

Education, children's social care and offending: multi-level modelling

Technical report

July 2023

Contents

List of figures	2
List of tables	3
Version control	4
1. Data sources and matching methodology	5
Data sources – MoJ/DfE data share	5
Matching process	6
Match rate	6
Representativeness of the matched dataset	7
Match rates by age	8
Caveats when using matched data	9
2. Definitions	10
Defining the total pupil cohort	10
Defining the offending cohort	11
3. Information about the pupil	13
Ethnicity data	13
Special educational needs data	15
Children known to children's social care data	16
4. Information about the local authority	18
Methodology	18
Assessment of local authority level variables	18
5. Multi-level modelling	26
Data clustering	26
Significance testing	27
Cross correlations	27
Limitations of the models	28
Model performance	30
6. Definition of serious violence	32

List of figures

Figure 1: Age breakdown of matched dataset compared to unmatched dataset7
Figure 2: Match rates by year of birth8
Figure 3: The average trend in children within tax credit recipient families as a proportion of children within child benefit recipient families, for local authorities in England between 2011/12 and 2017/18
Figure 4: The average trend in workless households as a proportion of all households, for local authorities in England between 2011/12 and 2017/1821
Figure 5: The average trend in the number of households on the waiting list for accommodation who are a) homeless, and/or b) occupying unsanitary or overcrowded housing, per 10,000 people, for local authorities in England between 2011/12 and 2017/18
Figure 6: The average trend in the number of people per square kilometre, for local authorities in England between 2011/12 and 2017/1824
Figure 7: The average trend in the proportion of children and young people with an EHC plan attending a mainstream school, for local authorities in England between 2011/12 and 2017/18

List of tables

Table 1: Officer identified ethnicity breakdown of matched data compared to unmatche data	ed 8
Table 2: Year group breakdown for the all-pupil cohort	.11
Table 3: Ethnic group major categories	14
Table 4: Ethnic group minor categories	.15
Table 5: Sample sizes, including number of observations and local authorities, for the multi-level regression models	7 29
Table 6: Adjusted r-squared measures for the fixed-effects factors, random-effects factors and the model in total, for each multi-level regression model	.30
Table 7: Home Office offence codes used to define serious violence	32

Version control

Publication date	Reason for update
10 March 2022	Technical report for Education, children's social care and offending: Descriptive Statistics ad-hoc statistics.
20 July 2023	Additional information added for Education, children's social care and offending: multi-level modelling research report.

1. Data sources and matching methodology

This section sets out details of the data share between the Ministry of Justice (MoJ) and Department for Education (DfE), and the matching process between the MoJ and National Pupil Database data. It refers to the most recent iteration of the share that occurred in 2022.

Data sources – MoJ/DfE data share

Data from several large datasets were brought together in this data share, as permitted by the Ministry of Justice's common law powers and various prescribed legislative data sharing powers available to DfE¹. A brief description of the two main datasets is included below:

National Pupil Database (NPD) – DfE

A wide range of information about pupils and students which provides evidence on educational performance and context. The data includes detailed information about pupils' test and exam results, prior attainment, and progression between each key stage for all state schools in England. It also includes information about the characteristics of pupils in the state sector and non-maintained special schools, such as their gender, ethnicity, first language, eligibility for free school meals, awarding of bursary funding for 16-19-year-olds, information about special educational needs, and detailed information about any absences and exclusions.

Police National Computer (PNC) – MoJ

This is an administrative data system used by police to monitor recordable offences, the offenders convicted or cautioned for them, and the outcomes received by these offenders. The system is live and subject to change e.g. following appeals. Recordable offences are defined as offences that can attract a custodial sentence plus some additional offences defined in legislation. Some non-recordable offences are also included on the PNC, particularly when they accompany recordable offences in the same case. The data analysed in this report is a subset of the total number of individuals on PNC and only some information (e.g. personal characteristics) will be available through the linked MoJ-DfE data. The linked share is based on an extract of PNC at a point in time. This report is based on offenders from the PNC that were successfully matched to the NPD, covering the period 2000 – 2020.

¹ How DfE shares personal data - GOV.UK (www.gov.uk)

Matching process

The methodology used to match the data sources together was similar to that used in other MoJ data linking projects, such as the data share between MoJ and the Department of Work and Pension (DWP) and Her Majesty's Revenue and Customs (HMRC)².

The data were matched using combinations of six demographic variables from the PNC and NPD: forename, surname, date of birth, gender, postcode, and the derived variable: full name.

Matching rules were agreed between MoJ and DfE and included combinations of at least four exact matches of the common variables. The majority of data was matched on rule 1 (exact names, date of birth, postcode and gender) accounting for 49% of all matches. In addition to full matching, partial matching was used to improve match rates when matching on forename, middle name and surname was not successful. As exact matching is very strict (either a word matches or it does not), partial matching improved match rates by including matches where the first two characters from forename, middle name or surname matched. Partial matching was also employed for date of birth (i.e. when date and month of birth were inverted) and postcode (i.e. by matching on the postcode sector, e.g. "SE14 5", rather than the full postcode).

Alias information – alternative names and dates of birth recorded for the same offender – from the PNC was also included in the data share. Previous data shares have indicated that this information plays a key role in data matching reports. As such, multiple names, dates of birth and postcodes were provided for some offenders.

Match rate

Not all offenders on the PNC were involved in the match to the National Pupil Database (NPD) as the NPD only began to record data from the 2001/02 academic year. Whilst attempting to match as many offenders on the PNC as possible, due to the limited time coverage of the NPD, it was only possible to match offenders between the ages of 10 and 35 as at December 2020. This meant the records of around 2.18 million offenders, aged between 10 and 35 years, from between 2000 and 2020 who were on the PNC were shared with DfE. Of those, around 1.67 million were matched and included in the final matched dataset after cleaning. A good match rate of around 77% was achieved. Figures in this publication are based on matched offenders only and, as a result, volumes will be lower than published statistics from individual data sources.

² Experimental statistics from the 2013 MoJ /DWP /HMRC data share - GOV.UK (www.gov.uk)

Representativeness of the matched dataset

The 1.67 million records in the final matched Police National Computer dataset were compared against the 2.18 million records shared with DfE for data matching that did not match. Overall, the matched dataset had similar characteristics to the unmatched dataset in terms of gender and age, with some noted differences for ethnicity.

