SECONDARY DATA ANALYSIS REPORT

Association between school exclusion, suspension, absence, and violent crime

Jasmine Rollings, Rosie Cornish, Iain Brennan and Alison Teyhan

May 2025

About the Youth Endowment Fund

The Youth Endowment Fund (YEF) is a charity with a mission that matters. We exist to prevent children and young people from becoming involved in violence. We do this by finding out what works and building a movement to put this knowledge into practice.

Children and young people at risk of becoming involved in violence deserve services that give them the best chance of a positive future. To make sure that happens, we'll fund promising projects and then use the very best evaluation to find out what works. Just as we benefit from robust trials in medicine, young people deserve support grounded in the evidence. We'll build that knowledge through our various grant rounds and funding activities.

And just as important, is understanding children and young people's lives. Through our Youth Advisory Board and national network of peer researchers, we'll ensure they influence our work and that we understand and are addressing their needs. But none of this will make a difference if all we do is produce reports that stay on a shelf.

Together, we need to look at the evidence and agree on what works, then build a movement to make sure that young people get the very best support possible. Our strategy sets out how we'll do it. At its heart, it says that we will fund good work, find what works and work for change. You can read it here.

For more information about the YEF or this report, please contact:

Youth Endowment Fund
1st Floor
64 Great Eastern Street
London
EC2A 3QR

www.youthendowmentfund.org.uk

hello@youthendowmentfund.org.uk

Registered Charity Number: 1185413

Contents

Acknowledgements

About the research team

This independent report, funded by the YEF, has been produced by:

Dr Jasmine Rollings (JR) – Analyst (University of Bristol) – has a background in psychological research and has experience working in criminal justice and analysing large, complex datasets.

Dr Rosie Cornish (RC) – Lead Statistician (University of Bristol) – is a medical statistician with substantial experience working with large, complex datasets.

Professor Iain Brennan (IB) – Lead Substantive Expert (University of Hull) – has been undertaking and publishing research on perpetrators and victims of crime and on criminal justice interventions for 15 years, publishing over 40 articles on violence prevention.

Dr Alison Teyhan (AT) – Principal Investigator (University of Bristol) – is an epidemiologist who has worked with the Avon Longitudinal Study of Parents and Children (ALSPAC) study for over a decade, often using linked health and administrative data, including crime data.

Further acknowledgements

We are extremely grateful to all the families who take part in ALSPAC, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

The authors would like to sincerely thank the Avon & Somerset Police staff who supported the linkage between ALSPAC and the local police data, the ALSPAC Data Linkage Team at the University of Bristol for their assistance with this project, and the members of our project steering group for sharing their expertise and insight with us.

Executive summary

About the project

The aim of this research is to test the association between school absence, suspensions and exclusions and subsequent offending and violent behaviour. The existing evidence suggests being in education can reduce the risk that a child will be involved in crime and violence (Department for Education, 2023; Ullman et al., 2024; YEF, 2024b). Despite this, there are limits to what we know. Studies that exclusively rely on administrative data (school, care and offending records) cannot fully control for the range of individual, family, interpersonal and community-level drivers that may also explain the correlation between absence from education and offending. Few existing studies look at the relationships between absence and exclusions and later violence using both police-recorded and self-report crime and violence measures.

To build on this, the study described in this report uses linked data between the Avon Longitudinal Study of Parents and Children (ALSPAC), the National Pupil Database (NPD) and Avon & Somerset Police Data. ALSPAC is a cohort study that recruited pregnant women with a due date between April 1991 and December 1992, living in and around Bristol. Almost 15,000 women were recruited, resulting in a sample of 14,901 children alive at one year of age. These children and their parents have been tracked continuously into adulthood. Suspension, exclusion and absence data (from the NPD) have been matched to the children in ALSPAC who attended state-funded schools in Key Stage 4 (when participants were 14-16 years old).

The main outcomes of interest for this study are self-reported violent behaviour, as measured by questionnaires at ages 17 and 18 years, and police-recorded crime (total and violent offending separately), as measured in official police records in the 24-month period following the end of either Year 10 or Year 11. The research uses logistic regression, testing the relationship between each exposure and outcome pair (e.g. having a police record and having been suspended) and controlling for a range of individual, behavioural, family and school factors.

Key findings

Being suspended or excluded and being absent from school more than 20% of the time are strongly correlated with self-reported and police-recorded offending	36% of suspended or excluded children self-reported violence perpetration, compared to 8% of children who hadn't been suspended or excluded. For children absent 20% or more of the time, this was 24%, and for those absent less than 20% of the time, it was 9%. 21% of children who'd been suspended or excluded had a police record for any offence, compared to 3% who hadn't been suspended or excluded. For children who were absent for 20% or more school sessions, this was 15%, and it was 4% for those who absent for less than 20%.
These associations are still present once individual, household and education factors are accounted for	Initially, the analysis controlled for individual, household and education-related factors. Once these were accounted for, children suspended or excluded were 4.06 times more likely to report perpetrating violence compared to those who hadn't been. Those who were absent 20% or more of the time were 2.85 times more likely to report violent behaviour. For police-recorded offending, children who were suspended or excluded were 5.16 times more likely to be involved. Those absent 20% or more of the time were 3.10 times more likely.
After controlling for behavioural difficulties, absence, and suspension and exclusion are still	Further analysis included additional controls for behavioural difficulties, smoking behaviours and childhood traumas. Once these factors were accounted for, children suspended or excluded were 2.36 times more likely to report violent behaviour compared to those who hadn't been suspended or excluded. For those absent 20% or more of the time, it was 2.09 times more

associated with violence and police involvement. However, these findings are more uncertain	likely. Children who were suspended or excluded were 4.68 times more likely to have a police record. For those absent 20% or more of the time, it was 3.47 times more likely. Greater caution is needed with these specific findings, given the smaller sample in this part of the analysis due to missing data.
The associations are stronger for involvement in more serious forms of crime and violence	After controlling for individual, household and education factors, children who had been suspended or excluded were 8.05 times more likely to have a police record for serious violence. This compares to 5.16 times more likely for any police record and 4.06 times more likely to self-report violence perpetration. Children absent 20% or more of the time were 4.79 times more likely to have a police record for serious violence, 3.10 times more likely to have a police record for any offence and 2.85 times more likely to self-report violence perpetration.

Interpretation and implications

This study shows that children who are suspended or excluded or absent for 20% or more of the time are more likely to become involved in violence and offending. This is a well-established finding supported by other evidence. However, what has previously been harder to establish is whether the act of being absent or suspended or excluded is what directly causes later involvement in crime or violence or whether it is just that excluded, suspended or absent children have other difficulties (such as behavioural problems) that make them more likely to offend. While other studies have attempted to control for other factors, this study goes further in controlling for a wider array of risk and contextual factors, including behavioural difficulties. Due to the nature of the study and the limitations around sample size, we cannot conclude that suspensions, exclusions and absences have a causal impact on later involvement in crime and violence. However, this study does provide a strong indication that suspensions, exclusions and absences are key risk factors for later involvement in violence, even once an array of other contextual and risk factors are accounted for.

There are limitations in this study. The size of the sample is too small to explore the relative impact of suspension compared to exclusions or how individual and family factors interact. The dataset exclusively covers children living in and around Bristol and spans education and justice outcomes in the mid-late 2000s. In addition, the study only focuses on absences and suspensions or exclusions in Years 10 and 11. Caution should, therefore, be taken when extrapolating these findings beyond this context. Furthermore, only 5% of the ALSPAC sample is made up of participants from ethnic minority backgrounds. While this aligns with the demographics of the study area at the time of recruitment, this means there were too few participants from these backgrounds to report findings by ethnicity.

Deciding whether to suspend or exclude a child is a difficult decision for a headteacher to make, as they need to balance the needs of individual children with the needs of the wider school community. In addition, schools and policymakers are currently making a concerted effort to improve attendance (following the post-COVID rise in school absence). However, what these findings indicate is that policymakers and schools should continue to target support towards children at risk of suspension and exclusion (attempting to reduce the need for suspension and exclusion), while high-quality support must be provided to children already suspended or excluded and to those who are absent from school. Targeting these children with evidence-based attendance and behaviour improvement interventions and violence reduction strategies could reduce their risk of later involvement in violence and offending.

Introduction

Rates of school absence and school disciplinary actions have been a topic of significant discussion and policy focus in recent years (Long and Roberts, 2024). The importance of these factors has been amplified since the COVID pandemic, which appears to have had a lasting effect on school engagement (Gibbons, McNally and Montebruno, 2024). Another topic that has been at the centre of social policy debate is rates of youth violence and knife crime, which, after reaching historically low levels ten years ago, has seen concerning increases across some measures (Allen, Carthew and Zayed, 2023). In addition to being important social problems independently, these two phenomena both tend to peak in mid-to-late adolescence. Understanding how they are related could provide valuable insight into how one or both could be better understood and prevented in school and community settings.

Weaker academic performance, special educational needs and disabilities and breaches of school disciplinary policy are correlated with later criminal justice outcomes (Department for Education (DfE), 2022). Strong academic performance is a protective factor against violent behaviour, while underperformance (Commission on Young Lives, 2022), special education needs status (ibid.), exclusion (McAra and McVie, 2010; DfE, 2022), persistent absence (DfE, 2022), having under-performing or delinquent peers (Miller, 2009) and area-level inequalities (Ejlskov et al., 2023) are correlated with violence and involvement with the criminal justice system. Of these risk factors for violence, absence and suspensions/exclusions are of particular interest because they are, in theory, modifiable (e.g. via changes in school actions and policies – as observed in Scotland (Scottish Government, 2024) – or through school-based interventions (Gaffney, Farrington and White, 2021; Mielke and Farrington, 2021). Further, both absence (Long and Roberts, 2024) and suspension (DfE, 2024a) rates are currently rising in the UK, particularly in secondary schools.

Understanding the relationship between absence from school – in the form of authorised (e.g. illness), unauthorised (e.g. truancy) or imposed absence (e.g. due to suspension or exclusion) – and violence is a pressing issue. There are beliefs among some policymakers and researchers that exclusions, in particular, are a risk factor for later offending (DfE, 2019; Gerlinger et al., 2021; Youth Endowment Fund, 2024). On this basis, some school areas have moved to significantly reduce exclusions (Southwark Council, 2022). While there are many harms associated with absence or suspensions/exclusions (Hirschfield, 2018), the decision to suspend or exclude a child requires a school to carefully balance the needs of the child and the wider school community (DfE, 2024b), and misconstruing the nature of the link between them and violence risks stigmatising those children and distracting attention from the underlying causes of violence.

In this work, we sought to estimate the absence/suspension/exclusion–violence connections in a way that would potentially guide the timing and context of violence prevention programmes for the future.

The association between absence, exclusion and offending

A review of 40 US studies found evidence that exposure to school exclusion or suspension was associated with an approximate doubling of the risk of delinquency (included studies measured a range of different outcomes, including criminal justice system contact, violent and non-violent offending, as well as antisocial and other behaviours; Gerlinger et al., 2021). Further, in the US, school exclusion – temporary or permanent, on-campus or off-campus – is seen as contributing to a 'school-to-prison pipeline' (Hirschfield, 2018) whereby a variety of policies, including the use of exclusion and formal methods of control (such as police in schools), are hypothesised to increase the risk of involvement in the criminal justice system (Cuellar and Markotvitz, 2015). This association has been shown to be more pronounced in people from particular groups, such as minoritised communities (Barnes and Motz, 2018) and students with disabilities and other vulnerabilities (Mallett, 2016). On the other hand, a US school-level analysis that modelled the relationship between disciplinary policy severity[1] and self-reported offending in adulthood via a longitudinal cohort study found that, in general, school disciplinary policies were only weakly associated with offending patterns (Matjasko, 2011). This observation is important, as it emphasises uncertainty over whether the school-to-prison pipeline is a causal mechanism or an explanation for observed disproportionality in who is subject to both school disciplinary procedures and criminal justice involvement. Resolving this is an important challenge for educational policy.

In the UK, several studies using different approaches have found individual-level associations in the same direction as those in the US, although of different magnitudes. Analysis by the DfE of approximately 1.63 million pupils found that exclusion from school, either temporarily (suspension) or permanently, is a relatively common event in the lives of teenagers convicted of serious violence: 15% of individuals who had been convicted or cautioned for a serious violent offence were permanently excluded from school during Key Stage 4, and 82% were suspended (DfE, 2022). For comparison, the rate of permanent exclusion across all individuals was 1%, and the rate of suspension was 15%, yielding unadjusted relative risks (RR)[2] of 15.0 (95% confidence interval (CI): 14.44-15.58) for permanent exclusion and 5.47 (95% CI: 5.42-5.51) for suspension. A subsequent multilevel, multivariable analysis using the same data and also conducted by the DfE found that being suspended across any year in secondary school, including Years 10 and 11, was associated with an increased risk of being convicted or cautioned for serious violence in the following two academic years, with odds ratios (ORs)

[1] The authors created a scale of readiness to use formal sanctions for disciplinary breaches. Schools that used formal sanctions for less serious breaches were rated as having more severe policies.

[2] Calculated with the following formula: $RR = \frac{Percentage\ in\ exposed\ group}{Percentage\ in\ unexposed\ group}$

ranging from 3.4 to 5.2, depending on the school year of exposure (DfE, 2023). In this same analysis, the ORs for permanent exclusion were smaller (ranging from 0.69 to 1.93). However, these results should be interpreted with caution as the models included all putative risk factors in the same model, including non-violent offending and attendance at a pupil referral unit, regardless of their temporal sequence, making the findings difficult to interpret. The analysis of the Edinburgh Study of Youth Transitions and Crime used linked police records and found a modest association between school exclusion and offending, albeit over a shorter period: OR = 1.5; 95% CI: 1.1-2.3 (McAra and McVie, 2012). Finally, a UK study using academisation (shown to result in an increased exclusion rate) as an instrumental variable in an attempt to estimate the causal effect of exclusion (including both permanent exclusions and suspensions) found that it was associated with a 33% greater risk of a student receiving a custodial sentence in the following two years (Cathro, Tagliaferri and Sutherland, 2023).

The first synthesis of research on the relationship between school attendance/absence and youth offending was published almost a century ago (Williams, 1927). The conclusions of that synthesis, replicated frequently over the next hundred years, were that mental health problems, such as anxiety, depression and conduct disorder (Farrington, 1980); personality (Farrington, 1996); peers (Cohen, 1955); families (Farrington, 1996); and socioeconomic factors (Klein, Socu and Dare, 2020) were all associated with increased absence from school in different groups: they have also all been found to predict offending, suggesting that any potential relationship between school attendance and offending could be explained by common causes.

Consistent with the international evidence, for many children in the UK who go on to be involved in violence, persistent absence is a common feature of their school record. In the 2012/13-2014/15 cohort of Key Stage 4 children in England, the unadjusted RR of subsequently being convicted or cautioned for serious violence was 1.93 (95% CI: 1.92-1.94) for those who were persistently absent compared to those who were not (DfE, 2022). When the reason for that absence was unauthorised other, which would include truancy, this unadjusted RR rose to 5.29 (95% CI: 5.18-5.39). In the subsequent multilevel analysis, persistent (unauthorised other) absence during Years 7 to 11 was associated with an increased risk of being convicted or cautioned for serious violence in the following two academic years, with ORs ranging from 1.1 to 2.0 (DfE, 2023). Again, these results are difficult to interpret due to the fact that the models included many possible risk factors, some of which could have occurred after the absence. Beyond studies using administrative data alone, two UK cohort studies have demonstrated an adverse association between school absence and offending (Rocque et al., 2017; McAra and McVie, 2005).

Weaknesses in the existing literature

The majority of research demonstrating these associations, reviewed comprehensively by Hirschfield, was derived from US self-report longitudinal surveys, which often lacked detailed information about school performance and suffered from selection and attrition bias. In

addition, their findings were most relevant to US education policy and may not be generalisable to the UK. However, as indicated above, a major limitation of the evidence as a whole, including that from the UK, is the potential for confounding. A vast array of other potentially contributing factors in the lives of young people have often not been available. There are several reasons for this: in some instances, the dataset involved primary data collection from individuals (or guardians/teachers) and factors such as school performance and special educational needs could not be verified; in others, the studies were based entirely on administrative data, so they lacked important personal information or data on family, interpersonal and community-level exposures that are likely to play a significant role in patterns of absence/exclusion and of violence. Finally, some factors – such as personality traits or peer relationships – are difficult to capture. There is no doubt that the strong associations observed in some studies are, at least partially, a result of confounding factors. For example, in the DfE analysis mentioned previously, perpetrating violence against another child or against an adult in the school was the reason for 18% of suspensions and 26% of exclusions from school in England in 2022/23 (DfE, 2022) – in other words, violent behaviour preceded the exclusion or suspension.

