures were actually less effective than the NPD SES measures. This might reflect the relative better standing of the Black Caribbean group compared to White British group on the traditional SES measures. For example the mean IDACI normal score for the Black Caribbean group differed by over 1.0 SD from the White British mean, while their average income was only 0.48 SD lower, average SEC was 0.28 SD lower and there was no difference at all in terms of average educational qualifications between the two groups.

Comparing the Akaike Information Criterion (AIC) for models 2 and 3, these were practically identical. Interestingly, therefore, FSM and IDACI were as efficient as traditional SES measures in accounting for variation in the identification of SEMH/BESD. This echoes conclusions from other recent research, which has indicated that entitlement to FSM is just as effective as traditional SES measures in account for educational attainment at age 11 and at age 16 (Ilie et al, 2017).

Model 4: SES controls with both sets of variables

Combining all five SES measures added very little to the overall power of the model, reducing the AIC by just 91 points relative to the NPD only model[23]. There was some small additional risk of identification in homes where parental SEC was lower supervisory or semi-routine occupations (ORs=1.8), where family income was £5,000

23. In a single level model the Nagelkerke Pseudo R^2 was 10.6% for Model 2 and 12.6% for Model 4.

per annum (OR> 3.0), and where neither parent had any educational qualifications (OR= 1.6), but the overall difference was extremely small. There was no change to the Black Caribbean OR which remained at OR=1.54, the same as in the NPD only model (OR=1.54).

ASD Identification

The rationale for the analysis

Our previous analysis has identified the under-representation of Asian groups as the key disproportionality issue for ASD. We know that socio-economic disadvantage is higher in Asian ethnic groups, and that socio-economic disadvantage is associated with increased risks of ASD identification. It is therefore extremely unlikely that socio-economic factors explain the under-representation of Asian students. However it is valuable to explore the more general relationship between traditional parental SES measures and the identification of ASD, so we attempted to repeat the same analysis we had completed for Ever SEMH/BESD using Ever ASD.

Methodological issues

SPSS GENLINMIXED could not even compute a two-level null model without generating the Hessian matrix error. MLWin would also not converge for a two-level null model with the preferred estimation method. The default 1st Order MQL estimated 64.4% of variance at the school level. The 1st order PQL estimated 28.8%. The preferred 2nd order estimation (either PQL or MQL) crashed. With such problematic results for multi-level models across multiple software packages we therefore moved to a single level model.

Even with a single level model the size of the longitudinal sample caused computational problems. There are blank cells for White Irish, GRT, Mixed Other groups, Bangladeshi and Black Other groups, since no pupils in any of these ethnic groups were ever identified with ASD (see Table 4.2). Hence the software could not compute standard errors for these cells. We recoded White Irish and GRT to White Other; Bangladeshi to a joint Pakistani/Bangladeshi category; Chinese to Asian Other groups; and Mixed Other and Black Other to any other ethnic group.

There was also a blank cell for the small number of pupils (n=54) with missing Parent Education. Given these students were near identical to the 'Other qualifications' group in terms of mean IDACI score and %FSM, we coded them with Other Qualifications.

Results: Regression analysis

Table 4.10 presents the results of four regression models, following the same logic as for Ever BESD. Model 1 tested ethnic group alone; Model 2 tested ethnicity, sex and the two SES measures from the NPD; Model 3 tested ethnicity, sex and the three SES measures from LSYPE2, and; Model 4 tested ethnicity, sex and all five SES measures.

Ethnic group

Despite very low ORs for several ethnic groups, the only statistically significant result was for the Pakistani/Bangladeshi group (OR=0.27) which was a relatively large group

(n=871). Again this reveals the limitations of sample size even in relatively large scale longitudinal studies. Looking across models the Pakistani/Bangladeshi under-representation remained consistent after SES controls. When all SES measures were included Black African and White Other groups were also seen to be under-represented, which is consistent with the full cohort analysis in Part 3.

Traditional SES measures

Before considering the joint impact of the SES variables, we consider the simple bivariate associations between each of the three LSYPE2 SES measures and Ever ASD identification, testing the statistically significance through univariate logistic regressions. Figure 4.4 plots the simple bivariate relationships.

- The association with SEC gave a significant omnibus test (p<.045) but none of the individual contrasts with the higher managerial & professional group were significant. Collapsing the eight category SEC into the three category version to increase sample size indicated that students from lower SEC homes (Semi-Routine, Routine & LT Unemployed) on average were significantly more likely to be identified than students from managerial & professional homes (Wald=4.09, p<.043, OR= 1.43).
- Omnibus tests indicated no significant relationship between family income and ASD. Collapsing the income variable from 15 bands into quartiles suggested a slight curvilinear relationship, but again these results were not statistically significant.
- There was no statistically significant relationship between Parent Educational Qualifications and Ever ASD.

When all three SES variables from LSYPE2 were included jointly in model 3, family income made no unique contribution. Parental SEC had an effect through raised odds of identification for students from low SES homes (Routine and Semi-Routine occupations). Parent Education had the greatest WALD statistics, with reduced odds for all categories below degree level. Given the consistently reduced odds across all categories, and the fact that the simple bivariate relationship was not significant, the most appropriate way to interpret this is that, after holding economic factors such as FSM and SEC constant, there are raised odds of identification in the most educated families (one or more parents with a Degree).

Comparing NPD and LSYPE2 SES measures

The LSYPE2 SES measures had marginally more explanatory power than the NPD measures (Nagelkerke Pseudo R^2 = 9.4% vs. 8.2%) but the difference was slight. Similarly, combining all five SES measures marginally improved prediction (Nagelkerke Pseudo R^2 = 10.5%).

An indication of the relative importance of each predictor can be gained from the Wald statistic in the final column of Table 4.10. Sex was the strongest predictor, followed by ethnic group, Parental Education, and then Family SEC or FSM joint fourth. Family income and neighbourhood deprivation (IDACI) were negligible.

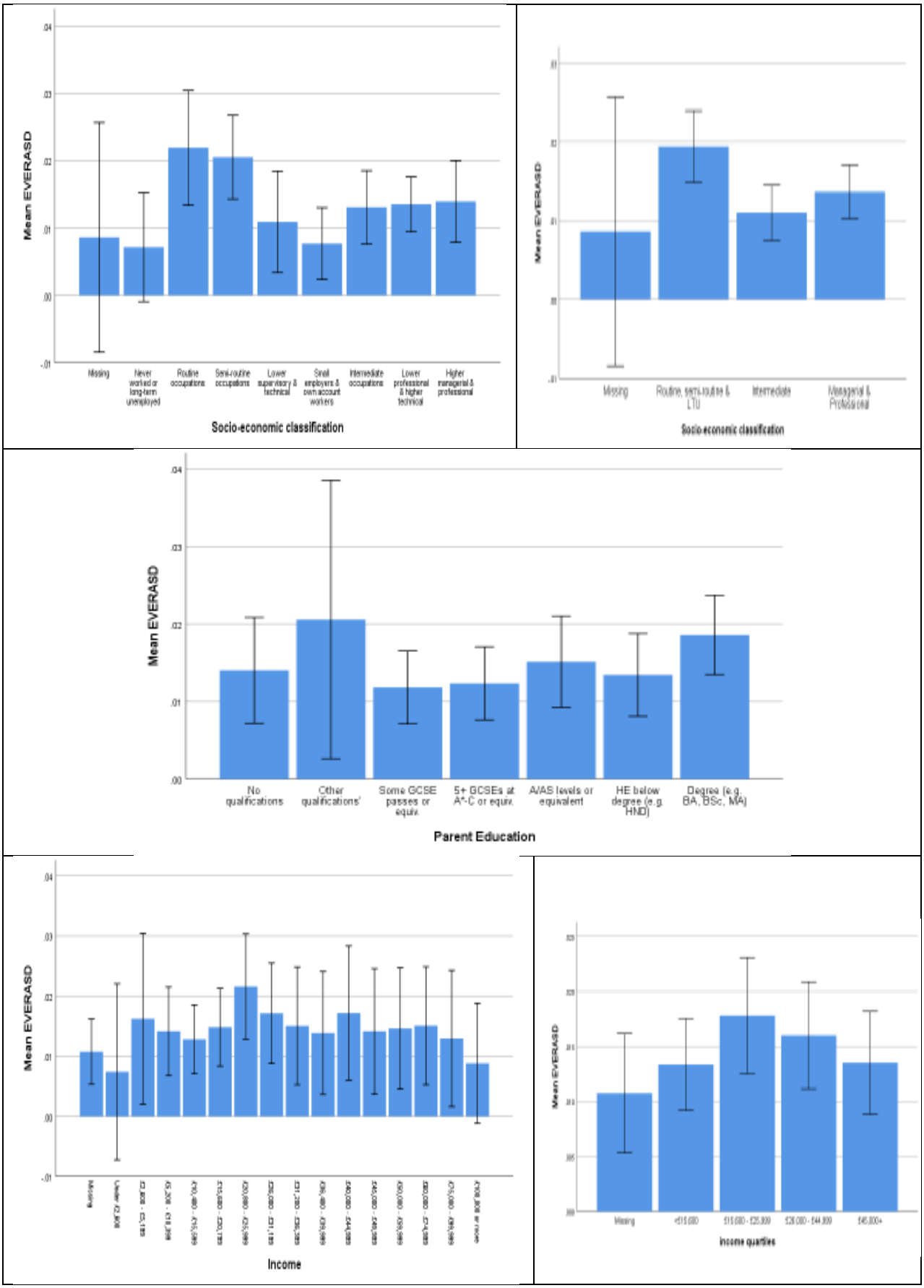


Figure 4-4: Traditional SES measures and Ever identified as ASD: LSYPE2 Sample

Table 4.10: Logistic regression models for Ever ASD: LSYPE2 Sample: LSYPE2 Sample

	Ethnic				NPD SES				LSYPE2 SES				All SES				
Variable	B	**SE**	**OR**		**B**	**SE**	**OR**		**B**	**SE**	**OR**		**B**	**SE**	**OR**		**Wald**
Constant	-4.08	0.08	0.017		-4.89	0.24	0.01		-5.27	0.31	0.01		-4.36	0.37	0.01		136.5
Ethnic group		16.27				20.0											**24.1**
White Other	-1.18	0.71	0.31		-1.27	0.72	0.28		-1.48	0.72	0.23	*	-1.48	0.72	0.23	*	4.2
MWBC	0.53	0.42	1.70		0.42	0.43	1.53		0.58	0.43	1.78		0.45	0.43	1.57		1.1
MWBA	-0.22	1.01	0.80		-0.35	1.02	0.71		-0.34	1.02	0.71		-0.45	1.02	0.64		0.2
Mixed White & Asian	0.05	0.72	1.06		-0.08	0.72	0.93		-0.10	0.73	0.90		-0.19	0.73	0.83		0.1
Indian	-1.45	1.01	0.24		-1.46	1.01	0.23		-1.61	1.01	0.20		-1.59	1.01	0.20		2.5
Pakistani / Bangladeshi	-1.33	0.58	0.27	*	-1.60	0.59	0.20	**	-1.47	0.59	0.23	*	-1.62	0.60	0.20	**	7.3
Asian Other	-0.43	0.72	0.65		-0.49	0.72	0.61		-0.71	0.73	0.49		-0.74	0.73	0.48		1.0
Black African	-0.82	0.51	0.44		-1.06	0.52	0.35	*	-1.15	0.52	0.32	*	-1.35	0.53	0.26	**	6.6
Black Caribbean	-0.24	0.46	0.78		-0.47	0.47	0.62		-0.37	0.46	0.69		-0.61	0.47	0.54		1.7
Any other ethnic group	-1.37	1.01	0.25		-1.58	1.01	0.21		-1.49	1.01	0.22		-1.67	1.01	0.19		2.7
Sex (Boy vs. girl)					1.80	0.22	6.03	***	1.81	0.22	6.14	***	1.80	0.22	6.06	***	**66.0**
Entitled FSM (vs. not entitled)					0.61	0.17	1.84	***					0.74	0.20	2.09	***	**13.3**
IDACI normal score (1SD)					0.10	0.09	1.10						0.16	0.10	1.18		**2.9**
Family SEC																	**13.8**
Missing									0.05	1.05	1.06		-0.20	1.05	0.82		0.0
Never worked/LT Unemployed									0.26	0.68	1.29		-0.28	0.68	0.76		0.2
Routine occupations									1.12	0.38	3.06	**	0.75	0.39	2.12		3.8
Semi-routine occupations									0.94	0.34	2.57	**	0.66	0.35	1.93		3.5
Lower supv & tech.									0.25	0.46	1.28		0.10	0.46	1.10		0.0
Small employers & own account									-0.17	0.45	0.85		-0.27	0.45	0.77		0.3
Intermediate occupations									0.27	0.35	1.32		0.17	0.35	1.18		0.2
Lower prof. & higher technical									0.04	0.29	1.05		-0.02	0.29	0.98		0.0
Parental Education																	**21.4**
No qualifications									-0.89	0.35	0.41	*	-1.06	0.35	0.35	**	9.2
Other qualifications									-0.36	0.51	0.70		-0.46	0.51	0.63		0.8
Some GCSE passes or equiv.									-1.15	0.30	0.32	***	-1.23	0.30	0.29	***	17.0
5+ GCSEs at A*-C or equiv									-0.99	0.28	0.37	***	-1.01	0.28	0.36	***	12.9
A/AS levels or equiv.									-0.62	0.27	0.54	*	-0.63	0.27	0.53	*	5.3
HE below degree (e.g. HND)									-0.59	0.26	0.55	*	-0.60	0.26	0.55	*	5.2
Family income																	**6.7**
Missing									-0.03	0.34	0.97		-0.23	0.35	0.79		0.4
<£15,600									0.07	0.30	1.07		-0.28	0.31	0.76		0.8
£15,600 - £25,999									0.38	0.28	1.46		0.16	0.28	1.18		0.3
£26,000 - £44,999									0.32	0.26	1.38		0.27	0.26	1.31		1.1
NagelKerke Pseudo R^2	1.3%				8.2%				9.4%				10.5%				

Notes: N=11,761 in all models. Reference categories were: Ethnic group = White British; IDACI= sample mean; SEC= Higher managerial and professional occupations; Parental Education= Degree; Family income= £45,000 per annum or above. SES variables sourced at age 14.

MLD Identification

The rationale for the analysis

Our previous analysis has identified the over-representation of Black Caribbean and Pakistani young people with MLD as a major issue. We have however seen that controls for socio-economic disadvantage eliminate this over-representation, although GRT do remain over-represented, and indeed in adjusted analyses the major feature is the under-representation of many ethnic groups (e.g. Bangladeshi, Black African, Indian, White Other, Mixed White & Asian). Given the SES measures from the NPD account for the Black Caribbean and Pakistani over-representation there is no specific need to draw on more traditional SES measures from the LSYPE2. However, as with ASD, there is interest in the more general relationship between parental SES measures and MLD identification, so we repeat the logistic regression models as we have done for BESD/SEMH and ASD.

Methodological issues

Unlike the outcome for BESD and ASD, GENLIN MIXED gave no error messages and was very stable across models, reflecting the higher incidence of MLD relative to ASD and the absence of empty cells. In a multi-level null model the proportion of variance at the school level was 16.0%[24]. The effect was robust. For example, in a model filtering just to schools with at least 10 observations, the school variance was still 15.2%. In a model for mainstream only schools (i.e. filtering out the 149 pupils (1.3%) in the 20 Special schools, Pupil Referral Units or Alternative Provision) the school variance was 14.7%. We therefore employed multi-level logistic regression models. Because GENLIN MIXED models were stable, we utilised SPSS.

Results: Multilevel regression

Table 4.11 presents the results of four regression models. Model 1 tested ethnic group alone; Model 2 tested ethnicity, sex and the two SES measures from the NPD; Model 3 tested ethnicity, sex and the three SES measures from LSYPE2, and Model 4 tested ethnicity, sex and all five SES measures simultaneously.

Ethnic group

The only statistically significant result was the over-representation of GRT students (OR=6.9). Again this reveals the limitations of sample size even in relatively large scale longitudinal studies. Looking across models adjusting for SES reduced but did not eliminate the over-representation of GRT students, but it is important to take into consideration that this is based on only 14 cases (5 of whom were identified as having MLD). After adjusting for SES, Black African and Bangladeshi groups were under-represented. This is consistent with the full cohort analysis in Part 3, although it underestimates the number of ethnic groups affected.

24. It was notable that the null model in MLWiN, estimated using Second order PQL, gave a higher estimate of school variance at 21.3%. However because GENLIN MIXED models were stable we utilised SPSS.

Traditional SES measures

Before considering the joint impact of the SES variables, we consider the simple bivariate associations between each of the three LSYPE2 SES measures and Ever MLD identification, testing the statistically significance through univariate logistic regressions. Figure 4.5 plots the simple bivariate relationships.

- There was a highly significant and consistent decrease in the proportion of pupils identified with MLD with higher Parental SEC, ranging from 17% of pupils from Never Worked/LTU households to around 2% in Higher Managerial & Professional homes.
- There were consistently higher levels of identification for homes where parents had fewer educational qualifications, ranging from 15% where parents had no educational qualifications to 3% in homes where one or more parent held a degree.
- There was a strong association between family income and MLD identification. Because of the large number of categories and the overlap in confidence intervals between many adjacent bands we worked with the quartile version, with MLD rates ranging from 10% in homes with family income <£15,600 to just 2% in homes with annual family income over £45,000.

When all three SES variables from LSYPE2 were included jointly in model 3, the strongest predictor was family SEC, with particularly raised OR's for long-term unemployment, routine and semi-routine occupations. In addition, having a parent with no educational qualifications raised the odds of identification to 2.3 times higher relative to a parent with a degree. Family income had a smaller association, but there were still raised odds for family income <£26,000 compared to £45,000 per annum.

Comparing NPD and LSYPE2 SES measures

We cannot compare the models directly because of the rescaling issue in logistic multi-level models.[25] In a single level model the NPD SES measures had marginal less explanatory power than the LSYPE2 measures (Nagelkerke Pseudo R^2 = 6.6% vs. 8.9%) but the difference was slight and came at the price of a much higher number of parameters to estimate. Consequently the AIC indicates the NPD model was the more efficient model, with an AIC of 64,081 compared to 66,812 for the LSYPE2 variables (See Table 4.11). Adding all five measures together the association with income ceased to be significant, although the income related NPD measures (entitlement to FSM and IDACI) remained significant. Overall the increase in the Nagelkerke Pseudo R^2 at 9.9% was only very slight, and again this was at the price of a raised AIC (67,021).

An indication of the relative importance of each predictor can be gained from the F statistic: Entitlement to a FSM was the strongest predictor (F=22.2), followed by sex (F=17.1), IDACI (F=11.1), SEC (F=5.0) and Parent Education (F=4.2). Ethnicity had a small effect only just statistically significant (F=1.7, p<.04) and income was not statistically significant (F=1.3, ns.).

---

[25]. See relevant discussion in Part 2.

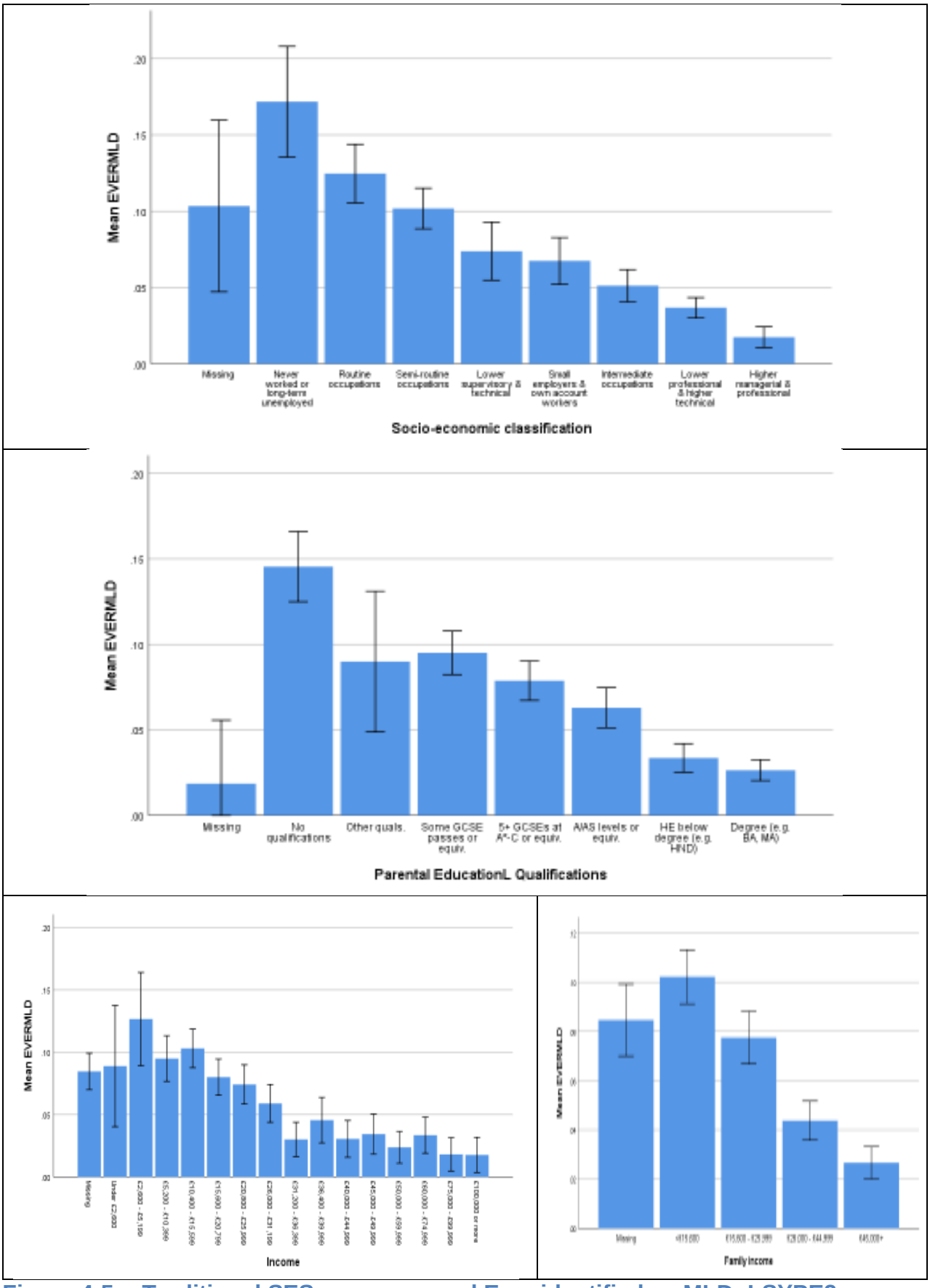


Figure 4-5: Traditional SES measures and Ever identified as MLD: LSYPE2 Sample

Table 4.11: Logistic regression models for Ever MLD: LSYPE2 Sample

	Ethnic				NPD				LSYPE				All SES			
Variable	B	SE	OR		B	SE	OR		B	SE	OR		B	SE	OR	
Constant	-2.71	0.05	0.07		-3.13	0.08	0.04		-4.39	0.22	0.01		-4.25	0.22	0.01	
White Irish	0.24	0.59	1.28		0.07	0.59	1.07		0.01	0.62	1.01		-0.03	0.62	0.97	
GRT	1.92	0.57	6.85	***	1.59	0.58	4.88	**	1.33	0.66	3.80	*	1.28	0.63	3.59	*
White Other	-0.11	0.23	0.90		-0.26	0.24	0.77		-0.31	0.24	0.74		-0.31	0.24	0.73	
MWBC	-0.03	0.27	0.97		-0.27	0.28	0.76		-0.19	0.29	0.83		-0.27	0.29	0.76	
MWBA	-1.02	0.72	0.36		-1.35	0.73	0.26		-1.20	0.75	0.30		-1.31	0.76	0.27	
Mixed White & Asian	-0.46	0.46	0.63		-0.64	0.47	0.53		-0.47	0.47	0.62		-0.55	0.47	0.58	
Mixed Other groups	0.29	0.34	1.34		-0.02	0.36	0.98		0.15	0.37	1.17		0.07	0.37	1.07	
Indian	-0.72	0.39	0.48		-0.74	0.40	0.48		-0.69	0.41	0.50		-0.69	0.42	0.50	
Pakistani	0.09	0.19	1.09		-0.23	0.19	0.80		-0.28	0.20	0.76		-0.32	0.20	0.73	
Bangladeshi	0.07	0.22	1.07		-0.44	0.24	0.64		-0.40	0.24	0.67		-0.53	0.25	0.59	*
Chinese	-0.27	1.03	0.76		-0.19	1.01	0.83		-0.52	1.01	0.59		-0.36	1.00	0.70	
Asian Other	-0.53	0.43	0.59		-0.80	0.43	0.45		-0.63	0.43	0.53		-0.73	0.44	0.48	
Black African	-0.29	0.20	0.75		-0.72	0.21	0.49	***	-0.50	0.21	0.60	*	-0.63	0.21	0.53	**
Black Caribbean	0.23	0.20	1.26		-0.15	0.20	0.86		0.14	0.21	1.15		-0.01	0.21	0.99	
Black Other	-1.09	1.05	0.34		-1.50	1.08	0.22		-0.98	1.08	0.37		-1.19	1.09	0.30	
Any other ethnic group	0.14	0.41	1.16		-0.43	0.40	0.65		-0.08	0.44	0.93		-0.30	0.43	0.74	
Sex (Boy vs. girl)					0.31	0.07	1.36	***	0.31	0.07	1.36	***	0.31	0.07	1.36	***
Entitled FSM (vs. not entitled)					0.83	0.09	2.30	***					0.43	0.09	1.54	***
IDACI normal score (1SD)					0.33	0.05	1.39	***					0.17	0.05	1.18	***
Family SEC																
Missing									1.30	0.36	3.65	***	1.12	0.37	3.06	**
Never worked/LT Unemployed									1.65	0.28	5.20	***	1.33	0.29	3.79	***
Routine occupations									1.25	0.26	3.50	***	1.04	0.26	2.84	***
Semi-routine occupations									1.14	0.25	3.13	***	0.96	0.25	2.62	***
Lower supv & tech.									0.87	0.28	2.39	**	0.77	0.28	2.16	**
Small employers & own acc.									0.83	0.26	2.30	**	0.78	0.25	2.18	**
Intermediate occupations									0.56	0.26	1.76	*	0.49	0.26	1.63	
Lower prof. & higher tech.									0.52	0.23	1.68	*	0.47	0.23	1.60	*
Parental Education																
Missing									-1.28	1.03	0.28		-1.35	1.04	0.26	
No qualifications									0.82	0.17	2.26	***	0.70	0.18	2.02	***
Other qualifications									0.44	0.28	1.55		0.36	0.28	1.43	
Some GCSE passes or equiv.									0.58	0.17	1.79	***	0.52	0.17	1.68	**
5+ GCSEs at A*-C or equiv.									0.53	0.16	1.69	**	0.50	0.16	1.64	**
A/AS levels or equiv.									0.40	0.17	1.49	*	0.38	0.17	1.46	*
HE below degree (e.g. HND)									-0.05	0.18	0.95		-0.05	0.18	0.95	
Family income																
Missing									0.47	0.19	1.60	*	0.31	0.19	1.37	
<£15,600									0.51	0.18	1.66	**	0.28	0.18	1.33	
£15,600 - £25,999									0.39	0.17	1.47	*	0.24	0.17	1.27	
£26,000 - £44,999									0.09	0.18	1.10		0.05	0.18	1.06	
Random Part																
School Variance (SE)	0.627	(.078)			0.489	(.072)			0.498	(.072)			0.478	(.072)		
School Variance %	16.0%				12.9%				13.1%				12.7%			
Akaike IC (Corrected)	65609				64081				66812				67021			

Key findings

SEMH/BESD

- Black Caribbean students were significantly over-represented (OR=2.20), while Indian, Pakistani, Bangladeshi, Asian other, White other and Black African students were all significantly under-represented (ORs ranging from 0.27 to 0.61), relative to White British students.