- The matched dataset was 73.5% male and 26.5% female (excluding unknown) which was broadly in line with the unmatched dataset (77.5% male and 22.5% female).
- The age breakdown of the matched dataset was similar to the unmatched dataset, although the matched dataset slightly under-represents older offenders and slightly over-represents young offenders. This is due to better matching rates for young offenders.



Figure 1: Age breakdown of matched dataset compared to unmatched dataset

Comparisons were also made between the matched and unmatched officer identified ethnicity. The comparisons in Table 1 show a higher proportion of offenders from a "Black" and "Asian" background (8.6% and 5.8% of the matched data, compared to 4.8% and 5.0% of the unmatched data) and a slightly lower proportion from a "White" background.

Table 1: Officer identified ethnicity breakdown of matched data compared to unmatched data

Ethnicity	Matched data	Unmatched data
White	85%	88%
Black	9%	5%
Asian	6%	5%
Other	1%	3%

Match rates by age

The overall match rate was around 77%, but a greater proportion of younger offenders were matched as they will have had a greater likelihood of being included in the National Pupil Database where matched data is available from 2002/03.



Figure 2: Match rates by year of birth

Caveats when using matched data

There are a number of caveats which should be considered when using the matched data:

- The matched data has been produced using administrative data sources whose main purposes are not solely statistical. Therefore, as with any large recording system, the data are subject to possible errors with data entry and processing. Quality assurance procedures, including cleaning of duplicated offender entries and checks for completeness and representativeness, have been applied to the matched data.
- The comparisons on representativeness provide some assurances that the matched data is broadly reflective of the offender cohort, but it should be made clear that this is not the full offender population.
- The analysis in this report is based only on the final matched PNC dataset. Around 23% of offenders aged 35 and under were not uniquely matched to the NPD. Reasons for this include:
 - They offended in England or Wales and were educated in Wales, Scotland, Northern Ireland or outside of the United Kingdom
 - Different names were recorded (potentially due to the offender changing their name or reporting a different name) on the NPD and the PNC
 - They have a common set of characteristics (i.e. the same name, date of birth and/or postcode) that make it difficult to determine a unique match across the datasets.

2. Definitions

Defining the total pupil cohort

The cohorts of children within this analysis includes all pupils who:

- Appear on a school census at the end of key stage 2 (KS2) and the end of key stage 4 (KS4). Pupils who do not have a KS4 record are not included.
- Finished KS2 in 2010/11, 2011/12 or 2012/13 and were aged ten at the start of these academic years
- Finished KS4 in 2015/16, 2016/17 or 2017/18³.

This amounts to approximately 1.53 million pupils⁴. For these children, all their records from Year 1 to Year 13 (or equivalent) are included (see Table 2).

Pupils who attended an independent primary or secondary school⁵ have been excluded from the data as their characteristics are not recorded, except for those registered in independent alternative provisions settings. The cohorts were selected in this way to maximise data coverage and balance data availability across each of the datasets provided in the share.

³ Allowances were made for children that moved ahead or were kept behind by one or more years.

⁴ An additional three cohorts of children were included within the analysis for patterns of offending over time in Section 1, to allow for comparison of the composition of cohorts used in analysis for the report 'Education, children's social care and offending: Descriptive Statistics'. These additional cohorts of children include all pupils who finished key stage 2 (KS2) in 2007/08, 2008/09 or 2009/10 and were aged 10 at the start of these academic years. Therefore, this cohort has a key stage 4 (KS4) academic year of 2012/13, 2013/14 or 2014/15 amounting to approximately 1.63 million pupils.

⁵ As measured by whether they appear in the KS2 or KS4 exam data under an independent school.

Table 2: Year group breakdown for the all-pupil cohort

Academic Year	KS4 Academic Year 2017/18	KS4 Academic Year 2016/17	KS4 Academic Year 2015/16
2019/20	Year 13		
2018/19	Year 12	Year 13	
2017/18	Year 11	Year 12	Year 13
2016/17	Year 10	Year 11	Year 12
2015/16	Year 9	Year 10	Year 11
2014/15	Year 8	Year 9	Year 10
2013/14	Year 7	Year 8	Year 9
2012/13	Year 6	Year 7	Year 8
2011/12	Year 5	Year 6	Year 7
2010/11	Year 4	Year 5	Year 6
2009/10	Year 3	Year 4	Year 5
2008/09	Year 2	Year 3	Year 4
2007/08	Year 1	Year 2	Year 3
2006/07		Year 1	Year 2
2005/06			Year 1

Results are provided for this total pupil cohort (including all children who had been cautioned or sentenced for an offence) for comparison purposes⁶.

Defining the offending cohort

We have identified two offending groups in the analysis:

1) Children that have been cautioned or sentenced for an offence (approximately 46,500 children, equivalent to 3% of the all-pupil cohort) – refers to all children and young people in the linked data who were in the academic years described above, and who had been

⁶ Offender numbers and pupil population figures quoted in this analysis may differ when analysing different education variables, since the information on these variables may not be recorded for all offenders and pupils.

cautioned or sentenced for any offence recorded in the PNC over the defined coverage period.

2) Children that have been cautioned or sentenced for a serious violence offence (approximately 14,900 children, equivalent to 1% of the all-pupil cohort) – refers to children who have been found guilty (in-court conviction) or cautioned (out-of-court caution) for an offence that falls under the following broad categories of offence groups and types: indictable only 'violence against the person' offences, indictable only 'robbery offences', and triable either way or indictable only 'possession of weapons offences'. See Section 6 for the full list of offences included within this definition of serious violence.

The analysis looks separately and independently at children who fall into these offending groups. Children who have been cautioned or sentenced for a serious violence offence are considered as subsets of children who have been cautioned or sentenced for an offence. Therefore, the findings for each of these groups are not additive as individuals can be considered within more than one group. Approximately 32% of the children who had been cautioned or sentenced for an offence group are also in the children who had been cautioned or sentenced for a serious violence offence group.

Limiting the possible time to offend

The DfE data was matched onto PNC data on a pupil level and covers offences in the period 2000 - 2020. However, in order to avoid skewing the offending data between the three sub-cohorts, we have limited the PNC records so that each cohort has an equivalent number of years of offending data. The coverage period is between: 2010 and 31 August 2018 for the offending group finishing KS4 in 2015/16, 2011 and 31 August 2019 for those finishing KS4 in 2016/17, and 2012 and 31 August 2020 for those finishing KS4 in 2017/18. This means that offences from age 10 and above are included only, and that the last year of offending data would be during Year 13 (or equivalent).