Theoretical explanations

As stated above, the observed associations could be explained by confounding factors. However, several plausible theories exist to explain why absence could be a cause of offending. These include routine activities theory (i.e. time spent outside of school creates more opportunity to commit crime; Cohen and Felson, 1979), social control (i.e. spending unauthorised time out of school with little supervision could result in offending), differential association (i.e. outside of school, a child's peer group may change to a more criminogenic one; Warr, 2005) and strain (i.e. absence from school will impact educational attainment, which harms the potential for the legitimate development of status, making illegitimate means more attractive; Agnew, 1992), among others. The pathway from absence from school to violence could be due to increased exposure to criminal opportunities for these children, as they do not have the daily routine, structure, social rules and protection provided by the school environment. School absence also hinders academic progress, which in the long term can exclude these young people from further educational and employment opportunities – again, potentially increasing their exposure to criminal opportunities and violence. Many of these explanations for a possible causal link between absence and violence are shared with the purported exclusion–violence link, although the mechanism by which this plays out may be different. In addition, the 'label' of being excluded may play a role in shaping an excluded child's identity as being deviant, which could be self-fulfilling.

Triangulation of research evidence

The large-scale linkage of administrative data in the UK is a welcome addition to the research landscape. Through the linkage of the National Pupil Database (NPD) and other data assets held by the DfE to the Police National Computer (PNC) database held by the Ministry of Justice

(MoJ), researchers can identify links between educational experience and criminal justice outcomes for almost 15 million people who entered state school from 1990 onwards. This allows an unprecedented ability to identify rare exposures and outcomes and to rigorously assess the association between them. A recent study used this linkage to compare serious violent offending in over 20,000 people who were permanently excluded to that of a matched sample with similar characteristics and who followed similar school disciplinary trajectories but who were not excluded (Cornish and Brennan, 2024). This study found that exclusion was associated with a doubling of risk of perpetrated serious violence within one year. However, despite the robustness of the study design, the model was limited to the use of official data about school experiences. It contained no data on individual experiences prior to or concurrent with their school careers, family factors or neighbourhood factors that could explain both the exclusion and offending, although a key strength of the study was the ability to control for prior, non-violent offending. In addition, the outcome, serious violence, was limited to that detected by police. While sensitivity analysis found that an unmeasured confounder would have to be very strongly associated with both exclusion and offending to neutralise the association, the authors suggested that factors such as parental involvement in education – not measured in administrative data but a noted important factor in educational success (Hill and Tyson, 2009; Wilder, 2014) – could explain the exclusion–violence link. One potential solution to this shortcoming is to triangulate evidence using different datasets. Although a far smaller sample with fewer events, the ALSPAC–police data linkage presented below includes a wide range of measures on individual experiences prior to and concurrent with school careers, family factors and neighbourhood factors that are not present in administrative data, as well as having both official and self-report measures of violence. This allows us, in part, to fill gaps in knowledge and to provide greater insight into the potential influence of confounders, such as parental involvement in education, and offer consensus or dissent from different perspectives, either of which moves us closer to understanding how school experiences and offending are connected.

Research aims

The aim of this research is to examine the association between school absence, suspensions and exclusions at ages 14-16 years and subsequent violent behaviour – both police-recorded and self-reported.

Research questions

Our four primary research questions are:

(1) What is the association between persistent school absence and self-reported violent behaviour?
(2) What is the association between persistent school absence and official sanction for any offence and for serious violence offences?

(3) What is the association between school suspension/exclusion and self-reported violent behaviour?
(4) What is the association between school suspension/exclusion and official sanction for any offences and for serious violence offences?

Hypotheses

Based on findings from previous studies, we expected the children in our study sample who are persistently absent and/or suspended/excluded from school in Key Stage 4 (Year 10 and Year 11, ages 14-16) to be at increased risk of violent behaviour in late adolescence/early adulthood. However, any observed relationship between absence/exclusion and violence could be a result of confounding. We did not have a preferred *a priori* hypothesis as to whether the relationships between school absence or suspension/exclusion and violence are causal or whether they result from confounding – this is what we aimed to investigate in our project. Potential confounders that we considered include deprivation (both family and neighbourhood), poor engagement with education and low attainment, substance misuse and other behavioural factors, and adverse childhood experiences.

In this report, we examined the 24 months following the end of the exposure period for each participant in order to study short-term associations between suspensions/absence and the outcomes.

Key concepts

We make frequent reference throughout this report to school years and Key Stages. These education-related terms are explained in this section and summarised in Table 2. Table 3 then summarises other key concepts relevant to this study.

Schooling of the Avon Longitudinal Study of Parents and Children cohort

In England, children usually enrol in primary school (Reception class) in the September following their fourth birthday and then move up a school year every September. Pupils attend primary schools from Reception to Year 6 and then enrol in secondary school in Year 7. The national curriculum is organised into a number of Key Stages (Table 2). Key Stage 3 (KS3) and Key Stage 4 (KS4) make up the compulsory component of secondary education. At the end of each Key Stage, a child's progress is assessed by compulsory national tests (national tests at the end of KS3 have now been discontinued but were still in place at the time the ALSPAC participants were in KS3). General Certificate of School Education (GCSE) exams are taken in several subjects by pupils at the end of KS4.

Table 1. Structure of the national curriculum in England (for pupils aged 4-16 years)[3]

Class	Age (years)	Key Stage	Tests
Reception	4-5	Foundation	Local entry assessments
Year 1	5-6	1	-
Year 2	6-7		National tests and tasks in English and maths
Year 3	7-8	2	-
Year 4	8-9		-
Year 5	9-10		-
Year 6	10-11		National tests in English, maths and science (science no longer included)
Year 7	11-12	3	-
Year 8	12-13		-
Year 9	13-14		National tests in English, maths and science (now discontinued)
Year 10	14-15	4	Some children take GCSEs
Year 11	15-16		Most children take GCSEs or other qualifications

Source: ALSPAC documentation

Table 2. Key concepts and definitions

Concept	Definition
Self-reported violence	Self-reported violence was defined as saying yes to either or both of the following: 'In the past 12 months, have you (i) hit/kicked/punched someone else on purpose with the intention of really hurting them or (ii) carried a knife or other weapon for protection or in case it was needed in a fight?' at 17.5 and/or 18.5 years of age.
Official sanction for any offence	The Avon and Somerset Police (A&SP) data linked to ALSPAC includes records of charges, cautions and other out of court disposals for crimes committed in Avon and Somerset. From these data, we studied the records related to any offence that occurred in the 24-month study period.

[3] The curriculum has changed since ALSPAC participants were in school. Of note, Year 9 Standard Assessment Tests (SATs) were discontinued in England in 2008, and from 2009, Year 6 SATs no longer included the science component.

Official sanction for serious violence	Serious violence is defined by the Home Office as (i) violence against the person, indictable only, (ii) robbery, indictable only, and (iii) possession of weapons, triable either way or indictable only. The A&SP data linked to ALSPAC includes Home Office offence codes for each offence. From this, we identified which records from the 24-month study period related to serious violence.
School session	School days are split into two sessions (morning and afternoon). Schools must take an attendance register for each session. Local authority–maintained schools must meet for a minimum of 380 sessions or 190 days (approximately 39 weeks) during any school year as set out in the Education (School Day and School Year) (England) Regulations 1999. There are six half-terms in a school year, and prior to 2012/13, reporting on attendance was limited to five out of six half-terms, with the second half of the summer term excluded.
Persistent school absence	Absence can be authorised by the school (e.g. if a child is unwell or has a medical appointment) or unauthorised (e.g. unexplained or unjustified absences or arrivals after registration has closed). Persistent absence is currently defined as being absent for ≥10% of school sessions in an academic year; however, at the time the ALSPAC cohort was in school, the threshold was ≥20%; therefore, this was the threshold used in this report.
Suspension	A suspension is when a pupil is excluded from school for a specific period of time. A pupil can be suspended for a maximum of 45 days in a single academic year. The NPD holds data on suspensions, and these have been linked to ALSPAC for the academic years that the cohort was in Year 10 (for 80% of the sample) and 11 (KS4). If a participant had a suspension in KS4, they had one or more records in the linked NPD suspension data (one record per suspension episode). The NPD also holds details on the reason for the suspension and the duration of each suspension episode (the number of sessions that each suspension lasted).
Permanent exclusion	A permanent exclusion is the most serious sanction a school can give a pupil. This means that the pupil can no longer attend the school and will be removed from the school roll. The NPD data on permanent exclusions have been linked to ALSPAC for the academic years that the cohort was in Year 10 (80% of the sample) and Year 11. Details are provided on the reason for permanent exclusion but not on what happened to the student following exclusion (for example, if they subsequently attended a new school or alternative provision).
Exposure period (data availability)	ALSPAC linked to absence and exclusions data for the academic years 2006/07 to 2008/09. The exposure period was the academic year from which absence and suspension data were available for each participant. For approximately 80% of the sample, these data were available for Year 10; for

	the remainder, it was Year 11 (the latter comprises those individuals who started KS4 in the academic year 2005/06).
Study period	The study period in this report was the 24 months from the 1st of August of the year in which the individual completed either Year 10 or Year 11, depending on which year was considered their exposure period (defined above).

Ethics

Ethical approval for ALSPAC was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for the use of questionnaires and clinic data was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time. When the study children turned 18, they were sent fair processing materials describing ALSPAC's intended use of their health and administrative records and were given a clear means to object. Data were not extracted for participants who objected or who did not receive fair processing materials. Linkage to criminal records was acceptable to the vast majority of ALSPAC participants, including many with criminal records (Boyd et al., 2022). This specific project was approved by the ALSPAC Executive (Reference: B4443).

Project team/stakeholders

Dr Alison Teyhan (University of Bristol) is the principal investigator and is responsible for managing the project. Alison has been involved in all elements of this project, including the research design, funding acquisition, advisory board engagement, ALSPAC executive approval application, project management, recruitment of JR supervision of JR, dataset curation and contribution to this report.

Dr Rosie Cornish (University of Bristol) is the lead statistician on the project. Rosie has been involved in the research design, funding acquisition, recruitment of JR, supervision of JR and contribution to this report.

Professor Iain Brennan (University of Hull) is the lead substantive expert on the project. Iain has been involved with the research design, funding acquisition and contribution to this report.

Dr Jasmine Rollings (University of Bristol) is the analyst for this project. Jasmine is responsible for data cleaning and preparation, data analysis and the initial drafting of the interim and final reports.

The UK Medical Research Council and Wellcome Trust (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. Linkage of ALSPAC to administrative data was supported by the Project to Enhance ALSPAC through Record Linkage (PEARL), a

programme of research funded by the Wellcome Trust (WT086118/Z/08/Z). Data on antisocial behaviours, including self-reported violence, at the Teen Focus 4 Clinic was supported by the Medical Research Council (G0800612/86812). This publication is the work of the authors, and JR, RC, IB and AT will serve as guarantors for the contents of this paper.

The authors have no competing interests to declare.

Advisory board

Marvin Riley, Co-Principal at a secondary school in Greenwich, South East London; Laurelle Brown, CEO of an equity and inclusion consultancy and Race Equity Advisor to the Youth Endowment Fund; Dr Matteo Sandi, a Research Economist at London School of Economics; Professor Alex Sutherland, an education and crime expert at the University of Oxford; and William Teager, an education and violence expert, former employee of the DfE and current Head of Data and Insight at Youth Endowment Fund, served as the advisory board for this project. All members of the advisory board took part in a hybrid workshop in February 2023 and contributed to discussions regarding the research questions and factors that may contribute to school absence, suspensions, exclusions and violent behaviour.

Study design

Overview of research design

Table 3: Research design

Research design		Prospective cohort study
Dataset(s) used		ALSPAC, NPD, A&SP data
Population of interest		ALSPAC participants were born in 1991/1992 to mothers who lived in a defined area in and around the city of Bristol. Our samples were restricted to those for whom we had permission to link to their education and crime records.
Size of sample population		In the initial ALSPAC cohort, there were 14,062 live births and 13,988 children alive at one year; this subsequently increased to 14,901 alive at one year with further recruitment of eligible children from age 7 years onwards.
Primary outcomes	**Variables**	Police record for any offence Police record for a serious violence offence
	Measure	Having a police record for any offence between the end of the exposure period and 24 months later, coded as 0 for having no offence and 1 as having an offence. Having a police record for a serious violence offence between the end of the exposure period and 24 months later, coded as 0 for having no serious violence offence and 1 as having a serious violence offence. From A&SP data.
	Variable	Self-reported violence
	Measure(s)	Having self-reported violence at either or both time points (17.5 or 18.5 years) based on the following questionnaire items: How often in the last year have you '…hit, kicked or punched someone else on purpose with the intention of really hurting them' or '…carried a knife or other weapon with you for protection or in case it was needed in a fight'? Coded as 0 if the answer was no to both and 1 if the answer was yes to either or both.
Main method to be used or tested		Logistic regression modelling (multilevel with clustering within schools or using robust standard errors to account for clustering within families)

Data sources

Dataset 1: Avon Longitudinal Study of Parents and Children

Name of dataset	Avon Longitudinal Study of Parents and Children (ALSPAC)
Data owner(s)	University of Bristol
Type of data	Birth cohort study
Population/geographic coverage or sampling frame	ALSPAC recruited pregnant women who had an expected due date between April 1991 and December 1992 and who lived in a defined area in and around the city of Bristol, UK. ALSPAC recruitment took place in the old administrative county of Avon, UK. The catchment area covered the three health administration districts within the South-West Regional Health Authority that became the Bristol & District Health Authority. This area includes the City of Bristol and the surrounding urban and rural areas. Of those initially recruited during pregnancy (core participants), there were 14,062 live births and 13,988 children alive at one year. The initial cohort was increased to 14,901 alive at one year through further recruitment from age 7 years onwards.
Years covered or survey waves	1991-present
Exclusion criteria	Participants not alive at one year. Participants who were not a singleton or a twin (triplets and quadruplets are excluded). Participants who do not have KS4 NPD data. For police-recorded offending analyses only: Participants who had opted out of linkage to crime records or who never had the opportunity to do so (e.g. if their address was unknown). Participants who did not live in the Avon and Somerset area on their 16th, 17th and 18th birthdays.
Expected population/sample size (following exclusion criteria)	Total number of individuals in the final analytical sample: 10,679 (self-reported outcome sample); 3,284 (complete case) 9,640 (police-recorded outcomes sample); 6,436 (complete case)
Documentation	Details of the ALSPAC study have been published (Boyd et al., 2013), and a searchable data dictionary is available (www.bristol.ac.uk/alspac).

Dataset 2: The National Pupil Database

Name of dataset	The National Pupil Database (NPD)
Data owner(s)	The DfE
Type of data	Administrative data
Population/geographic coverage or sampling frame	The NPD is a central repository of education data for children attending state schools in England.
Years covered or survey waves	This project used primarily KS4 data from the NPD. The ALSPAC sample spans three academic years. Participants started KS4 (i.e. started Year 10) in 2005, 2006 and 2007 and completed KS4 (i.e. reached the end of Year 11) in 2007, 2008 and 2009, respectively.
Documentation	https://www.find-npd-data.education.gov.uk/categories

Dataset 3: Avon and Somerset Police Data

Name of dataset	Avon and Somerset Police (A&SP) data
Data owner(s)	A&SP
Type of data	Administrative data
Population/geographic coverage or sampling frame	The A&SP data includes records of charges, cautions and other out of court disposals for crimes committed in Avon and Somerset (a geographical area that includes the original ALSPAC recruitment area).
Years covered or survey waves	Data are available from 2007-2021.
Documentation	A data note describing the linkage of ALSPAC to A&SP data has been published (Teyhan et al., 2023).