- Mixed White & Black Caribbean (MWBC) and Mixed White & Black African (MWBA) groups were both over-represented (OR=1.34 and 1.53 respectively), but the results were not statistically significant. This highlights the issue in working with sample data, even from a large nationally representative cohort of almost 12,000 young people, when trying to model relatively low incidence outcomes like type of SEN for ethnic minority groups.

- Three SES measures were drawn from LSYPE2 and all were strongly related to BESD/SEMH identification. For students from the four lowest categories of *Family Socio-economic Classification (SEC)* (parent in a lower supervisory, semi-routine occupation, routine occupation, or long-term unemployed) the odds of being identified were at least twice as high (OR range= 2.1 to 2.6) as pupils from the highest SEC group (higher managerial & professional occupations). *Family Income* was also strongly related to SEMH/BESD identification. Relative to the highest income group (£100,000+) the odds were significantly raised for all families with income<£26,000 by at least OR=3.0, up to OR=6.2 for families with income <£2,600. *Parental education qualifications* had a smaller association when SEC and income were also in the model, the only significant result being that the odds of identification for pupils from families were no parent held any educational qualifications were nearly twice as high (OR=1.94) as for pupils from families where one or more parents had a degree level qualification.

- The impact of controlling for finer-grained measures of SES upon Black-Caribbean over-representation was to reduce the OR from 2.20 to 1.87. The effect was more limited than for the SES controls in the NPD, which reduced the Black Caribbean OR to 1.54. This is likely to reflect the relatively smaller difference between the Black Caribbean and White British groups on the traditional SES measures. For example, the mean IDACI normal score for the Black Caribbean group differed by over 1.0 SD from the White British mean, while their average income was only 0.48 SD lower, average SEC was 0.28 SD lower and there was no difference at all in terms of average educational qualifications between the two groups.

- This echoes conclusions from other recent research, which has indicated that entitlement to FSM is just as effective as traditional SES measures in accounting for educational attainment at age 11 and at age 16 (Ilie, Sutherland & Vignoles, 2017).

ASD

- The low incidence of ASD identification caused significant problems in modelling the risk of identification, even with a sample as large as LSYPE2 with nearly 12,000 cases. To enable models to run some groups and explanatory variables had to be collapsed and single level rather than multi-level models had to be used.

- The combined Pakistani/Bangladeshi ethnic group were the ethnic grouping to be significantly and consistently under-represented in identification for ASD, although this partly reflects the low sample size for other ethnic groups. After controls for SES, White other and Black African groups were also under-represented.

- Gender was the strongest predictor of ASD identification (boys were six times more likely to be identified than girls), but ethnicity was the second strongest predictor, higher than any of the SES variables.

- There were raised risks of identification for students from low SEC homes (those where parents were in routine or semi-routine occupations). Family income did not have a significant association, but being entitled to a FSM (based on low income) doubled the odds of identification. After holding economic factors such as family SEC and entitlement to FSM constant, there were raised odds of identification in the most highly educated families (one or more parents holds a degree).

- The traditional SES measures derived from LSYPE2 had slightly more power than the NPD measures, accounting for 9.4% vs. 8.2% of the variance, but the difference was small.

MLD

- Our previous analysis using the national secondary cohort of 500,000+ students identified the over-representation of Black Caribbean and Pakistani young people with MLD as the major ethnic disproportionality (see Part 3). However we were not able to find statistically significant ethnic over-representation of these groups in the LSYPE2 sample, despite it being representative of the national cohort and being relatively large at nearly 12,000 cases. This illustrates the substantial problems that can arise in longitudinal samples as opposed to populations.

- The traditional SES measures were shown to have strong associations with the odds of identification with MLD. There was a consistent higher proportion of pupils identified with MLD in homes of lower *Parental SEC,* where parents had fewer *educational qualifications,* and there was lower reported *family income*. The traditional SES variables explain slightly more of the variation in identification of MLD than the NPD models, but requiring significantly more parameters to

estimate given the amount of missing data, so the NPD SES measures are arguably more efficient (reflected in lower AIC scores).

- An indication of the relative importance of each predictor showed that entitlement to FSM was the strongest predictor, followed by sex, IDACI, Family SEC and Parent Education. Ethnicity was barely significant after the inclusion of SES variables and family income was not statistically significant.

Conclusions

There are two main conclusions from this chapter, one substantive and one methodological.

First, the over-representation of Black Caribbean and Mixed White & Black Caribbean students for BESD/SEMH is not accounted for using detailed measures of SES such as parental occupation, parental education and family income. There is no indication that the measures in the NPD (entitlement to FSM and IDACI) are impoverished measures of SES. Indeed they account for a greater proportion of the Black Caribbean over-representation than measures of parental occupation, education and income. Combined with our finding regarding early attainment also not accounting for the over-representation of these groups (see part 3), there is strong evidence to suggest social processes are implicated in the over-representation of Black Caribbean and Mixed White & Black Caribbean students among those identified with BESD/SEMH.

Second, recent studies from the US based upon longitudinal surveys need to be interpreted with caution. Hibel et al (2010) analysed the US Early Childhood Longitudinal Study-Kindergarten (ECLS-K) tracking a sample of 11,000 students from age 3 to age 8/9. They report that Black and Latino students were actually under-represented for SEN after control for educational achievement and teacher's ratings of student's behaviour at Kindergarten entry. Morgan et. al. (2015) followed the same sample to age 11/12 and report the same result. It is notable though that these studies are unusual in that they report no significant over-representation of Black students for Intellectual Disability (ID), Emotional Disturbance (ED) or Learning Disability (LD) even *before* any adjustment for behaviour or attainment covariates, out of line with all other studies. The ECLS-K is representative and at 11,000 students is not a small sample, but given the US national incidence of ID at 0.7% and ED at 0.6% this represents just 77 and 66 students in the sample if it is representative, before considering any splits by gender, ethnicity, SES etc. The fact that we often did not find statistically significant relationships with ethnicity in the LSYPE2 sample, where we know these relationships do exist in the full cohort, indicates that even in large representative longitudinal studies there are substantial obstacles to accurate determination of disproportionality. Studies such as Hibel et. al. (2010) and Morgan et. al. (2015) need to be interpreted extremely cautiously and further studies such as ours, based on national population data, are urgently needed.

Overall conclusions

Key Findings

There is marked disproportionality for the following ethnic groups and SEN:

- Black Caribbean and Pakistani pupils are over-represented for MLD, Indian and Chinese pupils are under-represented;
- Black Caribbean and Mixed White & Black Caribbean pupils are substantially over-represented for SEMH;
- All Asian Groups (Indian, Pakistani, Bangladeshi and Other Asian) are substantially under-represented for SEMH and for ASD.

The over-representation for MLD can be accounted for by socio-economic factors, but the ethnic disproportionalities for SEMH and ASD remain substantial even after pupil background controls for age, sex and socio-economic deprivation. This is not because of the limited socio-economic measures available in the NPD, as we found the same results after control for parental social class, parental education and family income using the Second Longitudinal Study of Young People in England (LSYPE2).

Prior attainment/development also does not account for the ethnic disproportionality in SEMH and ASD. Literacy and mathematics measures from the Early Years Foundation Stage Profile at age 5 were strongly predictive of the likelihood of subsequent identification of MLD, and the Personal, Social and Emotional Development (PSED) measure was highly predictive of subsequent identification of SEMH and ASD. However, this did not remove the ethnic disproportionality for SEMH and ASD which remained substantial. The findings for the secondary cohort, accounting for age 11 English and mathematics attainment on-entry to secondary school, led to the same conclusion.

Local Authorities (LAs) account for little (2%-6%) of the variation in the identification of SEN. Patterns of disproportionality vary little in direction across LAs e.g. of 113 LAs with sufficient data for SEMH calculations, 84 show over-representation of Black Caribbean/Mixed White & Black Caribbean pupils, none show under-representation. Similarly, of 94 LAs with sufficient data for ASD calculations, 79 show under-representation of Asian pupils, only three show over-representation. This consistency suggests that variation in LA policy and practice plays a limited role in the over-representation of Black Caribbean/Mixed White & Black Caribbean pupils with SEMH or the under-representation of Asian pupils with ASD.

There is variation between schools in the frequency with which they identify SEN, but schools play a limited role in accounting for ethnic disproportionality, with the notable exception of identification of SEMH in secondary school. In

null models, around one-fifth of the variance in MLD is between schools (22%-25%) somewhat less for SEMH (13%-15%) and much less for ASD (11%-12%). Some of this variation can be explained by the socio-economic composition of the pupil intake, and by factors like school size and type (e.g. Grammar schools had very few SEN pupils). However, differences between schools played little role in accounting for ethnic disproportionality, with the notable exception of SEMH in secondary schools. Differences between secondary schools account for a substantial part of the over-representation of Black Caribbean and Mixed White and Black Caribbean pupils with SEMH. i.e. their over-representation occurs much more in some secondary schools than it does in others.

Longitudinal studies, even with large representative samples, can often be under-powered to detect relatively low incidence outcomes like type of SEN for ethnic minority groups. Results from sample studies need to be interpreted with caution and more population level studies, like those reported here, are required.

Detailed findings

Moderate Learning Difficulties (MLD)

Pakistani pupils (OR= 1.36) and Black Caribbean pupils (OR= 1.38) were over-represented for MLD relative to White British pupils. Indian (OR= 0.56) and Chinese (OR= 0.30) were substantially under-represented.

Despite changes in prevalence, the extent of the ethnic disproportionality noted above has not changed notably since 2005. Prevalence rates for MLD increased from 2.6% in 2005 to 4.0% in 2016. In part this reflects the fact that from 2015 onwards type of SEN was requested for all pupils on School Support, not just those on the former School Action Plus, so more pupils are recorded as having a specific type of need. However, the change in prevalence did not alter the extent of ethnic disproportionality.

The over-representation of Pakistani and Black Caribbean pupils could be accounted for by socio-economic factors. Pupils were more likely to be identified with MLD if they were entitled to a Free School Meal (OR= 2.4), lived in a deprived neighbourhood (OR= 1.9), were boys (OR= 1.7) and were young for their year group (summer-born pupils OR= 1.8). After controlling for these factors, Pakistani and Black Caribbean pupils were no more likely to be identified than White British pupils with similar characteristics.

Accounting for attainment and/or social development at the start of school made little difference to the results by ethnic group. Literacy and mathematics measures from the Early Years Foundation Stage Profile (EYFSP) at age 5 were very strong predictors of a pupil's likelihood of subsequent identification of MLD during primary school. However, it did not change the pattern of ethnic group difference, with many ethnic groups (particular Black African, Indian and Bangladeshi

pupils) less likely to be identified with MLD than White British pupils with the same prior attainment and socio-economic background. Similar conclusions apply when accounting for reading and mathematics test scores at age 11 for the secondary longitudinal cohort.

Differences between LAs and schools made little contribution to ethnic disproportionality for MLD. Local Authorities (LAs) account for very little (5%-6%) of the variation in identification. There is more variation at the school level (22%-26%), with some schools more likely to have pupils identified with MLD than others, and this partly reflects the characteristics of the pupils attending the school (e.g. more pupils identified in small schools and those with more deprived intakes). Importantly though, accounting for differences between schools did not materially alter the ethnic coefficients for under/over-representation**,** either at primary or secondary phases.

Social, Emotional and Mental Health (SEMH)

Black Caribbean (OR= 2.29) and Mixed White and Black Caribbean (OR= 1.94) pupils were substantially over-represented relative to White British pupils. Asian groups were all substantially under-represented, as was the White Other group (OR= 0.57).

The extent of ethnic disproportionality for the above groups has remained constant since 2005. Prevalence rates increased from 1.9% in 2005 to 2.8% in 2016, although as stated earlier this partly reflects the increase since 2015 in the number of pupils for whom data on type of need is requested. Importantly, though, the ethnic disproportionality identified above has not altered with the change in terminology from BESD to SEMH. This is perhaps not surprising since displaying "challenging, disruptive or disturbing behaviour" remains central to the description of SEMH (DFE, 2015), whatever the putative drivers of such behaviour.

Demographic and socio-economic variables had very strong associations with identification of SEMH, but controlling for these factors did not account for the ethnic over-representation. The odds of being identified with SEMH needs were much higher for boys than girls (OR= 3.2); for pupils entitled to FSM (OR= 3.1), for pupils from disadvantaged neighbourhoods (OR= 1.9) and for pupils in secondary school, particularly Y10 and Y11 (OR= 2.1 and OR= 2.4 compared to Y1). Controlling for these factors attenuated but did not eliminate the over-representation of Black Caribbean (OR= 1.43) and Mixed White & Black Caribbean (OR= 1.38) pupils.

Similarly, controlling for prior attainment/development at the start of school did not account for Black Caribbean and Mixed White & Black Caribbean over-representation. Literacy and mathematics scores at age 5 had little association with subsequent identification of SEMH, but a below average Personal, Social and Emotional Development (PSED) score at age 5 raised the odds substantially (HR=

2.54). The mean PSED scores for Black Caribbean and Mixed White & Black Caribbean pupils were lower than the national average, but even after adjusting for this Black Caribbean (HR= 1.42) and Mixed White & Black Caribbean (HR= 1.46) pupils were still over-represented. The findings for the secondary cohort, accounting for English and mathematics national test scores at age 11 on subsequent identification of SEMH during secondary school led to the same conclusion, with Black Caribbean (OR= 1.37) and Mixed White & Black Caribbean (OR= 1.53) pupils remaining over-represented.

Secondary schools seem to account for a significant part of the over-representation of Black Caribbean and Mixed White & Black Caribbean pupils with SEMH. In secondary schools the ORs for Black Caribbean and Mixed White & Black Caribbean pupils reduced substantially between single-level and multi-level models, from OR= 1.47 to 1.14 and from OR=1.47 to 1.29 respectively. This indicates that differences between schools play a part in the over-representation of these two specific ethnic groups. Our longitudinal analyses indicate that over-representation was reduced when account was taken of school composition factors, particularly in secondary schools. For example, schools in the top two quintiles of %FSM, and in the top two quintiles for % Black Caribbean pupils, had significantly raised odds of identification, and allowing for this did reduce the Black Caribbean and Mixed White & Black Caribbean over-representation. This suggests a particular focus on the context of, and processes occurring within, schools serving high deprivation communities and with large proportions of Black Caribbean and Mixed White & Black Caribbean pupils. What drives these associations is unknown, and could include unmeasured factors associated with high deprivation (e.g. high levels of crime, violence or gang culture), negative peer effects (such as disaffection or disengagement) or school policies (e.g. pre-emptive or zero tolerance disciplinary strategies).

Variation between LAs is minimal, accounting for <2% of variation in identification of SEMH. Of 113 LAs with sufficient data for SEMH calculations, 84 show over-representation for the combined Black Caribbean/MWBC group, none show under-representation. Nevertheless, there is a range in the risk ratios for 2016 from 0.77 in Newham to 3.15 in Barnsley. Data should be monitored annually to determine if any consistent LA patterns emerge.

Care needs to exercised in generalisations about 'Black' pupils. Black African pupils represent 3.7% of all pupils in England, a much larger group than either Black Caribbean (1.2%) or Mixed White and Black Caribbean (1.5%) pupils. They experience similar levels of socio-economic disadvantage yet they are not over-represented for SEMH, and are actually under-represented in the adjusted ORs, both in relation to socio-economic disadvantage and to prior attainment. This indicates that in the England context, care needs to be exercised in generalisations about 'Black' pupils. Similar differences have been reported for other outcomes such

as exclusion from school; attitudes, aspiration and motivation; and academic achievement, and may be related to recency of migration (e.g. Strand, 2011, 2012).

Autistic Spectrum Disorders (ASD)

There was substantial ethnic disproportionality for ASD. Black Caribbean and Black Other pupils were over-represented (both ORs= 1.34) compared to White British pupils. Asian groups were under-represented, particularly Indian (OR= 0.46) and Pakistani pupils (OR= 0.54) where the odds of identification were half those for White British pupils. White Other pupils (OR= 0.60) were also under-represented.

There was more variation in ethnic disproportionality over time than was the case for other SEN. Black Caribbean pupils were not over-represented 2005-2009 but have been consistently over-represented since 2011 (OR= 1.12 in 2005 to OR= 1.34 by 2016). White Other groups were not under-represented 2005-2009 but have been consistently under-represented since 2011 (OR= 0.96 in 2005 to OR= 0.60 by 2016). On a positive note, the under-representation of Bangladeshi pupils has steadily declined (from OR= 0.38 in 2005 to OR= 0.79 by 2016).

Demographic and socio-economic variables had strong associations with identification of ASD. Controlling for these factors accounted for the over-representation of Black Caribbean and Black-Other groups, but did not account for the under-representation of Asian pupils. The odds of being identified with ASD were much higher for boys than girls (OR= 5.4) and for pupils entitled to FSM (OR= 2.3), and were slightly raised for pupils from more disadvantaged neighbourhoods (OR= 1.2). Controlling for these factors had little or no impact on the under-representation of Asian pupils, who were still about half as a likely as White British pupils to have an identification of ASD. On the other hand, Black Caribbean and Black Other pupils were no longer over-represented (OR= 1.12 and OR= 1.13), suggesting that their over-representation was largely attributable to socio-economic factors.

Controlling for prior attainment/development at the start of school did not change the pattern of ethnic disproportionality. Below average Personal, Social and Emotional Development (PSED) scores at age 5 were associated with substantially increased odds of ASD identification (HR= 3.2), and the mean PSED scores for Pakistani and Bangladeshi pupils were below the national average, but after adjusting for these scores Asian groups remained under-represented. Similarly, for the secondary cohort, higher English and mathematics national test scores at age 11 were associated with lower odds of identification with ASD, but controlling for prior attainment did not alter the Asian under-representation.

LA and school variability was small, and school composition variables had little impact on ethnic disproportionality. Around 4% of variance was at the LA level and 11%-12% at the school level, much lower than for MLD or SEMH. Generally, school level factors had little impact on ethnic disproportionality. However,

both longitudinal cohorts suggested that pupils were somewhat more likely to be identified in schools in the top two quintiles for % Asian pupils, particularly among secondary schools, indicating that a high concentration of Asian pupils slightly moderated the effect, but overall Asian pupils remained substantially under-represented compared to White British pupils.

Parental education qualifications may be an important factor in identification of ASD. Some, predominantly US, research has suggested that high socio-economic families are *more* likely to receive an ASD diagnosis (e.g. Durkin et al, 2010) while our data indicates the opposite. Our NPD measures are of socio-economic *dis*advantage, which may be blunt in differentiating at the more advantaged end of the SES range. However, our analysis of LSYPE2, using parental occupation, educational qualifications and family income, broadly confirmed the NPD results, with pupils from low SES homes (parents in routine and semi-routine occupations) more likely to have an ASD identification than those in managerial and professional households (OR= 3.0 and 2.6 respectively). It may be that in England the NHS provides more equitable access to services with fewer financial barriers than in the US, and our study is based in schools where all children can be assessed rather than in clinics or other settings. Nevertheless, we note that once parental occupation was controlled, the odds of ASD identification were twice as high in homes where one or more parent held a degree compared to similar homes where parents' highest educational qualifications were below degree level. This does suggest that parental awareness and access to resources may be an issue.

The causes of ethnic disproportionality in identification of ASD are likely to be varied. Less extreme needs on the autistic spectrum can be subtle, identified by nuances in the use of language for social communication. These may be more difficult to identify if the first language of the assessor and pupil are not congruent, as might be the case for many pupils of Asian heritage. It may also be that these are communities with lower awareness of autism, parents' rights and relevant services; where having a child with SEN is particularly stigmatizing; where cultural or linguistic barriers impede access to services; or where the services available do not meet their needs (Corbett & Perapa, 2007). In any event, there is a need to raise awareness of ASD among Asian communities, improve outreach and review the extent to which services are configured appropriately.

Implications for policy and practice

- LAs, multi-academy trusts (MATs) and schools must have due regard to the Public Sector Equality Duty (PSED) requirements, and should monitor ethnic disproportionality in the identification of SEN.
- LAs or MATs with high levels of disproportionality should further investigate practices in their areas/trusts. Schools should identify priorities for the partnerships within which they work, so they can pool resources and develop effective responses.

- OFSTED should incorporate data on ethnic disproportionality into pre-inspection reports for LA SEND inspections, and include the issue of ethnic disproportionality within the LA SEND inspection.
- The original detailed guidance on data collection by type of SEN (DFE, 2005) no longer exists following the new SEND Code of Practice. The DFE should consider new guidance on definitions and criteria for defining different types of SEN.
- Secondary schools in particular should review their processes around the identification of SEMH needs, given variability between schools is a strong component of ethnic disproportionality in this domain.
- LAs and schools need to raise awareness of ASD among Asian communities, improve outreach and review the extent to which the services are configured appropriately for access by ethnic minority groups.
- Teachers need to be aware of the significant over-identification of summer born pupils for MLD and to consider carefully whether they are making sufficient allowance for the age of the child when forming their judgements.

References

Artiles, A. J., Kozleski, E. B., Trent, S. C., Osher, D. & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968-2008: A critique of underlying views of culture. *Exceptional Children, 76(3),* 279-299.

Austin, P. C., & Merlo, J. (2017). Intermediate and advanced topics in multilevel logistic regression analysis. *Statistics in Medicine*, *36(20),* 3257–3277. http://doi.org/10.1002/sim.7336

Bhasin, T. K., & Schendel, D. (2007). Sociodemographic Risk Factors for Autism in a US Metropolitan Area. *Journal of Autism and Developmental Disorders, 37(4), 667-677. doi:10.1007/s10803-006-0194-y*

Browne, W. J., Subramanian, S. V., Jones, K. and Goldstein, H. (2005), Variance partitioning in multilevel logistic models that exhibit overdispersion. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168*, 599-613. doi:10.1111/j.1467-985X.2004.00365.x

Centre for Disease Control and Prevention (CDCP) (2006). Mental health in the US: Parental reports of diagnosed autism in children aged 4-7 years- US 2003-2004. MMWR Morb Mortality Weekly Report 2006; 55(17): 481-486.

Centre for Disease Control and Prevention (CDCP) (2009). Prevalence of Autism Spectrum Disorders - Autism and Developmental Disabilities Monitoring Network, US, 2006. MMWR surveillance Summary 2009; 58(NoSS-10). http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5810a1.htm

Coard, B. (1971). *How the West Indian child is made educationally subnormal in the British school system: The scandal of the black child in schools in Britain.* London: New Beacon for the Caribbean Education and Community Workers' Association.

Coe, R. (2004). Issues arising from the use of effect sizes in analysing and reporting research. In I. Schagen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (80-100). Slough: National Foundation for Educational Research.

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* New York, NY: Routledge Academic.

Cooc, N. (2017). Examining Racial Disparities in Teacher Perceptions of Student Disabilities. *Teachers College Record, 119(7), 1-32.*

Corbett, C., & Perepa, P. (2007). *Missing out? Autism, education and ethnicity: The reality for families today*. London: The National Autistic Society.

Coutinho, M. J. & Oswald, D. P. (2000).Disproportionate representation in special education: A synthesis and recommendations. *Journal of Child and Family Studies, 9(2),* 135-156.

Coutinho, M. J., Oswald, D. P., & Best, A. M. (2002). The Influence of Sociodemographics and Gender on the Disproportionate Identification of Minority Students as Having Learning Disabilities. *Remedial & Special Education, 23(1), 49.*

Croen, L. A., Grether, J. K., & Selvin, S. (2002). Descriptive Epidemiology of Autism in a California Population: Who Is at Risk? *Journal of Autism and Developmental Disorders, 32(3), 217-224. doi:10.1023/a:1015405914950.*

Delobel-Ayoub, M., Ehlinger, V., Klapouszczak, D., Maffre, T., Raynaud, J.-P., Delpierre, C., & Arnaud, C. (2015). Socioeconomic Disparities and Prevalence of Autism Spectrum Disorders and Intellectual Disability. PLoS ONE, 10(11), e0141964-e0141964. doi:10.1371/journal.pone.0141964

DFE (2011). *Special Educational Needs in England: January 2011 (SFR 14/2011).* London: Department for Education.