3. Information about the pupil

Most of the DfE data used in the analysis is taken from the school census, which is a pupil-level data collection from primary, secondary, special and state-funded alternative provision (AP) schools (pupil referral units, AP academies and AP free schools). The school census takes place three times a year; in the Autumn, Spring and Summer terms.

Data from the Pupil Referral Unit (PRU) and the AP censuses is also included. The PRU census was a yearly Spring collection census of all state-funded AP settings which was incorporated into the school census in 2013/14. The AP census is also a yearly Spring census. Since the AP and PRU censuses are yearly census, missing termly data for Autumn and Summer terms was inferred from the Spring data collection of the same academic year, where appropriate. Additionally, some data is collected in the school census that is not collected in the AP and PRU census. Where appropriate, this missing data has been filled in from the school census.

To be aware; in most cases, where pupils are registered in two (or more) schools, the pupil's main record from the school census was used to obtain information about the pupil. However, in some cases, existence of a dual-subsidiary record was noted, and the student flagged as attending more than one educational setting. We have incorporated information from pupil's subsidiary records for school, local authority and SEN, in order to capture as much information as possible.

In addition to this, examination data was also included. This data was matched onto the school census base data at a pupil-level from the KS2 and KS4 examination data. Where duplicate results existed for students, the latest academic year was taken. If duplicates remained, the highest score was used.

Ethnicity data

Data on a child's ethnicity is taken from the School Census. As of 2011, information regarding ethnicity could only be provided by the child or their parent(s).

Table 3: Ethnic group major categories

Code	Ethnic group
AOEG	Any other ethnic group
ASIA	Asian
BLAC	Black
CHIN	Chinese
MIXD	Mixed
UNCL	Unclassified
WHIT	White

Table 4: Ethnic group minor categories

Code	Ethnic group
ABAN	Bangladeshi
AIND	Indian
AOTH	Any other Asian background
APKN	Pakistani
BAFR	Black African
BCRB	Black Caribbean
BOTH	Any other Black background
CHNE	Chinese
MOTH	Any other Mixed background
MWAS	Mixed White and Asian
MWBA	Mixed White and Black African
MWBC	Mixed White and Black Caribbean
NOBT	Information not yet obtained
OOTH	Any other ethnic group
REFU	Refused
WBRI	White British
WIRI	White Irish
WIRT	Traveller of Irish heritage
WOTH	Any other White background
WROM	Gypsy/Roma

Special educational needs data

Pupils identified with special educational needs (SEN) are classified as those that have an Education, Health and Care (EHC) plan (or, prior to reforms introduced in September 2014, a Statement of SEN) and those who are in the SEN Support category (or, prior to reforms introduced in September 2014, School Action or School Action Plus). The period for local authorities to transfer children and young people with Statements of SEN to EHC plans started in September 2014 and ended on 31 March 2018. Primary type of need is collected through the school census for those pupils on SEN Support, or EHC plan (or the pre-2014 equivalents). The coverage for January 2015 onwards is different to previous years. Pupils who were on School Action were not required to have a primary type of need recorded. From 2015 pupils who were on School Action who have transferred to SEN support will be recorded as having a primary type of need. This has led to an increase in the number of pupils recorded as having a primary type of need.

There were changes to the classification of type of need in 2015: the previous code of 'Behaviour, Emotional and Social Difficulties (BESD)' was removed. A new code 'Social, Emotional and Mental Health (SEMH)' was introduced, although this was not intended to be a direct replacement.

Children known to children's social care data

In Section 1 of the report, and in Models 1 and 2 in Section 2, the Longitudinal CIN Dataset (LCD) was provided internally and includes matched children in need (CIN) and children who were looked after (CLA) records from 2012/13 onwards. This dataset includes information on referrals, assessments, length of time with plan and periods of care, irrespective of when during the year they occurred. Children are included in the analysis for CIN and CLA if they have been recorded as such in any period between the ages of 10 and 17. Those matched to earlier years in the KS4 attainment data will as a result have less coverage than those matched to later years. For example: those with KS4 academic year 2015/16 have coverage from age 12 and above.

This differs from the measures used in the annual publication 'Children looked after in England including adoptions', which looks at whether the child was recorded as being looked after on 31st March in the previous year; whether the child started being looked after during the previous year ending 31st March; and for offending specifically, whether the child had been looked after for at least 12 months in the year ending 31st March. This also differs from the measures used in the previous publication⁷ that primarily looked at whether the child had ever been CIN or CLA on the 31st March in any given year.

Given that all periods during the year are considered, rather than just 31st March as in previous publications, figures provide a better estimate of those that have ever been CIN. However, figures are still likely an under-estimate of the number of children that have ever been a child in need, or a child who was looked after, and who have (or who have not) offended as the matched data only includes children of school age recorded as being CIN and/or CLA from 2012/13 onwards as held in the LCD. This means that any

⁷ Education, children's social care and offending (publishing.service.gov.uk)

child who started to be looked after and ceased to be looked after prior to school age will not appear in this matched data as ever being CLA.

The years 7 to 11 models use termly CIN and CLA datasets that were provided internally, which includes CIN data from academic year 2011/12 term 3, and CLA data from academic year 2004/05 term 3. Due to data coverage, one of the cohorts of children only has CIN data for term 3 in year 7, the other two cohorts of children have full CIN data in the year 7 model. The CIN census was introduced in 2008/09, and initially covered a reduced 6-month period. A full year collection was introduced in 2009 - 2010, however a number of local authorities were unable to provide a complete, clean children in need return for that year. Reviews were carried out on the CIN census and resulted in some data items being removed from the 2010 - 2011 collection onward.

4. Information about the local authority

Methodology

Details of the local authority the child attended school in were obtained via two different methods. These include:

1) The local authority of the school they attended during the school year they entered when they were aged 15 and completed KS4 examinations at. We selected this age as this is the most common age at which children who had been cautioned or sentenced for a serious violence offence commit their first serious violence offence⁸. Where this wasn't possible, the closest previous, or subsequent⁹, year has been used. Where duplicate results existed for students, the earliest academic year was taken. Duplicates that remained were primarily due to school changes as a result of academisation, therefore the most recent URN¹⁰ was chosen.