Linking datasets

The linkage between ALSPAC, A&SP data and NPD data had already been established. Some participants attended more than one secondary school. For the school-level data, IDs exist, which allowed participants to be matched to the correct school's data at the correct time point (e.g. the KS4 ALSPAC school ID variable identifies which school they attended at the end of KS4).

Data protection

The data for this project were provided by an ALSPAC data manager and included ALSPAC clinic and questionnaire variables, education variables (from the NPD) and crime variables (from A&SP). The variables were prepared for analysis and combined into one dataset.

Therefore, the rest of this section refers only to this one dataset and not the three datasets listed above.

1. Data protection

ALSPAC adheres to the principles of the Office for National Statistics' (ONS) Five Safes framework (safe data, safe projects, safe people, safe settings and safe outputs).

Safe data

In common with all projects using ALSPAC data:

(1) The dataset was minimised to only include variables necessary for this project.

(2) The data were de-identified.

(3) The ID variable in the dataset was unique to this project, meaning it could not be linked to any other ALSPAC data.

(4) Potentially disclosive variables (e.g. small cell counts, precise dates) were not released to researchers.

Safe projects

A proposal for this project was submitted to the ALSPAC Executive in October 2023 for approval. It was approved in November 2023, assigned 'B number' B4443 and is now listed on the study's Proposal Summaries webpage.

Safe people

All members of the research team are researchers experienced in working with sensitive, individual-level data. The data for this project were only accessible by three members of the research team, who are all ONS-accredited safe researchers (AT, RC and JR). AT, RC and JR have a Disclosure and Barring Service (DBS) certificate dated within the last 24 months.

Safe setting

Due to the sensitive nature of the linked education and crime data, the dataset was accessed via UKSeRP, a secure and controlled online data-sharing platform hosted by Swansea University. Only three members of the research team (AT, RC and JR) had access to the data. Their access to the data will cease at the end of this project.

Safe outputs

All outputs were disclosure checked by an ALSPAC Data Linkage Manager prior to release from UKSeRP.

2. Data processing roles

The Data Controller for the information directly collected by ALSPAC is the University of Bristol. ALSPAC is also the (joint) Data Controller for information about participants collected from routine administrative sources.

3. Legal basis for data processing

ALSPAC's purpose is to conduct scientific research that aims to improve the public good and improve scientific understanding. The legal basis for using participants' information under General Data Protection Regulation (GDPR) and the Data Protection Act 2018 is:

1) Performance of a task carried out in the public interest (Article 6(1)(e) in the GDPR), and where sensitive personal information is involved:

2) Scientific or historical research purposes or statistical purposes (Article 9(2)(j) in accordance with Article 89(1)).

The GDPR defines 'sensitive personal information' as information that reveals a person's racial or ethnic origin, political opinions, religious or philosophical beliefs; trade union membership; the processing of genetic data or biometric data for the sole purpose of identifying a person; data concerning health; data concerning a person's sex life or sexual orientation.

This legal basis within GDPR and the Data Protection Act 2018 is separate from and in addition to ALSPAC seeking consent to take part in the research process, which they use to help ensure that research is ethical and complies with other applicable laws.

4. Data access process

Access to ALSPAC data is via a managed open-access process. Researchers submit their research proposal to the ALSPAC Executive. Once the proposal has been approved, the researchers compile a detailed list of the variables they require and this is used to build the dataset for the project. AT is an ALSPAC Direct User and built the dataset of questionnaire and clinic variables required for this project. This dataset was then passed to an ALSPAC Data Buddy to check that it matched the approved research proposal, and then the Data Buddy passed it in turn to an ALSPAC Data Linkage Manager for upload to UKSeRP. The ALSPAC Data Linkage Manager prepared the police and education data required for this project and placed them, along with the uploaded file, into a specific project folder in UKSeRP. The ALSPAC Data Linkage Manager then performed an ID swap on all the datasets for this project so that the individual ID was unique to this project (meaning the data could not be linked with any other ALSPAC data). The project folder in UKSeRP could only be accessed by the project team (AT, RC and JR).

Variables and measurement

Outcomes measures

Our primary outcome measures were (1) having a police record for serious violence in the 24 months following the end of the exposure period, (2) having a police record for any offence in the 24 months following the end of the exposure period and (3) having self-reported violent behaviour at 17.5 years and/or 18.5 years.

(1) Police records for serious violence are from the linked A&SP data. The A&SP data include charges, offences taken into consideration (TICs) and out of court disposals. Serious violence was defined using Home Office offence codes and consisted of the following offence groups: (i) violence against the person, indictable only, (ii) robbery, indictable only, and (iii) possession of weapons, triable either way or indictable only (please see Appendix B of this published report (MoJ, 2022) for the Home Office offence codes categorised as serious violence).

(2) Police records for any offence (also from the linked A&SP data) include all offence types (including those categorised as serious violence) in the study period (derived as above for the 24 months following the end of the exposure period).

(3) Self-reported violence was measured by ALSPAC questionnaire items collected at approximately 17.5 years (ALSPAC study clinic) and 18.5 years (ALSPAC questionnaire). At both time points, ALSPAC participants were asked to report, 'How often in the last year have you hit, kicked or punched someone else on purpose with the intention of really hurting them?', and 'How often in the last year have you carried a knife or other weapon with you for protection or in case it was needed in a fight'? Answers were on a 4-point scale representing: 'Not at all', 'Just once', '2-5 times' and '6 or more times'. Self-reported violence was defined as saying that they had committed either or both acts at any frequency (just once, 2-5 times or 6+ times) at either or both time points. If they had data available at only one of the two time points, this information was used to define the outcome. In addition, for some participants, the time coverage of the question, which asks about the previous 12 months, overlapped with their exposure period. The exact age at which they attended the 17.5 years study clinic was used to determine whether the coverage period of the violence questions overlapped with the exposure period; if it did, only the age 18.5 years data were used.

Exposure measures

Our exposure measures are from the NPD, which holds data on pupils attending state schools in England. Our exposures were suspensions, exclusions and persistent absence. These data were available for school Year 10 for approximately 80% of those with KS4 data. Therefore, if the participant had Year 10 data on exposures available, these data were used; if they did not, Year 11 data were used. Throughout this report, we will refer to exposures in KS4 or to the

participant's exposure period – for 80% of the sample, this is based on Year 10 data, and for 20% of the sample, this is based on Year 11 data. This approach was taken to maximise the sample size. See Figure 1 for the timeline of the exposure and outcomes measures.

Suspensions and permanent exclusions

Initial analyses revealed that the number of individuals in the baseline sample who had been permanently excluded in an academic year was too small to analyse this outcome separately. Also, almost all pupils with a permanent exclusion also had a temporary suspension. Therefore, permanent exclusions were grouped with suspensions in the results presented in this report.

Henceforth (unless otherwise indicated), we use the term suspension to refer to both suspensions and permanent exclusions.

Suspensions and exclusions are recorded per occurrence in the NPD. Data were available on the reason for the suspension or exclusion (of 12 possible reasons) and, for suspensions only, the duration in terms of the number of sessions (there are two sessions per school day). From these data, we derived:

1. Whether a student had any suspensions in the academic year – a binary variable
2. The number of separate suspension episodes a student had in the academic year
3. The total number of sessions for which a student was suspended in the academic year (note that this applies to suspensions only and not to permanent exclusions)

Persistent absence

The NPD includes the total number of possible sessions the student could have attended per school year[4] and the number of sessions the student was absent for any reason per school year. The total absence data include both authorised and unauthorised absences. From this, we derived:

1. The percentage of sessions a student was absent for any reason
2. A binary variable indicating whether a student was persistently absent using the 10% threshold, i.e. they missed 10% or more of the total possible sessions for any reason
3. A binary variable indicating whether a student was persistently absent using the 20% threshold

[4] Prior to 2012/13, absence data were only collected by the DfE for five out of six half-terms. See Appendix A for more details.

4. A binary variable indicating whether a student was severely absent (≥50% of possible sessions)

In the analysis protocol, we said we would use the 10% threshold to indicate whether a student was persistently absent; this corresponds with the threshold currently used by the DfE. However, at the time of data collection, when ALSPAC participants were in secondary education, the threshold for being persistently absent was 20%; therefore, we used this threshold in the analyses for a time-relevant comparison.

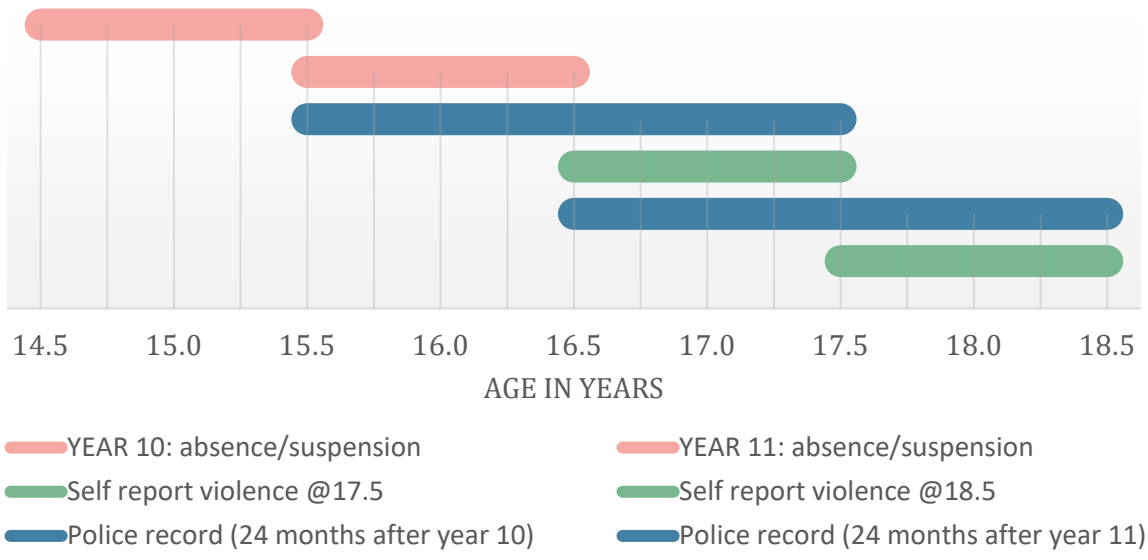
Persistent absence in this report refers to a pupil having missed 20% or more of their possible school sessions.

Initial analysis of severe absence (50% or more of total possible sessions missed) indicated that the sample with this exposure, any outcome and covariate data was too small to include this outcome in inferential analyses.

Unauthorised absence

The NPD provides the number of sessions in which student absence was unauthorised per academic year. From this, we derived the percentage of sessions from which a student was absent for unauthorised reasons.

Figure 1. Exposure and outcome timeline



Note. Exposure data were from Year 10 **or** Year 11; therefore, the study period for police-recorded outcomes was 24 months from the end of Year 10 **or** Year 11.

Other measures

The decisions on which covariates to include in the models were based on whether the variable was a potential confounder. Variables were investigated based on evidence from the

literature and/or discussions with the advisory panel. Following this, it was determined whether there were enough data available to include each covariate while retaining as many participants in the complete case samples as possible.

Here, we briefly outline the covariates: a thorough description of how these variables were derived, the data source and, where relevant, the reporting source (i.e. mother, child) is available in Appendix B. Note that full details of all ALSPAC variables are available in the searchable data dictionary (www.bristol.ac.uk/alspac).

Sex

ALSPAC provides child sex collected at birth (male/female).

Exposure period (data availability)

Whether the participant had data on suspensions and absences available for school Year 10 or Year 11.

Mother self-reported variables

Mother's age at the time of delivery (continuous), mother's highest level of qualification (dichotomised as: None/Vocational/O Level and A Level/Degree), birth parity (the number of previous births: none and one or more), mother's smoker status (yes or no) and housing tenure (owned/mortgaged and other).

Ethnicity

ALSPAC records indicate whether a child's ethnicity is 'White' or 'Non-White'. ALSPAC is predominately a White UK cohort (>95%), which largely reflects the demographics of the recruitment area at the time the study began in the early 1990s. As such, there are too few children from ethnic minority backgrounds within the dataset to disaggregate results by ethnicity. Therefore, our data are not suitable for examining ethnic differences in the relationship between absence/suspension and violence.

Neighboured deprivation

The Index of Multiple Deprivation (IMD) is a measure of relative neighbourhood deprivation and is available at the Lower Layer Super Output Area (LSOA) level. England is divided into over 30,000 LSOAs, which have a mean population of 1,500 people. The LSOAs are ranked from most to least deprived based on multiple aspects of deprivation: income; employment; education, skills and training; health and disability; crime; barriers to housing and services; and living environments. For ALSPAC participants, IMD data (based on residential postcode) were available for 2004 in the form of quintiles, with 1 being the least and 5 being the most deprived quintile.

Individual-level education variables

The NPD provided information on whether a child was eligible for free school meals (FSM) in KS4 (yes/no) and whether a child had a special educational need (SEN) (school) action plan or a SEN statement (the latter are now called Education, Health and Care plans) in KS4 (yes/no). The NPD resource also held pupil census data on whether a child had a SEN action plan/statement in the years 2001/2, 2002/3 or 2003/4 (yes/no). A variable was created to indicate whether the participant had any indicators of SEN (yes/no) from the pupil census or the KS4 data.

The NPD also provided KS3 scores for science, English and maths. Pupils could score a maximum total of 180 for science, 100 for English and 150 for maths. From this, we derived a KS3 educational attainment variable calculated as the sum of the three scores (out of a possible 430).

School-level variables

We used the following school-level NPD data: the number of students entering KS4 (a proxy for school size), the percentage of students who were eligible for FSM, the percentage of students who achieved a level 2 threshold at GCSE (5 or more A*-C GCSEs or equivalent) and, finally, the overall absence rate (percentage of total sessions missed). School-level data were selected from the academic year 2006/7 to align with the year in which the majority of participants were in KS4. If this was not available, data from other academic years (2007/8, 2005/6 or 2008/9) were used in that order.

Child and adolescent variables

'Mother's interest in child's education' was derived from the question, 'Are you interested in what your child does at school?' Responses were: 'Yes, very', 'Yes, mostly' and 'No, not really'. 'Parental involvement with the law' was derived from whether either or both of the mother and partner responded 'Yes' to the question 'Have you ever been in trouble with the law?' 'Smoking in adolescence' was coded as 'Yes' if participants reported at clinics at ages 12 and 13 years that they had smoked cigarettes at either time point. Mother-reported adolescent emotional well-being/behaviour was measured using the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). The SDQ domains, conduct issues, emotional difficulties, peer relationship problems and hyperactivity/inattention, are summed to derive a total difficulties score (TDS) ranging from zero to 40. Mothers completed the SDQ at two time points during adolescence (when their child was 11 and 13). An average score was calculated from these and used to determine if the participant fell into the top 10th percentile of scores (stratified by sex). The resulting cutoffs were 14 or more for girls (which aligns with the borderline category in the original research) and 12 or more for boys. This approach was taken as Goodman (1997) recommended that cutoffs may need to be adjusted based on age and sex. To maximize data availability, if only one questionnaire was completed, that single score was used. Any childhood trauma was a derived binary variable indicating whether the child had experienced any of the following six traumas or adversities, between the ages of 5

and 11 years (as reported by the mother): physical abuse, emotional abuse, sexual abuse, physical neglect, emotional neglect or bullying.

Pre-Key Stage 4 police record for any offence (exclusion variable)

The age at which the participant began KS4 was provided in the NPD data; therefore, police records prior to this age could be identified. A new binary variable was created to indicate whether a participant had a police record prior to the start of KS4. The potential coverage for this variable is age 10 to the age at the start of KS4 (14 years to under 16 years); however, very few records exist under the age of 14 years of age due to the retention of records when data were transferred from paper to digital format in 2007.