DFE (2013). *Schools, pupils and their characteristics January 2013 (SFR 21/2013).* London: Department for Education.

DFE (2014). *Special educational needs and disability: Research priorities and questions (DFE 0086-2014).* London: Department for Education.

DFE (2014). *Special Educational Needs in England: January 2014 (SFR 26/2014).* London: Department for Education.

DFE (2014). *Children with Special Educational Needs 2014: An analysis (SFR 31/2014).* London: Department for Education.

DFE (2015). *Special Educational Needs and Disability code of practice: 0 to 25 years (DFE-00205-2013).* London: Department for Education.

DFE (2015). *Special Educational Needs in England: January 2015 (SFR 25/2015).* London: Department for Education.

DFE (2015). *Statements of SEN and EHC plans: England 2015 (SFR 14/2015).* London: Department for Education.

DFE (2016). *Schools, pupils and their characteristics January 2016 (SFR 20/2016).* London: Department for Education.

DFE (2017). *Special Educational Needs in England: January 2017 (SFR 37/2017).* London: Department for Education.

Donovan, S. & Cross, C. (2002). Minority students in special and gifted education. Washington, D.C.: National Academy Press.

Dunn, L. M. (1968). Special education for the mildly retarded – Is much of it justifiable? *Exceptional Children, 23*, 5-21.

Durkin, M. S., Maenner, M. J., Meaney, F. J., Levy, S. E., di Guiseppi, C., Nicholas, J. S., … Schieve, L. A. (2010). Socioeconomic inequality in the prevalence of autism spectrum disorder: Evidence from a U.S. cross-sectional study. *PLoS*

ONE, *5*(7). https://doi.org/10.1371/journal.pone.0011551

Fischer Family Trust (2011). *KS1 estimates based on EYFSP (09/11).* Retrieved 27/09/17 from the world web at: http://csapps.norfolk.gov.uk/csshared/ecourier2/fileoutput.asp?id=11608

Flores, G. and the Committee on Pediatric Research. (2010). Racial and Ethnic Disparities in the Health and Health Care of Children. *Pediatrics, 125(4), 979-1020.*

Goldstein, H., Browne, W.J., and Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics 1*, 223-231.

Hibel, J., Farkas, G. & Morgan, P. L. (2010). Who is placed into special education? *Sociology of education, 83(4),* 312-332.

Ilie, S., Sutherland, A., & Vignoles, A. (2017). Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation. *British Educational Research Journal, 43(2),* 253-274. doi:10.1002/berj.3260

Hosp, J. L. & Reschly, D. J. (2004). Disproportionate representation of minority students in special education: Academic, demographic and economic predictors. *Exceptional Children, 70,* 185-199.

Jarquin, V., Wiggins, L., & Schieve, L. (2011). Racial disparitries in community identification of autism spectrum disorders over time: Metropolitan Atlanta, Georgi, 2000-2006. Lournal of Developmental & Behavioural Pediatrics, 32(3), 179-187.

Kogan, M. D., Blumberg, S. J., Schieve, L. A., Boyle, C. A., Perrin, J. M., Ghandour, R. M., . . . van Dyck, P. C. (2009). Prevalence of Parent-Reported Diagnosis of Autism Spectrum Disorder Among Children in the US, 2007. Pediatrics, 124(5), 1395-1403. doi:10.1542/peds.2009-1522.

Larsen, K. , Petersen, J. H., Budtz- Jørgensen, E. and Endahl, L. (2000), Interpreting Parameters in the Logistic Regression Model with Random Effects. *Biometrics, 56,* 909-914. doi:10.1111/j.0006-341X.2000.00909.x

Larsen, K. and Merlo, J. (2005) Appropriate Assessment of Neighborhood Effects on Individual Health: Integrating Random and Fixed Effects in Multilevel Logistic Regression. *American Journal of Epidemiology, 161, 81-88. doi: 10.1093/aje/kwi017.*

Larsson, H. J., Eaton, W. W., Madsen, K. M., Vestergaard, M., Olesen, A. V., Agerbo, E., . . . Mortensen, P. B. (2005). Risk Factors for Autism: Perinatal Factors, Parental Psychiatric History, and Socioeconomic Status. *American Journal of Epidemiology, 161(10), 916-925. doi:10.1093/aje/kwi123.*

Leonard, H., Glasson, E., Nassar, N., Whitehouse, A., Bebbington, A., Bourke, J., . . . Stanley, F. (2011). Autism and Intellectual Disability Are Differentially Related to Sociodemographic Background at Birth. *PLoS ONE, 6(3), e17875.*

doi:10.1371/journal.pone.0017875.

Lindsay, G., & Strand, S. (2016). Children with Language Impairment: Prevalence, Associated Difficulties, and Ethnic Disproportionality in an English Population. *Frontiers in Education, 1(2), 1-14. doi:10.3389/feduc.2016.00002.*

Mandell, D., & al, e. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health, 99(3), 493-498.*

McCoy, S., Banks, J., & Shevlin, M. (2012). School matters: how context influences the identification of different types of special educational needs. *Irish Educational Studies, 31(2), 119-138. doi:10.1080/03323315.2012.669568*

Merlo, J., Chaix, B., Ohlsson, H., Beckman, A., Johnell, K., Hjerpe, P., … Larsen, K. (2006). A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health*, *60*(4), 290–297. http://doi.org/10.1136/jech.2004.029454

Merlo, J., Wagner, P., Ghith, N., & Leckie, G. (2016) An Original Stepwise Multilevel Logistic Regression Analysis of Discriminatory Accuracy: The Case of Neighbourhoods and Health. *PLoS ONE 11(4): e0153778. https://doi.org/10.1371/journal.pone.0153778*

Morgan, P. L., Farkas, G., Hilemeier, M. M., Mattison, R., Maczuga, S., Li, H. & Cook, M. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher*, *44)5, 278-292.*

Morgan, P. L., & Farkas, G. (2016). Are We Helping All the Children That We Are Supposed to Be Helping? *Educational Researcher, 45(3), 226-228. doi:10.3102/0013189x16644607*

Morgan, P., Farkas, G., Cook, M., Strassfeld, N., & Hillemeier, M. (2017). Are Black Children Disproportionately Overrepresented in Special Education? A Best-Evidence Synthesis. *Exceptional Children, 83(2), 181-198., 83*(2), 181-198.

Morgan, P. L., Farkas, G., Cook, M., Strassfeld, N. M., Hillemeier, M. M., Pun, W. H., . . . Schussler, D. L. (2018). Are Hispanic, Asian, Native American, or Language-Minority Children Overrepresented in Special Education? *Exceptional Children, 84(3), 261-279. doi:10.1177/0014402917748303.*

Morrier, M. J., & Hess, K. L. (2012). Ethnic Differences in Autism Eligibility in the United States Public Schools. *The Journal of Special Education, 46(1), 49-63. doi:10.1177/0022466910372137.*

Newschaffer, C. J., Croen, L. A., Daniels, J., Giarelli, E., Grether, J. K., Levy, S. E., . . . Windham, G. C. (2007). The Epidemiology of Autism Spectrum Disorders. *Annual Review of Public Health, 28(1), 235-258.*

doi:doi:10.1146/annurev.publhealth.28.021406.144007.

Norwich, B. & Eaton, A. (2015). The new Special Educational Needs (SEN) legislation in England and implications for services for children and young people with social, emotional and behavioural difficulties. Emotional and Behavioural Difficulties, 20(2), 117-132.

Perepa, P. (2014). Cultural basis of social 'deficits' in autism spectrum disorders. *European Journal of Special Needs Education, 29(3), 313-326. doi:10.1080/08856257.2014.908024*

Rai, D., Lewis, G., Lundberg, M., Araya, R., Svensson, A., Dalman, C., . . . Magnusson, C. (2012). Parental Socioeconomic Status and Risk of Offspring Autism Spectrum Disorders in a Swedish Population-Based Study. *Journal of the American Academy of Child & Adolescent Psychiatry, 51(5), 467-476.e466. doi:10.1016/j.jaac.2012.02.012*

Rice, C. (2009). Prevalence of Autistic Spectrum Disorders - Autism and Developmental Disabilities Monitoring Network, United States, 2006. Atlanta: Centers for Disease Control and Prevention. MMWR Surveill Summ 2009;58 (NoSS-10). Available from: http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5810a1.htm.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15, 351-357.*

Saunders, C., Brunton-Smith, I., Garrod, B., Stepanek, M., Vignoles, A., Sutherland, A. (2016). *Missing Key Stage 2 data in LYSPE2: technical report.* London: Department for Education.

Shifrer, D., Muller, C. & Callahan, R. (2011). Disproportionality and learning disabilities: Parsing apart race, socioeconomic status and language. *Journal of Learning Disabilities, 44(3),* 246-257.

Skiba, R., Poloni-Staudinger, L., Simmons, A., Feggins-Azziz, L. & Chung, C.-G. (2005). Unproven links: Can poverty explain ethnic disproportionality in special education? *Journal of Special Education, 39(3),* 130-144.

Skiba, R., Simmons, A., Ritter, S., Gibb, A., Rausch, M., Cuadrado, J. & Chung, C.-G. (2008). Achieving equity in special education: History, status and current challenges. *Exceptional Children, 74(3),* 264-288.

Skiba, R. J., Artiles, A. J., Kozleski, E. B., Losen, D. J., & Harry, E. G. (2016). Risks and Consequences of Oversimplifying Educational Inequities: A Response to Morgan et al. (2015). *Educational Researcher, 45(3), 221-225. doi:10.3102/0013189x16644606*

Snowling, M. J., Hulme, C., Bailey, A., Stothard, S., & Lindsay, G. (2011). *Better communication research programme: Language and Literacy Attainment of Pupils during Early Years and through KS2: Does teacher assessment at five*

provide a valid measure of children's current and future educational attainments? (DFE RR172a). London: Department for Education.

Solon, G., Haider, S. J., & Woodridge, J, M. (2015). What are we weighting for? *Journal of Human Resources, 50(2),* 301-316. doi:10.3368/jhr.50.2.301

Strand, S. (2010). Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement, 21(3),* 289-314. doi: 10.1080/09243451003732651.

Strand, S. (2011). The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal, 37*(2), 197-229. doi:doi:10.1080/01411920903540664.

Strand, S. (2012). The White British–Black Caribbean achievement gap: tests, tiers and teacher expectations. *British Educational Research Journal, 38(1), 75-101. doi:10.1080/01411926.2010.526702*

Strand, S. (2014a). Ethnicity, gender, social class and achievement gaps at age 16: Intersectionality and 'Getting it' for the white working class. *Research Papers in Education, 29(2),* 131-171.

Strand, S. (2014b). Doubly disadvantaged? Being poor in an affluent school. Paper presented to the *Annual Conference of the British Educational Research Association (BERA), University of London Institute of Education, 23-25 September 2014.*

Strand, S. (2014c). Doubly disadvantaged? Being poor in an affluent school. *Paper presented at the Annual Conference of the British Educational Research Association, 23-25 September 2014, University of London Institute of Education, London.*

Strand, S. (2015). *Ethnicity, deprivation and educational achievement at age 16 in England: Trends over time (DFE Research Report 439B).* London: Department for Education.

Strand, S. & Fletcher, J. (2014). A quantitative analysis of exclusions from English secondary schools. Oxford: University of Oxford, Department of Education.

Strand, S. & Hessel, A. (2018). *English as an Additional Language, Proficiency in English and pupils' educational achievement: An analysis of Local Authority data*. Cambridge: The Bell Foundation.

Strand, S. & Lindsay, G. (2009). Evidence of ethnic disproportionality in special education in an English population. *Journal of Special Education, 43(3),* 174-190.

Strand, S. & Lindsay, G. (2012). *Ethnic disproportionality in the identification of Speech Language and Communication Needs (SLCN) and Autistic Spectrum Disorders (ASD): 2005-2011.* London: Department for Education.

Strand, S., Malmberg, L. & Hall, J. (2015). *English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database*. Oxford: University of Oxford, Department of Education.

Sullivan, A. (2013). School-based autism identification: Prevalance, racial disparities and systemic correlates. *School Psychology Review, 42(1), 298-316.*

Sullivan, A. & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. Exceptional Children, 79(4), 475-494.

Thomas, P., Zahorodny, W., Peng, B., Kim, S., Jani, N., Halperin, W., & Brimacombe, M. (2012). The association of autism diagnosis with socioeconomic status. *Autism, 16(2), 201-213. doi:10.1177/1362361311413397*

TNS-BRMB (2015). *LSYPE2 User Guide to the Datasets: Wave 1.* Available from the WWW at: http://doc.ukdataservice.ac.uk/doc/7838/mrdoc/pdf/7838user_guide_wave_1.pdf

Travers, J. C., Tincani, M., & Krezmien, M. P. (2011). A Multiyear National Profile of Racial Disparity in Autism Identification. *The Journal of Special Education, 47*(1), 41-49.

Travers, J. C., Krezmien, M. P., Mulcahy, C., & Tincani, M. (2014). Racial Disparity in Administrative Autism Identification Across the United States During 2000 and 2007. The Journal of Special Education, 48(3), 155-166. doi:10.1177/0022466912454014.

Travers, J., & Krezmien, M. (2018). Racial Disparities in Autism Identification in the United States During 2014. Exceptional Children, 84(4), 403-419. doi:10.1177/0014402918771337.

Waitoller, F. R., Artiles, A. J., & Cheney, D. A. (2010). The miner's canary: A review of overrepresentation research and explanations. *Journal of Special Education, 44(1),* 29-49.

Williams, E., Thomas, K., Sidebotham, H., & Emond, A. (2008). Prevalence and characteristics of autistic spectrum disorders in the ALSPAC cohort. *Developmental Medicine & Child Neurology, 50(9), 672-677. doi:doi:10.1111/j.1469-8749.2008.03042.x.*

UK Data Forum (2013). *The UK strategy for data resources for social and economic research 2013-2018.* London: UK Data Forum.

US Department of Education (2014). *36th Annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2014.* Washington, D.C.: US Department of Education. Available at http://www2.ed.gov/about/reports/annual/osep/index.html [Accessed 23 February 2016].

Research Team

Steve Strand is Professor of Education at the University Oxford, Department of Education.

Steve Strand has been Professor of Education at the University of Oxford since January 2013. Previously he was Professor of Education at the University of Warwick (2005-2012). Prior to that Steve was Senior Assessment Consultant at GL Assessment, the UK's leading educational test and assessment publisher (1998-2005) and Head of Research and Evaluation at Wandsworth and at Croydon Local Education Authorities (1988-1998). In these roles he has been responsible for pioneering work on 'value added' analyses of school performance. Steve has particular skills in statistical modelling and analysis of large scale longitudinal datasets including the first and second Longitudinal Studies of Young People in England (LSYPE) and the National Pupil Database (NPD). He leads the Quantitative Methods (QM) hub at the Oxford University Department of Education. He is the author of over 100 peer-reviewed journal articles, book chapters and research reports. He has been an adviser to the DFE, a specialist adviser to the House of Common Education Select Committee inquiry into White Working Class achievement and is a member of the REF 2021 panel for Education. For further details see: http://www.education.ox.ac.uk/people/steve-strand/

Ariel Lindorff is Research Fellow at the University of Oxford, Department of Education.

Ariel's research involves the analysis of large-scale secondary datasets using quantitative methods, as well as mixed methods approaches to investigating educational effectiveness and equity at the classroom, school and system levels. Her DPhil, completed at the University of Oxford in 2016, involved a mixed methods study of networks of schools in a USA school district using a combination of multilevel analysis of secondary achievement and administrative data and case study methodology. Before coming to Oxford, she earned a Master's degree in Applied Mathematics and Statistics from Hunter College, City University of New York. She previously worked as a secondary mathematics teacher in the USA for approximately eight years, and holds QTS in the UK. Her particular research skills include secondary analysis of large-scale cross-sectional and longitudinal datasets, mixed methods research design, and subject-specific and generic approaches to the observation of classroom practice. Ariel leads the MSc Education (Research Design and Methodology) course in the Oxford University Department of Education, and has previously co-convened the Quantitative Methods (QM) Hub seminar series in the same department. For further details see: http://www.education.ox.ac.uk/people/ariel-lindorff/

Appendices

Appendix A: Cross-tabulation of primary and secondary SEN, 2016

Table A.1: Cross-tabulation of primary and secondary types of SEN (2016, Y1-11)

		Secondary need														
		None	SpLD	MLD	SLD	PMLD	SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA	Total
Primary need	None	5470700	0	0	0	0	0	0	0	0	0	0	0	0	0	5470700
	SpLD	121750	467	2585	77	23	7183	4910	884	373	433	83	825	2397	407	142397
	MLD	212751	3326	726	194	26	15803	16831	2324	986	783	137	2138	3420	850	260295
	SLD	12619	160	92	24	81	758	3688	4014	445	496	71	1430	500	11	24389
	PMLD	4354	28	24	137	7	63	518	456	105	805	228	963	183	0	7871
	SEMH	144752	5491	12303	290	32	382	7800	3305	356	295	147	820	3809	681	180463
	SLCN	130608	5883	17226	931	98	10661	386	3105	1022	450	227	1959	3505	748	176809
	ASD	56278	2132	4212	3221	126	6850	9635	148	237	253	138	700	2035	53	86018
	HI	13230	450	954	40	10	499	1689	138	32	155	43	270	383	39	17932
	VI	7188	312	686	91	37	303	402	155	161	10	36	296	276	30	9983
	MSI	1288	53	81	26	35	95	121	31	37	20	1	102	50	4	1944
	PD	18343	841	2553	520	215	963	2369	302	293	458	116	41	649	87	27750
	Other	41381	1051	1650	113	23	2144	1715	460	176	142	47	358	742	269	50271
	NSA	31557	215	689	3	1	525	414	69	23	21	4	34	222	16	33793
	Total	6266799	20409	43781	5667	714	46229	50478	15391	4246	4321	1278	9936	18171	3195	6490615
	Row percentages															
	None	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
	SpLD	85.5	0.3	1.8	0.1	0.0	5.0	3.4	0.6	0.3	0.3	0.1	0.6	1.7	0.3	100.0
	MLD	81.7	1.3	0.3	0.1	0.0	6.1	6.5	0.9	0.4	0.3	0.1	0.8	1.3	0.3	100.0
	SLD	51.7	0.7	0.4	0.1	0.3	3.1	15.1	16.5	1.8	2.0	0.3	5.9	2.1	0.0	100.0
	PMLD	55.3	0.4	0.3	1.7	0.1	0.8	6.6	5.8	1.3	10.2	2.9	12.2	2.3	0.0	100.0
	SEMH	80.2	3.0	6.8	0.2	0.0	0.2	4.3	1.8	0.2	0.2	0.1	0.5	2.1	0.4	100.0
	SLCN	73.9	3.3	9.7	0.5	0.1	6.0	0.2	1.8	0.6	0.3	0.1	1.1	2.0	0.4	100.0
	ASD	65.4	2.5	4.9	3.7	0.1	8.0	11.2	0.2	0.3	0.3	0.2	0.8	2.4	0.1	100.0
	HI	73.8	2.5	5.3	0.2	0.1	2.8	9.4	0.8	0.2	0.9	0.2	1.5	2.1	0.2	100.0
	VI	72.0	3.1	6.9	0.9	0.4	3.0	4.0	1.6	1.6	0.1	0.4	3.0	2.8	0.3	100.0
	MSI	66.3	2.7	4.2	1.3	1.8	4.9	6.2	1.6	1.9	1.0	0.1	5.2	2.6	0.2	100.0
	PD	66.1	3.0	9.2	1.9	0.8	3.5	8.5	1.1	1.1	1.7	0.4	0.1	2.3	0.3	100.0
	Other	82.3	2.1	3.3	0.2	0.0	4.3	3.4	0.9	0.4	0.3	0.1	0.7	1.5	0.5	100.0
	NSA	93.4	0.6	2.0	0.0	0.0	1.6	1.2	0.2	0.1	0.1	0.0	0.1	0.7	0.0	100.0
	Total	96.6	0.3	0.7	0.1	0.0	0.7	0.8	0.2	0.1	0.1	0.0	0.2	0.3	0.0	100.0

Appendix B: Adjusted ORs for ethnic groups and additional pupil background characteristics after including EAL as a predictor

Table B.2: ORs for ethnic groups by SEN type, 2016, including EAL

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
	SpLD		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*		*ASD*		*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
White Irish	1.12	*	0.77	*	0.86		1.29		0.86	*	0.93		1.01		1.21		0.74		1.42		0.95		0.99		1.18	
Traveller Irish	***2.05***	*	***2.45***	*	1.01		1.21		***1.53***	*	1.47	*	***0.31***	*	1.28		1.49		1.13		0.94		***2.02***	*	***3.16***	*
Traveller Gypsy/Roma	***1.70***	*	***2.61***	*	***1.63***	*	0.86		1.41	*	***1.62***	*	***0.41***	*	***2.14***	*	1.39	*	***0.31***		0.95		1.42	*	***2.30***	*
White other groups	0.68	*	***0.66***	*	0.82	*	1.01		0.71	*	0.94	*	0.88	*	***0.52***	*	***0.63***	*	0.76		***0.65***	*	0.84	*	0.89	*
Mixed White & African	0.77	*	***0.62***	*	0.85	*	1.15		0.99		0.86	*	0.93		***0.53***	*	***0.41***	*	0.67		***0.65***	*	0.77	*	0.77	*
Mixed White & Caribbean	0.96	*	0.90	*	0.72	*	1.01		1.38	*	0.90	*	0.97		***0.65***	*	0.84	*	***0.65***	*	0.76	*	0.95		1.13	*
Mixed White & Asian	***0.56***	*	***0.62***	*	0.83	*	0.94		0.71	*	0.79	*	0.87	*	0.67	*	0.68	*	***0.53***	*	0.70	*	0.68	*	0.70	*
Any other mixed	0.71	*	***0.62***	*	1.01		1.34	*	0.95	*	0.86	*	1.09	*	0.67	*	0.70	*	0.86		0.72	*	0.87	*	0.82	*
Indian	***0.33***	*	***0.52***	*	0.89	*	1.02		***0.30***	*	0.71	*	***0.64***	*	***0.57***	*	0.75	*	***0.48***	*	0.68	*	***0.60***	*	***0.51***	*
Pakistani	***0.42***	*	0.96	*	1.33	*	***2.24***	*	***0.48***	*	0.93	*	0.68	*	1.32	*	***1.93***	*	0.81		1.22	*	0.82	*	1.04	
Bangladeshi	***0.39***	*	***0.50***	*	0.94		***1.56***	*	***0.37***	*	0.94	*	0.98		0.82	*	0.80	*	***0.58***	*	***0.64***	*	***0.60***	*	0.77	*
Any other Asian	***0.37***	*	***0.47***	*	1.09		1.36	*	***0.35***	*	0.85	*	0.84	*	***0.62***	*	***0.61***	*	***0.39***	*	0.70	*	***0.62***	*	***0.62***	*
Black African	***0.46***	*	***0.51***	*	1.11	*	1.25	*	***0.65***	*	0.99		1.27	*	***0.41***	*	***0.62***	*	***0.59***	*	***0.61***	*	0.67	*	0.76	*
Black Caribbean	0.90	*	0.89	*	0.92		1.21		1.44	*	1.34	*	1.12	*	0.72	*	0.70	*	***0.65***		0.73	*	1.01		1.19	*
Black other groups	***0.60***	*	***0.65***	*	1.28	*	1.35	*	0.94	*	1.09	*	1.30	*	***0.49***	*	***0.55***	*	0.81		0.68	*	0.82	*	0.92	
Chinese	***0.30***	*	***0.29***	*	0.82		0.80		***0.28***	*	0.95		1.39	*	***0.59***	*	***0.49***	*	0.73		***0.33***	*	***0.46***	*	***0.51***	*
Any other ethnic group	***0.53***	*	***0.58***	*	0.84	*	1.29	*	***0.54***	*	0.92	*	0.89	*	***0.65***	*	***0.64***	*	0.69		***0.66***	*	0.77	*	0.96	
Unknown	0.89	*	0.79	*	1.20	*	1.17		1.01		0.98		1.21	*	0.76	*	0.98		1.20		0.89		1.02		1.26	*

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.096. *Adjusting for: Normalised IDACI, DSM eligibility, gender, birth season, Year Group, EAL.