2) The local authority of the school they were recorded in the school census as having a main registration at (where possible) each school year. Where duplicate results existed for students, the local authority recorded in the latest academic term was taken.

Assessment of local authority level variables

Local authority level characteristics were generally selected to align with the domains of deprivation in the index of multiple deprivation¹¹ (IMD). The following provides a description of the local authority level characteristics that have been chosen, including how they compare with the results of the IMD for each domain, as well as a high-level analysis of the average trend¹² for each characteristic for the timeframes covered by the models. Figures may differ from nationally reported statistics as these measures, where necessary, have been converted from counts into rates to account for the population size of the local authority.

Income deprivation domain

⁸ Education, children's social care and offending: local authority level dashboard, Academic year 2019/20 – Explore education statistics – GOV.UK (explore-education-statistics.service.gov.uk)

⁹ Up to age 17.

¹⁰ Unique reference number.

¹¹ The English Indices of Deprivation 2019 (publishing.service.gov.uk)

¹² We have calculated and displayed the average trend for the 150 local authorities included in the underlying data.

Children¹³ within tax credit¹⁴ recipient families¹⁵ as a proportion of children¹⁶ within child benefit recipient families. Tax credits provide income related support for families with children, and in-work support for people of low incomes¹⁷, and therefore tax credit recipient families have been used in this instance as a proxy for families on low income. Child benefit recipient families have been used in this instance as a proxy for the number of households with children in England¹⁸. Both measures are likely to be an underestimate of the true number of families as they refer to the number of claimants, which is a subset of those eligible to claim.

On average, the proportion of children in families claiming tax credits fell between 2011/12 - 2017/18, with the biggest drop occurring between 2011/12 and $2012/13^{19}$. There were significant overlaps between local authorities with the highest levels of income deprivation in 2015^{20} and local authorities with the highest proportion of children in families claiming tax credits, suggesting this is a good indicator of income deprivation, especially among households with children.

¹⁶ Includes qualifying young people up to age 19.

¹³ Includes qualifying young people up to age 19.

¹⁴ Working tax credits and/or child tax credits. This measure also includes out of work families with children who receive the same level of support as provided by child tax credits, but where it is paid as child allowances in Income Support or income-based Jobseeker's Allowance (IS/JSA).

¹⁵ Formed of at least one person aged at least 16 responsible for a child or young person(s).

¹⁷ For more information on eligibility for tax credits, please see: <u>Personal tax credits statistics - GOV.UK</u> (www.gov.uk)

 ¹⁸ Child Benefit is received if responsible for bringing up a child who is under 16, or is under 20 and in education or training. For more details, visit: <u>Child Benefit: How it works - GOV.UK (www.gov.uk)</u>
¹⁹ Note that welfare reforms were introduced in 2013, including the introduction of Universal Credit. Please see here for more information: <u>2010 to 2015 government policy: welfare reform - GOV.UK (www.gov.uk)</u>
²⁰ Indices of deprivation 2015 (publishing.service.gov.uk)

Figure 3: The average trend in children within tax credit recipient families as a proportion of children within child benefit recipient families, for local authorities in England between 2011/12 and 2017/18



Source: HMRC (2011 - 2018) 'Personal tax credits statistics' (Personal tax credits statistics)

Employment deprivation domain

Workless households²¹ as a proportion of all households²². This provides an indication of the level of unemployment in a local authority. On average, the proportion of workless households fell between 2011/12 – 2017/18. Local authorities in the North East tended to have consistently higher proportions of households that were workless. There were significant overlaps between local authorities with the highest levels of employment deprivation in 2015²³ and local authorities with the highest proportion of workless households, suggesting this is a good indicator of employment deprivation.

²¹ Defined as no individuals aged 16 and over are in employment.

²² Defined as a single person, or a group of people living at the same address who have the address as their only or main residence and either share one main meal a day or share living accommodation (or both).

²³ Indices of deprivation 2015 (publishing.service.gov.uk)



Figure 4: The average trend in workless households as a proportion of all households, for local authorities in England between 2011/12 and 2017/18

Source: ONS (2011 – 2018) 'Official Census and Labour Market Statistics' (Annual population survey)

Barriers to housing and services domain

The number of households²⁴ on the waiting list for accommodation as they are homeless²⁵ per 10,000 people²⁶, and the number of households on the waiting list for accommodation as they are occupying unsatisfactory or overcrowded housing per 10,000 people²⁷.

These measures provide an indication of issues relating to access to, and quality of, housing in a local authority. These measures are not mutually exclusive and as such, households can fall under more than one measure. To understand whether this will be a

²⁶ Refers to all ages. Underlying data available here: <u>Estimates of the population for the UK, England,</u> Wales, Scotland and Northern Ireland - Office for National Statistics (ons.gov.uk)

²⁴ Defined as one person or a group of people (not necessarily related) who have the accommodation as their only or main residence, and (for a group) share cooking facilities and share a living room or sitting room or dining area. Please see here for more information: <u>Housing statistics and English Housing Survey glossary - A to Z - Guidance - GOV.UK (www.gov.uk)</u>

²⁵ Homeless within Part 7 of the Housing Act 1996. For more information, please see: <u>Local authority</u> housing statistics guidance 2020-21 (publishing.service.gov.uk)

²⁷ Refers to all ages. Underlying data available here: <u>Estimates of the population for the UK, England,</u> <u>Wales, Scotland and Northern Ireland - Office for National Statistics (ons.gov.uk)</u>

concern for the model, a Spearman's correlation coefficient²⁸ was calculated for these two measures. The output was between 0.20 and 0.51²⁹, demonstrating a positive, but weak to moderate monotonic association³⁰ between the two.

The average rate of households on the waiting list for accommodation who are homeless has remained relatively steady, whereas the rate of households on the waiting list for accommodation who are occupying unsanitary or overcrowded housing has fallen significantly between 2011/12 – 2017/18. Local authorities in London consistently had some of the highest rates across the two measures, with local authorities in the South East also having some of the higher rates of households on the waiting list for accommodation who are occupying unsanitary or overcrowded housing. There were more overlaps between local authorities with the highest rates of households on the waiting list for accommodation who are homeless and local authorities with the highest levels of barriers to housing and services deprivation in 2015³¹, than those with the highest rates of households on the waiting list for accommodation, suggesting these to be fair indicators of barriers to housing and services.