Changes from the analysis protocol

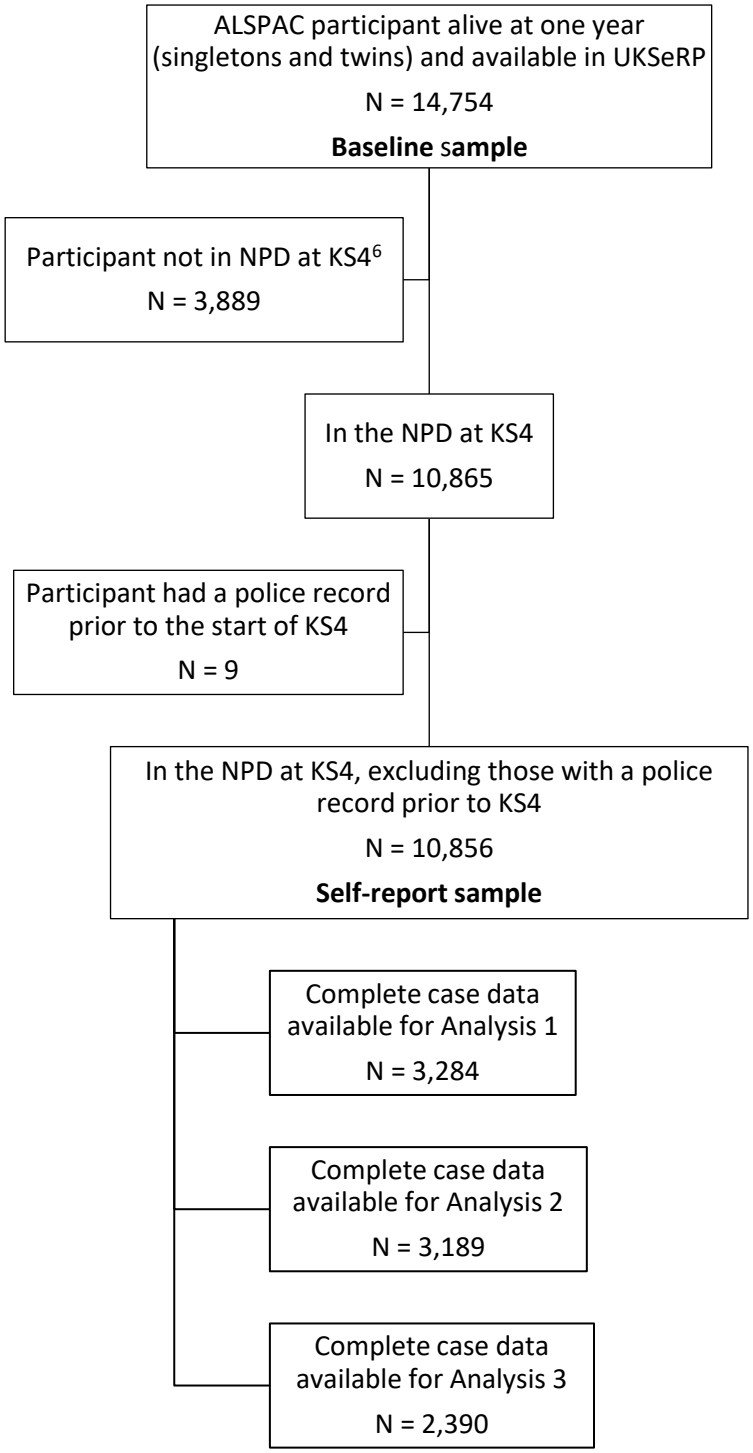
- Exposure data from Year 10 were not available for all participants. Therefore, Year 10 data were used if they were available; otherwise, Year 11 data were used. Consequently, the 24-month study period for police outcomes refers to the 24 months following the individual's exposure period (either Year 10 or Year 11). In this report, we use the phrase KS4 suspensions or KS4 persistent absence rather than Year 10 or Year 11 suspensions/persistent absence.
- Persistent absence was set at a threshold of missing ≥20% of possible sessions rather than ≥10% (which is the current threshold used by the DfE) because ≥20% was the threshold used at the time ALSPAC participants were in KS4.
- The outcome of 'having a police record for any offence' was included as an additional outcome because having a police record for a serious violence offence was rare in this dataset. Using this additional measure allowed for the analysis of additional confounders in the complete case samples.

Sample size

Participants are the study children from the ALSPAC birth cohort study, who are now adults in their early 30s. There were 14,062 live births and 13,988 children alive at one year. The initial cohort was increased to 14,901 alive at one year with the further recruitment of eligible children from age 7 years onwards. Figures 2(a) and 2(b) illustrate how the samples for this particular study were defined. Only participants who had a KS4 record in the NPD were eligible for inclusion, as this was the source of the suspension and absence data. Having a KS4 NPD record was the only requirement to be included in the self-report sample used to examine the self-reported violence outcomes. The police record sample had additional inclusion criteria based on consent to crime linkage and area of residence. This sample was used to examine the police-recorded violence outcome.

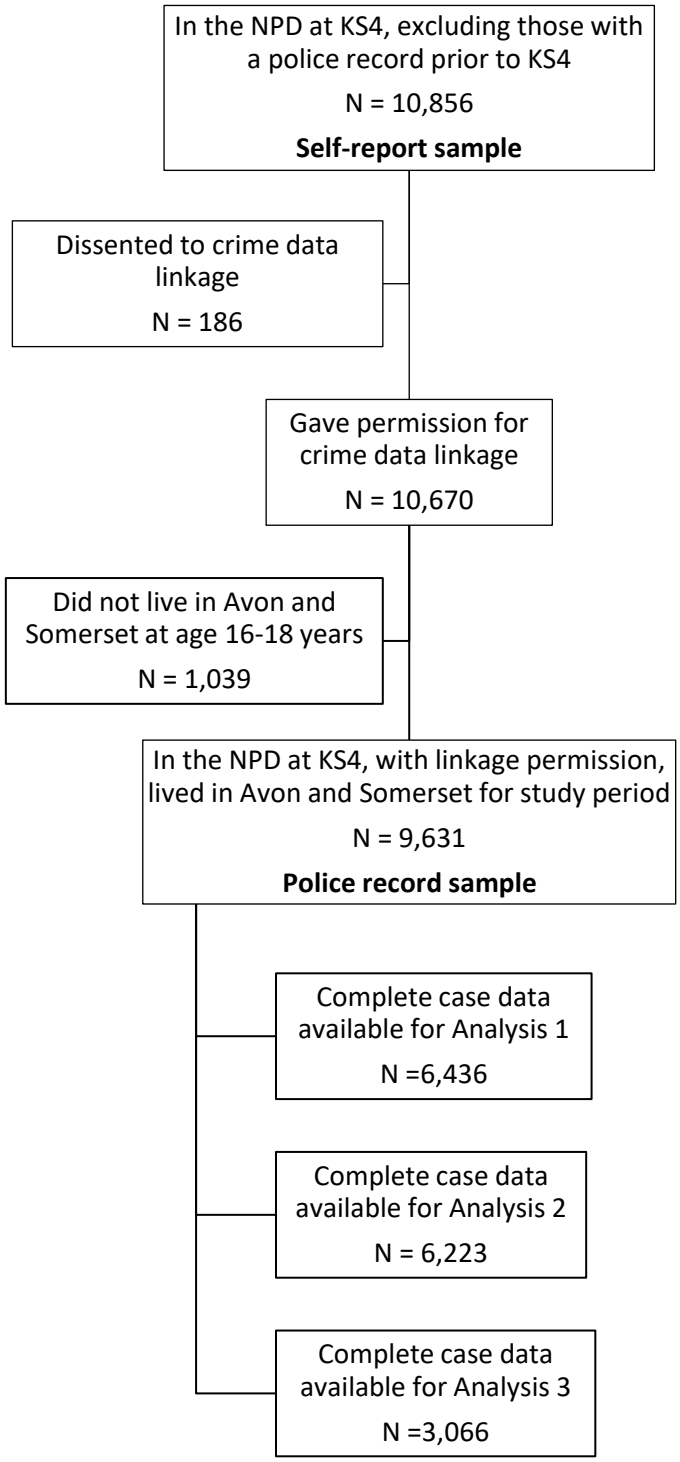
From the self-report sample and the police-record sample, analysis samples were comprised of participants with complete data for all variables included in the analytical models. The three analyses (see below) used different variables; therefore, the samples with complete data were different in each.

Figure 2(a). Study sample for self-reported violence[5]



[5] Reasons for a participant to not be in the NPD data: they did not receive fair processing materials, they opted out of education data linkage or they did not attend a state school in England during KS4 (for example, they attended private school, attended a school outside England or were home-schooled).

Figure 2(b). Study sample for police-recorded outcomes[6,6]



[6] Reasons for a participant to not be in the police data sample: they did not receive fair processing materials or they opted out of crime data linkage (either before the linkage took place in July 2021 or subsequently).

Research methodology

There are three analyses in this report:

Analysis 1. Main analyses

The main analyses studied the association between exposures and outcomes for a restricted set of covariates, taking account of clustering within siblings to account for non-independent data from sibling participants. In these main analyses (Analysis 1), the models contained the same covariates for each exposure/outcome set (see Table 4). For each exposure/outcome set, there was a slightly different sample with complete data due to differences in missing data on the exposures and outcomes.

Analysis 2. School-level variables analyses

Analysis 2 included the same covariates as Analysis 1, with the addition of school-level variables and a random intercept for school. The random intercept models the differences between schools in the baseline level of the outcome while assuming that the relationship between the predictors and the outcome is the same for all schools. The school-level variables had missing data; therefore, the sample sizes for the school-level analyses were slightly lower than those in Analysis 1 (see Figures 2a and 2b).

Analysis 3. Further analyses

Self-reported violence and any police-recorded offence were less rare than police-recorded serious violence. Therefore, we were able to include additional variables (Table 4) into models for these outcomes. As these additional variables had missing data, these samples were smaller than those in Analysis 1.

Descriptive analysis

The descriptive analysis consisted of sample characteristics for exposures, outcomes and covariates, along with cross-tabulations of exposures and outcomes within the relevant samples.

Inferential analysis

The analytical models were logistic regressions with clustering within siblings (Analysis 1 and 3) or multilevel logistic regressions with individuals nested within schools (Analysis 2).

Table 4. Control variables included in the three analyses

Grouping	Variables	Analysis 1	Analysis 2	Analysis 3
Basic demographics	Sex	x	x	x
	Exposure period	x	x	x
Early life family socioeconomic position	Mother's age at delivery	x	x	x
	Birth parity	x	x	x
	Mother's highest education	x	x	x
	Housing tenure	x	x	x
	Mother's smoking status	x	x	x
Neighbourhood factors	Neighbourhood IMD	x	x	x
Individual education factors	Eligibility for FSM	x	x	x
	SEN status	x	x	x
	KS3 attainment	x	x	x
School-level factors	Per cent of total sessions absent		x	
	Per cent of students eligible for FSM		x	
	Per cent of students attaining Level 2 threshold at GCSE		x	
	Number of students entering KS4		x	
Child and adolescent psychosocial and behavioural factors	Childhood traumas			x
	Parental involvement with the law			x
	SDQ Total Difficulties Score (TDS)			x
	Smoking behaviour			x
	Mother's interest in child's education			x

Missing data

There were missing data on self-reported violence, and for exposures and covariates (see Table 5). The decisions on which covariates to include in the models were based on whether the variable was identified as a potential confounder and whether there were enough data available for that covariate to retain as many participants in the samples as possible.

Due to missing data, the samples differed according to a range of characteristics (see Table 5). The self-report analysis sample, which depended on having complete data for self-reported violence, contained a smaller proportion of male participants, and participants were more economically advantaged, had higher KS3 attainment and were less likely to be suspended or persistently absent than the baseline self-report sample. In contrast, the police record analysis sample, which utilised police-recorded outcomes, was more similar to the baseline sample. Both further analysis samples (Analysis 3) relied on self-reported data being

available from questionnaires administered during childhood and adolescence; therefore, the differences between these and the baseline samples were most pronounced. See Appendix C for a comparison of the characteristics of the ALSPAC baseline sample to those included or not included in any analysis in this report.

Clustering within the ALSPAC sample

ALSPAC individuals were clustered within schools as well as within families (some ALSPAC mothers had two children who were recruited to the study – as a result of a twin pregnancy or two separate pregnancies (and births) within the 18-month recruitment period). However, we found that less than 1% of the variability in the outcomes was accounted for by between-school variability once individual-level factors were taken into account (see Appendix D for details). Accordingly, we used standard logistic regression models with robust standard errors to account for the clustering of individuals within families. The exceptions to this were the models that incorporated school-level variables (Analysis 2), where we used multilevel modelling with individuals clustered within schools. This was the only analysis in which we did not take account of clustering within siblings.

Table 5. Characteristics across samples

Characteristic		**Self-report sample N = 10,856**	**Analysis 1: complete case self-report sample N = 3,284**	**Analysis 3: complete case self-report sample N = 2,390**	**Police record sample N = 9,631**	**Analysis 1: complete case police record sample N = 6,436**	**Analysis 3: complete case police record sample N = 3,066**
Any police record in the study period	*Yes*	-	-	-	564	271	82
					5.7%	**4.2%**	**2.7%**
Police record for serious violence in the study period	*Yes*	-	-	-	143	68	11
					1.5%	**1.1%**	**0.4%**
Self-reported violent behaviour		N = 4,366			N = 3,786		
	Yes	404	306	211	354	269	185
		9.3%	**9.3%**	**8.8%**	**9.4%**	**9.4%**	**8.8%**
Suspension in the exposure period		N = 10,853			N = 9,630		
	Yes	727	103	66	676	395	124
		6.7%	**3.1%**	**2.8%**	**7.0%**	**6.1%**	**4.0%**
Number of sessions suspended (for those with a suspension)	*Mean*	10	8	7	10	8	8
Persistently absent in the exposure period (20% threshold)		N = 10,222			N = 9,105		
	Yes	756	70	40	700	307	77
		7.4%	**2.1%**	**1.7%**	**7.7%**	**4.8%**	**2.5%**
Persistently absent in the exposure period (10% threshold)		N = 10,222			N = 9,105		
	Yes	2,543	485	318	2,335	1,369	475
		24.88%	**14.78%**	**13.31%**	**25.65%**	**21.30%**	**15.49%**
Sex	*Female*	5,478	1,941	1,354	4,834	3,276	1,584
		50.5%	**59.1%**	**56.7%**	**50.2%**	**50.9%**	**51.7%**

Ethnicity		N = 8,958			N = 7,866		
	Other than white	404	108	66	350	237	92
		4.51%	**3.34%**	**2.79%**	**4.45%**	**3.79%**	**3.04%**
Exposure period		N = 10,853					
	Year 10	8,011	2,583	1,877	7,135	5,070	2,431
		73.8%	**78.7%**	**78.5%**	**74.1%**	**78.8%**	**79.3%**
	Year 11	2,842	701	513	2,495	1,366	635
		26.2%	**21.4%**	**21.5%**	**25.9%**	**21.2%**	**20.7%**
Birth parity		N = 9,493			N = 8,368		
	1 +	5,274	1,730	1,233	4,701	3,583	1,632
		55.6%	**52.7%**	**51.6%**	**56.2%**	**55.7%**	**53.2%**
Mother's age at delivery (years)	*Mean (SD)*	28 (5)	29 (5)	29 (4)	28 (5)	28 (5)	29 (4)
Housing tenure		N = 9,544			N = 8,412		
	Rent/council	2,393	506	311	2,119	1,383	416
	housing/other	**25.1%**	**15.4%**	**13.0%**	**25.2%**	**21.5%**	**13.6%**
Mother's education		N = 9,188			N = 8,077		
	A Level/degree	3,058	1,400	1,115	2,480	1,956	1,236
		33.3%	**42.6%**	**46.7%**	**30.7%**	**30.4%**	**40.3%**
Mother's smoking		N = 9,655			N = 8,515		
	Yes	2,457	570	339	2,230	1,549	507
		25.5%	**17.4%**	**14.2%**	**26.2%**	**24.1%**	**16.5%**
IMD		N = 10,806			N = 9,626		
	Lowest (Q5)	1,503	215	125	1,409	700	181
		13.91%	**6.55%**	**5.23%**	**14.64%**	**10.88%**	**5.90%**
Eligible for FSM		N = 10,268			N = 9,136		
	Yes	640	83	53	574	274	73
		6.2%	**2.5%**	**2.2%**	**6.3%**	**4.3%**	**2.4%**
		N = 10,667			N = 9,475		

SEN statement, action plan or action plan plus	*Yes*	2,659	457	312	2,401	1,330	464
		24.9%	**13.9%**	**13.1%**	**25.3%**	**20.7%**	**15.1%**
KS3 attainment (quartiles)		N = 9,109			N = 8,134		
	Lowest (Q4)	2,099	426	273	1,964	1,416	462
		23.0%	**13.0%**	**11.4%**	**24.2%**	**22.0%**	**15.1%**

Note. Police record data was not presented for the self-report sample, as not all participants in this sample have current consent for crime data linkage. Denominators vary because the variables come from different sources/questionnaires and have different completion rates.