Table B.3: ORs for additional pupil controls, 2016, including EAL

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
Pupil variables	*SpLD*		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*		*ASD*		*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
FSM																										
Eligible	***1.68***	*	***2.42***	*	***3.51***	*	***2.75***	*	***3.07***	*	***2.10***	*	***2.30***	*	***1.81***	*	***1.99***	*	***1.74***	*	***2.28***	*	***2.04***	*	***1.97***	*
Gender																										
Boy	***1.78***	*	***1.71***	*	***2.06***	*	1.46	*	***3.17***	*	***2.53***	*	***5.38***	*	1.16	*	1.38	*	***1.98***	*	***1.56***	*	***1.73***	*	***1.70***	*
Birth Season																										
Summer	***1.52***	*	***1.83***	*	1.26	*	1.10	*	1.16	*	***1.64***	*	1.09	*	1.12	*	1.07	*	1.27	*	1.18	*	1.43	*	***1.71***	*
Spring	1.24	*	1.35	*	1.09	*	1.05		1.07	*	1.29	*	1.04	*	1.07	*	1.00		1.06		1.09	*	1.19	*	1.27	*
Year Group																										
Y2	***1.97***	*	***1.65***	*	1.14	*	0.95		1.32	*	0.89	*	1.09	*	1.18	*	1.17	*	1.09		1.12	*	1.33	*	1.36	*
Y3	***2.80***	*	***1.99***	*	1.16	*	0.93		1.49	*	0.72	*	1.15	*	1.23	*	1.19	*	1.07		1.14	*	1.40	*	1.31	*
Y4	***3.74***	*	***2.29***	*	1.20	*	0.98		***1.70***	*	***0.62***	*	1.23	*	1.28	*	1.29	*	0.95		1.12	*	***1.55***	*	1.27	*
Y5	***4.53***	*	***2.49***	*	1.28	*	0.90	*	***1.86***	*	***0.53***	*	1.33	*	1.26	*	1.34	*	0.97		1.11	*	***1.65***	*	1.17	*
Y6	***5.25***	*	***2.71***	*	1.44	*	0.88	*	***1.98***	*	***0.49***	*	1.42	*	1.37	*	1.37	*	***0.63***	*	1.13	*	***1.72***	*	1.20	*
Y7	***5.39***	*	***2.15***	*	1.30	*	0.86	*	***1.77***	*	***0.36***	*	***1.54***	*	***1.53***	*	1.42	*	0.72	*	1.12	*	***2.15***	*	1.02	
Y8	***5.40***	*	***2.05***	*	1.29	*	0.76	*	***1.79***	*	***0.33***	*	***1.53***	*	***1.57***	*	***1.59***	*	***0.56***	*	1.05		***1.99***	*	0.84	*
Y9	***5.29***	*	***1.87***	*	1.34	*	0.75	*	***1.91***	*	***0.30***	*	***1.53***	*	***1.56***	*	***1.57***	*	***0.54***	*	1.04		***1.82***	*	***0.60***	*
Y10	***5.39***	*	***1.75***	*	1.35	*	0.74	*	***2.06***	*	***0.27***	*	1.46	*	***1.62***	*	***1.61***	*	***0.59***	*	1.02		***1.80***	*	***0.53***	*
Y11	***5.41***	*	***1.72***	*	1.39	*	0.70	*	***2.34***	*	***0.24***	*	1.48	*	***1.55***	*	***1.58***	*	***0.53***	*	1.07	*	***1.90***	*	***0.62***	*
Deprivation																										
Normalised IDACI (2SD)	1.13	*	***1.90***	*	1.18	*	1.00		***1.94***	*	***1.71***	*	1.19	*	1.16	*	1.25	*	1.07		1.15	*	1.42	*	***1.64***	*
Combined deprivation (IDACI+FSM)	***1.78***	*	***3.33***		***3.82***		***2.75***		***4.28***		***2.74***		***2.51***		***1.95***		***2.22***		***1.80***		***2.44***		***2.43***		***2.53***	
Language																										
EAL	0.78	*	1.06	*	0.97		1.12	*	0.68	*	1.19	*	***0.61***	*	***1.58***	*	0.96		0.83		0.79	*	0.98		1.18	*

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.096

*Adjusting for: Normalised IDACI, DSM eligibility, gender, birth season, Year Group, EAL.

Appendix C: Unadjusted and adjusted ORs for ethnic groups, separated by phase of schooling (primary and secondary)

Table C.4: Unadjusted ORs by ethnic group for types of primary SEN 2016 - Primary (Y1-6) only

Unadjusted Ratios	Cognition & Learning				Social, Emotional & Mental Health	Communication & Interaction		Sensory & Physical				Unspecified/Other	
Ethnic Group	SpLD	MLD	SLD	PMLD	SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA
White Irish	1.11	0.89	1.08	***1.61***	0.89	1.00	1.04	1.21	0.96	1.20	1.01	0.98	1.31
Traveller Irish	***3.04***	***4.30***	***1.51***	1.21	***2.50***	***2.32***	***0.47***	1.44	***1.96***	***1.71***	0.95	***2.87***	***4.84***
Traveller Gypsy/Roma	***2.09***	***3.67***	***1.87***	1.11	1.34	***2.32***	***0.42***	***3.66***	***1.56***	***0.42***	1.11	***1.59***	***3.48***
Other White	***0.56***	0.74	0.78	1.04	***0.57***	1.20	0.72	0.83	***0.61***	***0.62***	***0.54***	0.73	1.06
Mixed White & African	0.77	0.83	1.17	***1.53***	1.20	1.12	1.06	***0.61***	***0.50***	***0.66***	0.71	0.95	0.96
Mixed White & Caribbean	0.99	1.24	0.96	1.18	***1.86***	1.12	1.18	0.78	0.97	***0.67***	0.89	1.08	1.36
Mixed White & Asian	***0.53***	***0.67***	0.90	1.00	0.69	0.90	0.94	0.88	***0.66***	***0.38***	0.70	***0.65***	0.80
Other mixed	0.70	0.79	1.16	***1.62***	1.06	1.06	1.15	0.83	0.70	0.81	0.73	0.90	0.96
Indian	***0.31***	***0.53***	0.88	1.09	***0.24***	0.84	***0.62***	0.80	***0.65***	***0.43***	***0.54***	***0.50***	***0.61***
Pakistani	***0.45***	1.33	1.48	***2.49***	***0.47***	1.37	0.69	***2.11***	***2.15***	***0.64***	1.19	0.95	***1.63***
Bangladeshi	***0.39***	0.81	1.27	***1.84***	***0.44***	***1.56***	1.14	***1.60***	0.93	***0.58***	***0.61***	0.83	1.27
Other Asian	***0.35***	***0.58***	1.13	***1.51***	***0.34***	1.11	0.83	0.98	***0.64***	***0.28***	***0.65***	***0.59***	0.77
Black African	***0.45***	0.81	1.49	***1.56***	0.91	***1.55***	***1.52***	***0.66***	0.74	***0.56***	0.68	0.81	1.11
Black Caribbean	1.05	1.36	1.16	1.44	***2.26***	***1.67***	1.46	0.96	0.81	0.81	0.86	1.33	***1.64***
Other Black	***0.65***	1.05	***1.61***	***1.70***	1.29	***1.56***	***1.56***	0.69	***0.66***	0.80	0.72	1.04	1.36
Chinese	***0.28***	***0.31***	0.78	0.70	***0.22***	1.16	1.17	0.85	***0.40***	0.84	***0.24***	***0.38***	***0.59***
Any other	***0.47***	0.86	1.00	1.34	***0.61***	1.40	0.93	1.06	0.78	***0.53***	***0.65***	0.83	1.50
Unknown	0.81	0.89	***1.55***	1.32	1.01	1.10	1.41	0.86	0.78	1.00	0.87	0.99	1.38

Table C.5: Adjusted ORs by ethnic group for types of primary SEN 2016 - Primary (Y1-6) only

Adjusted Ratios	Cognition & Learning				Social, Emotional & Mental Health	Communication & Interaction		Sensory & Physical				Unspecified/Other	
Ethnic Group	SpLD	MLD	SLD	PMLD	SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA
White Irish	1.09	0.87	1.03	***1.60***	0.87	1.00	1.04	1.20	0.96	1.19	1.00	0.98	1.30
Traveller Irish	***2.43***	***2.57***	0.85	0.77	1.37	1.45	***0.32***	1.05	1.36	1.35	***0.64***	***1.86***	***3.26***
Traveller Gypsy/Roma	***1.95***	***2.81***	1.48	0.94	1.01	***1.85***	***0.36***	***3.21***	1.32	***0.39***	0.94	1.34	***2.84***
Other White	***0.60***	0.73	0.83	1.11	***0.58***	1.12	0.75	0.85	***0.62***	***0.62***	***0.56***	0.73	1.03
Mixed White & African	0.76	0.69	0.98	1.37	0.98	0.92	0.96	***0.57***	***0.45***	***0.62***	***0.64***	0.84	0.82
Mixed White & Caribbean	0.91	0.94	0.75	1.00	1.39	0.88	1.02	0.69	0.83	***0.61***	0.75	0.90	1.11
Mixed White & Asian	***0.53***	***0.65***	0.87	0.98	***0.67***	0.86	0.91	0.87	***0.65***	***0.37***	0.68	***0.64***	0.78
Other mixed	0.69	0.69	1.06	***1.52***	0.93	0.93	1.08	0.79	***0.65***	0.78	0.68	0.83	0.87
Indian	***0.33***	***0.57***	0.98	1.21	***0.26***	0.88	***0.65***	0.85	0.69	***0.44***	***0.57***	***0.53***	***0.64***
Pakistani	***0.44***	1.13	1.34	***2.39***	***0.40***	1.19	***0.63***	***1.98***	***1.97***	***0.62***	1.11	0.86	1.45
Bangladeshi	***0.36***	***0.60***	1.06	***1.69***	***0.32***	1.24	0.99	1.43	0.80	***0.56***	***0.54***	0.69	1.04
Other Asian	***0.35***	***0.54***	1.11	***1.53***	***0.32***	1.04	0.81	0.96	***0.62***	***0.28***	***0.64***	***0.56***	0.73
Black African	***0.42***	***0.58***	1.20	1.37	***0.65***	1.17	1.31	***0.57***	***0.62***	***0.53***	***0.57***	***0.66***	0.88
Black Caribbean	0.94	0.97	0.91	1.24	***1.61***	1.29	1.25	0.82	0.67	0.77	0.72	1.07	1.30
Other Black	***0.60***	0.75	1.29	1.46	0.93	1.18	1.34	***0.60***	***0.56***	0.75	***0.61***	0.84	1.07
Chinese	***0.30***	***0.32***	0.87	0.76	***0.23***	1.13	1.24	0.89	***0.42***	0.86	***0.26***	***0.39***	***0.60***
Any other	***0.44***	***0.67***	0.83	1.21	***0.47***	1.11	0.81	0.95	0.69	***0.50***	***0.57***	0.70	1.25
Unknown	0.84	0.86	1.48	1.28	0.97	1.01	1.37	0.85	0.77	0.98	0.84	0.97	1.32

Table C.6: Unadjusted ORs by ethnic group for types of primary SEN 2016 - Secondary (Y7-11) only

Undjusted Ratios	Cognition & Learning				Social, Emotional & Mental Health	Communication & Interaction		Sensory & Physical				Unspecified/Other	
Ethnic Group	SpLD	MLD	SLD	PMLD	SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA
White Irish	1.16	0.73	0.73	0.85	0.94	0.93	1.01	1.25	***0.57***	***1.88***	0.93	1.06	1.07
Traveller Irish	***1.81***	***4.33***	***3.01***	***5.05***	***4.32***	***3.01***	***0.48***	***2.59***	***2.64***	***0.00***	***3.05***	***3.71***	***5.05***
Traveller Gypsy/Roma	1.20	***3.93***	***2.53***	1.03	***2.17***	***2.24***	***0.37***	***2.56***	***1.70***	***0.00***	0.83	***1.96***	***2.56***
Other White	***0.56***	0.83	0.75	1.03	***0.59***	1.30	***0.47***	***0.66***	***0.65***	0.79	***0.52***	1.06	1.28
Mixed White & African	0.78	0.78	0.77	0.82	1.23	1.24	0.84	0.72	***0.40***	0.79	0.68	0.83	1.12
Mixed White & Caribbean	1.10	1.22	0.87	1.18	***2.06***	1.38	1.08	0.70	0.99	0.80	0.88	1.26	***1.66***
Mixed White & Asian	***0.56***	0.70	0.83	0.96	0.78	0.85	0.76	***0.58***	0.75	0.96	0.70	0.77	0.69
Other mixed	0.71	0.75	1.08	1.17	1.12	1.17	1.00	0.80	0.82	1.10	0.76	1.07	1.13
Indian	***0.25***	***0.60***	0.71	0.86	***0.24***	0.81	***0.31***	0.76	0.81	***0.37***	***0.59***	0.71	***0.57***
Pakistani	***0.33***	1.41	***1.58***	***2.80***	***0.54***	1.34	***0.40***	***2.10***	***2.13***	0.99	1.05	0.98	1.15
Bangladeshi	***0.34***	0.95	1.10	***2.07***	***0.48***	***1.81***	***0.47***	1.29	1.00	***0.43***	***0.65***	0.77	1.37
Other Asian	***0.29***	***0.60***	1.07	1.43	***0.28***	1.15	***0.43***	0.88	***0.63***	***0.53***	***0.55***	0.78	0.90
Black African	***0.46***	0.89	1.30	1.47	0.76	***1.84***	0.78	***0.66***	0.73	***0.62***	***0.55***	0.95	1.45
Black Caribbean	0.98	1.39	1.23	1.32	***2.29***	***2.48***	1.20	0.78	0.90	***0.48***	0.88	1.33	***1.71***
Other Black	***0.60***	1.00	***1.67***	1.50	1.36	***2.00***	1.15	***0.67***	0.69	0.92	0.79	1.10	1.40
Chinese	***0.21***	***0.30***	***0.65***	1.00	***0.18***	1.13	***0.64***	0.79	***0.52***	***0.00***	***0.30***	***0.56***	***0.64***
Any other	***0.52***	0.97	1.10	***2.08***	***0.63***	***1.72***	***0.53***	1.15	0.67	0.95	***0.63***	1.13	1.43
Unknown	0.91	0.93	1.04	1.14	1.13	1.28	1.06	0.87	1.20	***1.51***	0.93	1.19	***1.57***

Table C.7: Adjusted ORs by ethnic group for types of primary SEN 2016 - Secondary (Y7-11) only

Adjusted Ratios	Cognition & Learning				Social, Emotional & Mental Health	Communication & Interaction		Sensory & Physical				Unspecified/Other	
Ethnic Group	SpLD	MLD	SLD	PMLD	SEMH	SLCN	ASD	HI	VI	MSI	PD	Other	NSA
White Irish	1.15	0.72	0.71	0.84	0.91	0.92	1.00	1.24	***0.57***	***1.89***	0.90	1.05	1.06
Traveller Irish	1.40	***2.28***	***1.53***	***3.04***	***2.10***	***1.78***	***0.31***	***1.94***	***1.85***	***0.00***	***2.02***	***2.42***	***2.98***
Traveller Gypsy/Roma	1.07	***2.65***	***1.82***	0.86	1.45	***1.68***	***0.31***	***2.22***	1.43	***0.00***	0.70	***1.55***	***1.84***
Other White	***0.55***	0.74	0.75	1.09	***0.54***	1.21	***0.47***	***0.65***	***0.64***	0.81	***0.53***	1.00	1.15
Mixed White & African	0.73	***0.60***	***0.63***	0.74	0.95	1.02	0.75	***0.67***	***0.36***	0.75	***0.61***	0.71	0.89
Mixed White & Caribbean	1.02	0.92	0.69	1.03	***1.51***	1.12	0.95	***0.63***	0.87	0.74	0.78	1.06	1.32
Mixed White & Asian	***0.55***	***0.66***	0.78	0.92	0.73	0.80	0.72	***0.56***	0.73	0.94	0.68	0.74	***0.65***
Other mixed	0.67	***0.63***	0.94	1.09	0.94	1.02	0.92	0.75	0.77	0.97	0.71	0.96	0.97
Indian	***0.25***	***0.61***	0.76	0.93	***0.25***	0.81	***0.31***	0.78	0.82	***0.38***	***0.62***	0.72	***0.56***
Pakistani	***0.30***	1.09	1.29	***2.54***	***0.41***	1.10	***0.35***	***1.92***	***1.89***	0.94	0.95	0.84	0.93
Bangladeshi	***0.30***	***0.62***	0.80	***1.76***	***0.30***	1.33	***0.39***	1.12	0.83	***0.39***	***0.55***	***0.60***	0.98
Other Asian	***0.28***	***0.53***	1.02	1.41	***0.25***	1.05	***0.40***	0.85	***0.59***	***0.53***	***0.53***	0.72	0.81
Black African	***0.42***	***0.59***	0.98	1.30	***0.50***	1.38	***0.66***	***0.58***	***0.61***	***0.58***	***0.48***	0.75	1.04
Black Caribbean	0.90	0.97	0.94	1.17	***1.56***	***1.92***	1.05	0.68	0.77	***0.45***	0.77	1.07	1.26
Other Black	***0.54***	0.67	1.27	1.32	0.90	***1.51***	0.97	***0.59***	***0.56***	0.85	0.69	0.87	1.02
Chinese	***0.21***	***0.31***	0.72	1.10	***0.20***	1.16	0.68	0.82	***0.54***	***0.00***	***0.32***	***0.55***	***0.65***
Any other	***0.46***	0.68	0.82	***1.79***	***0.42***	1.31	***0.44***	1.02	***0.57***	0.87	***0.53***	0.91	1.07
Unknown	0.87	0.80	0.94	1.07	0.99	1.12	0.98	0.84	1.12	1.44	0.88	1.08	1.35

Appendix D: Unadjusted and adjusted ORs for ethnic groups and other pupil background characteristics for SEN support only

Table D.8: Unadjusted ORs for ethnic groups - SEN support only (2016)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
	SpLD		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*	*ASD*	*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
White Irish	1.16	*	0.82	*	1.18		***2.42***		0.94		0.98	1.08	1.24		0.68		1.46		1.06		1.04		1.22	*
Traveller Irish	***2.22***	*	***4.33***	*	***4.39***	*	***0.00***		***2.67***	*	***2.88*** *	***0.23*** *	1.38		***2.23***	*	1.03		1.17		***2.98***	*	***5.27***	*
Traveller Gypsy/Roma	***1.58***	*	***4.00***	*	***8.03***	*	***2.53***		***1.65***	*	***2.66*** *	***0.18*** *	***2.87*** *		***1.70*** *		***0.44***		1.14		***1.77***	*	***3.37***	*
White other groups	***0.55***	*	0.83	*	1.12		***1.63***	*	***0.63***	*	1.41 *	***0.49*** *	***0.65*** *		***0.56*** *		***0.59*** *		***0.54*** *		0.88	*	1.18	*
Mixed White & African	0.73	*	0.83	*	1.06		***2.03***		1.20	*	1.25 *	0.68 *	***0.63*** *		***0.42*** *		0.70		0.78 *		0.90		1.04	
Mixed White & Caribbean	1.01		1.26	*	1.31		***2.00***	*	***1.91***	*	1.22 *	1.11 *	0.68 *		0.87		0.82		0.91		1.14	*	1.46	*
Mixed White & Asian	***0.53***	*	0.71	*	0.86		1.50		0.77	*	0.98	0.67 *	0.71 *		***0.60*** *		***0.58***		***0.62*** *		0.69	*	0.81	*
Any other mixed background	0.67	*	0.78	*	1.25		***1.86***	*	1.10	*	1.15 *	0.78 *	0.71 *		***0.66*** *		0.96		0.75 *		0.95		1.04	
Indian	***0.28***	*	***0.58***	*	***0.66***	*	***1.85***	*	***0.27***	*	0.91 *	***0.31*** *	0.68 *		***0.59*** *		***0.37*** *		***0.52*** *		***0.61***	*	***0.61***	*
Pakistani	***0.38***	*	1.42	*	***1.54***	*	***2.06***	*	***0.56***	*	1.48 *	***0.37*** *	***1.65*** *		***1.81*** *		***0.40*** *		0.96		0.96		***1.52***	*
Bangladeshi	***0.36***	*	0.90	*	1.45	*	***2.49***	*	***0.52***	*	***1.64*** *	***0.45*** *	1.16 *		0.84		***0.61*** *		***0.55*** *		0.82	*	1.31	*
Any other Asian	***0.32***	*	***0.60***	*	0.94		***0.66***		***0.34***	*	1.18 *	***0.32*** *	0.78 *		***0.56*** *		***0.17*** *		***0.50*** *		0.68	*	0.83	*
Black African	***0.44***	*	0.87	*	1.36	*	***2.45***	*	0.92	*	***1.72*** *	***0.54*** *	***0.45*** *		***0.66*** *		***0.47*** *		***0.65*** *		0.87	*	1.24	*
Black Caribbean	0.99		1.39	*	***1.69***	*	1.03		***2.32***	*	***1.79*** *	0.88 *	***0.65*** *		0.81		0.74		0.83 *		1.29	*	***1.60***	*
Black other groups	***0.59***	*	1.06	*	***2.01***	*	***3.32***	*	1.36	*	***1.71*** *	0.73 *	***0.58*** *		***0.63*** *		0.69		0.74 *		1.04		1.38	*
Chinese	***0.24***	*	***0.31***	*	***0.60***		***0.55***		***0.23***	*	1.22 *	***0.55*** *	0.90		***0.43*** *		0.71		***0.25*** *		***0.44***	*	***0.63***	*
Any other ethnic group	***0.48***	*	0.95	*	1.45	*	***1.91***	*	0.68	*	***1.60*** *	***0.44*** *	0.87 *		***0.65*** *		0.67		***0.60*** *		0.98		***1.53***	*
Unknown	0.90	*	0.85	*	1.02		0.76		1.06	*	0.98	0.88 *	0.76 *		1.03		1.24		0.80 *		1.12	*	1.35	*

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.011

*SEN type for SEN support only for outcome variable; Ethnic group, no additional pupil predictors/controls.

Table D.9: Adjusted ORs for ethnic groups - SEN support only (2016)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
	SpLD		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*	*ASD*	*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
White Irish	1.13	*	0.78	*	1.14		***2.34***		0.88	*	0.96	1.05	1.21		***0.65***		1.47		1.03		0.99		1.18	
Traveller Irish	***2.05***	*	***2.46***	*	***2.32***	*	***0.00***	*	1.46	*	***1.51***	* ***0.18***	* 1.14		***1.64***		0.75		0.78		***2.16***	*	***3.14***	*
Traveller Gypsy/Roma	1.50	*	***2.85***	*	***5.77***	*	***1.75***		1.18	*	***1.88***	* ***0.16***	* ***2.53***	*	1.42	*	***0.38***		0.93		1.46	*	***2.51***	*
White other groups	***0.58***	*	0.73	*	1.03		1.36		***0.58***	*	1.14	* ***0.48***	* ***0.63***	*	***0.54***	*	***0.54***	*	***0.51***	*	0.84	*	1.03	
Mixed White & African	0.73	*	***0.63***	*	0.81		1.45		0.92	*	0.89	* ***0.62***	* ***0.58***	*	***0.37***	*	***0.61***		0.67	*	0.79	*	0.81	*
Mixed White & Caribbean	0.95	*	0.92	*	0.97		1.42		1.37	*	0.90	* 0.98	***0.60***	*	0.73	*	0.73		0.76	*	0.94		1.13	*
Mixed White & Asian	***0.55***	*	***0.65***	*	0.78		1.31		0.72	*	0.84	* ***0.66***	* 0.70	*	***0.58***	*	***0.54***	*	***0.59***	*	0.67	*	0.73	*
Any other mixed	0.67	*	***0.63***	*	1.02		1.43		0.90	*	0.89	* 0.73	* 0.67	*	***0.61***	*	0.86		0.67	*	0.86	*	0.85	*
Indian	***0.28***	*	***0.55***	*	***0.66***	*	***1.74***	*	***0.26***	*	0.84	* ***0.31***	* 0.68	*	***0.59***	*	***0.36***	*	***0.52***	*	***0.60***	*	***0.58***	*
Pakistani	***0.35***	*	1.05	*	1.21	*	***1.53***	*	***0.41***	*	1.12	* ***0.33***	* 1.46	*	***1.55***	*	***0.37***	*	0.83	*	0.80	*	1.19	*
Bangladeshi	***0.32***	*	***0.55***	*	0.97		***1.56***		***0.31***	*	1.08	* ***0.37***	* 0.95		***0.65***	*	***0.54***	*	***0.44***	*	***0.61***	*	0.91	*
Any other Asian	***0.31***	*	***0.50***	*	0.83		***0.55***		***0.29***	*	0.98	***0.30***	* 0.73	*	***0.52***	*	***0.16***	*	***0.47***	*	***0.62***	*	0.71	*
Black African	***0.41***	*	***0.54***	*	0.92		***1.54***	*	***0.57***	*	1.12	* ***0.46***	* ***0.38***	*	***0.53***	*	***0.40***	*	***0.51***	*	***0.66***	*	0.85	*
Black Caribbean	0.87	*	0.91	*	1.19		0.70		1.46	*	1.31	* 0.73	* ***0.54***	*	***0.64***	*	0.69		0.68	*	0.97		1.18	*
Black other groups	***0.54***	*	0.68	*	1.38		***2.13***	*	0.87	*	1.14	* ***0.62***	* ***0.49***	*	***0.51***	*	***0.60***		***0.59***	*	0.80	*	0.98	
Chinese	***0.25***	*	***0.30***	*	***0.60***		***0.50***		***0.23***	*	1.07	***0.57***	* 0.91		***0.43***	*	0.67		***0.25***	*	***0.45***	*	***0.59***	*
Any other ethnic group	***0.45***	*	***0.64***	*	1.03		1.27		***0.45***	*	1.10	* ***0.37***	* 0.75	*	***0.54***	*	***0.58***	*	***0.49***	*	0.78	*	1.12	*
Unknown	0.83	*	0.76	*	0.96		0.72		0.92	*	0.96	0.80	* 0.70	*	0.95		1.26		0.76	*	0.99		1.29	*

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared = 0.083; SEN type for SEN support only, adjusting for additional pupil variables.