Figure 5: The average trend in the number of households on the waiting list for accommodation who are a) homeless, and/or b) occupying unsanitary or

²⁸ A full table of Spearman correlation coefficients and associated p values can be found in the accompanying tables.

²⁹ This range of values was calculated across the seven models, comparing the mean number of households on the waiting list for accommodation who are homeless with the mean number of households on the waiting list for accommodation who are occupying unsanitary or overcrowded housing, and the deviation from the mean of both variables.

³⁰ <u>spearmans.pdf (statstutor.ac.uk)</u>

³¹ Indices of deprivation 2015 (publishing.service.gov.uk)

overcrowded housing, per 10,000 people, for local authorities in England between 2011/12 and 2017/18



Source: DLUHC (2011 – 2018) 'Local authority housing statistics' (Local authority housing statistics) and ONS (2011 – 2018) 'Estimates for the population for the UK' (Population estimates)

Living environment deprivation domain

The number of people³² per square kilometre in a local authority. This is an indication of how densely populated a local authority is. The number of people per square kilometre has gradually grown between 2011/12 - 2017/18, with local authorities in London being some of the most densely populated. There were significant overlaps between local authorities with the highest levels of living environment deprivation in 2015^{33} and the most densely populated local authorities, suggesting this is a good indicator of living environment deprivation.

³² Refers to all ages.

³³ Indices of deprivation 2015 (publishing.service.gov.uk)



Figure 6: The average trend in the number of people per square kilometre, for local authorities in England between 2011/12 and 2017/18

Source: ONS (2011 – 2018) 'Estimates for the population for the UK' (Population estimates)

Additionally, we have included the proportion of children and young people³⁴ with an EHC plan attending a mainstream school³⁵ as a local authority area level characteristic.

The proportion of children and young people with an EHC plan in mainstream settings slightly fell between 2011/12 and 2017/18 (see Figure 7).

³⁴ Up to age 25.

³⁵ Note that this measure refers to state funded mainstream schools only. This measure differs to that collected in SEN2, which is collected by the local authority that maintains the EHC plan and therefore may include out of area placements ie: the EHC plan may be maintained by one local authority, however the child attends school in a different local authority area. The measure used in this analysis refers strictly to the proportion of children with an EHC plan that attend school within that local authority area, irrespective of which local authority maintains the EHC plan.

Figure 7: The average trend in the proportion of children and young people with an EHC plan attending a mainstream school, for local authorities in England between 2011/12 and 2017/18



Source: DfE (2011 - 2018) 'Statistics: special educational needs' (Special educational needs)

5. Multi-level modelling

Models	What is estimated?
Model 1	The association between <i>ever</i> having been identified with a characteristic and the outcome of <i>ever</i> having been cautioned or sentenced for a serious violence offence.
Model 2	The association between ever being identified with a characteristic <i>before</i> a serious violence offence and the outcome of <i>ever</i> having been cautioned or sentenced for a serious violence offence.
Years 7 – 11 models	The association between being identified with a characteristic <i>during one academic year</i> , and the outcome of being cautioned or sentenced for a serious violence offence at any point in the <i>following two academic years</i> .

Data clustering

To understand whether there is clustering in the data, and thereby determining whether a single or multilevel model would be most appropriate for this analysis, we ran a variance components model (a model with the outcome variable and a local authority random intercept, but no explanatory variables) to calculate the variance partition coefficient (VPC). In a variance components model, the VPC describes the degree of clustering observed in the outcome variable, in this case the propensity for a child to be cautioned or sentenced for a serious violence offence. We calculated the VPC to be between 0.06 and 0.09 across each of the 7 models, which means between 6-9% of the variation in the propensity to be cautioned or sentenced for a serious violence offer a serious violence offence lies between local authorities. In this instance, the VPC is calculated using the same formula as the intraclass correlation coefficient (ICC), which here offers the interpretation that the expected correlation between two children attending school in the same local authority is between 0.06 and 0.09.

When we included both school and local authority in the variance components model, we calculated the VPC for schools to be between 0.15 and 0.44, and the VPC for local authorities to be between 0.04 and 0.08. This means between 15-44% of the variation in the propensity to be cautioned or sentenced for a serious violence offence lies between schools, and between 4-8% of the variation lies between local authorities, when schools and local authorities are accounted for as levels.

Significance testing

In previous analysis³⁶, it was identified that children who were cautioned or sentenced for a serious violence offence were more likely to have received multiple suspensions, compared to the all-pupil cohort. We conducted significance testing³⁷ to test whether there was a distinct number of suspensions received that was statistically different between the two groups. We identified that for any given number³⁸ of suspensions received³⁹, the difference between the proportion of children cautioned or sentenced for a serious violence offence receiving that number of suspensions, and the proportion of the all-pupil cohort receiving the same number of suspensions was statistically significant⁴⁰. We have therefore included whether the child has received at least two suspensions to capture the cumulative effect of receiving multiple suspensions.

Cross correlations

We calculated Spearman's correlation coefficients for most of the variables included in the multilevel models⁴¹, to assess whether any collinear relationships existed⁴². Whether the child had been cautioned or sentenced for a non-serious violence offence had the highest correlation coefficient with the outcome variable (0.37). This was calculated for Model 1, however once sequencing of events was accounted for by the other models, this reduced considerably. When looking at pairwise comparisons amongst the pupil level explanatory variables, there were two pairs of variables that had a strong correlation coefficient (greater than absolute 0.6). These were: suspended and two or more suspensions (0.62 - 0.72), and EHC plan and special school (0.65 - 0.67). This suggests that these variables could be correlated, however these relationships are not unexpected. A full table of cross correlations and their associated p values can be found in the accompanying tables.

³⁶ Education, children's social care and offending (publishing.service.gov.uk)

³⁷ This is an analytical method which tells us if the difference in a finding at the population level (in this case all pupils) and the sample level (in this case serious violence offenders) is due to random chance (variation) in the data, or if it is a true difference between the two groups.

³⁸ Between 1 and 6. This is due to there being low proportions of the all-pupil cohort that had been suspended more than 6 times.

³⁹ This refers to an exact number of suspensions received, for example the difference between the proportion of children cautioned or sentenced for a serious violence offence, and the proportion of the all-pupil cohort, that had been suspended *exactly* 4 times, was statistically significant, and cumulative number of suspensions received, for example the difference between the proportion of children cautioned or sentenced for a serious violence offence, and the proportion of suspensions received, for example the difference between the proportion of children cautioned or sentenced for a serious violence offence, and the proportion of the all-pupil cohort, that had been suspended *at least* 4 times.