Results

Descriptive results

Suspensions

There are several reasons why a pupil can be suspended from school. Prior to 2020/21, there were 12 reasons reported in the NPD. In state-funded secondary schools in 2006/7 in the City of Bristol, the most common reasons for suspension were persistent disruptive behaviour, verbal abuse/threatening behaviour against an adult, physical assault against a pupil and other. Our self-report sample had the same most common reasons for suspension across all school years (see Appendix E).

In the self-report sample, 727 (6.7%) participants had at least one suspension in the exposure period; of those who had a suspension, the majority (59.3%) had just one suspension in that period.

Persistent absence

In the self-report sample, 756 (7.4%) participants were persistently absent at the 20% threshold. Of those who were persistently absent, 27% also had a suspension during the exposure period. Similarly, of those with a suspension during the exposure period, 29% were persistently absent (see Table 6).

Table 6. Overlap between persistent absence and suspensions in the self-report sample

		Persistently absent		
		No	***Yes***	**Total**
Suspended	***No***	8,959 (87.6%)	550 (5.4%)	9,509 (93%)
	Yes	507 (5%)	206 (2%)	713 (7%)
	Total	9,466 (92.6%)	756 (7.4%)	10,222 (100%)

Self-reported violence

Data on self-reported violence were available for 4,365 participants in the self-report sample. Participants were more likely to report that in the past year that they had 'Hit, kicked or punched someone else on purpose with the intention of really hurting them' than that they had 'Carried a knife or other weapon for protection or in case it was needed in a fight'. Overall, of those who answered these questions, 404 participants (9.3%) reported that they had done at least one of these violent behaviours. Fewer participants reported weapon carrying; therefore, the violence variable is largely made up of participants who reported that they had hit, kicked or punched someone (Table 7).

Table 7. Overlap between self-reported violence questions at 17 or 18 (hit, kicked, punched or weapon-carrying) in the self-report sample

		Response to weapon-carrying		
		No	***Yes***	***Total***
Response to 'Hit, kicked or punched'	***No***	3,962 (90.8%)	52 (1.2%)	4,014 (92%)
	Yes	310 (7.1%)	41 (0.9%)	351 (8%)
	Total	4,272 (97.9%)	93 (2.1%)	4,365 (100%)

Note. These figures excluded participants who were missing responses (for example, if they responded to the hit, kicked, punched question at either time point but did not answer the weapon-carrying question at either time point).

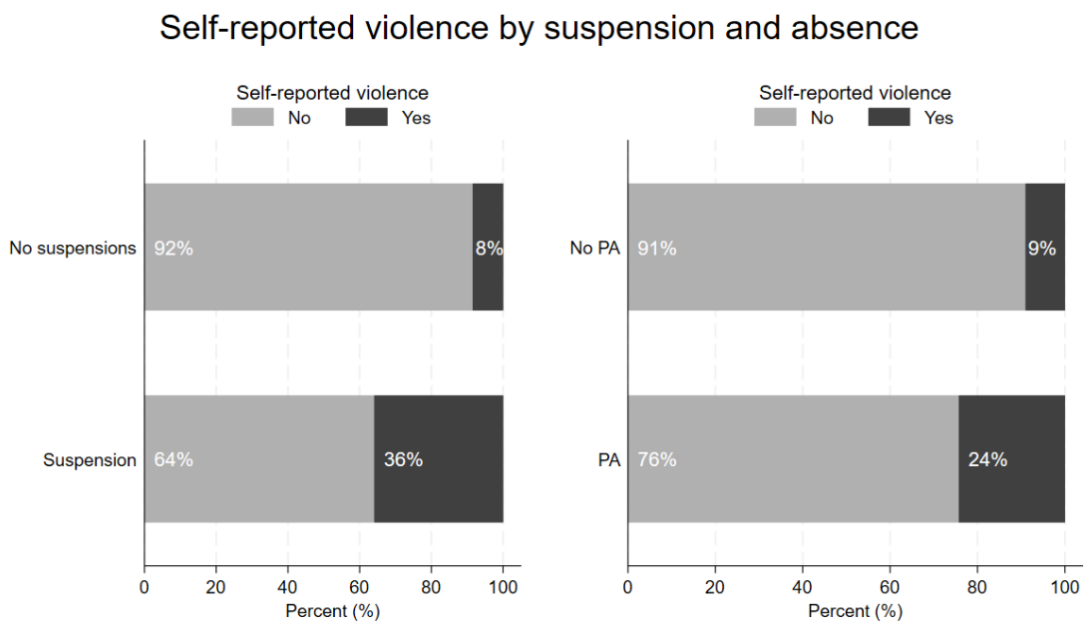
Police-recorded outcomes

Of the 9,631 ALSPAC participants in the police record sample, 5.9% had at least one A&SP record related to a charge, offence TIC, caution or other out of court disposal in the 24-month study period. Of these, 143 (1.5% of the police record sample) had a record for a serious violence offence in the 24-month study period.

Self-reported violence and exposures

In the self-report analysis sample, 37 (35.9%) participants who had a suspension during the exposure period self-reported violent behaviour, compared to 269 (8.5%) of those who were not suspended (Figure 3). Similarly, participants who were persistently absent were more likely to report violent behaviour, with 17 (24.2%) persistent absentees reporting such behaviour, compared to 289 (9.0%) of those who were not persistently absent.

Figure 3. The proportion of participants in the self-report analysis sample who reported violent behaviour according to suspension and persistent absence status during the exposure period



Note. Self-report analysis sample. PA = persistently absent; No PA = not persistently absent. N = 3,284

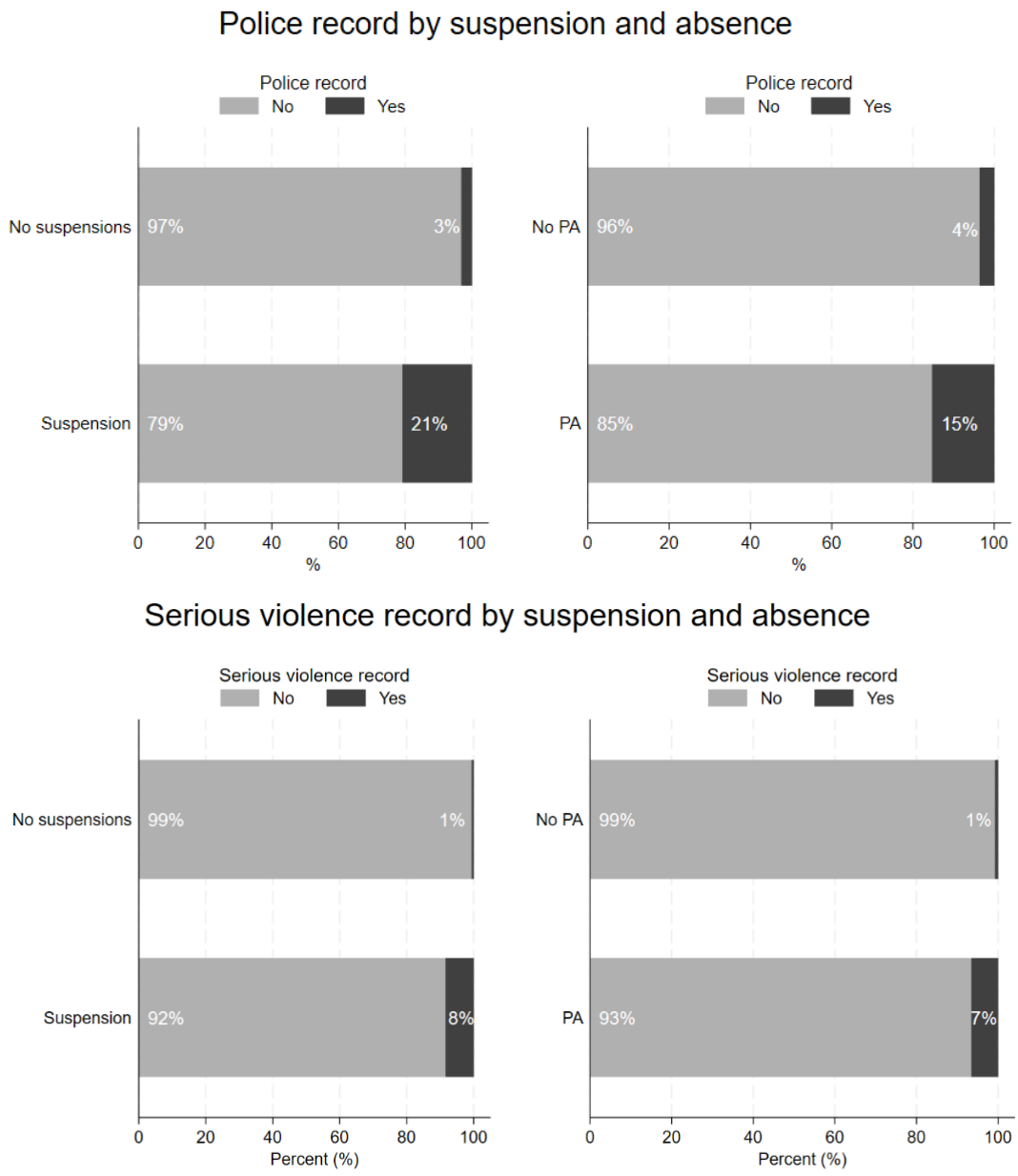
Of those who self-reported violent behaviour, 8.3% also had a police record in the study period, compared to 1.2% who did not self-report violent behaviour.

Police-recorded outcomes and exposures

In the police record analysis sample, 83 (21.0%) of the 395 participants who were suspended during the exposure period had a police record, in contrast to 188 (3.1%) of the 6,041 who were not suspended (Figure 4). For serious violence offences, of those who were suspended, 33 participants (8.4%) had a record for serious violence, compared to 35 (0.6%) of those who were not suspended.

Of those who were persistently absent, 47 (15.3%) had a police record for any offence, and 20 (6.5%) had a police record for a serious violence offence. This is compared to 220 (3.6%) and 46 (0.8%) for any offence and serious violence, respectively, among those who were not persistently absent.

Figure 4. The proportion of participants in the police record sample who had a police record for (a) any offence or (b) a serious violence offence according to suspension and persistent absence status during the exposure period



Note. Police record sample. PA = persistently absent; No PA = not persistently absent. N=6,436

Of those who had a police record in the study period, only 60 participants had self-reported violence data available. Of this small group, 41.7% had self-reported that they had committed violent behaviours. Of those with a police record for any offence, 25.1% had a record for serious violence.

Results by research question

Analysis 1: main analysis

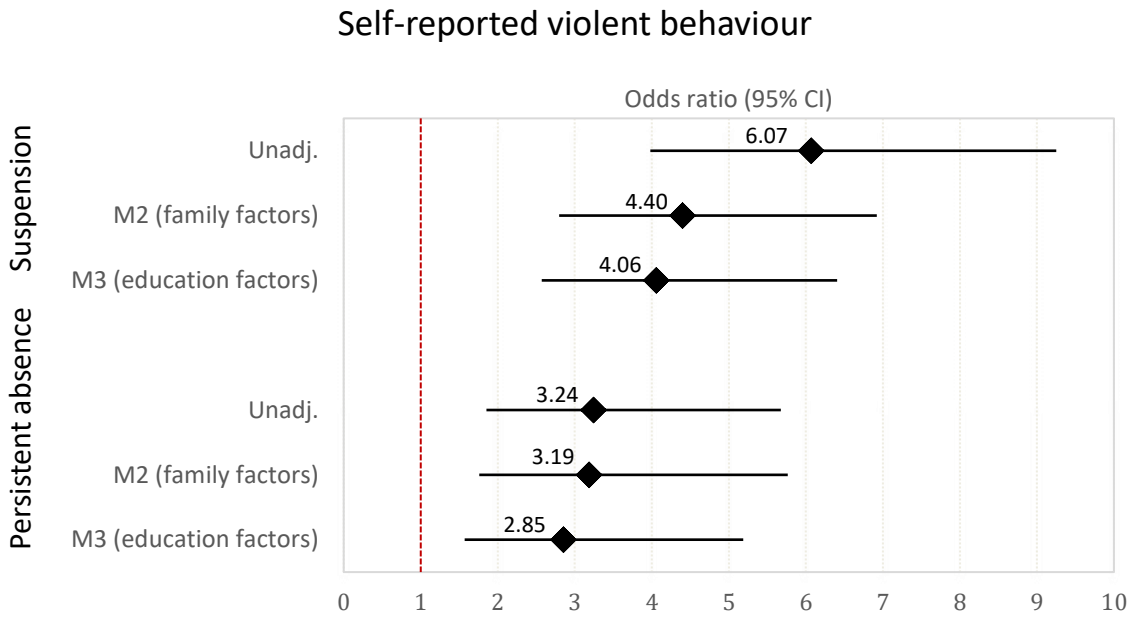
Model name	***Variables included in Analysis 1 models***
Unadjusted	Exposure
Model 2 (M2)	Unadjusted model + sex + exposure period + maternal smoking + housing tenure + mother's highest education + birth parity + mother's age at delivery + neighbourhood IMD
Model 3 (M3)	Model 2 + school attainment + eligibility for FSM + SEN

Note: Full model results for Analysis 1 are provided in Appendices F to K.

Self-reported violence

There was a strong association between school suspension during the exposure period and self-reported violent behaviour. Even once the model was adjusted for family and individual factors, those who were suspended were still 4.06 times more likely to report violence (see Figure 5, M3 – the fully adjusted model for suspension, which controls for the most covariates). Being persistently absent was also associated with self-reported violence. Those who were persistently absent in the exposure period were 2.85 times more likely to report violent behaviour than those who were not persistently absent (see Figure 5, M3 – the fully adjusted model for persistent absence).

Figure 5. Association between school suspension and persistent absence and self-reported violence (N = 3,284)



Sensitivity analyses were carried out to assess the associations between the two exposures and the self-reported violence domains separately. The associations between suspension and 'hit, kicked or punched' in the fully adjusted model were slightly lower than the combined outcome measure (OR = 3.53), whereas the associations between suspension and 'carried a weapon' were slightly stronger (OR = 4.81). In contrast, for persistent absence, the fully adjusted model for 'hit, kicked or punched' gave a higher OR than the combined outcome measure (OR = 3.51), but for weapon carrying, it was lower (OR = 1.71). However, due to small numbers in the outcomes, particularly for weapon carrying, the CIs are wide, and caution needs to be taken when interpreting these results.