Table D.10: ORs for additional pupil control variables, SEN support only (2016)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
Pupil variables	*SpLD*		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*		*ASD*		*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
FSM																										
Eligible	***1.53***	*	***2.14***	*	***2.51***	*	***2.15***	*	***2.66***	*	***1.88***	*	***1.79***	*	1.44	*	***1.71***	*	1.38	*	***1.86***	*	***1.85***	*	***1.84***	*
Gender																										
Boy	***1.66***	*	***1.60***	*	***1.80***	*	***1.85***	*	***2.69***	*	***2.37***	*	***4.44***	*	1.05	*	1.32	*	***2.22***	*	***1.54***	*	***1.63***	*	***1.62***	*
Birth Season																										
Summer	1.50	*	***1.87***	*	***1.82***	*	1.35	*	1.14	*	***1.67***	*	1.05	*	1.12	*	1.05		1.32	*	1.22	*	1.45	*	***1.70***	*
Spring	1.23	*	1.37	*	1.28	*	1.03		1.07	*	1.32	*	1.03	*	1.08	*	0.97		1.08		1.11	*	1.20	*	1.27	*
Year Group																										
Y2	***2.01***	*	***1.65***	*	1.24	*	0.81		1.28	*	0.87	*	1.07	*	1.12	*	1.16	*	1.00		1.09	*	1.37	*	1.36	*
Y3	***2.86***	*	***1.97***	*	1.20	*	***0.66***	*	1.40	*	0.68	*	1.14	*	1.18	*	1.16	*	0.99		1.10	*	1.44	*	1.30	*
Y4	***3.83***	*	***2.24***	*	1.38	*	0.69	*	***1.57***	*	***0.56***	*	1.27	*	1.23	*	1.26	*	0.82		1.05		***1.60***	*	1.26	*
Y5	***4.61***	*	***2.40***	*	1.44	*	***0.46***	*	***1.67***	*	***0.47***	*	1.42	*	1.15	*	1.18	*	0.83		0.98		***1.70***	*	1.16	*
Y6	***5.27***	*	***2.56***	*	***1.70***	*	***0.40***	*	***1.72***	*	***0.41***	*	1.46	*	1.24	*	1.20	*	***0.53***	*	0.99		***1.76***	*	1.18	*
Y7	***5.36***	*	***1.92***	*	1.06		***0.29***	*	1.48	*	***0.28***	*	***1.60***	*	1.40	*	1.28	*	***0.57***	*	0.94		***2.20***	*	0.99	
Y8	***5.29***	*	***1.78***	*	0.84		***0.35***	*	1.47	*	***0.24***	*	***1.52***	*	1.45	*	1.37	*	***0.44***	*	0.85	*	***2.05***	*	0.81	*
Y9	***5.11***	*	***1.58***	*	0.75	*	***0.36***	*	***1.54***	*	***0.21***	*	1.49	*	1.45	*	1.34	*	***0.37***	*	0.84	*	***1.87***	*	***0.57***	*
Y10	***5.15***	*	1.43	*	0.73	*	***0.28***	*	***1.65***	*	***0.18***	*	1.41	*	1.43	*	1.34	*	***0.43***	*	0.77	*	***1.83***	*	***0.51***	*
Y11	***5.15***	*	1.38	*	***0.66***	*	***0.29***	*	***1.92***	*	***0.16***	*	1.37	*	1.39	*	1.32	*	***0.34***	*	0.83	*	***1.93***	*	***0.59***	*
Deprivation																										
Normalised IDACI (2SD)	1.11	*	***1.95***	*	***1.61***	*	***1.94***	*	***1.94***	*	***1.83***	*	1.24	*	1.28	*	1.36	*	1.17	*	1.31	*	1.46	*	***1.65***	*
Combined deprivation (IDACI+FSM)	***1.61***		***2.99***		***3.18***		***3.00***		***3.71***		***2.54***		***1.99***		***1.63***		***2.00***		1.50		***2.13***		***2.24***		***2.37***	

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.083

*SEN type for SEN support only, after adjusting for: Normalised IDACI, DSM eligibility, gender, birth season, Year Group.

Appendix E: Unadjusted and adjusted ORs for ethnic groups and other pupil background characteristics for Statement/EHC only

Table E.11: Unadjusted ORs for ethnic groups - Statement/EHC only (2016)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
	SpLD		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*		*ASD*		*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
White Irish	1.15		0.79	*	0.87		1.23		0.85		0.86		1.01		1.26		1.05		1.30		0.82		1.06		1.36	
Traveller Irish	1.43		***2.57***	*	1.22		***1.70***		***2.64***	*	1.44	*	***0.48***	*	***1.89***		1.04		***2.58***		1.41		***0.56***		***0.00***	*
Traveller Gypsy/Roma	1.26		1.38	*	1.20		0.85		1.24	*	1.24	*	***0.47***	*	***3.26***	*	1.04		***0.00***		0.67	*	0.80		***2.39***	
White other groups	***0.40***	*	***0.42***	*	0.74	*	1.05		***0.30***	*	0.83	*	0.72	*	1.10		0.82	*	1.03		***0.56***	*	0.73	*	***0.63***	
Mixed White & African	0.81		0.72	*	1.00		1.31	*	1.15	*	1.12		1.18	*	0.70		***0.55***		0.82		***0.59***	*	***0.66***	*	***0.57***	
Mixed White & Caribbean	1.10		0.91		0.84	*	1.10		***1.91***	*	1.06		1.07		0.88		1.23		***0.39***		0.82	*	1.12		***2.17***	*
Mixed White & Asian	***0.50***	*	***0.58***	*	0.90		1.02		***0.51***	*	0.83	*	1.03		0.87		1.01		***0.47***		0.87		0.86		***0.33***	
Any other mixed background	0.72	*	0.72	*	1.12	*	1.50	*	0.94		1.09	*	1.32	*	1.07		0.98		0.71		0.73	*	1.00		1.27	
Indian	***0.35***	*	***0.53***	*	0.87	*	1.04		***0.11***	*	0.71	*	***0.64***	*	1.20	*	1.20		***0.61***		0.68	*	***0.64***	*	***0.43***	
Pakistani	***0.44***	*	0.92	*	***1.52***	*	***2.66***	*	***0.19***	*	0.96		0.69	*	***3.47***	*	***3.05***	*	***1.84***	*	1.42	*	1.00		***0.56***	
Bangladeshi	***0.43***	*	***0.62***	*	1.20	*	***1.95***	*	***0.13***	*	***1.60***	*	1.13	*	***2.40***	*	1.36	*	***0.34***		0.77	*	***0.59***	*	***0.24***	
Any other Asian	***0.36***	*	***0.58***	*	1.19	*	***1.64***	*	***0.17***	*	1.14	*	0.95		1.47	*	0.88		1.00		0.84	*	0.68	*	0.70	
Black African	***0.48***	*	***0.62***	*	1.43	*	***1.51***	*	***0.38***	*	1.38	*	***1.71***	*	1.26	*	0.93		1.00		***0.60***	*	0.85	*	***0.33***	
Black Caribbean	1.26	*	0.96		1.07		1.32	*	***1.81***	*	***1.69***	*	***1.62***	*	1.41	*	0.92		***0.47***		0.85		1.45	*	***1.97***	
Black other groups	0.77		0.68	*	***1.57***	*	1.49	*	0.98		***1.53***	*	***1.86***	*	0.94		0.70		1.33		0.76	*	1.04		***0.56***	
Chinese	***0.26***	*	***0.32***	*	0.80		0.90		***0.09***	*	1.27	*	1.33	*	***0.63***		***0.57***		***0.47***		***0.32***	*	***0.51***	*	***0.00***	
Any other ethnic group	***0.55***	*	***0.58***	*	1.00		***1.62***	*	***0.29***	*	1.18	*	1.00		***1.81***	*	0.98		0.67		0.72	*	0.79		1.39	
Unknown	1.32	*	1.28	*	1.34	*	1.26	*	1.31	*	1.30	*	***1.56***	*	1.31		1.08		0.83		1.05		1.28		***1.73***	

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.006. *SEN type for Statement/EHC only for outcome variable; Ethnic group, no additional pupil predictors/controls.

Table E.12: Adjusted ORs for ethnic groups - Statement/EHC only (2016)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
	SpLD		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*		*ASD*		*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
White Irish	1.10		0.73	*	0.84		1.23		0.77	*	0.84		0.99		1.24		1.03		1.30		0.82		1.04		1.32	
Traveller Irish	1.31		***1.64***	*	0.70		1.08		1.42	*	0.94		***0.33***	*	1.39		0.82		***1.87***		0.98		***0.38***		***0.00***	*
Traveller Gypsy/Roma	1.22		1.08		0.94		0.72		0.89		1.02		***0.41***	*	***2.94***	*	0.96		***0.00***		***0.60***	*	0.70		***2.31***	
White other groups	***0.45***	*	***0.47***	*	0.81	*	1.13	*	***0.33***	*	0.87	*	0.77	*	1.23	*	0.92		1.18		***0.64***	*	0.80	*	0.75	
Mixed White & African	0.85		***0.66***	*	0.89		1.21		0.99		1.02		1.12	*	0.69		***0.56***		0.82		***0.57***	*	***0.62***	*	***0.61***	
Mixed White & Caribbean	1.03		0.71	*	0.68	*	0.97		1.40	*	0.90	*	0.95		0.80		1.14		***0.37***		0.74	*	1.00		***2.05***	*
Mixed White & Asian	***0.54***	*	***0.60***	*	0.89		0.99		***0.51***	*	0.81	*	1.02		0.88		1.04		***0.47***		0.88		0.86		***0.35***	
Any other mixed	0.76	*	0.68	*	1.04		1.44	*	0.85	*	1.02		1.28	*	1.07		0.99		0.72		0.73	*	0.98		1.35	
Indian	***0.39***	*	***0.59***	*	0.99		1.15		***0.13***	*	0.76	*	0.69	*	1.33	*	1.33	*	0.70		0.77	*	0.71	*	***0.50***	
Pakistani	***0.43***	*	0.77	*	1.38	*	***2.63***	*	***0.15***	*	0.87	*	***0.65***	*	***3.46***	*	***3.05***	*	***1.98***	*	1.45	*	0.98		***0.58***	
Bangladeshi	***0.39***	*	***0.45***	*	1.01		***1.90***	*	***0.09***	*	1.36	*	1.02		***2.33***	*	1.31	*	***0.38***		0.77	*	***0.57***	*	***0.24***	
Any other Asian	***0.38***	*	***0.58***	*	1.21	*	***1.71***	*	***0.16***	*	1.12	*	0.97		***1.55***	*	0.93		1.10		0.90		0.71	*	0.77	
Black African	***0.46***	*	***0.48***	*	1.21	*	1.42	*	***0.27***	*	1.18	*	***1.57***	*	1.23	*	0.91		1.07		***0.60***	*	0.81	*	***0.35***	
Black Caribbean	1.08		0.67	*	0.86	*	1.22	*	1.15	*	1.40	*	1.41	*	1.29	*	0.84		***0.48***		0.80	*	1.31	*	***1.85***	
Black other groups	0.73	*	***0.53***	*	1.31	*	1.39	*	0.69	*	1.30	*	***1.67***	*	0.90		0.68		1.38		0.74	*	0.97		***0.57***	
Chinese	***0.30***	*	***0.38***	*	0.91		0.98		***0.11***	*	1.38	*	1.48	*	0.71		***0.64***		***0.53***		***0.37***	*	***0.57***		***0.00***	
Any other ethnic group	***0.53***	*	***0.46***	*	0.85	*	***1.52***	*	***0.21***	*	1.02		0.91	*	***1.75***	*	0.96		0.69		0.70	*	0.75	*	1.39	
Unknown	1.16		1.06		1.24	*	1.25		1.07		1.21	*	1.45	*	1.25		1.03		0.83		1.02		1.22		***1.58***	

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.061

*SEN type for Statement/EHC only, after adjusting for: Normalised IDACI, DSM eligibility, gender, birth season, Year Group.

Table E.13: ORs for additional pupil control variables - Statement/EHC only (2016)

	Cognition & Learning								Social, Emotional & Mental Health		Communication & Interaction		Sensory & Physical								Unspecified/ Other			
Pupil variables	*SpLD*		*MLD*		*SLD*		*PMLD*		*SEMH*		*SLCN*	*ASD*	*HI*		*VI*		*MSI*		*PD*		*Other*		*NSA*	
FSM																								
Eligible	***2.06***	*	***3.33***	*	***3.14***	*	***2.43***	*	***3.97***	*	***2.27***	*	***2.24***	*	***2.23***	*	***2.14***	*	***2.29***	*	***2.48***	*	***2.33***	*	***2.64***	*
Gender																								
Boy	***2.59***	*	***1.86***	*	***1.87***	*	1.29	*	***6.49***	*	***2.67***	*	***5.47***	*	1.23	*	1.26	*	1.22	*	1.34	*	***1.78***	*	***5.84***	*
Birth Season																								
Summer	1.46	*	1.34	*	1.12	*	1.02		1.06	*	1.32	*	1.04	*	1.03		1.03		1.06		1.03		1.11	*	1.25	
Spring	1.24	*	1.14	*	1.04	*	1.02		1.03	*	1.13	*	1.01		1.03		1.02		0.98		1.02		1.06		1.17	
Year Group																								
Y2	1.25	*	1.39	*	1.09	*	0.95		***1.74***	*	1.11	*	1.07	*	1.30	*	1.16		***1.53***	*	1.12	*	1.00		0.76	
Y3	***1.55***	*	***1.97***	*	1.12	*	0.93		***2.65***	*	1.28	*	1.12	*	1.33	*	1.24	*	1.43		1.16	*	1.01		1.13	
Y4	***1.74***	*	***2.60***	*	1.12	*	0.98		***3.55***	*	1.35	*	1.14	*	1.35	*	1.30	*	***1.55***	*	1.18	*	1.07		***0.59***	
Y5	***2.20***	*	***3.34***	*	1.20	*	0.92		***4.44***	*	1.32	*	1.19	*	1.49	*	***1.77***	*	***1.64***	*	1.29	*	1.11		0.81	
Y6	***3.18***	*	***4.32***	*	1.32	*	0.90	*	***5.52***	*	1.47	*	1.29	*	***1.65***	*	***1.79***	*	1.11		1.33	*	1.13		1.19	
Y7	***4.23***	*	***5.76***	*	1.30	*	0.90	*	***6.35***	*	1.44	*	1.46	*	***1.75***	*	***1.77***	*	1.46		1.44	*	1.39	*	***1.64***	
Y8	***5.01***	*	***6.44***	*	1.32	*	0.79	*	***6.97***	*	***1.52***	*	1.49	*	***1.75***	*	***2.22***	*	1.22		1.40	*	1.20	*	***1.77***	
Y9	***5.92***	*	***6.93***	*	1.40	*	0.79	*	***7.77***	*	***1.52***	*	***1.53***	*	***1.74***	*	***2.22***	*	1.46		1.41	*	1.11		***2.47***	*
Y10	***6.62***	*	***7.54***	*	1.42	*	0.79	*	***8.53***	*	1.44	*	1.48	*	***2.04***	*	***2.42***	*	1.45		***1.52***	*	1.22	*	***2.31***	*
Y11	***6.91***	*	***7.89***	*	1.47	*	0.75	*	***8.90***	*	1.37	*	***1.52***	*	***1.90***	*	***2.36***	*	***1.55***	*	1.50	*	1.28	*	***3.22***	*
Deprivation																								
Normalised IDACI (2SD)	1.02		1.29	*	1.04	*	0.88	*	1.50	*	1.12	*	1.02		0.90	*	0.91	*	0.74	*	0.83	*	0.92	*	0.80	
Combined deprivation (IDACI+FSM)	***2.08***		***3.79***		***3.21***		***2.28***		***4.85***		***2.41***		***2.26***		***2.12***		***2.04***		***1.97***		***2.26***		***2.23***		***2.37***	

* indicates significance at the p<0.05 level; Nagelkerke Pseudo R-squared =0.061

*SEN type for Statement/EHC only, after adjusting for: Normalised IDACI, DSM eligibility, gender, birth season, Year Group,

Appendix F: Interpretation of reported statistics

Single-level models (Part 1): Odds Ratios (OR)

An explanation of effect size measures

We are interested in this report in establishing the size of the differences in outcomes between groups of students. Where the outcome of interest is in a readily interpretable or meaningful scale this can be relatively straightforward. For example, if the outcome were average income we might feel this metric is of itself meaningful. If the average weekly earnings of one group of students was £160 and the average weekly earnings of another group was £200 then the second group on average earn £40 per week more than EVER6 students. However, metrics in educational research are often not inherently meaningful in this way.

Cohen's D

Suppose rather than £, Kg or cm we are measuring achievement as indicated by GCSE average points score. It can be difficult to interpret what constitute a large or a small gap in terms of points scores. It is also difficult to compare the size of the gap in GCSE points score at age 16 with the size of the gap measured in National Curriculum levels at age 7, or in KS2 test marks at age 11.

One way to estimate the absolute size of the gap, and to do this in a form that is consistent across many different measures, is to calculate Cohen's D. Cohen's D is an effect size measure for use with continuous variables. It is calculated as:

$$\textit{Cohen's D} = \frac{\text{[Mean of comparator group]} - \text{[Mean of reference group]}}{}$$

There is no restriction on which is the comparator and which the reference group as the absolute value of the difference between the two groups is the same whichever is defined as the reference group, though the sign of the difference (+/-) will change. The important thing is this expresses the difference between the groups in standard deviation (SD) units. The ‘standard deviation’ is a measure of the spread of a set of values and here it refers to the pooled standard deviation of the whole sample. The interpretation is therefore consistent whatever units the outcome is measured in since the Cohen's D gives the gap as the number of SD units, and so is comparable across many different measures.

Cohen's D effect sizes are generally given labels of “small”, “moderate”, or “large”. The most frequent guidelines from Cohen (1988) are 0.2 is small, 0.5 is medium and 0.8 is large. However, these are rough guidelines not cut-off values. What constitutes a small, moderate, or large effect does depend on the area of research and should be interpreted relative to typical results in the particular field of enquiry.

Odds Ratios

Odds Ratios are an effect size measure used where the outcome is dichotomous or binary, for example a student achieves 5 or more GCSE passes at A*-C including English and Mathematics (5EM) or does not achieve this outcome, or a student is identified with a particular type of SEN or is not identified. We can report these percentages achieving the outcome for the two groups and the figures can be directly compared. However, we sometimes want to go further, we might want to:

(i) compare across different measures, for example if 1% of group 1 and 5% of group 2 achieve outcome X, how does this gap compare in size to the gap for outcome Y which is achieved by 10% of group 1 and 15% of group2?

(ii) (ii) compare changes in the percentages achieving a particular measure over time, for example if the proportion of Group 1 students achieving outcome X increases from 10% to 30% and the proportion of Group 2 increases from 25% to 50% has the gap widened, closed or stayed the same?

(iii) (iii) Explore how other variables (like SES) may impact on or change the probabilities of the outcome occurring for the two groups through a technique called logistic regression.

For these reasons the Odds Ratio (OR) is a particularly useful effect size measure. The OR compares the odds of the outcome occurring for the comparison group (say FSM) divided by the odds of the outcome occurring for a reference group (say Non-FSM). The OR can range from 0 to infinity where:

- OR >1 indicates the odds of the outcome occurring are higher for the comparison group relative to the reference group
- OR =1 indicates the odds of the outcome occurring are equal for both groups
- OR <1 indicates the odds of the outcome occurring are lower for the comparison group relative to the reference group

The OR is contingent on which group is defined as the reference group. For example, if the odds of an outcome are twice as high for girls as boys (OR=2.0) this is equivalent to saying the odds of the outcome are half as high for boys as for girls (OR=0.50). The ratios are equivalent, they just vary depending on whether it is the boys or the girls who are defined as the reference group. Any OR can be converted to its complement by dividing the OR into 1 (e.g. 1/2 = 0.50, and 1/0.5= 2.0).

To illustrate the process, consider the odds of achieving Level 2 or above for KS1 reading. The odds for Non FSM pupils achieving this threshold are .931/(1-.931)=13.5. The odds for EVER6 pupils achieving this threshold are .838/(1-.838)=5.2. So the ratio of the two odds (the odds ratio) is 2.6.

We should note that effect sizes do not imply causality, they are just a measure of the association between two variables.

For further references on Effect sizes see Cohen (1988) and Coe (2004).

Multilevel logistic regression models (Part 2): Area Under the receiving operator Characteristic Curve (AUC), Variance Partition Coefficients (VPC), Median Odds Ratio (MOR)

Area Under the receiving operator Characteristic Curve (AUC)

While for linear multilevel models we would traditionally use a model fit statistic to justify a particular structure (e.g. multilevel rather than single-level regression), such fit statistics are generally unreliable for logistic multilevel models. To address this issue and to quantify the classification accuracy of models using only pupil-level versus pupil-, school- and LA-level information, we report a statistic used in epidemiology research literature, the area under the receiving operator characteristic curve (AUC; see Merlo et al., 2016)[26]. The AUC is a measure of discriminatory accuracy that varies from 0.5 to 1; values closer to 1 indicate that a model correctly classifies individuals with or without SEN identification (for a given focal SEN type) based on the predicted probabilities from that model, and values closer to 0.5 are closer to random predictions (i.e. a model's predictions are closer to those provided by flipping a coin). We report the AUCs for models with the same set of pupil predictors but with and without accounting for clustering (i.e. single- versus multilevel models) in order to quantify the extent to which accounting for school and LA clustering improves discriminatory accuracy for our focal primary SEN types.

Variance Partition Coefficient (VPC)

Additionally, the notion of pupil-level variance in the context of a binary outcome variable is not necessarily substantively meaningful, but is implicit to discussions of variance partitioning and variance explained at higher levels in the context of traditional linear multilevel analysis (Austen & Merlo, 2017). We report Variance Partition Coefficients (VPC) calculated using the latent variable approach, which assumes that the binary SEN type identification indicators are dichotomized from a meaningful underlying distribution (Goldstein et al., 2002; Browne et al., 2005) and uses a correction factor ($\frac{\pi^2}{3}$) in place of level-1 (pupil-level) variance to modify the VPC formula used for linear multilevel models.[27]

[26] The receiving operator characteristic curve uses predicted probability values from a model and observed values to plot the 'true positives fraction' (sensitivity) against the 'false positive fraction' (1-specificity) across the full range of dichotomous classification thresholds.

[27] The VPC formulae used here, extended to the three-level case from the latent variable approach given by Goldstein et al. (2002) are: $VPC_{LA} = \frac{\sigma^2_{LA}}{\sigma^2_{LA}+\sigma^2_{School}+\frac{\pi^2}{3}}$ and $VPC_{School} = \frac{\sigma^2_{School}}{\sigma^2_{LA}+\sigma^2_{School}+\frac{\pi^2}{3}}$

Median Odds Ratio (MOR)

Because of the complications described above in assessing cluster-level (i.e. school- and LA-level) heterogeneity in logistic multilevel regression models, we report another statistic commonly used in the field of epidemiology, the *Median Odds Ratio* (MOR) (Austin & Merlo, 2017; Larsen & Merlo, 2005; Larsen et al., 2000)[28], which provides an effect size for higher-level heterogeneity on an odds-ratio scale that is conceptually in keeping with the logic of a logistic regression model. An MOR at the LA level, for example, is the median of odds ratios between any two randomly-selected pupils in different LAs but with otherwise similar predictor values (always comparing the LA with higher odds of identification to the LA with lower odds of identification). Thus, an MOR can take values from 1 to infinity (theoretically); MORs close to 1 suggest little heterogeneity, and the higher the MOR, the greater the heterogeneity at the relevant level. We interpret MORs in relation to the corresponding VPCs as these provide complementary information; the former gives an effect size of the heterogeneity across schools/LAs, and the latter gives a measure of the extent of clustering in schools/LAs (Merlo et al., 2006). Because MORs are on the odds ratio scale, they can be compared with ORs expressing the effects of pupil and higher-level predictors.

Specific effects of pupil/school variables: Interval Odds Ratio (IOR) and Proportion of Opposed Odds Ratio (POOR)

As in the previous section on single-level results, we report ORs as effect sizes for the effects of each predictor. In a multilevel framework, the exponentiated coefficients for any pupil-level covariate or factor are interpreted as within-cluster (i.e. within-school) ORs. So, for example, in a model that has pupils nested within schools, and an exponentiated coefficient of 2 for FSM eligibility, this means that within the average/typical school the odds of identification are twice as high for a pupil eligible for FSM.

Interval Odds Ratio (IOR)

The coefficients for predictors at higher levels are slightly more complex to interpret. To facilitate interpretations specific to the effects of school-level variables, in addition to the ORs representing the average effect (e.g. across all schools, for a school-level variable) we report *Interval Odds Ratios* (IOR) (Larsen et al., 2000)[29]. These provide

---

[28] Formula for calculating the MOR at the LA level: $MOR = exp\left(\sqrt{2\sigma^2_{LA}} \times 0.6745\right)$,

and at the school level: $MOR = exp\left(\sqrt{2(\sigma^2_{LA} + \sigma^2_{School})} \times 0.6745\right)$,

where σ^2_{LA} and σ^2_{School} are the LA- and school-level variances and 0.6745 is an approximate value of the 75th percentile of the cumulative distribution function of the normal distribution with mean=0 and standard deviation=1.

[29] Formula used to calculate the IOR for an LA-level variable:

a way to assess the substantive importance of a cluster-level (here, school-level) variable in relation to the heterogeneity across schools for the effect of a particular predictor. For a school level variable, the IOR for that variable is an interval covering the middle 80% of odds ratios between two randomly-selected pupils in schools with different values for the relevant variable – the width of the interval provides a measure of how much school-level heterogeneity exists in the effect of that contextual (e.g. school-level) variable, and if the interval contains 1 this suggests that the fixed effect of the variable of interest may be trivial compared to the variability of the effect across schools (because for some schools, the association will be in the opposite direction to the overall OR).

Proportion of Opposed Odds Ratios (POOR)

We additionally report the *Proportion of Opposed Odds Ratios (POOR[30])* for each higher-level (school) coefficient; this is expressed as a percentage (which can vary from 0-50%) representing how many schools would have effects (for a given school predictor) of the opposite sign to the overall OR for that predictor. A higher POOR value indicates greater heterogeneity across schools in the effect of a particular school-level predictor.