⁴⁰ We have used a 5% measure, meaning we can conclude that there is less than a 5% chance that the results are due to random chance alone.

⁴¹ Discrete variables with more than two categories with no natural ordering (eg: ethnicity) were not included.

⁴² Note that Spearman's correlation coefficients are a measure of a monotonic relationship, and therefore if the coefficient equals 0, it does not necessarily mean that there is no relationship between the two variables, it simply implies there is no *monotonic* relationship (ie: the relationship could be quadratic).

Limitations of the models

Model 1 refers to any instance of a child being identified with a characteristic, irrespective of when that characteristic was identified. This has implications for children cautioned or sentenced for a serious violence offence, as characteristics identified both before and after their serious violence offence are included, which increases the risk of endogeneity between certain characteristics and serious violence. For example, the introduction of the Legal Aid, Sentencing and Punishment of Offenders Act 2012 (LASPOA) meant that, from 3rd December 2012, children up to the age of 18 who are remanded to youth detention accommodation as a result of being charged with or convicted of an offence will be 'looked after' by the designated local authority⁴³. Therefore, a child may have become CLA due to being cautioned or sentenced for a serious violence offence. A comparison of the results between Models 1 and 2 could help to gauge the severity and direction of any bias caused by endogenous relationships between any of the explanatory variables and the outcome variable.

It is important to note that the measurement of local authority level characteristics in Models 1 and 2 do not account for characteristics prior to 2015/16 and therefore only depict a snapshot of the characteristics of the local authority the child has attended school in. This makes interpretation of the local authority level estimates in these models limited; however, we can still estimate which aspects of local authority level deprivation are significantly associated with serious violence.

Models 1 and 2 take no account of the number of serious violence offences that children in the data have been cautioned or sentenced for, only that they have ever been cautioned or sentenced for a serious violence offence. The years 7-11 models improve on this by identifying whether a child has been cautioned or sentenced for a serious violence offence within five separate 2-year periods, thereby accounting for whether a child has been cautioned or sentenced for a serious violence offence on more than one occasion, provided that those occasions spanned across multiple years. However, the years 7-11 models take no account of children being cautioned or sentenced for multiple serious violence offences within the given 2-year period.

The Central Limit Theorem dictates that a sample size greater than or equal to 30 is generally considered sufficient to assume a normal distribution of the sample means. Therefore, for Models 1 and 2 we have limited the data to only include local authorities that have 30 or more children that have ever been cautioned or sentenced for a serious violence offence. For the years 7-11 models, we have limited the data to only include local or sentenced for a serious sentenced for a serious violence offence that, in one year, have 30 or more children that have been cautioned or sentenced for a serious violence offence in the following two years.

⁴³ Please see here for more information: <u>Children looked after return 2022 to 2023: guide - GOV.UK</u> (www.gov.uk)

For Models 1 and 2, eight local authorities, and approximately 32,000 observations, are dropped from the analysis.

However, serious violence offending is a rare event, and peaks in the mid-teens. Therefore, there are fewer instances of serious violence offending when modelling younger year groups. In the year 7 model, only 22 out of 150 local authorities⁴⁴ and approximately 420,000 out of 1.53 million observations remain in the model, reducing the sample size by 72%. In the year 8 model, 58 out of 150 local authorities and approximately 823,000 out of 1.52 million observations remain in the model, reducing the sample size by 46% (see Table 5).

The problem is not as pronounced for the remaining models, however they are still being estimated using data that is 22-26% smaller than the original sample size. Caution should be taken when considering the results of these 5 models, *especially the year 7 model*, as they have been estimated using a reduced sample.

Given that we seek to understand how significant characteristics are at the local authority level in association with serious violence at the individual level, rather than to identify and target individual local authorities, the number of individual level observations is of greater weight than the number of local authorities included in the models. Therefore, although the models are estimated using reduced samples, the number of observations included is still deemed sufficient.

Table 5: Sample sizes	, including number	of observations a	Ind local	authorities,	for
the 7 multi-level regre	ssion models				

	Model 1	Model 2	Year 7	Year 8	Year 9	Year 10	Year 11
Local authorities	142	142	22	58	88	95	87
All pupils (limited sample)	1,498,603	1,498,603	419,536	822,696	1,138,658	1,190,533	1,114,727
All pupils (original sample)	1,530,943	1,530,943	1,525,224	1,524,149	1,523,321	1,520,369	1,514,045

⁴⁴ City of London, Isles of Scilly and BFPO Overseas Establishments have been removed from the analysis due to inconsistent reporting of local authority level characteristics.

Model performance

Adjusted R-squared

R-squared values in a linear regression measure the amount of variation in an outcome explained by a model (0 being none, 1 being all). As these models use a multilevel binary logistic regression approach there is not an equivalent R-squared measure, though adjusted pseudo r-squared measures are provided in Table 6 and provide adjusted r-squared measures for the fixed effects factors only, the random effects factors only and the model overall (including both fixed and random effects factors). Adjusted here means that the r-squared has been modified to account for the number of predictors in the model and reduce the optimism in the model produced by overfitting. The adjusted pseudo r-squared is generally used to compare goodness of fit for the *same model* with an increase indicating an improvement in model fit. Given that not all the multi-level models have been estimated using the exact same sample, the adjusted pseudo r-squared is not directly comparable across the models. It does however still give an indication as to how well the models explain the observed variance in the outcome measures.

factors a	and the mod	del in total,	for each mi	ulti-level reg	gression mo	odel	

Table 6: Adjusted r-squared measures for the fixed-effects factors, random-effects

R ²	Model 1	Model 2	Year 7	Year 8	Year 9	Year 10	Year 11
Fixed	0.21	0.09	0.03	0.04	0.05	0.06	0.04
Random	0.004	0.003	0.0002	0.002	0.003	0.002	0.001
Model	0.22	0.09	0.03	0.04	0.05	0.05	0.04

Variance inflation factors

Variance inflation factors (VIF) were calculated for each of the explanatory variables in the seven models to assess the severity of multicollinearity by measuring how much the regression coefficients have been inflated by multicollinearity. Generally, where a VIF is equal to 1, variables are assumed to not be correlated; where a VIF is greater than 5, variables are assumed to be highly correlated, and where a VIF is greater than 10, there is significant multicollinearity⁴⁵.