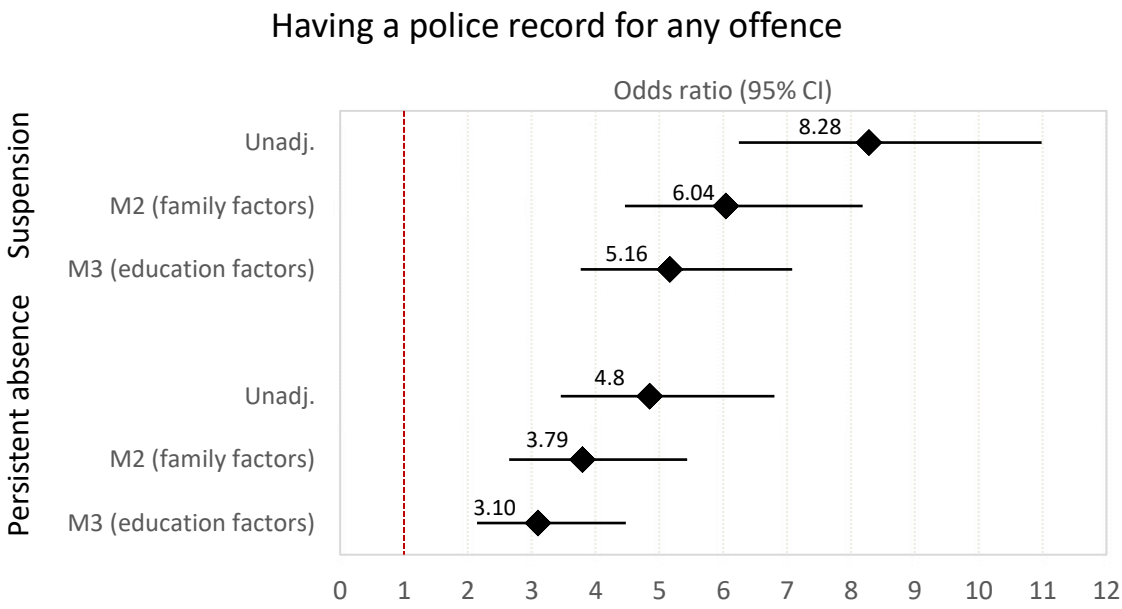
Additional analyses were carried out to establish whether persistent absence at the current threshold of 10% of sessions missed was also associated with self-reported violence. The OR for self-reporting violent behaviour in the unadjusted model was 1.95 (95% CI: 1.47-2.59); this was only slightly attenuated to 1.86 (95% CI: 1.38-2.50) by the inclusion of family and education factors (the fully adjusted model) (see Appendix L for model results).

Police record for any offence

The strong association between school suspension during the exposure period and having a police record for any offence during the study period was partially reduced by the inclusion of family and individual education-related factors. As can be seen in Figure 6, the OR decreased as family (M2) and then individual education factors (M3) were added to the model. However, even in the fully adjusted model, individuals who were suspended remained 5.16 times more likely to have a police record than those who were not suspended.

Similarly, the inclusion of family and individual factors also partially attenuated the relationship between persistent absenteeism and having a police record for any offence. Despite this adjustment, the fully adjusted model shows that individuals who were persistently absent during the exposure period were still 3.10 times more likely to have a police record during the study period compared to those who were not persistently absent.

Figure 6. Association between school suspension and persistent absence and having a police record for any offence (N = 6,436)



Additional analyses were carried out to study the association between 10% persistent absence and having a police record. The OR for having a police record in the unadjusted model was 2.68 (95% CI: 2.09-3.45); this reduced to 1.97 (95% CI: 1.50-2.58) in the fully adjusted model (see Appendix L for model results).

Police record for serious violence

The strong association between school suspension and having a police record for serious violence was reduced by the inclusion of family and individual education factors (Figure 7). However, even in the fully adjusted model, individuals who were suspended remained 8.05 times more likely to have a serious violence record than those who were not suspended.

The inclusion of family and individual factors also attenuated the relationship between persistent absenteeism and having a police record for serious violence. Despite this adjustment, the fully adjusted model shows that individuals who were persistently absent during the exposure period were still 4.79 times more likely to have a serious violence record during the study period compared to those who were not persistently absent.

Figure 7. Association between school suspension and persistent absence and having a police record for a serious violence offence (N = 6,436)

Analysis 2: school-level variables

Model name	***Variables included in Analysis 2 models***
Unadjusted (M1)	Exposure + random effects (school ID)
Covariates controlled for in all models	sex + exposure period + maternal smoking + housing tenure + mother's highest education + birth parity + mother age at delivery + neighbourhood IMD
Model 2 (M2)	Model 1 + covariates + school-level absence
Model 3 (M3)	Model 1 + covariates + school-level eligibility for FSM
Model 4 (M4)	Model 1 + covariates + school-level attainment
Model 5 (M5)	Model 1 + covariates + school size

Self-reported violence and school-level variables

There was little evidence for an association between the school-level variables (absence, eligibility for FSM, attainment and school size) and self-reported violence. The addition of school-level variables and school level as a random effect did not substantially alter the associations between either of the exposures and self-reported violence. Model results from these analyses are available in Appendices M and N.

Police record and school-level variables

As with self-reported violence, there was little evidence for an association between the school-level variables and having a police record. The addition of school-level variables did not alter the associations between either of the exposures and having a police record (Appendices O and P).

A comparison of model fit statistics (AIC and log-likelihood) indicated that the addition of each school-level variable did not improve model fit once the other covariates were included. Some of the models in these analyses had convergence issues, as noted in Appendices N to P.

Analysis 3: Further analysis: psychosocial and behavioural factors

Model name	***Variables included in Analysis 3 models***
Unadjusted (M1)	Exposure
Model 2 (M2)	Model 1 + sex + exposure period + maternal smoking + housing tenure + mother's highest education + birth parity + mother's age at delivery + neighbourhood IMD + parental involvement with the law + mother's interest in child's education + childhood traumas (age 5-11) + smoking behaviour (age 12 or 13) + SDQ

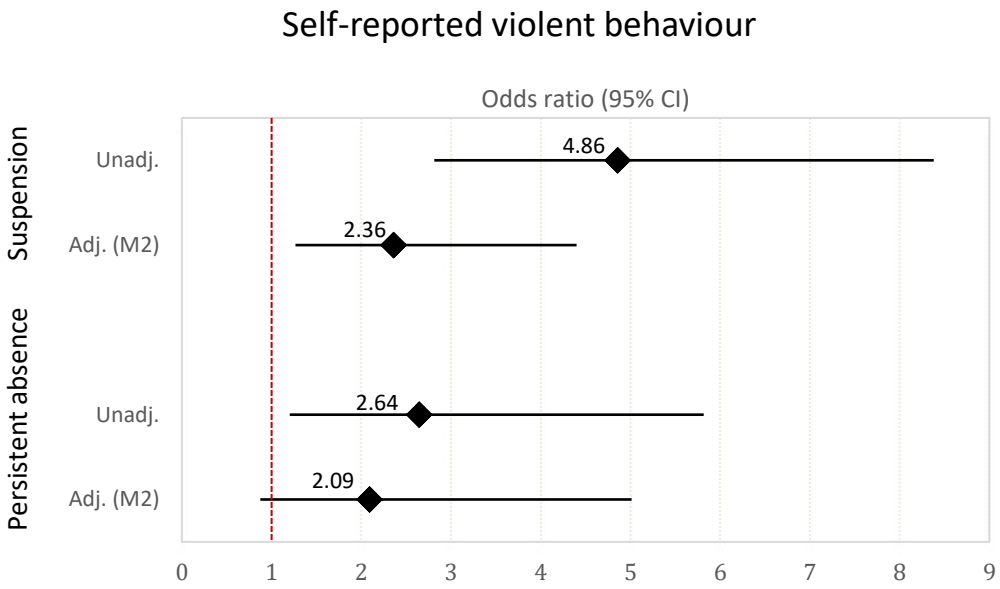
Note that the samples in this analysis were smaller than in Analyses 1 and 2 due to missing data in the additional variables. Full models for Analysis 3 are provided in Appendices Q to T.

Self-reported violence

The OR for self-reported violent behaviour comparing those suspended to those not suspended was 4.86 in the unadjusted model. In the final model, this decreased to 2.36 (Figure 8).

Controlling for these additional variables also resulted in a reduction in the OR for persistent absence, albeit to a lesser extent. After including these additional variables, the CI was wide and crossed 1. The wide CI suggests a degree of statistical uncertainty.

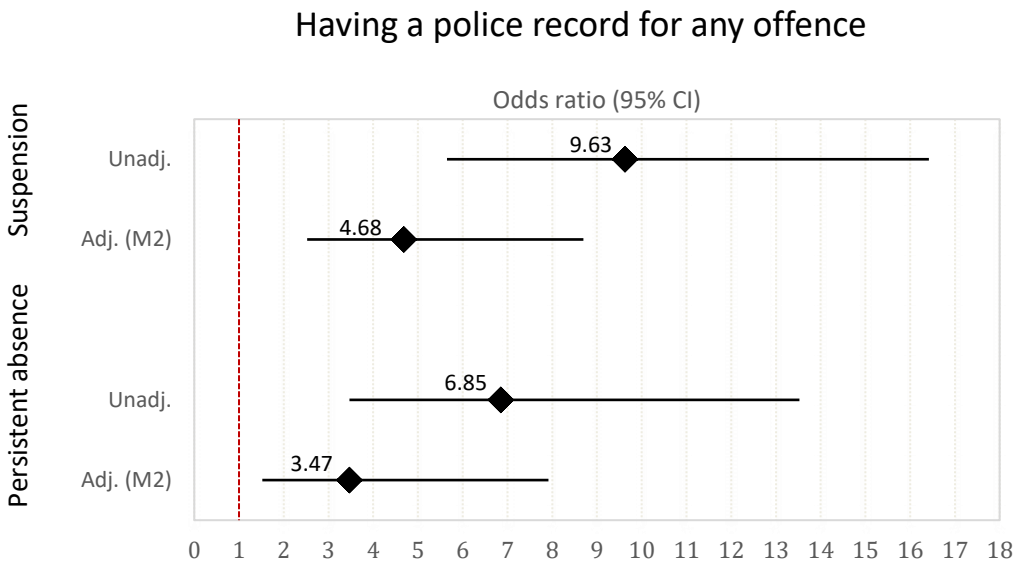
Figure 8. Further analysis: association between school suspension and persistent absence and self-reported violence (N = 2,390)



Police record

The ORs for having a police record were greatly reduced between the unadjusted and final models for suspensions and persistent absence (Figure 9). However, the likelihood of having a police record was still 4.68 times greater if an individual had a suspension in the exposure period and 3.47 times more likely if the individual had been persistently absent.

Figure 9. Further analysis: association between school suspension and persistent absence and having a police record for any offence (N = 3,066)



Conclusions and implications

Summary of findings

The results of this report demonstrate that both suspensions and persistent absence during KS4 (either Year 10 or 11) are associated with having a police record and with violent behaviour. These associations remain even after adjusting for several important individual and family factors, although the CI for the association between persistent absence and self-reported violence includes the null after adjusting for adolescent psychosocial and behavioural factors, so the results are also consistent with there being no association. The association between the two exposures appears stronger for having a police record for a serious violence offence than for the other two, arguably less serious, outcomes. It should be noted that, for all analyses, the CIs from the models are wide, so there is considerable uncertainty regarding the likely magnitude of the associations, and caution is needed in the interpretation of these results.

Individuals who were suspended (and/or permanently excluded) were five times more likely to have a police record in the following 24-month period and four times more likely to self-report violent behaviour at 17.5 or 18.5 years of age.

To a lesser extent, the same pattern is apparent for persistent absentees. Individuals who were persistently absent (for any reason) during the exposure period were three times more likely to have a police record in the succeeding 24 months and just under three times more likely to self-report violent behaviour.

The results of this study identify suspension and persistent absence, independent of a range of individual, family, school and neighbourhood factors, as predictors of adverse criminal and both self-reported and police-recorded violent outcomes in the short term. The strengths of the associations between these exposures and outcomes are substantially reduced by the inclusion of confounding factors. This suggests the presence of shared underlying determinants and indicates that while these risk factors may contribute to adverse outcomes in crime or violence, they are unlikely to be the primary cause of such outcomes.

Discussion

Individual and family factors attenuate the associations between the exposures and outcomes but do not fully explain these relationships

In this report, we conducted three analyses. In the first, we controlled for indicators of family socioeconomic status and individual-level education-related variables. These models included multiple socioeconomic indicators spanning from early childhood to adolescence. Even after accounting for these variables, the associations between the two exposures (absence and suspension) and the outcomes remained strong, particularly for suspensions. The second analysis incorporated school-level factors, such as the percentage of students eligible for FSM and the school-level absence rate. None of these variables were strongly linked to the outcomes, nor did they affect the associations between the exposures and outcomes. In the third analysis, which used a smaller sample, we further controlled for parental involvement with the law, mother's interest in the child's education, childhood traumas, socioemotional difficulties and adolescent smoking behaviour. While these additional controls further attenuated the associations, suspension remained the strongest predictor of both outcomes during the study period.

Associations between suspensions, persistent absence and offending or violence remain consistent, even after adjusting for a larger number and variety of covariates than previous studies

This study advances the literature on the link between exclusion and absence from school through the use of multiple outcomes and a wider range of potential confounders than previously used. In so doing, our results suggest that the association between these absence types and offending may not simply be the result of confounding factors, although this cannot be ruled out, and it is important to note how the inclusion of confounders markedly reduced the association between the exposures and outcomes. This suggests that offending and both types of school absence have a range of common determinants and that, while harmful, being absent from school on a voluntary or involuntary basis will rarely be the sole cause of an individual's violent/offending behaviour. Nevertheless, the experience of suspensions or absence could represent a 'tipping point' for some, with the initiation of a pattern of behaviour that leads to violence or crime.

The literature on school-level effects on violence and crime is mixed. Some school-level characteristics have been identified as potential risk factors for violent behaviour (Pauwels, Weerman, Bruinsma and Bernasco, 2015). However, when these factors are studied alongside individual-level characteristics, results indicate that the latter have a stronger influence on violent or criminal outcomes than school-level factors (Brookmeyer, Fanti and Henrich, 2006). Our findings align with this perspective: while school environments can contribute to shaping student behaviour, individual and family-level factors are often more powerful predictors of violent or criminal behaviour.

Strengths and limitations

A key strength of ALSPAC is the ability to include detailed information on a participant's family and schools, which are often missed from analyses that only use routine administrative data. Further, ALSPAC is a prospective cohort study, meaning it follows participants over time to observe how exposures influence subsequent outcomes. Unlike some studies that retrospectively examine specific groups that have already experienced an outcome (such as those with a police record), ALSPAC uses a general population sample, providing broader insights into a range of outcomes across different individuals. In this report, we were able to look at two different types of outcomes – self-reported violence and police records. One of the benefits of using police data was that the proportion of participants with these data was relatively high, largely because crime linkage in ALSPAC was acceptable to the majority of participants, including many with a police record, as indicated by the low opt-out rate (Boyd et al., 2022). However, police-recorded outcomes rely on criminal behaviour being identified and an individual being arrested; therefore, police data do not capture all crimes and are more likely to capture offences of greater severity and those perpetrated by strangers (Brennan, 2017; Brennan, 2011). On the other hand, self-reported data captures lower-level offending behaviour but relies on individuals responding to surveys and being honest in their answers, and questions can be misinterpreted. For example, in this report, one of the items for self-reported violence asks if they had 'hit, kicked or punched someone with the intention to harm'; one could assume this only includes serious harm, such as injury, but they could assume harm is anything that might hurt someone else, such as a schoolyard scuffle. Further, since the self-reported data asked participants to report on behaviour over the past 12 months, this has the potential to introduce misclassification due to recall error. Finally, and similar to previous research using the same data (Cornish et al., 2024), we found that participants with a police record for violence were less likely to have provided self-reported data, suggesting that those missing information on self-reported violence differed systematically from those with these data in terms of their levels of violent behaviour. There are also limitations relating to the time coverage of the outcomes. Police records were limited to 24 months from the end of the participant's exposure period and began on the 1st of August in the year in which the participant finished either Year 10 or Year 11. Participants could have been suspended or absent at any point in that academic year, so the exposures could have

occurred up to 11 months earlier than the start of the outcome period. This creates a gap between the exposure and outcome measurement, potentially underestimating the influence of suspensions or absences. In summary, the outcome measures in this report are affected by distinct challenges. A strength of this study is that we have both types of data, and the associations between the two exposures and the outcomes are apparent with each.

A major limitation is the availability of absence and suspension data. These data were only collected by the DfE from academic year 2005/06 onwards, and ALSPAC only linked to these data for the years 2006/07 to 2008/09. For most of the individuals in our sample, absence and suspension data were available for Year 10, whereas for others, it was Year 11. Critically, we did not have data on absences or suspensions prior to this. It is likely that there are individuals in our sample who had previously been persistently absent or suspended – i.e. including those not suspended or persistently absent in either Year 10 or 11. In addition, we cannot determine from the data whether the individuals had been violent prior to their suspensions or absences. In terms of the police-recorded outcomes, an issue with early local police records is that pre-2007, the A&SP used paper records, and the majority of these were not transferred to an electronic form. Furthermore, as per Management of Police Information rules, many older records for Category 3 offences where the individual was not involved in any further crime will likely have been deleted and will therefore not have been included in this linkage. Group 1 offences are the most serious; Group 2 covers sexual, violent and serious offences not included in Group 1; and Group 3 covers all other offences. In the present study, we excluded participants who had a police record prior to starting KS4, but based on the data retention limitations outlined above, this will not capture all participants who had a police record prior to this time. From the previous pilot linkage of ALSPAC to the PNC, researchers identified several pre-2007 records linked to ALSPAC participants. Therefore, pre-KS4 A&SP records are incomplete and we will not have excluded all people with a police record for any offence or a violence offence prior to the exposure period. Similarly, it is likely that at least some (if not many) individuals who reported violent behaviour at 17.5/18.5 years had previously been violent, including before their exposure period. In these cases, the absences or suspensions could not be the cause of subsequent violence.