Precedent for our approach to reporting logistic multilevel regression results, including VPC, MOR, AUC, IOR and POOR, can be found in the epidemiology literature (e.g. Merlo, 2016).

Cox proportional hazards regression (Part 3): Hazard Ratio (HR)

Cox proportional hazards regression results are reported in terms of Hazard Ratios (HR). These have an interpretation somewhat similar to ORs, but with a time aspect. An HR gives the 'risk' of identification *per unit time* (in the context of this research, 1 year) for a given condition (e.g. a particular ethnic group) relative to the reference group (e.g. White British pupils). So, for example, an HR of 2 for a minority ethnic group would indicate that this group has twice the probability of identification in any particular year, relative to the White British majority reference group, while an HR of 0.5 would indicate half the probability of identification compared to the White British majority.

$\left(IOR_{lower}, IOR_{upper}\right) = \exp\left[\beta + \left(\sqrt{2\sigma_{LA}^2} \times (\pm 1.2816)\right)\right]$;

And for a school-level variable: $\left(IOR_{lower}, IOR_{upper}\right) = \exp\left[\beta + \left(\sqrt{2(\sigma_{LA}^2 + \sigma_{School}^2)} \times (\pm 1.2816)\right)\right]$;

where β is the LA- or school-level coefficient, and ± 1.2816 approximates the 10th and 90th percentiles of the normal distribution with mean=0 and variance=1.

[30] Formula used to calculate the POOR (expressed in %): For an LA-level variable: $POOR = \Phi\left(-\left|\frac{\beta}{\sqrt{2\sigma_{LA}^2}}\right|\right) * 100$, and for a school-level variable: $POOR = \Phi\left(-\left|\frac{\beta}{\sqrt{2(\sigma_{LA}^2+\sigma_{School}^2)}}\right|\right) * 100$, where Φ is the cumulative distribution function for the standard normal distribution with mean=0 and standard deviation=1.

Appendix G: Frequency and % (within each ethnic group, of those identified with each focal SEN) enrolled in mainstream/non-mainstream settings 2016

Table G.1: Frequency and percent in mainstream/non-mainstream settings by focal SEN type and ethnic group, 2016 (Y1-11)

	MLD					SEMH					ASD				
Ethnic group	*Mainstream*	*%*	*Special/ PRU/AP*	*%*	*Total*	*Mainstream*	*%*	*Special/ PRU/AP*	*%*	*Total*	*Mainstream*	*%*	*Special/ PRU/AP*	*%*	*Total*
White Irish	603	94.5	35	5.5	638	442	83.7	86	16.3	528	221	78.9	59	21.1	280
Traveller Irish	640	96.7	22	3.3	662	244	76.5	75	23.5	319	16	69.6	7	30.4	23
Traveller Gypsy/Roma	2709	97.1	82	2.9	2791	731	82.6	154	17.4	885	54	55.1	44	44.9	98
White other groups	11772	97.3	325	2.7	12097	5975	92.5	481	7.5	6456	2390	74.6	815	25.4	3205
Mixed White & African	1449	95.3	71	4.7	1520	1443	89.7	165	10.3	1608	406	68.0	191	32.0	597
Mixed White & Caribbean	4400	95.3	218	4.7	4618	4545	85.7	761	14.3	5306	1111	77.0	332	23.0	1443
Mixed White & Asian	2200	95.2	110	4.8	2310	1618	91.4	153	8.6	1771	722	74.4	248	25.6	970
Any other mixed background	3746	95.1	195	4.9	3941	3447	86.8	522	13.2	3969	1320	71.6	524	28.4	1844
Indian	4259	95.5	203	4.5	4462	1323	96.4	49	3.6	1372	851	67.9	402	32.1	1253
Pakistani	14919	96.9	470	3.1	15389	3833	93.2	281	6.8	4114	1306	62.3	791	37.7	2097
Bangladeshi	3827	97.1	115	2.9	3942	1436	94.4	85	5.6	1521	755	60.9	484	39.1	1239
Any other Asian	2639	93.5	182	6.5	2821	1036	94.5	60	5.5	1096	624	60.3	411	39.7	1035
Black African	7885	96.4	294	3.6	8179	5603	94.7	314	5.3	5917	2333	60.7	1508	39.3	3841
Black Caribbean	4008	96.8	131	3.2	4139	4376	86.7	671	13.3	5047	948	68.5	436	31.5	1384
Black other groups	1860	96.4	70	3.6	1930	1595	89.0	198	11.0	1793	497	57.5	368	42.5	865
Chinese	326	92.6	26	7.4	352	164	96.5	6	3.5	170	242	67.0	119	33.0	361
Any other ethnic group	4057	96.7	138	3.3	4195	1907	92.1	164	7.9	2071	738	64.0	416	36.0	1154
Unknown	2001	90.1	221	9.9	2222	1626	81.3	373	18.7	1999	698	66.5	352	33.5	1050
White British	172629	93.8	11458	6.2	184087	117181	87.1	17340	12.9	134521	48454	76.6	14825	23.4	63279
Total	245929	94.5	14366	5.5	260295	158525	87.8	21938	12.2	180463	63686	74.0	22332	26.0	86018

Appendix H: Alternative filtering for multilevel models – School-level descriptive information

Table H.2: Alternative filter for 2016 MLD multilevel models: School descriptive information

Alternative filter for MLD		Primary						Secondary					
		N	*%*	*M*	*SD*	*Min*	*Max*	*N*	*%*	*M*	*SD*	*Min*	*Max*
School type	Foundation	532	4.2	--	--	--	--	252	7.8164	--	--	--	--
	Academy - Converter	1612	12.7	--	--	--	--	1232	38.2	--	--	--	--
	Academy - Sponsored	802	6.3	--	--	--	--	562	17.4	--	--	--	--
	Church	3466	27.2	--	--	--	--	304	9.4	--	--	--	--
	Grammar	--	--	--	--	--	--	163	5.1	--	--	--	--
	Other (Free/CTC/UTC)	103	0.8	--	--	--	--	154	4.8	--	--	--	--
	Community	6222	48.8	--	--	--	--	557	17.3	--	--	--	--
School FSM	Highest	2552	20.0	31.8	7.5	22.9	78.1	643	19.9	30.20824	7.702397	21.7	62.7
	Average-High	2546	20.0	18.2	2.4	14.4	22.8	644	20.0	17.5	2.2	14.2	21.6
	Average	2540	19.9	11.3	1.6	8.8	14.3	650	20.2	11.6	1.4	9.4	14.1
	Low-Average	2590	20.3	6.6	1.2	4.8	8.7	635	19.7	7.4	1.1	5.7	9.3
	Lowest	2509	19.7	2.9	1.3	0.0	4.7	652	20.2	3.7	1.3	0.3	5.6
	OVERALL	12737	100.0	14.2	10.8	0.0	78.1	3224	100.0	14.1	10.0	0.3	62.7
School % Asian (excl. Pakistani; for MLD model)	Highest	2542	20.0	24.1	16.2	11.0	100.0	642	19.9	24.4	17.2	10.4	97.2
	Average-High	2558	20.1	8.0	1.5	5.8	10.9	644	20.0	7.2	1.6	4.9	10.3
	Average	2511	19.7	4.4	0.7	3.4	5.7	641	19.9	3.5	0.7	2.5	4.8
	Low-Average	2583	20.3	2.6	0.4	2.0	3.3	650	20.2	1.8	0.3	1.3	2.4
	Lowest	2543	20.0	1.4	0.4	0.3	1.9	647	20.1	0.8	0.3	0.2	1.2
	OVERALL	12737	100.0	8.1	11.0	0.3	100.0	3224	100.0	0.2	97.2	7.5	11.6
School size (roll)	Smallest	2570	20.2	115.7	33.1	10	160	650	20.2	421.9	150.6	13	615
	Small-Average	2440	19.2	175.0	5.2	165	180	629	19.5	734.8	62.5	620	835
	Average	2562	20.1	224.1	26.0	185	265	648	20.1	942.1	60.9	840	1045
	Average-Large	2713	21.3	322.5	28.2	270	355	653	20.3	1166.3	71.8	1050	1295
	Largest	2452	19.3	477.2	198.8	360	2750	644	20.0	1544.0	217.1	1300	2750
	OVERALL	12737	100.0	262.5	154.6	10	2750	3224	100.0	962.4	402.0	13	2750

*Filtered out schools with <2 pupils in the combined ethnic group of interest (Asian excluding Pakistani); N=3,206,749 pupils.

Table H.3: Alternative filter for 2016 SEMH multilevel models: School descriptive information

Alternative filter for SEMH		Primary						Secondary					
		N	*%*	*M*	*SD*	*Min*	*Max*	*N*	*%*	*M*	*SD*	*Min*	*Max*
School type	Foundation	355	4.1	--	--	--	--	219	7.5	--	--	--	--
	Academy - Converter	1077	12.4	--	--	--	--	1147	39.5	--	--	--	--
	Academy - Sponsored	613	7.1	--	--	--	--	503	17.3	--	--	--	--
	Church	2134	24.6	--	--	--	--	273	9.4	--	--	--	--
	Grammar	--	--	--	--	--	--	151	5.2	--	--	--	--
	Other (Free/CTC/UTC)	77	0.9	--	--	--	--	124	4.3	--	--	--	--
	Community	4425	51.0	--	--	--	--	485	16.7	--	--	--	--
School % FSM	Highest	1737	20.0	32.8	7.4	24.1	78.1	584	20.1	30.0	7.8	21.4	62.7
	Average-High	1739	20.0	19.4	2.4	15.5	24.0	575	19.8	17.3	2.2	14.1	21.3
	Average	1721	19.8	12.5	1.6	9.9	15.4	579	20.0	11.5	1.4	9.3	14.0
	Low-Average	1752	20.2	7.6	1.3	5.5	9.8	588	20.3	7.3	1.1	5.5	9.2
	Lowest	1732	20.0	3.4	1.4	0.0	5.4	576	19.8	3.6	1.3	0.3	5.4
	OVERALL	8681	100.0	15.1	11.0	0.0	78.1	2902	100.0	13.9	10.0	0.3	62.7
School % Black Car. /Mixed Wh. & Car.	Highest	1738	20.0	11.4	5.9	5.8	47.6	579	20.0	10.4	6.1	4.5	48.8
	Average-High	1708	19.7	4.1	0.8	3.0	5.7	603	20.8	2.9	0.7	1.9	4.4
	Average	1716	19.8	2.4	0.3	1.9	2.9	602	20.7	1.3	0.2	1.0	1.8
	Low-Average	1773	20.4	1.5	0.2	1.2	1.8	504	17.4	0.7	0.1	0.6	0.9
	Lowest	1746	20.1	0.8	0.2	0.1	1.1	614	21.2	0.4	0.1	0.1	0.5
	OVERALL	8681	100.0	4.0	4.7	0.1	47.6	2902	100.0	3.2	4.6	0.1	48.8
School size (roll)	Smallest	1701	19.6	132.1	35.4	10	170	578	19.9	455.1	162.2	13	660
	Small-Average	1784	20.6	185.3	11.6	175	215	586	20.2	775.3	61.8	665	875
	Average	1725	19.9	261.8	25.8	220	310	583	20.1	979.1	59.1	880	1080
	Average-Large	1755	20.2	345.2	13.5	315	360	575	19.8	1195.5	68.7	1085	1315
	Largest	1716	19.8	511.3	221.3	365	2750	580	20.0	1564.7	216.2	1320	2750
	OVERALL	8681	100.0	286.8	166.5	10	2750	2902	100.0	993.5	397.1	13	2750

*Filtered out schools with <2 pupils in the combined ethnic group of interest (Black Caribbean and Mixed White & Caribbean); N=2,370,685 pupils.

Table H.4: Alternative filter for 2016 ASD multilevel models: School descriptive information

Alternative filter for ASD		Primary						Secondary					
		N	*%*	*M*	*SD*	*Min*	*Max*	*N*	*%*	*M*	*SD*	*Min*	*Max*
School type	Foundation	436	4.2	--	--	--	--	228	7.5	--	--	--	--
	Academy - Converter	1328	12.9	--	--	--	--	1171	38.7	--	--	--	--
	Academy - Sponsored	685	6.7	--	--	--	--	530	17.5	--	--	--	--
	Church	2571	25.0	--	--	--	--	290	9.6	--	--	--	--
	Grammar	--	--	--	--	--	--	162	5.4	--	--	--	--
	Other (Free/CTC/UTC)	91	0.9	--	--	--	--	147	4.9	--	--	--	--
	Community	5187	50.4	--	--	--	--	500	16.5	--	--	--	--
School FSM	Highest	2067	20.1	32.5	7.3	23.7	78.1	3028	100.0	30.4	7.7	21.9	62.7
	Average-High	2056	20.0	19.1	2.4	15.3	23.6	611	20.2	17.6	2.3	14.2	21.8
	Average	2062	20.0	12.2	1.7	9.5	15.2	607	20.0	11.6	1.4	9.4	14.1
	Low-Average	2032	19.7	7.2	1.2	5.2	9.4	600	19.8	7.3	1.1	5.6	9.3
	Lowest	2081	20.2	3.1	1.3	0.0	5.1	607	20.0	3.6	1.3	0.3	5.5
	OVERALL	10298	100.0	14.8	10.9	0.0	78.1	3028	100.0	14.1	10.1	0.3	62.7
School % Asian (Indian/Pakistani/ Bangladeshi/Other) for ASD model)	Highest	2064	20.0	42.2	23.1	16.5	100.0	606	20.0	39.5	22.9	15.0	98.7
	Average-High	2049	19.9	10.6	2.6	7.0	16.4	600	19.8	9.1	2.7	5.3	14.9
	Average	2060	20.0	4.9	1.0	3.4	6.9	606	20.0	3.4	0.9	2.1	5.2
	Low-Average	1978	19.2	2.5	0.5	1.8	3.3	576	19.0	1.4	0.3	0.9	2.0
	Lowest	2147	20.8	1.2	0.4	0.2	1.7	640	21.1	0.5	0.2	0.1	0.8
	OVERALL	10298	100.0	12.3	18.5	0.2	100.0	3028	100.0	10.8	18.0	0.1	98.7
School size (roll)	Smallest	2100	20.4	133.6	34.1	10	170	604	19.9	443.7	162.1	13	645
	Small-Average	2028	19.7	182.4	8.1	175	205	609	20.1	763.9	60.7	650	865
	Average	2012	19.5	250.7	24.2	210	295	604	19.9	969.1	59.0	870	1070
	Average-Large	2270	22.0	341.0	17.9	300	360	599	19.8	1184.8	69.0	1075	1310
	Largest	1888	18.3	508.0	216.1	365	2750	612	20.2	1556.1	215.8	1315	2750
	OVERALL	10298	100.0	280.4	160.5	10	2750	3028	100.0	984.3	399.0	13	2750

*Filtered out schools with <2 pupils in the combined ethnic group of interest (Indian, Pakistani, Bangladeshi, and Asian Other); N=2,760,426 pupils.

Appendix I: Alternative filtering for multilevel models including school variables (robustness checks)

Table I.5: MLD with additional filtering: OR comparisons

MLD		**Primary**					**Secondary**			
		Main		*>2 filter*			*Main*		*>2 filter*	
		Exp(B)		*Exp(B)*			*Exp(B)*		*Exp(B)*	
Pupil ethnic group	White Irish	1.03		1.03			0.88		0.88	
	Traveller Irish	***2.74***	*	***2.72***	*		***2.32***	*	***2.30***	*
	Traveller Gypsy/Roma	***2.58***	*	***2.48***	*		***2.03***	*	***2.02***	*
	White other groups	0.81	*	0.81	*		0.79	*	0.80	*
	Mixed White & African	0.74	*	0.74	*		0.69	*	0.69	*
	Mixed White & Caribbean	0.92	*	0.92	*		0.90	*	0.90	*
	Mixed White & Asian	0.67	*	0.68	*		0.70	*	0.70	*
	Any other mixed	0.75	*	0.75	*		0.71	*	0.71	*
	Indian	***0.51***	*	***0.50***	*		***0.58***	*	***0.58***	*
	Pakistani	0.90	*	0.89	*		0.88	*	0.88	*
	Bangladeshi	***0.66***	*	***0.66***	*		***0.62***	*	***0.62***	*
	Any other Asian	***0.54***	*	***0.54***	*		***0.57***	*	***0.58***	*
	Black African	***0.62***	*	***0.62***	*		0.67	*	0.68	*
	Black Caribbean	0.96		0.97			0.96		0.96	
	Black other groups	0.80	*	0.80	*		0.73	*	0.73	*
	Chinese	***0.35***	*	***0.34***	*		***0.40***	*	***0.40***	*
	Any other	0.70	*	0.70	*		0.74	*	0.74	*
	Unknown	0.94		0.93			0.87	*	0.88	*
Pupil FSM	Eligible	***2.03***	*	***2.00***	*		***1.92***	*	***1.92***	*
Pupil gender	Boy	***1.61***	*	***1.61***	*		1.46	*	1.46	*
Birth season	Summer	***2.05***	*	***2.06***	*		***1.53***	*	***1.53***	*
	Spring	1.43	*	1.43	*		1.25	*	1.25	*
Pupil year group	Primary: Y6	***2.83***	*	***2.82***	*	Y11	0.75	*	0.75	*
	Y5	***2.62***	*	***2.60***	*	Y10	0.78	*	0.78	*
	Y4	***2.40***	*	***2.40***	*	Y9	0.86	*	0.86	*
	Y3	***2.09***	*	***2.09***	*	Y8	0.95	*	0.95	*
	Y2	***1.67***	*	***1.67***	*					
Pupil IDACI	(Normalised, 2SD)	1.50	*	***1.50***	*		***1.65***	*	***1.65***	*
	Combined deprivation (FSM+IDACI)	***2.48***	*	***2.46***	*		***2.47***	*	***2.47***	*
	AUC									

*=significant at the level p<0.05;highlighting = ***OR<0.67***; OR<0.75; OR>1.33; ***OR>1.50***

Alternative filter version Primary pupil N=3,206,749; Secondary pupil N=2,621,052; excludes schools with <2 Asian (excl. Pakistani) pupils

Table I.6: MLD with additional filtering: Specific school context/composition effects

MLD School contextual effects - additional filter		PRIMARY				SECONDARY			
		Exp(B)		*IOR*	*POOR*	*Exp(B)*		*IOR*	*POOR*
School type	Foundation	1.10		(0.15,7.75)	0.476	1.07		(0.16,7.03)	0.482
	Academy - Converter	0.92	*	(0.13,6.53)	0.479	0.92		(0.14,6.06)	0.478
	Academy - Sponsored	0.84	*	(0.12,5.93)	0.454	0.92		(0.14,6.03)	0.477
	Church	0.90	*	(0.13,6.35)	0.472	0.89		(0.14,5.84)	0.468
	Grammar	--		--	--	***0.06***	*	(0.01,0.36)	0.024
	Other (Free/CTC/UTC)	0.68	*	(0.1,4.79)	0.399	0.79	*	(0.12,5.17)	0.435
School FSM	Highest	***1.65***	*	(0.23,11.68)	0.371	***1.81***	*	(0.27,11.88)	0.344
	Average-High	***1.52***	*	(0.21,10.74)	0.392	***1.50***	*	(0.23,9.88)	0.391
	Average	1.32	*	(0.19,9.32)	0.428	1.37	*	(0.21,9.03)	0.415
	Low-Average	1.22	*	(0.17,8.61)	0.449	1.15	*	(0.18,7.56)	0.462
School ethnic group %	Highest	1.01		(0.14,7.11)	0.498	0.86	*	(0.13,5.65)	0.459
	Average-High	0.97		(0.14,6.88)	0.493	0.83	*	(0.13,5.47)	0.450
	Average	0.92	*	(0.13,6.54)	0.479	0.89		(0.14,5.88)	0.470
	Low-Average	0.97		(0.14,6.86)	0.492	0.97		(0.15,6.35)	0.491
School size (roll)	Smallest	1.33	*	(0.19,9.38)	0.427	1.14	*	(0.17,7.51)	0.464
	Small-Average	1.16	*	(0.16,8.19)	0.462	1.10		(0.17,7.21)	0.475
	Average	1.15	*	(0.16,8.1)	0.465	1.10		(0.17,7.25)	0.473
	Average-Large	1.08	*	(0.15,7.61)	0.481	1.00		(0.15,6.58)	0.500
Variance/heterogeneity	LA Variance	0.216				0.223			
	LA (residual) VPC	0.049				0.051			
	LA MOR	1.56				1.57			
	School Variance	0.948				0.856			
	School (residual) VPC	0.213				0.196			
	School MOR	2.80				2.69			

Notes: prop.=proportion; VPC=Variance Partition Coefficient; MOR=Median Odds Ratio; * indicates significance at p<0.05
Pupil level variables (not including EAL) are also controlled for in these models but coefficients for these are not reported here.

Combined ethnic group of interest for school composition: Asian (excluding Pakistani) groups

Table I.7: SEMH with additional filtering: OR comparisons

SEMH		**Primary**					**Secondary**			
		Main		*>2 filter*			*Main*		*>2 filter*	
		Exp(B)		*Exp(B)*			*Exp(B)*		*Exp(B)*	
Pupil ethnic group	White Irish	0.82	*	0.86			0.83	*	0.83	*
	Traveller Irish	0.92		0.96			1.27		1.22	
	Traveller Gypsy/Roma	0.68	*	***0.64***	*		0.81	*	0.85	
	White other groups	***0.50***	*	***0.50***	*		***0.44***	*	***0.43***	*
	Mixed White & African	0.91	*	0.92	*		0.90	*	0.90	*
	Mixed White & Caribbean	1.32	*	1.31	*		1.28	*	1.28	*
	Mixed White & Asian	***0.66***	*	***0.65***	*		0.72	*	0.71	*
	Any other mixed	0.88	*	0.89	*		0.77	*	0.77	*
	Indian	***0.24***	*	***0.24***	*		***0.22***	*	***0.23***	*
	Pakistani	***0.33***	*	***0.31***	*		***0.32***	*	***0.32***	*
	Bangladeshi	***0.23***	*	***0.23***	*		***0.22***	*	***0.22***	*
	Any other Asian	***0.28***	*	***0.27***	*		***0.22***	*	***0.22***	*
	Black African	***0.58***	*	***0.59***	*		***0.43***	*	***0.43***	*
	Black Caribbean	1.36	*	1.37	*		1.11	*	1.11	*
	Black other groups	0.82	*	0.84	*		0.73	*	0.73	*
	Chinese	***0.23***	*	***0.23***	*		***0.24***	*	***0.24***	*
	Any other	***0.39***	*	***0.38***	*		***0.31***	*	***0.31***	*
	Unknown	0.89	*	0.84	*		0.89	*	0.88	*
Pupil FSM	Eligible	***2.35***	*	***2.31***	*		***2.45***	*	***2.45***	*
Pupil gender	Boy	***3.19***	*	***3.13***	*		***2.28***	*	***2.27***	*
Birth season	Summer	1.10	*	1.09	*		1.07	*	1.07	*
	Spring	1.05	*	1.05	*		1.04	*	1.03	*
Pupil year group	Primary: Y6	***1.81***	*	***1.84***	*	Y11	1.03		1.03	
	Y5	***1.74***	*	***1.77***	*	Y10	1.03		1.03	
	Y4	***1.61***	*	***1.61***	*	Y9	1.01		1.01	
	Y3	1.44	*	1.44	*	Y8	0.97		0.97	
	Y2	1.27	*	1.27	*					
Pupil IDACI	(Normalised, 2SD)	1.37	*	1.35	*		***1.52***	*	***1.51***	*
	Combined deprivation (FSM+IDACI)	***2.75***	*	***2.68***	*		***3.01***	*	***3.02***	*
	AUC									

*=significant at the level p<0.05;highlighting = ***OR<0.67***; OR<0.75; OR>1.33; ***OR>1.50***

Alternative filter version primary pupil N=2,370,685; Secondary pupil N=2,431,489; excludes schools with <2 Black Car./Mixed White & Car. pupils

Table I.8: SEMH with additional filtering: Specific school context/composition effects

SEMH School contextual effects - additional filter		PRIMARY				SECONDARY			
		Exp(B)		*IOR*	*POOR*	*Exp(B)*		*IOR*	*POOR*
School type	Foundation	0.94		(0.28,3.21)	0.475	0.97		(0.27,3.52)	0.488
	Academy - Converter	0.90	*	(0.26,3.06)	0.454	0.90	*	(0.25,3.28)	0.460
	Academy - Sponsored	0.95		(0.28,3.23)	0.478	0.89	*	(0.24,3.21)	0.452
	Church	0.91	*	(0.27,3.1)	0.461	0.84	*	(0.23,3.06)	0.433
	Grammar	--		--	--	***0.42***	*	(0.12,1.52)	0.193
	Other (Free/CTC/UTC)	0.97		(0.28,3.3)	0.486	0.97		(0.27,3.54)	0.490
School FSM	Highest	***1.59***	*	(0.47,5.41)	0.315	***1.62***	*	(0.45,5.9)	0.315
	Average-High	***1.52***	*	(0.44,5.17)	0.332	1.43	*	(0.39,5.18)	0.362
	Average	1.41	*	(0.41,4.81)	0.360	1.36	*	(0.37,4.92)	0.381
	Low-Average	1.18	*	(0.35,4.02)	0.432	1.09		(0.3,3.96)	0.465
School ethnic group %	Highest	0.99		(0.29,3.37)	0.494	1.23	*	(0.34,4.47)	0.418
	Average-High	1.04		(0.3,3.54)	0.485	1.18	*	(0.33,4.3)	0.433
	Average	1.02		(0.3,3.48)	0.492	1.15	*	(0.32,4.18)	0.444
	Low-Average	1.01		(0.3,3.46)	0.494	1.14	*	(0.31,4.14)	0.448
School size (roll)	Smallest	1.17	*	(0.34,3.99)	0.435	1.12	*	(0.31,4.05)	0.457
	Small-Average	0.99		(0.29,3.37)	0.496	1.01		(0.28,3.67)	0.495
	Average	1.07	*	(0.31,3.63)	0.474	0.97		(0.27,3.53)	0.488
	Average-Large	0.98		(0.29,3.35)	0.492	1.04		(0.29,3.76)	0.486
Variance/heterogeneity	School Variance	0.458				0.506			
	School (residual) VPC	0.122				0.133			
	School MOR	1.91				1.97			

Notes: prop.=proportion; VPC=Variance Partition Coefficient; MOR=Median Odds Ratio; * indicates significance at p<0.05

Pupil level variables (not including EAL) are also controlled for in these models but coefficients for these are not reported here.