When considering the individual level variables, the VIFs that were estimated to be greater than 5 were relating to an individual explanatory variable and its interaction (AP and the interaction between SEN and AP, and less frequently, CLA and the interaction

⁴⁵ Variance Inflation Factor (VIF) - Spark By {Examples} (sparkbyexamples.com)

between SEN and CLA). This could be expected and structural multicollinearity in these instances can often have no adverse consequences⁴⁶. The variables highlighted as having strong correlation coefficients in the previous 'Cross correlations' section had VIFs that did not surpass 2.5 and therefore were not of concern.

However, when looking at AP and the interaction between SEN and AP, the VIFs were extremely high in the year 7 model; greater than 10 in the year 8 model and models 1 / 2; and greater than 5 in the years 9, 10 and 11 models. The VIFs calculated for these two variables in the year 7 model were of particular concern⁴⁷ due to being excessively large with a very low tolerance. However, the VIF for SEN on its own was well below 5. This suggests that there is a large overlap of SEN within AP pupils in younger years, (year 7 especially) however it cannot be said that the majority of SEN pupils attend AP in younger years. The Spearman's correlation coefficient between AP and SEN was estimated to be between 0.16 and 0.18 across the models, which suggests a weakly positive relationship. Given that Spearman's correlation coefficients are a measure of a monotonic relationship, this could be an indicator that the relationship between AP and SEN is non-linear.

When looking at CLA and the interaction between SEN and CLA, the VIFs were greater than 10 in the year 7 model and models 1 / 2, and VIFs greater than 5 in the year 8 model. The VIF for SEN on its own was well below 5. Again, these results suggest that there is an overlap of SEN within CLA in younger years, (year 7 especially) however it cannot be said that the majority of SEN pupils were also CLA in younger years. The Spearman's correlation coefficient between the two was estimated to be between 0.10 and 0.12.

Several of the between effects local authority area level variables generated VIFs between 5 and 10 in the year 7 model only⁴⁸. All VIFs can be found in the accompanying tables.

⁴⁶ When Can You Safely Ignore Multicollinearity? | Statistical Horizons

⁴⁷ Caution has been taken when interpreting any findings from the year 7 model due to low sample size and outcome detection.

⁴⁸ Caution has been taken when interpreting any findings from the year 7 model due to low sample size and outcome detection.

6. Definition of serious violence

The definition of serious violence used in this analysis is broadly based on the following categories of offence groups and offence types:

- Violence against the person offences, indictable only
- Robbery offences, indictable only
- Possession of weapons offences, triable either way or indictable only

However, analysts and policy officials in the DfE and MoJ assessed the full list of all offences to ensure that the list of offences included in the definition is as exhaustive as possible – this means that some offences that fit into the above categorisation have been removed from the definition of serious violence used for the analysis, whilst some offences which do not fit into the above categorisation have been defined as serious violence for the purpose of this work. The rationale applied for omitting or incorporating offences was challenged and applied consistently in all cases.

Homo Offico	
offence code	Offence description
00100	Violence against the person
00101	Murder of persons aged 1 year or over
00102	Murder of infants under 1 year of age
00200	Attempted murder
00301	Making threats to kill
00302	Conspiracy or soliciting, etc., to commit murder
00303	Assisting offender by impeding his apprehension or prosecution in a case of murder
00304	Intentionally encouraging or assisting commission of murder
00305	Encouraging or assisting in the commission of murder believing it will be committed
00306	Encouraging or assisting in the commission of one or more offences of murder believing one or more will be committed
00401	Manslaughter
00402	Infanticide
00403	Child destruction
00501	Wounding etc. with intent to do grievous bodily harm etc. or to resist apprehension
00504	Attempting to choke, suffocate etc. with intent to commit an indicta- ble offence (garrotting)
00505	Using chloroform, etc., to commit or assist in committing an indicta- ble offence
00506	Burning, maiming, etc. by explosion
00507	Causing, explosions or casting corrosive fluids with intent to do grievous bodily harm

Table 7: Home Office offence codes used to define serious violence

Home Office	
offence code	Offence description
00509	Placing, etc. explosives in or near ships or buildings with intent to do bodily harm. etc.
00510	Endangering life or causing harm by administering poison
00513	Possession etc. of explosives with intent to endanger life
00514	Possession of firearms etc., with intent to endanger life (Group I)
00515	Possession of firearms etc. with intent to endanger life (Group II)
00516	Possession of firearms etc. with intent to endanger life (Group III)
00517	Using etc. firearms or imitation firearms with intent to resist arrest etc. (Group I)
00518	Using etc. firearms or imitation firearms with intent to resist arrest etc. (Group II)
00519	Using etc. firearms or imitation firearms with intent to resist arrest etc. (Group III)
00520	Use etc. of chemical weapons
00521	Use of premises or equipment for producing chemical weapons
00522	Use, threat of use, production or possession of a nuclear weapon
00527	Torture
00802	Administering poison with intent to injure or annoy
00804	Causing bodily harm by furious driving
00805	Assaults on person preserving wreck
00806	Assaults occasioning actual bodily harm
00833	Racially aggravated wounding or inflicting grievous bodily harm (in- flicting bodily injury with or without weapon)
00840	Religiously aggravated malicious wounding or grievous bodily harm
00846	Racially or religiously aggravated malicious wounding or grievous bodily harm (GBH)
00859	Racially or religiously aggravated wounding or grievous bodily harm
03401	Robbery
03402	Assault with intent to rob
03410	Robbery
02802	Burglary in a dwelling with intent to inflict grievous bodily harm - in- dictable only
05601	Arson endangering life
00803	Setting spring guns etc. to injure trespassers
00811	Possession of offensive weapons without lawful authority or reason- able excuse
00813	Possessing firearm or imitation firearm at time of committing or be- ing arrested for an offence specified in Schedule 1 of the Act (Group I)
00814	Possessing firearm or imitation firearm at time of committing or be- ing arrested for an offence specified in Schedule 1 of the Act (Group II)
00815	Possessing firearm or imitation firearm at time of committing or be- ing arrested for an offence specified in Schedule 1 of the Act (Group III)