We were only able to examine the exposure variables broadly rather than in detail. We combined suspensions and permanent exclusions into one variable because the number of permanent exclusions was small. It is possible that permanent exclusions may lead to different risks than suspensions. For example, students who have been permanently excluded may be placed into alternative provision or, if they are of an age to leave full-time education, may not return to school. There are many reasons a student could have been suspended or excluded,[7] but these were not analysed in the present study due to sample size limitations.

[7] Prior to 2020/21, a single reason could be recorded for each suspension and permanent exclusion. From 2020/21, up to three reasons could be recorded.

Finally, for similar reasons, we did not examine the associations between outcomes and the frequency or total duration of suspensions. Recent evidence indicates that multiple suspensions have a cumulative effect on early adult outcomes, with students suspended ten or more times showing outcomes as poor as those who had been permanently excluded (Education Policy Institute, 2024). Similarly, there are many reasons a student may be absent or persistently absent. The most common reason for absence is illness (DfE, 2024b). It is unlikely that persistent absence due to long-term illness would result in an increased risk of violent or criminal outcomes. However, since we did not have data on reasons for absence, we were not able to explore this. Additionally, we have not differentiated between authorised and unauthorised absences. The number of participants who had been persistently absent (at the 20% threshold) and with either of the outcomes studied was small; therefore, disaggregation into authorised and unauthorised absence was not feasible. Further, the NPD data linked to ALSPAC did not have information on unauthorised (other) absence, which is unauthorised absence without a given reason (i.e. not including unauthorised family holidays, etc.). This measure mainly captures absence due to truancy and has, therefore, been important in the study of the absence–violence link. It is likely that the associations between persistent absence and violent outcomes observed in our study are weaker than what we might have observed had we been able to examine this specific category of absence.

At a population level, the outcomes studied in this report are rare. An MoJ report found that in a cohort of individuals born in 1988, 8% of males and under 2% of females had been convicted by age 18 (MoJ, 2010). Serious violent offences are even more uncommon. Here, in the police record sample, 5.7% had a police record for any offence and 1.5% had a record for serious violence in the study period. While the exposures studied in this report were more common, when combined with the outcomes, numbers often became small. This highlights the challenges of studying relatively rare events using general population cohort studies. The power to detect modest but meaningful associations between exposures and outcomes is low, and CIs are wide, resulting in considerable uncertainty regarding the likely magnitude of any association. In our study, the problem was further exacerbated by missing data. In order to maximise sample size while retaining as much information as possible on potential confounders, we limited the number of covariates included in the analysis models. Undoubtedly, some of the variables we could not include due to insufficient data (e.g. child's enjoyment of school) and unmeasured variables (e.g. personality, peer relationships) contributed to residual confounding in our models. The variables we could not include due to insufficient data were from adolescence. This omission means that we lacked information on certain factors that were proximal to the exposure period, potentially limiting our understanding of key influences on an individual's life. One potential source of residual confounding is peer influence, which is particularly important during adolescence (Steinberg and Monahan, 2007). Peer group factors have been found to be associated with problematic behaviours, such as drug use, offending behaviour and truancy (Aston, 2015; Rambaran et al., 2017). Arguably, peer groups can have a direct effect on these behaviours but may also be

indicative of an individual being susceptible to peer influence. This leads to the consideration of individual differences in more immediate factors, such as personality. Various aspects of personality have been linked to violence and criminal behaviour, with traits like conscientiousness acting as a protective factor (Joliffe and Farrington, 2022), while impulsivity has been shown to contribute to a number of harmful behaviours (Bechtold et al., 2014). In addition, parental supervision (parental monitoring and awareness of their child/children's activities and location and setting rules in relation to this) has been shown to be associated with later offending (Flanagan, Auty and Farrington, 2019). Although we had some parental measures that would likely be associated with this (i.e. the mother's interest in the child's education and parental involvement with the law), we did not have any direct measurement of the extent to which parents were monitoring their child's activities outside of school. This is another source of potential unmeasured confounding.

Our dataset did not contain neighbourhood-level identifiers, so we did not take account of clustering within neighbourhoods. Neighbourhood effects could be driven by differences in police presence, which in turn could affect the detection of crime. Additionally, there may be an interplay between neighbourhood and school-level effects in that individuals from disadvantaged neighbourhoods but 'affluent' schools may be most at risk of perpetrating violence (Pinchak and Swisher, 2022); this aligns with the relative deprivation hypothesis, which suggests that unfavourable social comparisons with economically advantaged peers can lead to negative outcomes, including increased likelihood of criminal behaviour. However, while some research has found that neighbourhoods can influence crime and violence outcomes in young people, this influence has been found to be largely diluted by the inclusion of individual factors (McVie and Norris, 2006).

The NPD only collects data on pupils attending state schools in England, meaning that participants who attended private schools lack data on key exposures and are therefore excluded from this report. Further, participants who are missing the most ALSPAC questionnaire data and therefore not included in the complete case samples were more likely to be economically disadvantaged and have lower school attainment. In addition, and as stated above, those with police records, including for serious violence, were more likely to be missing from our complete case samples, as were individuals who were suspended or persistently absent. We have previously shown that this association between having a police record for serious violence and missing data in ALSPAC is only partially explained by socio-demographic factors known to be associated with ALSPAC study participation (Cornish et al., 2024). As such, it is likely that self-reported violence in ALSPAC is missing not at random (MNAR). It is also likely that certain covariates, particularly adolescent smoking and others, are MNAR in ALSPAC. We decided *a priori* to use a complete case analysis because, in this situation, multiple imputation would produce bias, whereas a complete case logistic regression gives unbiased estimates of the exposure odds ratio unless the chance of being a complete case depends jointly on the exposure and outcome (Bartlett, Harel and Carpenter, 2015; Hughes et al., 2019; i.e. in our case unless the relationship between absence or

suspensions and the likelihood of being a complete case were influenced by whether or not an individual went on to offend, which we thought unlikely). Thus, although the final samples are missing both the most economically disadvantaged and the most advantaged individuals, we do not believe missing data will have had a major impact on our findings in terms of bias, although this possibility cannot be ruled out.

It is important to note that local police records only provide information about crimes within the Avon and Somerset area. Although the pilot linkage of ALSPAC to the PNC, which covers the whole of England and Wales, revealed that most offences by ALSPAC participants had happened locally, there were instances of crimes in other parts of England and Wales (and potentially elsewhere, though we could not confirm the latter through the PNC; Boyd et al., 2022). Therefore, the absence of a police record in our dataset does not necessarily mean a participant has no police record. This highlights the importance of our approach of restricting analysis to those believed to be residing in the Avon and Somerset area during the relevant study period.

Implications and future directions

Our results indicate that having been suspended may be associated with a greater risk of having criminal or violent outcomes than having been persistently absent from school. While both suspensions and absences result in an individual not being present in school, an important distinction is that absences are due to reasons led by the individual or their family, whereas suspensions are initiated by the school. Educational policymakers and school management should consider the evidence presented in this report when determining the best course of action for behavioural interventions within schools. Both types of students, suspended students and persistent absentees, may be at greater risk of adverse outcomes than the general student population and, therefore, need increased support. Even if there is not a causal link between suspension or absence and violence, these factors are important indicators of risk and may help identify students for targeted interventions. In this discussion, we have emphasised the need to better understand how school policy influences the impact of the exposures. There is limited evidence on how discipline severity in schools impacts student outcomes: a comprehensive review of school policies across the UK would be beneficial to assess their impact. Alternatively, conducting a smaller-scale study of a select group of schools could offer the advantage of collecting richer, more detailed data, help identify nuanced patterns that may be lost in larger studies and inform more targeted interventions.

A larger sample would enable the study of interactions between covariates and exposures. The current models do not include interaction terms and, therefore, assume that the associations of suspensions or absence with the outcomes are the same in different subgroups of individuals. It is possible that certain factors, such as socioeconomic status, interact with the exposures in terms of their relationship with violence – for example, the

negative impact of being suspended may be greater for an individual from a more disadvantaged background. This knowledge could also help identify students who are most at risk of negative consequences.

One of the limitations we have noted is that police records for the samples here are incomplete: records prior to 2007 are limited, and we do not have records for crimes that occurred outside of Avon and Somerset. A potential solution to address this is to link ALSPAC data with the PNC. This linkage would provide access to national data, allowing researchers to identify individuals with criminal records prior to the exposure period and capture all recorded crimes across England and Wales. Such a dataset would offer a more complete picture of police records, enhancing the robustness of future analyses.

Because the exposures and outcomes in this study occurred up to fifteen years ago, there is potential that the relationship between suspensions, absence and offending has changed. For example, school disciplinary policies have responded to a more neoliberal approach to education funding in recent years, resulting in the emergence of alternatives to formal suspensions or exclusions through strategies such as managed moves, directing off-site and off-rolling. These new options may have changed how school behaviour, attendance and offending interact and are measured, observed or potentially obscured in administrative data. The emergence of this new lexicon of student management emphasises the importance of triangulating datasets to gain a full understanding of student exposures while administrative records adapt to accommodate these formal and informal strategies. The COVID pandemic represents a similar and arguably more significant influence on school discipline and attendance (Gibbons et al., 2024) whose association with offending should be monitored closely.

The study has further strengthened the evidence that suspension and absence from school are associated with later offending and violence. The strength of the associations indicates that they share mechanisms and that suspension and absence are warning signs for future offending, particularly violence. Naturally, this has relevance for violence prevention policy, which may achieve some of its goals by advocating for school funding. In turn, the evidence should serve to strengthen the commitment of education policymakers at school, local authority, regional and national levels to a whole-system, public health–oriented approach to violence prevention that focuses on early prevention.

References

Agnew R (1992) Foundation for a general strain theory of crime and delinquency. Criminology 30(1):47-88. Doi:10.1111/j.1745-9125.1992.tb01093

Allen G, Carthew H and Zayed Y (2023) Knife crime statistics: England and Wales. House of Commons Library. Available at: https://researchbriefings.files.parliament.uk/documents/SN04304/SN04304.pdf

Arnez J and Condry R (2021) Criminological perspectives on school exclusion and youth offending. Emoti Behav Diffic 26(1):87-100. Doi:10.1080/13632752.2021.1905233

Aston EV (2015) Are risk factors for drug use and offending similar during the teenage years? Int J Drug Policy 26(4):396-403. Doi:10.1016/j.drugpo.2014.12.0084o

Barnes JC and Motz RT (2018) Reducing racial inequalities in adulthood arrest by reducing inequalities in school discipline: evidence from the school-to-prison pipeline. Devel Psychol 54(12): 2328-2340. Doi:10.1037/dev0000613

Bartlett JW, Harel O and Carpenter JR (2015) Asymptotically unbiased estimation of exposure odds ratios in complete records logistic regression. Am J Epidemiol 182(8):730-736. Doi:10.1093/aje/kwv114

BBC News (2022) Southwark schools agree not to exclude students, BBC News 19 July 2022. Available at: https://www.bbc.co.uk/news/uk-england-london-62213675

Bechtold J, Cavanagh C, Shulman EP and Cauffman E (2014) Does mother know best? Adolescent and mother reports of impulsivity and subsequent delinquency. J Youth Adolesc 43:1903-13. Doi:10.100 7/s109640013-0080-9

Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A et al. (2013) Cohort profile: the “children of the 90s”; the index offspring of The Avon Longitudinal Study of Parents and Children (ALSPAC). Int J Epidemiol 42(1):111-127. Doi:10.1093/ije/dys064

Boyd A, Teyhan A, Cornish RP, Croft J, Thomas R et al. (2022) The potential for linking cohort participants to official criminal records: a pilot study using the Avon Longitudinal Study of Parents and Children (ALSPAC) [version 2; peer review: approved]. Wellcome Open Research 5:271. Doi:10.12688/wellcomeopenres.16328.2

Brennan, IR (2011) In vino silentium? Individual, situational, and alcohol-related factors in reporting violence to the police. Violence and Victims, 26(2):191-207. Doi:10.1007/s00787-022-01964-1

Brennan I (2017) Victims of serious violence in England and Wales, 2011–2017. College of Policing. Available at: https://library.college.police.uk/docs/college-of-policing/Victim_of_serious_violence.pdf

Brookmeyer KA, Fanti KA and Henrich CC (2006) Schools, parents, and youth violence: a multilevel, ecological analysis. J Clin Child Adolesc Psychol 35(4):504–514. Doi:10.1207/s15374424jccp3504_2

Cathro C, Tagliaferri G and Sutherland A (2023) Exclusions and youth custody: understanding the links. Behavioural Insights Team. Available at: https://www.bi.team/wp-

content/uploads/2023/01/Nuffield-Foundation-Exclusions-and-Youth-Custody-Report-vFinal-2023-01-17.pdf

Centre for Social Justice (2023) School absence risks tidal wave of youth crime, CSJ analysis reveals. Available at: https://www.centreforsocialjustice.org.uk/newsroom/school-absence-risks-tidal-wave-of-youth-crime-csj-analysis-reveals

Cohen AK (1955) Delinquent boys: the culture of the gang. New York: Free Press.

Commission on Young Lives (2022) All together now: inclusion not exclusion – supporting all young people to succeed in school. Available at: https://thecommissiononyounglives.co.uk/wp-content/uploads/2022/04/COYL-Education-report-FINAL-APR-29-2022.pdf

Cornish R and Brennan I (2024) Exclusion from school and risk of serious violence: a target trial emulation study. SocArXiv. Doi:10.31235/osf.io/ghpqn

Cornish RP, Teyhan A, Macleod J, Tilling K and Brennan I. (2025) Measuring serious violence: a comparison of self-reported and police-recorded outcomes in a UK birth cohort linked to local police data. Int J Popul Data Sci, 10(1). Doi: 10.23889/ijpds.v10i1.2391

Cuellar AE and Markowitz S (2015) School suspension and the school-to-prison pipeline. Int Rev Law Econ 43:98-106. Doi:10.1016/j.irle.2015.06.001

Department for Education (2008) Permanent and fixed period exclusions from schools in England 2006/07. Available at: https://webarchive.nationalarchives.gov.uk/ukgwa/20120505112011/http:/www.education.gov.uk/rsgateway/DB/SFR/s000793/index.shtml

Department for Education (2019) Timpson review of school exclusion. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/807862/Timpson_review.pdf

Department for Education (2022) Education, children's social care, and offending: descriptive statistics. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1059556/Education_children_s_social_care_and_offending_descriptive_stats_FINAL.pdf

Department for Education (2023) Education, children's social care and offending: multi-level modelling. Available at: https://assets.publishing.service.gov.uk/media/64b67bfb71749c001389ed72/Education_childrens_social_care_and_offending_multi-level_modelling.pdf

Department for Education (2024a) Suspensions and permanent exclusions in England. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/suspensions-and-permanent-exclusions-in-england

Department for Education (2024b) Suspension and permanent exclusion from maintained schools, academies and pupil referral units in England, including pupil movement. Guidance for maintained schools, academies and pupil referral units in England. Available at: https://assets.publishing.service.gov.uk/media/66be0d92c32366481ca4918a/Suspensions_and_permanent_exclusions_guidance.pdf

Department for Education (2024c) Pupil absence in schools in England: autumn and spring term 2023/24. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/pupil-absence-in-schools-in-england

Education Policy Institute (2021) Examining post-pandemic absences in England. Available at: https://epi.org.uk/publications-and-research/absence/

Education Policy Institute (2024) Early adult outcomes for suspended and excluded pupils. Available at: https://epi.org.uk/wp-content/uploads/2024/08/Early-adult-outcomes-for-suspended-pupils-FINAL.pdf

Ejlskov L, Antonsen S, Wulff JN, Agerbo E, Plana-Ripoll O et al. (2023) Multilevel interactions between family and neighbourhood socioeconomic indices in childhood and later risks of self-harm and violent criminality in Denmark: a national cohort study. Lancet Public Health 8(2):e99–e108. Doi:10.1016/S2468-2667(22)00292-4

Farrington D (1980) Truancy, delinquency, the home and the school, p49-64 in Berg I and Hersov LA (eds.) Out of school: modern perspectives in truancy and school refusal. Chichester: Wiley.