Combined ethnic group of interest for school composition: Black Caribbean and Mixed White & Caribbean

Table I.9: ASD with additional filtering: OR comparisons

ASD		**Primary**					**Secondary**			
		Main		*>2 filter*			*Main*		*>2 filter*	
		Exp(B)		*Exp(B)*			*Exp(B)*		*Exp(B)*	
Pupil ethnic group	White Irish	1.00		0.94			0.96		0.97	
	Traveller Irish	***0.24***	*	***0.27***	*		***0.21***	*	***0.22***	*
	Traveller Gypsy/Roma	***0.23***	*	***0.20***	*		***0.13***	*	***0.11***	*
	White other groups	***0.66***	*	***0.65***	*		***0.45***	*	***0.44***	*
	Mixed White & African	0.78	*	0.79	*		***0.65***	*	***0.65***	*
	Mixed White & Caribbean	0.89	*	0.89	*		0.85	*	0.85	*
	Mixed White & Asian	0.88	*	0.88	*		0.71	*	0.71	*
	Any other mixed	0.94		0.93			0.82	*	0.82	*
	Indian	***0.62***	*	***0.62***	*		***0.27***	*	***0.27***	*
	Pakistani	***0.54***	*	***0.55***	*		***0.28***	*	***0.28***	*
	Bangladeshi	0.76	*	0.75	*		***0.30***	*	***0.30***	*
	Any other Asian	***0.64***	*	***0.63***	*		***0.27***	*	***0.28***	*
	Black African	0.97		0.97			***0.43***	*	***0.43***	*
	Black Caribbean	0.90	*	0.90	*		0.70	*	0.70	*
	Black other groups	0.91		0.92			***0.53***	*	***0.53***	*
	Chinese	*1.04*		*1.00*			***0.56***	*	***0.57***	*
	Any other	***0.66***	*	***0.64***	*		***0.34***	*	***0.34***	*
	Unknown	1.07		1.04			0.85	*	0.85	*
Pupil FSM	Eligible	***1.64***	*	***1.61***	*		***1.75***	*	***1.75***	*
Pupil gender	Boy	***4.95***	*	***4.93***	*		***4.69***	*	***4.69***	*
Birth season	Summer	0.96	*	0.96	*		1.06	*	1.06	*
	Spring	0.99		0.99			1.01		1.01	
Pupil year group	primary: Y6	1.39	*	1.31	*	Y11	0.95	*	0.95	*
	Y5	1.31	*	1.24	*	Y10	0.94	*	0.94	*
	Y4	1.22	*	1.18	*	Y9	0.98		0.98	
	Y3	1.12	*	1.09	*	Y8	0.98		0.97	
	Y2	1.04		1.02						
Pupil IDACI	(Normalised, 2SD)	0.97		0.98			*1.03*		1.02	
	Combined deprivation (FSM+IDACI)	***1.61***	*	***1.59***	*		***1.77***	*	***1.77***	*
	AUC									

*=significant at the level p<0.05;highlighting = ***OR<0.67***; OR<0.75; OR>1.33; ***OR>1.50***

Alternative filter primary pupil N=2,760,426; Secondary pupil N=2,514,358; excludes schools with <2 Asian (Indian/Pakistani/Bangladeshi/Asian Other) pupils

Table I.10: ASD with additional filtering: Specific school context/composition effects

ASD School contextual effects - additional filter		PRIMARY				SECONDARY			
		Exp(B)		*IOR*	*POOR*	*Exp(B)*		*IOR*	*POOR*
School type	Foundation	0.98		(0.24,3.99)	0.491	1.08		(0.29,4.02)	0.468
	Academy - Converter	0.93	*	(0.23,3.8)	0.473	0.97		(0.26,3.61)	0.490
	Academy - Sponsored	0.96		(0.24,3.93)	0.486	0.93		(0.25,3.44)	0.472
	Church	0.86	*	(0.21,3.52)	0.446	0.98		(0.26,3.62)	0.491
	Grammar	--		--	--	***0.60***	*	(0.16,2.23)	0.310
	Other (Free/CTC/UTC)	0.87		(0.21,3.57)	0.451	1.14		(0.31,4.23)	0.448
School FSM	Highest	1.10	*	(0.27,4.49)	0.466	0.95		(0.26,3.52)	0.481
	Average-High	1.15	*	(0.28,4.71)	0.448	1.02		(0.28,3.79)	0.491
	Average	1.17	*	(0.29,4.77)	0.444	1.10		(0.3,4.08)	0.462
	Low-Average	1.09	*	(0.27,4.43)	0.470	1.09	*	(0.29,4.04)	0.466
School ethnic group %	Highest	0.94		(0.23,3.83)	0.476	1.05		(0.28,3.87)	0.482
	Average-High	1.04		(0.25,4.25)	0.486	1.06		(0.29,3.92)	0.477
	Average	1.06		(0.26,4.32)	0.480	1.04		(0.28,3.83)	0.486
	Low-Average	1.01		(0.25,4.13)	0.496	1.00		(0.27,3.7)	0.5
School size (roll)	Smallest	1.11	*	(0.27,4.53)	0.462	1.15	*	(0.31,4.26)	0.445
	Small-Average	1.10	*	(0.27,4.5)	0.465	1.12	*	(0.3,4.16)	0.454
	Average	1.02		(0.25,4.17)	0.492	1.07		(0.29,3.95)	0.475
	Average-Large	0.98		(0.24,3.98)	0.491	1.03		(0.28,3.79)	0.49
Variance/heterogeneity	LA Variance	0.193				0.157			
	LA (residual) VPC	0.050				0.041			
	LA MOR	1.52				1.46			
	School Variance	0.410				0.365			
	School (residual) VPC	0.105				0.096			
	School MOR	2.10				1.99			

Notes: prop.=proportion; VPC=Variance Partition Coefficient; MOR=Median Odds Ratio; * indicates significance at p<0.05
Pupil level variables (not including EAL) are also controlled for in these models but coefficients for these are not reported here.

Combined ethnic group of interest for school composition: Asian (Indian/Pakistani/Bangladeshi/Asian Other) groups

Appendix J: The Early Years Foundation Stage Profile (EYFSP)

Contents of the EYFSP

The Early Years Foundation Stage Profile (EYFSP) is a statutory assessment completed for all pupils in England in receipt of a government funded early education place at the end of Reception Year (aged approximately 5 years[31]). The EYFSP was first introduced in 2003 and has undergone a number of developments over time, being revised quite substantially in Summer 2013. For our cohort the majority of pupils (99.8%) were assessed in summer 2009, so the description here refers to the form of the EYFSP used between 2003-2012.

In this period the EYFSP measured the achievement of pupils against 13 assessment scales. Each scale consists of multiple statements and is scored by the teacher from 1-9 (see further below). The 13 assessment items are grouped into six areas of learning as shown below.

- Personal, Social and Emotional Development (PSED) - 3 scales
  Personal development (1-9)
  Social development (1-9)
  Emotional development (1-9)

- Communication, Language & Literacy (CLL) - 4 scales
  Language for communication & thinking (1-9)
  Linking sounds and letters (1-9)
  Reading (1-9)
  Writing (1-9)

- Problem Solving, Reasoning & Numeracy[32] (PSRN) - 3 scales
  Numbers as labels for counting (1-9)
  Calculating (1-9)
  Shapes, space and measures (1-9)

- Knowledge & Understanding of the World - 1 scale
- Physical Development - 1 scale
- Creative Development -1 scale

Scoring each individual scale

Each of the 13 assessment scales consists of nine items and is scored from 1-9. The **first three points (1-3**), are hierarchical and describe the achievement of a child

[31] . This is the last term before statutory school age (the term after the child has their fifth birthday), so some pupils may be home educated, but the numbers are very low with the vast majority of children starting Reception Year in schools/early years settings in the September of the school year in which they will turn five. Pupils educated in private schools/settings (around 7% of the population) do not have to submit assessment data, as is true for all national assessment in England.
[32]. This area of learning was known as Mathematical Development prior to September 2009.

who is still progressing towards the Early Learning Goals. Children should achieve all of these three points before they achieve any of points 4-8. A child who does not achieve Scale Point 1 is likely to have significant developmental delay. The **next five points (4-8)** describe the achievement of a child in the context of the Early Learning Goals themselves. These are not hierarchical and can be achieved in any order. A child may attain a later point without having attained some or all of the earlier points. The **final point (9)** in each scale describes a child who has achieved all the points from 1-8 on that scale, has developed further both in breadth and depth, and is working consistently *beyond the level of the Early Learning Goals.* Children who achieve a scale score of **six points or more** are classified as working securely within the Early Learning Goals.

Excluded cases

A total score of **(0)** for the EYFS profile indicates a child with significant and complex special educational needs for whom it has not been possible to record an assessment. Whilst the EYFSP was developed to be inclusive, for a small number of children it may not be appropriate to make an assessment against some of the EYFS profile scales. These children were included in the analysis with a score of 0. A small number of children were recorded as **(N)** where there was insufficient evidence to make an assessment, for example where a child has recently arrived from abroad. These student were excluded from the analysis.

Concurrent and Predictive Validity

Previous research (Snowling et al, 2011) has indicated that the language components of the EYFSP have good concurrent validity against formal tests of language abilities. Snowling et al (2011) report a correlation between the Communication, Language and Literacy (CLL) score and the Language Link test of r=0.63, and conclude that teachers can make valid judgments of children's development in language and literacy and can accurately monitor their pupils' progress in key reading skills. There is also good evidence for predictive validity against subsequent educational achievement at age 7. The Fischer Family Trust (FFT; 2011) report correlations between EYFSP total score across all 13 items and KS1 overall achievement at age 7 of 0.71. Further work indicated that the CLL and PSRN had the strongest relations, but that adding the PSED score did provide some improvements in overall correlations and improvements in consistency in their models. In relation to SEN, we are not exclusively interested in the prediction of academic achievement. Pupils may have high academic achievement and still have SEN (for example SEMH needs). It therefore seems appropriate to use not only the academic scales (CLL and PSRN) but also to included PSED. We use all three scales as separate independent variables in our models so the relative influence of academic (CLL and PSRN) and PSE domains can be evaluated.

Appendix K: Treatment of Key Stage 2 results

Working out various approaches to using measures of KS2 attainment

Around 23% of pupils were missing a KS2 test scores because of the school boycott in 2010. Here we considered three approaches to the issue. For reference shown below are the distributions of KS2 finely graded levels in English and in Mathematics in 2010.

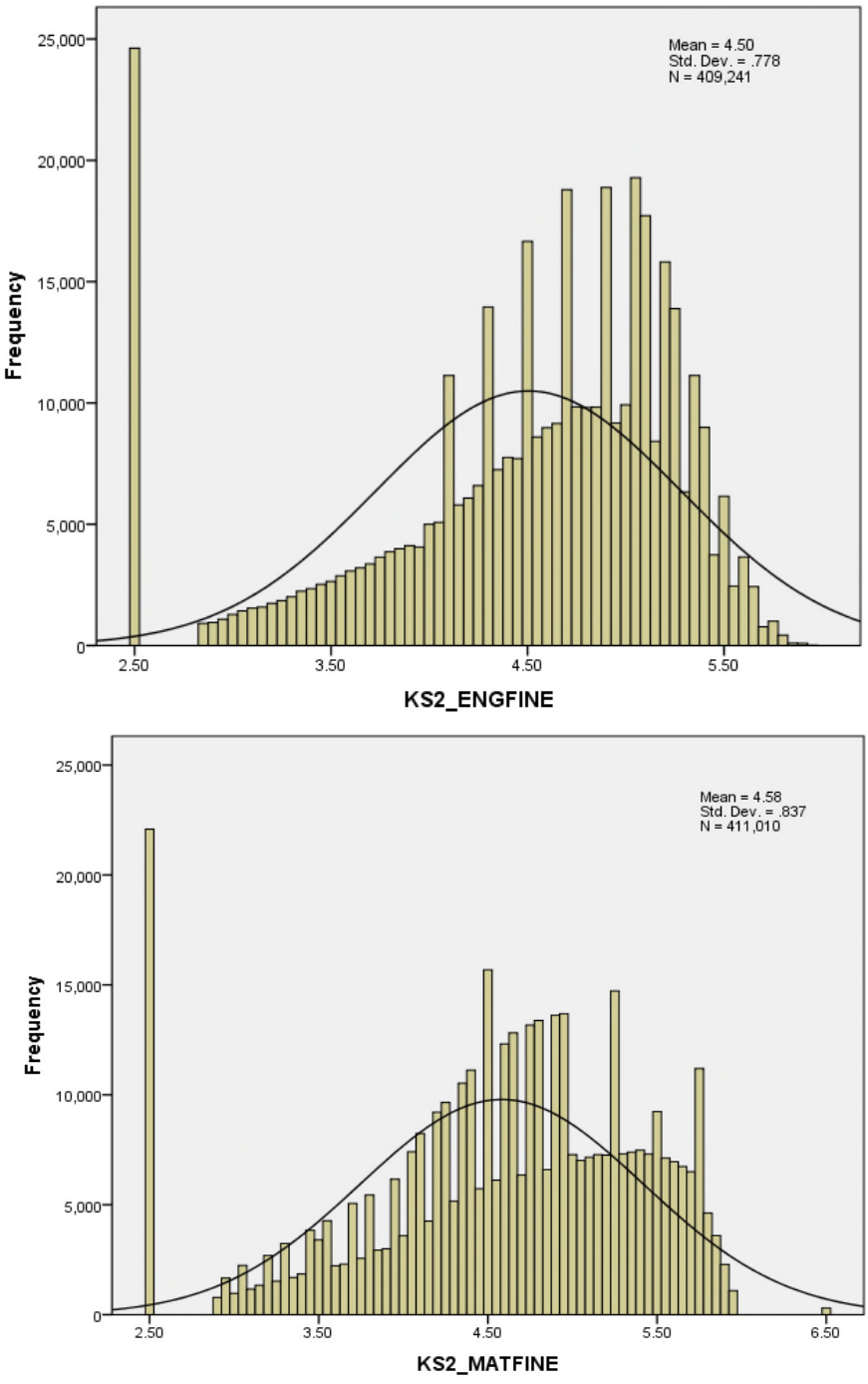


Figure K-1: Key Stage 2 English and Maths finely-graded level distributions

Option 1 – Use Teacher Assessment (TA) instead.

We created an average Teacher Assessment (TA) score:

RECODE ks2_ENGlevTA KS2_MATlevTA KS2_SCILevTA ('W'=0) (CONVERT) into enta mata scta.
COMPUTE AvTA=mean.3(enta,mata,scta).

On the positive side this includes 99% of all pupils (see Frequency table below).

Table K.1: Frequency table: Key Stage 2 Teacher Assessment levels

AvTA

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	2346	.4	.4	.4
	.33	317	.1	.1	.5
	.67	498	.1	.1	.6
	1.00	1300	.2	.2	.8
	1.33	1191	.2	.2	1.0
	1.67	1681	.3	.3	1.3
	2.00	5888	1.0	1.1	2.4
	2.33	6293	1.1	1.1	3.5
	2.67	9898	1.8	1.8	5.3
	3.00	31394	5.6	5.6	10.9
	3.33	30500	5.4	5.5	16.4
	3.67	47443	8.4	8.5	24.9
	4.00	155648	27.6	27.9	52.8
	4.33	73662	13.1	13.2	66.0
	4.67	68996	12.3	12.4	78.4
	5.00	118450	21.0	21.2	99.6
	5.33	1808	.3	.3	100.0
	5.67	205	.0	.0	100.0
	6.00	40	.0	.0	100.0
	Total	557558	99.0	100.0	
Missing	System	5589	1.0		
Total		563147	100.0		

Further, the new measure does correlate very highly with ENGfine and MATfine (r=0.84 & 0.86 respectively, Spearman's rho= 0.82 & 0.85 respectively) and shown below. However, it is not very fine grained (and negatively skewed).

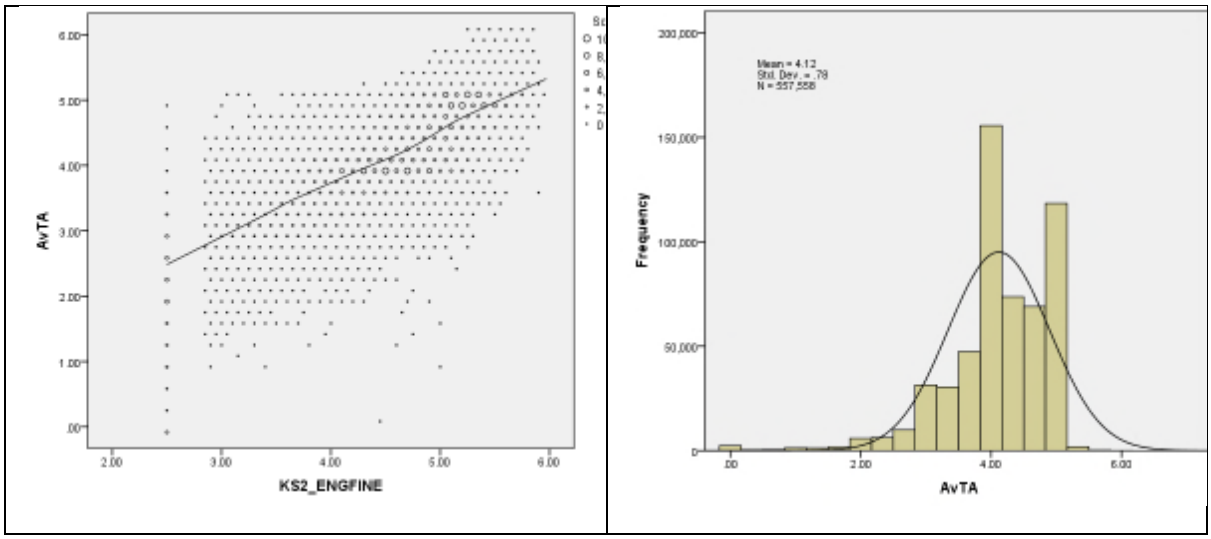


Figure K-2: Key Stage 2 English finely-graded level and Teacher Assessment correlation and Teacher Assessment distribution

Option 2. Exclude missing cases listwise

We took as a base population all those with valid average TA scores (99%). We then compared the distribution of pupil characteristics for the total Y6 population against the distribution within the boycott group. We identified the boycott group as pupils with neither a KS2 English nor a KS2 Maths fine grade. This is over-inclusive, because some pupils without KS2 test scores may have severe SEN or be in special schools rather than boycott schools, but it provide an approximate test. The table below shows the number and proportion of pupils who were tested or boycott.

Table K.1: Cases included with listwise deletion based on KS2 Teacher Assessment

Base is all with valid AVTA

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	tested	410718	72.9	73.3	73.3
	boycott	149709	26.6	26.7	100.0
	Total	560427	99.5	100.0	
Missing	-9.00	2720	.5		
Total		563147	100.0		

Of primary interest were representativeness for our key variables of ethnicity and SEN type. Apart from some under-representation of Pakistani pupils (2.9% of tested vs. 3.7% of total population) and Bangladeshi (1.2% vs. 1.5%) the proportion of ethnic minorities in the three-quarters of pupils who have KS2 results is similar to the whole population.

Table K.2: Comparison of valid 2010 (Y7) and only non-missing TA records: Ethnic group %

EthnicGroupMinor_SPR10 * Base is all with valid AVTA Crosstabulation

% within Base is all with valid AVTA

		Base is all with valid AVTA		
		tested	boycott	Total
EthnicGroupMinor_SPR10	ABAN	1.2%	2.2%	1.5%
	AIND	2.2%	2.6%	2.3%
	AOTH	1.2%	1.2%	1.2%
	APKN	2.9%	5.8%	3.7%
	BAFR	2.8%	3.0%	2.8%
	BCRB	1.3%	1.6%	1.4%
	BOTH	0.5%	0.6%	0.5%
	CHNE	0.3%	0.3%	0.3%
	MOTH	1.4%	1.4%	1.4%
	MWAS	0.8%	0.8%	0.8%
	MWBA	0.4%	0.4%	0.4%
	MWBC	1.3%	1.4%	1.3%
	NOBT	0.2%	0.1%	0.2%
	OOTH	1.2%	1.5%	1.3%
	REFU	0.5%	0.5%	0.5%
	WBRI	77.8%	72.4%	76.3%
	WIRI	0.4%	0.3%	0.3%
	WIRT	0.1%	0.1%	0.1%
	WOTH	3.4%	3.4%	3.4%
	WROM	0.2%	0.2%	0.2%
Total		100.0%	100.0%	100.0%

The tested group is also fairly representative with regard to SEN type. 88.2% of the tested sample had no SEN compared to 88.8% of the population. For the individual types the difference between the proportion in the tested sample and the population never differed by more than 0.1%.

Table K.3: Comparison of valid 2010 (Y7) and only nonmissing Teacher Assessment records: primary SEN type

PrimarySENtype_SPR10 * Base is all with valid AVTA Crosstabulation

% within Base is all with valid AVTA

		Base is all with valid AVTA		
		tested	boycott	Total
PrimarySENtype_SPR10		88.2%	90.2%	88.8%
	ASD	1.0%	0.6%	0.9%
	BESD	2.3%	2.1%	2.2%
	HI	0.2%	0.2%	0.2%
	MLD	3.5%	3.2%	3.4%
	MSI	0.0%	0.0%	0.0%
	OTH	0.4%	0.4%	0.4%
	PD	0.4%	0.3%	0.4%
	PMLD	0.1%	0.0%	0.1%
	SLCN	1.5%	1.2%	1.5%
	SLD	0.5%	0.1%	0.4%
	SPLD	1.7%	1.5%	1.7%
	VI	0.1%	0.1%	0.1%
Total		100.0%	100.0%	100.0%

There is a slight under-representation for FSM (16.6% vs. 17.6%), but this is not large.

Table K.4: Comparison of valid 2010 (Y7) and only nonmissing Teacher Assessment records: FSM

FSMeligible_SPR10 * Base is all with valid AVTA Crosstabulation

% within Base is all with valid AVTA

		Base is all with valid AVTA		
		tested	boycott	Total
FSMeligible_SPR10	0	83.4%	79.8%	82.4%
	1	16.6%	20.2%	17.6%
Total		100.0%	100.0%	100.0%

Option 3. Impute missing KS2 fine grades

We were reluctant to do this for a missing proportion as large as 25% of the population.

Appendix L: Primary cohort, logistic regression results

Table L.1: Primary cohort (age 5-11) MLD: Odds Ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 3 Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	0.74	*	0.68	*	0.71	*	0.73	*
	Traveller Irish	***3.90***	*	***2.64***	*	1.00		0.92	
	Traveller Gypsy/Roma	***3.54***	*	***2.74***	*	1.02		0.95	
	White Other	0.79	*	0.73	*	***0.51***	*	***0.50***	*
	Mixed White & African	0.76	*	***0.59***	*	***0.61***	*	***0.62***	*
	Mixed White & Caribbean	1.20	*	0.87	*	0.93		0.94	
	Mixed White & Asian	***0.60***	*	***0.57***	*	***0.56***	*	***0.57***	*
	Mixed Other	0.83	*	***0.67***	*	***0.66***	*	***0.67***	*
	Indian	***0.51***	*	***0.49***	*	***0.47***	*	***0.47***	*
	Pakistani	1.27	*	0.97		***0.67***	*	***0.66***	*
	Bangladeshi	0.81	*	***0.54***	*	***0.37***	*	***0.37***	*
	Asian Other	***0.60***	*	***0.53***	*	***0.41***	*	***0.41***	*
	Black African	0.84	*	***0.51***	*	***0.44***	*	***0.44***	*
	Black Caribbean	1.34	*	0.88	*	0.87	*	0.87	*
	Black Other	1.10		0.70	*	***0.61***	*	***0.62***	*
	Chinese	***0.26***	*	***0.25***	*	***0.20***	*	***0.20***	*
	Any other group	0.88	*	***0.61***	*	***0.45***	*	***0.44***	*
	Unknown	1.08	*	0.91	*	0.86	*	0.87	*
FSM	Entitled to FSM			***2.07***	*	1.50	*	1.43	*
Gender	Boy			***1.61***	*	1.19	*	1.18	*
Birth season	Spring			1.33	*	1.00		0.97	
	Summer			***1.83***	*	1.02		0.97	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.85***	*	1.38	*	1.24	*
Combined Depr.	(IDACI 1SD + FSM)			***2.82***	*	***1.76***	*	***1.59***	*
EYFSP	CLL 1SD					***0.39***	*	***0.40***	*
	PSRN 1SD					0.72	*	0.69	*
	PSE 1SD					1.36	*	1.28	*
School type	Foundation							1.16	*
	Academy							0.77	
	Church							0.97	*
	Special/PRU/AP							***0.07***	*
School % FSM	Highest							1.23	*
	Average-High							1.16	*
	Average							1.09	*
	Low-Average							1.01	
School % Asian	Highest							0.98	
(except Pakistani)	Average-High							1.00	
	Average							1.01	
	Low-Average							0.96	
School Size	Smallest							1.11	*
	Small-Average							1.14	*
	Average							1.10	*
	Average-Large							1.06	*
	Nagelkerke Pseudo R Squared	0.005		0.059		0.225		0.235	
	Model-specific N (# of pupils)	553264		553264		553264		553264	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 3 includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 4 additionally includes school variables. Note: *=Significant at the level of p<0.05.