Home Office	
offence code	Offence description
00816	Possessing firearm or imitation firearm with intent to commit an in-
	dictable offence or resist arrest etc (Group I)
00817	Possessing firearm or imitation firearm with intent to commit an in-
	dictable offence, or resist arrest etc (Group II)
00818	Possessing firearm or imitation firearm with intent to commit an in-
00000	dictable offence, or resist arrest etc (Group III)
00823	Possession of a firearm or imitation firearm, with intent to cause fear
00824	Possession of a firearm or imitation firearm with intent to cause fear
00024	of violence
	(Group II)
00825	Possession of a firearm or imitation firearm with intent to cause fear
	of violence (Group III)
00826	Having an article with a blade or point in a public place
00827	Having an article with a blade or point on school premises
00828	Possession of offensive weapons without lawful authority or reason-
	able excuse on school premises
00853	Using another to look after, hide or transport a dangerous weapon -
	offensive weapon, knife or bladed weapon
00854	Using another to look after, hide or transport a dangerous weapon -
	a firearm
00861	I hreaten with an offensive weapon in a public place
00862	Threaten with a blade or sharply pointed article on school premises
00863	Threaten with an offensive weapon on school premises
00864	Threaten with blade/sharply pointed article in a public place
08101	Possession of weapons
08103	Possessing etc. firearm or ammunition without firearm certificate
00404	(Group I)
08104	Possessing etc. shotgun without certificate (Group II)
08107	Trading in firearms without being registered as a firearms dealer
00400	(Group I) Tradium in finanzas with suit hainer namistanad as a finanzas da slan
08108	Calling in lifearms without being registered as a lifearms dealer
08109	Selling lifearm to person without a certificate (Group I)
08110	Selling firearm to person without a certificate (Group II)
08111	Repairing, testing etc. firearm for person without a certificate (Group
00110	1) Repairing testing at firearm for parean without a partificate (Croup
00112	Repairing, testing etc. meann for person without a certificate (Group
08113	Falsifying certificate etc. with a view to acquisition of firearm (Group
08114	Falsifying certificate etc. with a view to acquisition of firearm (Group
08115	Shortening a shotgun or other smooth bore gun (Group I)
08116	Conversion of firearms (Group I)

Home Office	
offence code	Offence description
08117	Possessing or distributing prohibited weapons or ammunition (Group
09126	1) Corruing firearm in public place etc. (Croup I)
00120	Carrying loaded firearm in public place etc. (Group II)
00127	Tracessing with freerm or imitation freerm in a building (Croup I)
00129	Trespassing with firearm or imitation firearm in a building (Group I)
00130	Decessing with means of finances by persons provide the convicted of evines
08135	(Group I)
08136	Possession of firearms by persons previously convicted of crime (Group II)
08137	Possession of firearms by persons previously convicted of crime (Group III)
08138	Supplying firearms to person denied them under Section 21 (Group I)
08139	Supplying firearms to person denied them under Section 21 (Group II)
08140	Supplying firearms to person denied them under Section 21 (Group III)
08142	Failure to transfer firearms or ammunition in person (Group I)
08143	Failure to give notice in writing to the Chief Officer of Police of trans- fers involving firearms (Group I)
08144	Failure by certificate holder to notify in writing Chief Officer of Police of deactivation, destruction or loss of firearms or ammunition (Group I)
08145	Failure by certificate holder to notify in writing Chief Officer of Police of events taking place outside Great Britain involving firearms and ammunition (sold or otherwise disposed of, lost etc) (Group I)
08169	Possession of weapons
08170	Possessing or distributing prohibited weapons designed for dis- charge of noxious liquid etc. (Group I)
08171	Possessing or distributing firearm disguised as other object (Group I)
08172	Possessing or distributing other prohibited weapons
08173	Offence in relation to the unlawful IMPORTATION of any weapon or ammunition of a kind mentioned in S.5(1)(a),(ab),(aba),(ac), (ad),(ae),(af) or (c) of the Firearms Act 1968
08174	Offence in relation to the unlawful EXPORTATION of any weapon or ammunition of a kind mentioned in S.5(1)(a) (ab),(aba),(ac),(ad),(ae), (af) or (c) of the Firearms Act 1968
08176	Selling or transferring an air weapon unlawfully
08177	Carrying a loaded or unloaded or imitation firearm or air weapon in public place
08178	Knowingly being concerned in activity prohibited by Parts 2, 3 or 4 of the Order with intent to evade the relevant prohibition
08179	Unship / unload prohibited weapon / ammunition with intent to evade prohibition / restriction

Home Office	
offence code	Offence description
08180	Remove prohibited weapons / ammunition from their place of impor-
	tation with intent to evade prohibition / restriction
08181	Import prohibited weapons / ammunition with intent to evade a prohi- bition / restriction
08182	Export prohibited weapon / ammunition with intent to evade prohibi- tion / restriction
08183	Carry / remove / deposit etc. prohibited weapons / ammunition with intent to evade a prohibition / restriction
08184	Knowingly concerned in fraudulent evasion of prohibition / restriction on prohibited weapon / ammunition
08185	Manufacture weapon / ammunition specified in section 5(1) of the Firearms Act 1968
08186	Sell / transfer prohibited weapon / ammunition
08187	Possess prohibited weapon / ammunition for sale / transfer
08188	Purchase / acquire prohibited weapon / ammunition for sale / trans- fer
08189	Offences under Explosives Precursors Regulations 2014
08190	Manufacture an offensive weapon; Possess article for use in con- nection with conversion of imitation firearm
08191	Make / sell / give as gift defectively deactivated weapon - Police and Crime Act 2017
09001	Unlawful marketing of knives (selling or hiring)
09002	Unlawful marketing of knives (offering or exposing to sell or hire)
09003	Unlawful marketing of knives - having in possession for the purpose of sale or hire
09004	Publication of any written, pictorial or other material in connection with the marketing of any knife - the material suggesting or indicat- ing knife suitable for combat
09005	Publication of any written, pictorial or other material in connection with the marketing of any knife - the material is otherwise likely to stimulate or encourage violent behaviour involving use of the knife as a weapon
05914	Manufacture, possession or control of explosives under suspicious circumstances
05915	Possessing or making an explosive substance, a noxious or danger- ous thing, a machine, engine or instrument with intent to commit an offence under the Offences against the Person Act 1861
06906	Unauthorised possession in prison of knife or offensive weapon



© Department for Education copyright 2023

For any enquiries regarding this publication, contact us at: <u>sv.RESEARCH@education.gov.uk</u>This document is available for download at <u>www.gov.uk/government/publications</u>