Table L.2: Primary cohort (age 5-11) SEMH/BESD: Odds Ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 3 Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	0.96		0.89		1.02		1.03	
	Traveller Irish	***2.19***	*	1.42		1.15		1.04	
	Traveller Gypsy/Roma	***1.52***	*	1.13		0.91		0.83	
	White Other	0.81	*	0.75	*	0.71	*	0.68	*
	Mixed White & African	1.29	*	1.05		1.10		1.07	
	Mixed White & Caribbean	***1.90***	*	1.39	*	***1.51***	*	1.44	*
	Mixed White & Asian	***0.63***	*	***0.60***	*	***0.63***	*	***0.62***	*
	Mixed Other	1.17	*	0.95		0.99		0.97	
	Indian	***0.26***	*	***0.25***	*	***0.26***	*	***0.25***	*
	Pakistani	***0.46***	*	***0.35***	*	***0.31***	*	***0.30***	*
	Bangladeshi	***0.34***	*	***0.22***	*	***0.21***	*	***0.20***	*
	Asian Other	***0.37***	*	***0.33***	*	***0.30***	*	***0.28***	*
	Black African	1.06		***0.64***	*	***0.61***	*	***0.58***	*
	Black Caribbean	***2.35***	*	***1.57***	*	***1.51***	*	1.39	*
	Black Other	1.44	*	0.91		0.86		0.82	*
	Chinese	***0.27***	*	***0.27***	*	***0.25***	*	***0.23***	*
	Any other group	***0.67***	*	***0.45***	*	***0.43***	*	***0.40***	*
	Unknown	0.95	*	0.80	*	0.80	*	0.78	*
FSM	Entitled to FSM			***2.30***	*	***2.01***	*	***1.89***	*
Gender	Boy			***3.36***	*	***2.54***	*	***2.54***	*
Birth season	Spring			1.05	*	0.94	*	0.92	*
	Summer			1.09	*	0.87	*	0.83	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.79***	*	***1.54***	*	1.30	*
Combined Depr.	(IDACI 1SD + FSM)			***3.08***	*	***2.49***	*	***2.16***	*
EYFSP	CLL 1SD					0.93	*	0.95	*
	PSRN 1SD					1.39	*	1.31	*
	PSE 1SD					***0.43***	*	***0.42***	*
School type	Foundation							1.05	
	Academy							1.43	*
	Church							0.90	*
	Special/PRU/AP							***0.02***	*
School % FSM	Highest							1.35	*
	Average-High							1.27	*
	Average							1.13	*
	Low-Average							1.06	*
School % Black Caribbean	Highest							1.16	*
/Mixed White & Caribbean	Average-High							1.06	*
	Average							1.05	*
	Lowest							--	
School Size	Smallest							1.19	*
	Small-Average							1.16	*
	Average							1.11	*
	Average-Large							1.06	*
	Nagelkerke Pseudo R Squared	0.010		0.090		0.162		0.172	
	Model-specific N (# of pupils)	553264		553264		553264		553264	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 3 includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 4 additionally includes school variables. Note: *=Significant at the level of p<0.05.

Table L.3: Primary cohort (age 5-11) ASD: Odds Ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 3 Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	0.76		0.76		0.96		0.92	
	Traveller Irish	***0.53***		***0.50***		***0.19***	*	***0.20***	*
	Traveller Gypsy/Roma	0.77		0.72		***0.34***	*	***0.35***	*
	White Other	0.85	*	0.83	*	0.68	*	***0.66***	*
	Mixed White & African	1.01		0.97		0.95		0.90	
	Mixed White & Caribbean	1.04		0.98		1.04		0.98	
	Mixed White & Asian	0.70	*	0.71	*	0.73	*	0.70	*
	Mixed Other	1.19	*	1.13		1.13		1.08	
	Indian	***0.40***	*	***0.39***	*	***0.37***	*	***0.34***	*
	Pakistani	***0.47***	*	***0.43***	*	***0.30***	*	***0.28***	*
	Bangladeshi	0.71	*	***0.63***	*	***0.51***	*	***0.48***	*
	Asian Other	0.67	*	***0.65***	*	***0.49***	*	***0.46***	*
	Black African	1.17	*	1.04		0.87		0.84	*
	Black Caribbean	1.40	*	1.26	*	1.05		1.01	
	Black Other	1.29		1.14		0.98		0.93	
	Chinese	0.76		0.78		0.68		***0.66***	
	Any other group	0.73	*	***0.66***	*	***0.54***	*	***0.52***	*
	Unknown	0.95		0.91	*	0.86	*	0.85	*
FSM	Entitled to FSM			1.11	*	0.79	*	0.82	*
Gender	Boy			***5.09***	*	***3.24***	*	***3.26***	*
Birth season	Spring			1.00		0.80	*	0.81	*
	Summer			1.00		***0.64***	*	***0.65***	*
Neighbourhood Depr.	Normalised IDACI 2SD			1.21	*	0.86	*	0.93	*
Combined Depr.	(IDACI 1SD + FSM)			1.22	*	0.74	*	0.79	*
EYFSP	CLL 1SD					0.90	*	0.89	*
	PSRN 1SD					1.27	*	1.31	*
	PSE 1SD					***0.30***	*	***0.31***	*
School type	Foundation							0.97	
	Academy							1.38	
	Church							0.99	
	Special/PRU/AP							***1.70***	*
School % FSM	Highest							0.85	*
	Average-High							0.84	*
	Average							1.02	
	Low-Average							1.07	
School % Asian (Indian/Pakistani/	Highest							1.26	*
./Bangladeshi/Asian Other)	Average-High							1.27	*
	Average							1.11	*
	Low-Average							1.01	
School Size	Smallest							0.97	
	Small-Average							1.03	
	Average							1.09	*
	Average-Large							1.03	
	Nagelkerke Pseudo R Squared	0.003		0.053		0.237		0.239	
	Model-specific N (# of pupils)	553264		553264		553264		553264	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 3 includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 4 additionally includes school variables. Note: *=Significant at the level of p<0.05.

Appendix M: Any SEN (SAP+) results for Cox and logistic regression models

Table M.4: Primary cohort (age 5-11) Any SEN (SAP+): Hazard ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 3 Exp(B)	
Ethnic group	White Irish	0.90		0.86	*	0.96	
	Traveller Irish	***2.61***	*	***1.90***	*	0.77	*
	Traveller Gypsy/Roma	***2.35***	*	***1.90***	*	0.75	*
	White other groups	0.93	*	0.87	*	***0.63***	*
	Mixed White & African	0.97		0.82	*	0.88	*
	Mixed White & Caribbean	1.24	*	0.98		1.07	*
	Mixed White & Asian	0.69	*	***0.66***	*	0.69	*
	Any other mixed	1.02		0.87	*	0.90	*
	Indian	***0.51***	*	***0.49***	*	***0.48***	*
	Pakistani	0.95	*	0.77	*	***0.55***	*
	Bangladeshi	0.77	*	***0.57***	*	***0.42***	*
	Any other Asian	***0.64***	*	***0.58***	*	***0.47***	*
	Black African	1.06	*	0.73	*	0.67	*
	Black Caribbean	***1.51***	*	1.11	*	1.08	*
	Black other groups	1.27	*	0.90	*	0.81	*
	Chinese	***0.59***	*	***0.58***	*	***0.50***	*
	Any other ethnic group	0.92	*	0.69	*	***0.53***	*
	Unknown	1.02	*	0.90	*	0.88	*
FSM	Entitled to FSM			***1.76***	*	1.27	*
Gender	Boy			***2.04***	*	***1.52***	*
Birth Season	Spring			1.16	*	0.88	*
	Summer			1.40	*	0.80	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.54***		1.11	*
Combined Deprivation	(IDACI 1SD+FSM)			***2.18***		1.34	*
EYFSP	CLL 1SD					***0.54***	*
	PSRN 1SD					0.84	*
	PSE 1SD					0.82	*
Deviance (-2LL)	Initial (null): 2906375.56	2904532.6		2875303.89		2772878.31	
Δ-2LL from empty model		1842.94		29228.72		102425.58	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI; Model 3 additionally includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); N=553264 pupils are included in all models. Deviance=-2*Log-Likelihood (-2LL) and deviance change from previous model (Δ-2LL) are used to assess model fit. *=significant at the level of p<0.05.

Table M.5: Primary cohort (age 5-11) Any SEN (SAP+): Odds ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 3 Exp(B)	
Ethnic group	White Irish	0.85	*	0.79	*	0.91	
	Traveller Irish	***3.09***	*	***2.25***	*	0.74	*
	Traveller Gypsy/Roma	***2.86***	*	***2.33***	*	0.80	*
	White Other	0.85	*	0.79	*	***0.52***	*
	Mixed White & African	0.95		0.77	*	0.81	*
	Mixed White & Caribbean	1.31	*	0.99		1.09	*
	Mixed White & Asian	***0.64***	*	***0.61***	*	***0.59***	*
	Mixed Other	0.99		0.82	*	0.84	*
	Indian	***0.46***	*	***0.43***	*	***0.36***	*
	Pakistani	0.94	*	0.73	*	***0.42***	*
	Bangladeshi	0.75	*	***0.51***	*	***0.29***	*
	Asian Other	***0.57***	*	***0.50***	*	***0.33***	*
	Black African	1.02		***0.65***	*	***0.55***	*
	Black Caribbean	***1.65***	*	1.15	*	1.13	*
	Black Other	1.28	*	0.86	*	0.73	*
	Chinese	***0.52***	*	***0.50***	*	***0.36***	*
	Any other group	0.84	*	***0.60***	*	***0.39***	*
	Unknown	1.01		0.87	*	0.82	*
FSM	Entitled to FSM			***2.03***	*	1.46	*
Gender	Boy			***2.26***	*	***1.74***	*
Birth season	Spring			1.19	*	0.84	*
	Summer			1.49	*	0.73	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.70***	*	1.18	*
Combined Depr.	(IDACI 1SD + FSM)			***2.66***	*	***1.58***	*
EYFSP	CLL 1SD					***0.47***	*
	PSRN 1SD					0.72	*
	PSE 1SD					0.75	*
	Nagelkerke Pseudo R Squared	0.006		0.093		0.343	
	Model-specific N (# of pupils)	553264		553264		553264	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 3 includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE). Note: *=Significant at the level of p<0.05.

Table M.6: Secondary cohort (age 11-16) Any SEN (SAP+): Odds ratios

		Model 1 Exp(B)		Model 2 Exp(B)		Model 3 Exp(B)		Model 4 Exp(B)	
Ethnic group	White Irish	1.00		0.91		0.90		1.32	*
	Traveller Irish	***4.92***	*	***2.54***	*	***1.95***	*	0.87	
	Traveller Gypsy/Roma	***3.28***	*	***2.33***	*	***1.94***	*	***0.40***	*
	White Other	0.88	*	0.77	*	0.76	*	***0.52***	*
	Mixed White & African	1.06		0.80	*	0.81	*	1.02	
	Mixed White & Caribbean	1.44	*	1.08	*	1.08	*	1.29	*
	Mixed White & Asian	0.82	*	0.75	*	0.75	*	0.99	
	Mixed Other	1.01		0.82	*	0.82	*	0.99	
	Indian	***0.45***	*	***0.41***	*	***0.42***	*	***0.46***	*
	Pakistani	0.96		***0.66***	*	***0.66***	*	***0.52***	*
	Bangladeshi	0.78	*	***0.42***	*	***0.43***	*	***0.45***	*
	Asian Other	***0.53***	*	***0.45***	*	***0.45***	*	***0.39***	*
	Black African	1.04		***0.59***	*	***0.60***	*	***0.63***	*
	Black Caribbean	***1.74***	*	1.15	*	1.17	*	1.30	*
	Black Other	1.30	*	0.83	*	0.84	*	0.77	*
	Chinese	***0.39***	*	***0.36***	*	***0.37***	*	***0.47***	*
	Any other group	0.90	*	***0.58***	*	***0.59***	*	***0.47***	*
	Unknown	1.10		1.02		1.01		0.98	
FSM	Entitled to FSM			***2.34***	*	***2.29***	*	***1.56***	*
Gender	Boy			***1.93***	*	***1.93***	*	***1.65***	*
Birth season	Spring			1.11	*	1.11	*	0.93	*
	Summer			1.23	*	1.24	*	0.91	*
Neighbourhood Depr.	Normalised IDACI 2SD			***1.76***	*	***1.75***		1.24	*
Combined Depr.	(IDACI 1SD + FSM)			***3.11***	*	***3.02***		***1.74***	*
Attendance	Persistent absence (>63 days/yr)					***3.88***	*	***2.16***	*
Prior attainment	KS2 English Finely Graded Level							***0.39***	*
	KS2 Maths Finely Graded Level							***0.64***	*
	Nagelkerke Pseudo R Squared	0.007		0.089		0.060		0.381	
	Model-specific N (# of pupils)	392708		392708		392465		392465	

Models correspond to stages of analysis: Model 1 has only ethnic group as a predictor; Model 2 additionally includes pupil background factors (FSM, gender, birth season, IDACI); Model 3 includes a persistent absence indicator (absent>63 days over a year); Model 4 additionally includes Reception attainment/development (EYFSP scores in CLL, PSRN, PSE); Model 5 additionally includes an interaction between KS2 scores and attendance, and Model 6 additionally includes school variables. Note: *=Significant at the level of p<0.05.

Appendix N: Filtered samples/interactions robustness of school composition effects

Table N.1: Primary cohort alternative filter results for ethnic composition variables

Alternative filtering - Primary		MLD Exp(B)		SEMH/BESD Exp(B)		ASD Exp(B)	
Ethnic group	White Irish	0.80		1.06		1.49	
	Traveller Irish	1.01		1.35		***0.00***	
	Traveller Gypsy/Roma	***0.62***	*	1.05		***0.14***	*
	White other groups	***0.48***	*	***0.66***	*	0.80	*
	Mixed White & African	***0.63***	*	1.09		1.13	
	Mixed White & Caribb.	0.96		1.36	*	1.02	
	Mixed White & Asian	***0.60***	*	***0.66***	*	0.82	
	Any other mixed	0.68	*	0.95		1.20	
	Indian	***0.48***	*	***0.25***	*	***0.41***	*
	Pakistani	***0.60***	*	***0.33***	*	***0.31***	*
	Bangladeshi	***0.36***	*	***0.22***	*	***0.63***	*
	Any other Asian	***0.41***	*	***0.30***	*	***0.67***	*
	Black African	***0.45***	*	***0.62***	*	1.06	
	Black Caribbean	0.91		1.29	*	1.07	
	Black other groups	***0.64***	*	0.85	*	1.39	*
	Chinese	***0.19***	*	***0.35***	*	0.75	
	Any other ethnic group	***0.45***	*	***0.43***	*	***0.51***	*
	Unknown	0.85	*	0.75	*	0.92	
FSM	Entitled to FSM	1.37	*	***1.69***	*	0.81	*
Gender	Boy	1.23	*	***2.47***	*	***3.12***	*
Birth Season	Spring	0.94	*	0.93	*	0.83	*
	Summer	0.90	*	0.78	*	***0.60***	*
Neighbourhood Depr.	Normalised IDACI 2SD	1.14	*	1.21	*	0.94	
Combined Deprivation	(IDACI 1SD+FSM)	1.46	*	***1.86***	*	0.78	*
EYFSP	CLL 1SD	***0.41***	*	0.97		0.91	*
	PSRN 1SD	***0.66***	*	1.19	*	1.17	*
	PSE 1SD	1.14	*	***0.41***	*	***0.33***	*
School Type	Foundation	1.14	*	1.13	*	1.08	
	Academy	1.03		***1.92***	*	1.07	
	Church	0.99		0.92	*	0.93	
	Special	***0.48***	*	***0.07***	*	***2.66***	*
School % FSM	Highest	1.18	*	1.28	*	0.72	*
	Average-High	1.13	*	1.32	*	***0.63***	*
	Average	1.06	*	1.14	*	0.80	*
	Low-Average	1.05	*	1.04		0.91	
School ethnic group quintile [a]	Highest	1.05	*	1.09	*	0.87	*
	Average-High	1.04		1.12	*	1.23	*
	Average	1.10	*	1.02		1.16	*
	Low-Average	1.05	*	0.95		1.00	
School Size	Smallest	1.16	*	1.23	*	0.92	
	Small-Average	1.09	*	1.19	*	1.05	
	Average	1.05	*	1.10	*	0.95	
	Average-Large	1.04	*	1.10	*	0.91	
Deviance (-2LL)		543840.78		284516.69		62509.16	
Model-specific N		422330.00		295861.00		363652.00	

*([a]) For MLD, the relevant combined ethnic groups are Asian, excluding Pakistani. For SEMH/BESD, the relevant combined groups are Black Caribbean and Mixed White and Caribbean. For ASD, the relevant combined groups are Asian restricted to Indian, Pakistani, Bangladeshi, and Asian Other. These groupings were defined according to earlier findings regarding which groups were persistently disproportionately identified for a given focal type of SEN. Filtering here excludes pupils in schools with fewer than 2 individuals in the particular combined ethnic grouping. *=significant at the level p<0.05.*

Table N.2: Secondary cohort alternative filter for MLD, SEMH/BESD, ASD ethnic composition variables

Alternative filtering -- Secondary		MLD Exp(B)		SEMH/BESD Exp(B)		ASD Exp(B)	
Ethnic group	White Irish	0.96		0.84		1.44	
	Traveller Irish	1.02		1.18		***0.32***	
	Traveller Gypsy/Roma	***0.53***	*	0.90		***0.08***	*
	White Other	***0.47***	*	***0.61***	*	***0.60***	*
	Mixed White & African	***0.64***	*	0.95		1.20	
	Mixed White & Caribbean	0.94		1.37	*	1.11	
	Mixed White & Asian	0.87		0.84	*	1.04	
	Mixed Other	0.73	*	1.03		1.04	
	Indian	0.77	*	***0.26***	*	***0.43***	*
	Pakistani	0.75	*	***0.37***	*	***0.33***	*
	Bangladeshi	***0.60***	*	***0.27***	*	***0.32***	*
	Asian Other	***0.55***	*	***0.36***	*	***0.40***	*
	Black African	***0.59***	*	***0.50***	*	0.72	*
	Black Caribbean	0.84	*	1.14	*	1.31	*
	Black Other	0.73	*	0.71	*	0.83	
	Chinese	***0.39***	*	***0.25***	*	***0.46***	*
	Any other group	***0.48***	*	***0.44***	*	***0.37***	*
	Unknown	1.05		0.98		0.94	
FSM	Entitled to FSM	1.29	*	***2.19***	*	0.77	*
Gender	Boy	1.09	*	***1.83***	*	***4.00***	*
Birth season	Spring	1.02		0.92	*	0.96	
	Summer	1.07	*	0.89	*	0.94	
Neighbourhood Depr.	Normalised IDACI 2SD	1.26	*	1.49	*	0.93	
Combined Depr.	(IDACI 1SD + FSM)	1.45	*	***2.67***	*	0.74	*
Attendance	Persistent absence (>63 days)	***1.58***	*	***2.19***	*	***2.26***	*
Prior attainment	KS2 English Finely Graded Level	***0.38***	*	0.79	*	0.69	*
	KS2 Maths Finely Graded Level	***0.60***	*	0.78	*	0.73	*
AttendanceXPrior attainment	KS2 English X Persistent abs.	1.17	*				
	KS2 Maths X Persistent abs.	1.18	*			***1.69***	*
School type	Foundation	0.99		0.99		1.02	
	Academy-Converter	0.73	*	0.96		1.18	
	Academy-Sponsored	0.98		0.97		1.09	
	Church	0.96		0.85	*	1.19	*
	Selective/Grammar	***0.43***	*	0.67	*	1.15	
	Special/PRU/AP	***0.51***	*	***0.50***	*	***6.35***	*
School % FSM	Highest	1.09	*	1.25	*	0.88	
	Average-High	1.11	*	1.26	*	0.95	
	Average	0.95		1.13	*	1.07	
	Low-Average	0.98		0.96		1.05	
School ethnic group %(*)	Highest	1.10	*	1.34	*	1.16	*
	Average-High	1.05		1.18	*	1.08	
	Average	1.05		1.09	*	1.14	*
	Low-Average	0.96		1.06	*	0.99	
School Size	Smallest	0.86	*	0.98		1.00	
	Small-Average	1.08	*	1.02		1.04	
	Average	1.01		1.08	*	1.02	
	Average-Large	1.07	*	1.03		1.04	
	Nagelkerke Pseudo R Squared	0.329		0.126		0.166	
	Model-specific N (# of pupils)	378459		344474		353280	

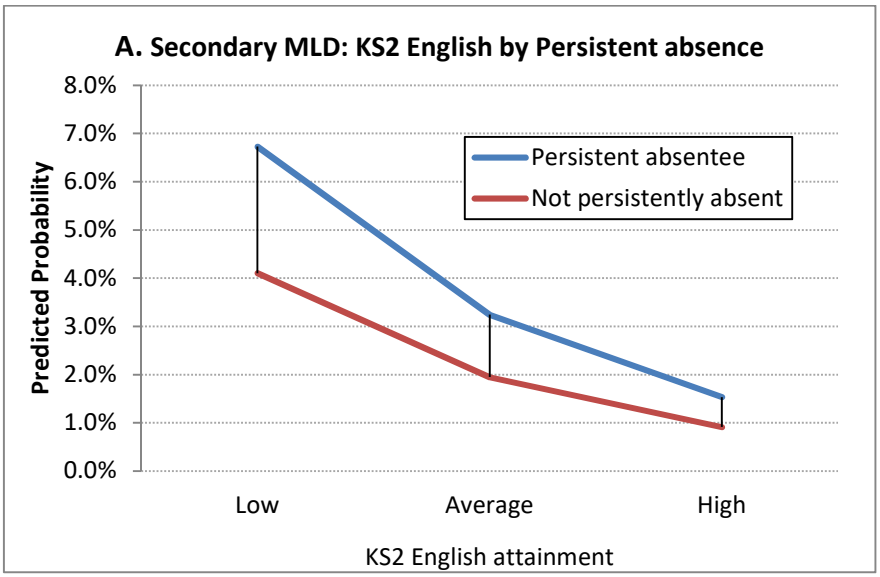
*(*) For MLD, the relevant combined ethnic groups are Asian, excluding Pakistani. For SEMH/BESD, the relevant combined groups are Black Caribbean and Mixed White and Caribbean. For ASD, the relevant combined groups are Asian restricted to Indian, Pakistani, Bangladeshi, and Asian Other. These groupings were defined according to earlier findings regarding which groups were persistently disproportionately identified for a given focal type of SEN. Filtering here excludes pupils in schools with fewer than 2 individuals in the particular combined ethnic grouping. *=significant at the level p<0.05.*

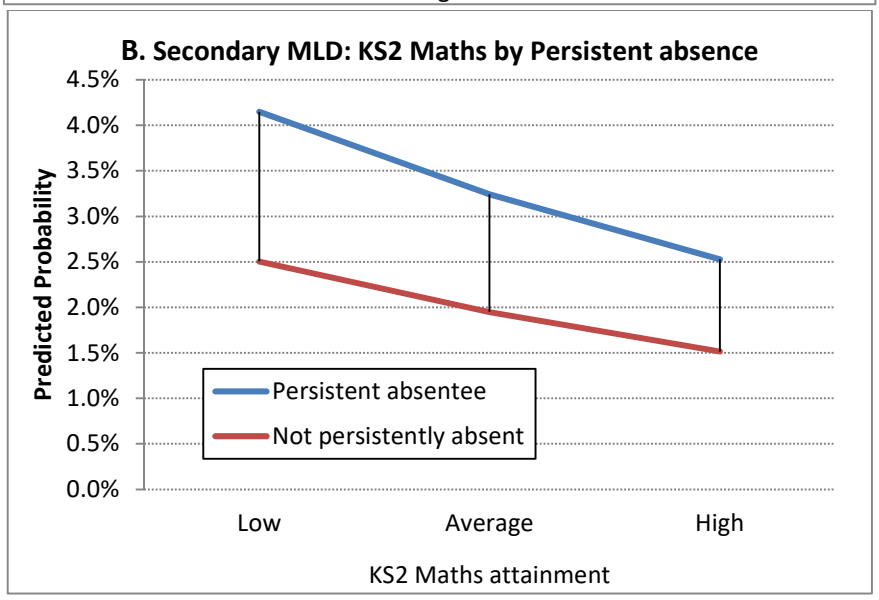
Appendix O: Persistent absence and KS2 attainment interactions (Secondary cohort)

MLD: KS2 English/maths attainment and persistent absence interaction

Figures O1 A and B provide visualisations of the interactions between KS2 English and Maths attainment with persistent absence (noted on p43) in terms of predicted probabilities. As the plots show, persistent absence in Y6 had a greater association with the probability of ever being identified with MLD in Secondary school for the lowest performers in terms of their KS2 attainment in both English and Maths, although the effect appears more dramatic for the KS2 English interaction.

Figure O-1: Secondary cohort MLD: Interactions between attendance and KS2 English and Maths attainment (predicted probabilities)





ASD: KS2 Maths attainment and persistent absence interaction

Figure O-2 below visualises this interaction (noted on p51) in terms of predicted probabilities; as was the case for MLD in the Secondary cohort, persistent absence was more strongly associated with higher odds of ever being identified with ASD for those pupils with the lowest prior attainment in Maths at KS2. This does not appear dramatic on the scale given; this is in part a consequence of the underlying very low incidence of ASD overall. The corresponding KS2 English and persistent absence interaction was not significant.

Figure O-2: Secondary cohort ASD: Interactions between attendance and KS2 Maths attainment (predicted probabilities)