Farrington D (1996) Later life outcomes of truants in the Cambridge Study. p96-118 in Berg I and Nursten J (eds.) Unwillingly to school. 4th edn. London: Royal College of Psychiatrists.

Farrington DP and Ttofi MM (2011) Bullying as a predictor of offending, violence and later life outcomes. Crim Behav Ment Health 21(2):90-98. Doi:10.1002/cbm.801

Flanagan IML, Auty KM and Farrington D (2019) Parental supervision and later offending: a systematic review of longitudinal studies. Aggress Violent Behav 47:215-29 Doi:10.1016/j.avb.2019.06.003

Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J et al. (2013) Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers' cohort. Int J Epidemiol 42(1):97-110. Doi: 10.1093/ije/dys066

Gaffney H, Farrington DP and White H (2021) Interventions to prevent school exclusion. Toolkit technical report. Youth Endowment Fund. Available at: https://youthendowmentfund.org.uk/wp-content/uploads/2021/06/School-exclusions-technical-report-june.pdf

Gerlinger J, Viano S, Gardella JH, Fisher BW, Curran FC et al. (2021) Exclusionary school discipline and delinquent outcomes: a meta-analysis. J Youth Adolescence 50:1493-1509. Doi:10.1007/s10964-021-01459-3

Gibbons, S, McNally, S and Montebruno, P (2024) Absence and attainment: evidence from pandemic police. Centre for Economic Performance. Available at: https://cep.lse.ac.uk/pubs/download/special/cepsp47.pdf

Goodman R (1997) The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry 38(5):581-586. Doi:10.1111/j.1469-7610.1997.tb01545

Hersov LA (1960) Persistent non-attendance at school. J Child Psychol Psychiatry 1:130-136. Doi:10.1111/j.1469-7610.1960.tb01987

Hirschfield PJ (2018) Schools and crime. Annual Rev Criminol 1(1):149-169. doi:10.1146/annurev-criminol-032317-092358

Hughes RA, Heron J, Sterne JAC and Tilling K (2019) Accounting for missing data in statistical analyses: multiple imputation is not always the answer. Int J Epidemiol 48(4):1294-1304. Doi:10.1093/ije/dyz032

Joliffe D and Farrington DP (2022) The promotive relationship between personality and self-reported offending. Psychol Crime Law 30(4):319-37. Doi:10.1080/1068316X.2022.2092620

Klein M, Sosu EM and Dare S (2020) Mapping inequalities in school attendance: the relationship between dimensions of socioeconomic status and forms of school absence. Child Youth Serv Rev 118:105432. Doi:10.1016/j.childyouth.2020.105432

Long R and Roberts N (2024) School attendance in England. Research Briefing Number 09710. House of Commons Library. Available at: https://researchbriefings.files.parliament.uk/documents/CBP-9710/CBP-9710.pdf

Mallett CA (2016) The school-to-prison pipeline: disproportionate impact on vulnerable children and adolescents. Educ Urban Soc 49(6):563-592 Doi:10.1177/0013124516644053

Matjasko JL (2011) How effective are severe disciplinary policies? School policies and offending from adolescence into young adulthood. J Sch Psychol 49(5):555-572. Doi:10.1016/j.jsp.2011.05.002

McAra L and McVie S (2005) The usual suspects?: Street-life, young people and the police. Crim Justice 5(1):5-36. Doi:10.1177/1466802505050977

McAra L and McVie S (2010) Youth crime and justice: key messages from the Edinburgh Study of Youth Transitions and Crime. Criminol Crim Just 10(2):179-209. Doi:10.1177/1748895809360971

McAra L and McVie S (2012) Negotiated order: the groundwork for a theory of offending pathways. Criminol Crim Justice 12(4):347-375. Doi:10.1177/1748895812455810

McVie S and Norris P (2006) Neighbourhood effects on youth delinquency and drug use. The Edinburgh Study of Youth Transitions and Crime, Research Digest No 10. Centre for Law and Society, University of Edinburgh. Available at: https://www.edinstudy.law.ed.ac.uk/wp-content/uploads/sites/36/2019/10/10Drugs.pdf

Merlo J, Chaix B, Yang M, Lynch J and Råstam L (2005) A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. J Epidemiol Comm Health 59:443-449. Doi:10.1136/jech.2004.023473

Mielke M and Farrington DP (2021) School-based interventions to reduce suspension and arrest: a meta-analysis. Aggress Violent Behav 56:101518. Doi:10.1016/j.avb.2020.101518

Miller HV (2009) If your friends jumped off of a bridge, would you do it too? Delinquent peers and susceptibility to peer influence. Justice Quarterly 27(4):473-491. Doi:10.1080/07418820903218974

Ministry of Justice (2010) Conviction histories of offenders between the ages of 10 and 52: England and Wales. Statistics Bulletin. Available at:

https://assets.publishing.service.gov.uk/media/5a7c98db40f0b65b3de09ea1/criminal-histories-bulletin.pdf

Northstone K, Ben Shlomo Y, Teyhan A, Hill A, Groom A et al. (2023) The Avon Longitudinal Study of Parents and Children ALSPAC G0 partners: a cohort profile [version 2; peer review: approved]. Wellcome Open Research 8:37. Doi:10.12688/wellcomeopenres.18782.2

Pauwels LJR, Weerman FM, Bruinsma GJN and Bernasco W (2015) How much variance in offending, self-control and morality can be explained by neighbourhoods and schools? An exploratory cross-classified multi-level analysis. Eur J Crim Policy Res 21:523-537. Doi:10.1007/s10610-014-9262-6

Pinchak NP and Swisher RR (2022). Neighborhoods, schools, and adolescent violence: ecological relative deprivation, disadvantage saturation, or cumulative disadvantage? J Youth Adolescence 51(2):261-277. Doi:10.1007/s10964-021-01551-8

Rambaran JA, Hopmeyer A, Schwartz D, Steglich C, Badaly D et al. (2017). Academic functioning and peer influences: a short-term longitudinal study of network–behavior dynamics in middle adolescence. Child Devel 88:523-543. Doi:10.1111/cdev.12611

Rocque M, Jennings WG, Piquero AR, Ozkan T and Farrington DP (2017) The importance of school attendance: findings from the Cambridge Study in Delinquent Development on the life-course effects of truancy. Crime Delinquency 63(5):592-612. Doi:10.1177/0011128716660520

Ullman R, Lereya ST, Glendinnin F, Deighton J, Labno A, Liverpool S and Edbrooke-Childs J (2024) Constructs associated with youth crime and violence amongst 6-18 year olds: a systematic review of systematic reviews. Aggression and Violent Behavior, 75. https://doi.org/10.1016/j.avb.2023.101906.

Scottish Government (2024). School exclusion statistics. Available at: https://www.gov.scot/publications/school-exclusion-statistics/

Southwark Council (2022) Schools Inclusion Charter. Available at: https://moderngov.southwark.gov.uk/ieDecisionDetails.aspx?ID=7666

Steinberg L and Monahan KC (2007). Age differences in resistance to peer influence. Devel Psychol 43(6):1531-1543. Doi:10.1037/0012-1649.43.6.1531

Tiiri E, Uotila J, Elonheimo H, Sillanmäki L, Klomek AB et al. (2023) Bullying at 8 years and violent offences by 31 years: the Finnish nationwide 1981 birth cohort study. Eur Child Adolesc Psychiatry 32(9):1667-1678. Doi:10.1007/s00787-022-10964-1

Warr M (2005) Making delinquent friends: adult supervision and children's affiliations. Criminology 43(1):1-30. Doi:10.1111/j.0011-1348.2005.00003.

Wilder S (2014) Effects of parental involvement on academic achievement: a meta-synthesis. Educ Rev 66(3):377-397. Doi:10.1080/00131911.2013.780009

Youth Endowment Fund (2024a). Education, children and violence. Guidance for school, college and alternative provision leaders to help prevent children's involvement in violence. Available at: https://youthendowmentfund.org.uk/wp-content/uploads/2024/06/YEF.-Education-Children-and-Violence.-May-2024-5.pdf

Youth Endowment Fund (2024b). Evidence to decision framework: education practice guidance. Available at: https://youthendowmentfund.org.uk/wp-content/uploads/2024/06/YEF.-Education-Children-and-Violence.-May-2024-5.pdf

Appendices

Appendix descriptions

Appendix A: Absence data methodology

This appendix provides more detail on the DfE absence data methodology for defining persistent absence.

Appendix B: Variable descriptions

In this appendix the detailed description of variables mentioned in this report are provided. This includes the source of the data (ALSPAC, NPD, A&S Police), and for ALSPAC data the reporter (e.g. child, mother) and the file name are also provided.

Appendix C: Characteristics of the ALSPAC baseline sample, those included or not included in any analysis

Table with sample characteristics of the ALSPAC baseline sample (excluding those not alive at 1 year), those included in any analysis and those not included in any analysis in this report.

Appendix D: Clustering in the ALSPAC data and school-level variance

The school-level variance metrics and model fit statistics are presented for the outcomes self-reported violence and having a police record. Presented are the empty models (comprised of only the random effects – school-level), the unadjusted model (comprised of the exposure and random effects) and the fully adjusted model (comprised of the exposure, covariates and random effects).

Appendix E: Reason for suspension / exclusion (per suspension period) in the self-report sample (for all school years)

Appendices F-K: Analysis 1 tables

Regression tables for Analysis 1 (multiple logistic regressions) with each exposure/outcome pairing. The unadjusted models (with just the exposure as a predictor) are presented, alongside adjusted models with additional co-variates.

Appendix L: Additional analysis of persistent absence at the 10% threshold

Appendix M-P : Analysis 2 tables

Regression tables for Analysis 2 (multiple logistic regressions) with each exposure/outcome pairing. The unadjusted models (with just the exposure as a predictor) are presented, alongside adjusted models with additional co-variates.

Appendix Q-T: Analysis 3 tables

Regression tables for Analysis 3 (multiple logistic regressions) with each exposure/outcome pairing. The unadjusted models (with just the exposure as a predictor) are presented, alongside adjusted models with additional co-variates.

Appendix A: Absence data methodology

Total sessions

Prior to 2012/13 the DfE collected absence data on 5 out of 6 school half-terms. Therefore the total possible sessions is a maximum of 340 sessions, rather than the full school year figure with a minimum requirement of 390 sessions.

Some schools or individuals do not have complete absence data within a school year. Based on the assumption that the school has reported the absence data for the same period as the total sessions data, a percentage of sessions missed was still calculated.

However, there were some exceptions. For some schools or individuals the total possible sessions exceeded the total number possible in 5/6 half terms (and exceeded a full academic year in most cases), therefore these were capped at 340 possible sessions as it likely to be data entry error. For some schools or individuals the total sessions possible was less than one school term (which would be a minimum of 50 sessions), therefore this data was not included.

Persistent Absence

Persistent absence has been included in DfE data since 2005/6. The method for classifying persistent absence has changed over time as can be seen in the table below. In our study we have used a 20% of sessions missed threshold to be representative of the era in which ALSPAC participants were in school. However, methodologically we have used the current method of basing this threshold on each pupil's possible sessions to account for data only being available for 5/6 half-terms and for individuals who have incomplete absence data.

Persistent absence measures since 2005/06

Academic years	Description of persistent absence measure
2015/16 onwards	Academic years 10% or more of sessions missed (based on each pupil's possible sessions)
2010/11 to 2014/15	Around 15% or more of sessions missed (based on a standard threshold)
2005/06 to 2009/10	Around 20% or more of sessions missed (based on a standard threshold)

From: Department for Education 2024b

Appendix B: Variable descriptions

Variable	Data source (Respondent)	Description	Age	Grouping
Age at outcome	A&S police data	Age at time of first offence and the first serious violence offence during the study period (only relevant to those with offences)		Demographics
Exposure period	NPD	Binary: Year 10 or Year 11. Exposure data was taken from Year 10 where available and Year 11 if Year 10 data was unavailable.		Demographics
Sex	ALSPAC (Mother)	Binary: sex reported at birth coded female (0) or male (1)	Birth	Demographics
Ethnicity	ALSPAC (Mother)	Binary: ethnicity coded as White (0) or Non-White (1)	Pregnancy	Demographics
Mothers age at delivery	ALSPAC (Mother)	Continuous	Birth	Family factors
Birth parity	ALSPAC (Mother)	Categorical: “None”, “1”, “2+” Binary: “None”, “1+”	Pregnancy	Family factors
Mother’s highest level of education	ALSPAC (Mother)	Categorical: “Degree”, “A Level”, “O Level”, “None/vocational” Binary: “Degree” or “A Level” (0); “O Level” or “None/vocational” (1)	Pregnancy	Family factors
Housing tenure	ALSPAC (Mother)	Categorical: “Owned/mortgaged”, “Private rented”, “Council/Housing association rented”, “Other” Binary:	Pregnancy	Family factors
Mother or partner has even been involved with the law	ALSPAC (Mother & Partner)	Binary: derived from whether either the mother or partner responded “Yes” to the question “Have you ever been in trouble with the law?”	12yrs	Family factors

IMD neighbourhood		Categorical: quintiles	2004	Neighbourhood
Mother interest in education	ALSPAC (Mother)	Binary: derived from the questions “Are you interested in what your child does at school?”. Responses “Yes, very” coded as 0; “Yes, mostly” and “No, not really” coded as 1.	11yrs8m	School experience
Eligibility for free school meal (FSM) in KS4	NPD	Binary variable.	KS4	School variables
KS3 attainment	NPD	Quartiles of total score for Maths, English and Science.	End of KS3	School variables
Special educational needs	NPD Pupil Census	Binary variable whether SEN was indicated in the education data. Two sources of data were used to indicate whether the participant had SEN: Firstly, the NPD indicated whether school action or school action plus was in place during KS4. Coded as “Yes” if either was in place. Secondly, linked Pupil Census data indicated whether the participant had a SEN statement in each academic year. We coded this as “Yes” if SEN statement was indicated in the pupil census data from any of the three timepoints (2001/2, 2002/3 or 2003/4). Our SEN variable is derived from whether either of these variables (school action/action plus in KS4 or SEN statement in pupil census) is “Yes”, otherwise coded as “No”.	KS4 2001-4	School variables
Strengths & Difficulties Questionnaire (SDQ) total difficulties score	ALSPAC (Mother)	Binary variables: whether a participant’s score was in the top 10th percentile of scores, derived from a possible score of 0-40 on the SDQ stratified by sex (evidence indicates boys have higher SDQ scores on average). The total difficulties score (TDS) is the sum of	11yrs8m 13yr1m	Psychosocial and behavioural

		the Hyperactivity, Emotional Symptoms, Conduct Problems and Peer Problems scores. Calculated for KW and TA separately, then to maximise sample size a mean for KW and TA scores. For those missing data at either KW or TA only one value was be used.		
Any childhood traumas / adversities	ALSPAC (Mother)	Binary: derived variable already available in ALSPAC indicates whether mother reported any childhood traumas (physical abuse, emotional abuse, sexual abuse, physical neglect, emotional neglect, or bullying) on questionnaires between 5-11years.	5-11 years (from multiple questionnaires)	Psychosocial and behavioural
Smoking	ALSPAC (Child)	Binary: coded as 'Yes' if participants reported they had smoked cigarettes at either time point.	12yrs6m / 13yrs6m	Psychosocial and behavioural
School academic year size	NPD	Quartiles of the number of pupils in the year group entering KS4.	KS4	School-level variables
School-level attainment	NPD	Quartiles of the percent of pupils who passed the L2 threshold at GCSE level.	KS4	School-level variables
School-level eligibility for free school meals (FSM)	NPD	Quartiles of the percent of pupils eligible for FSM in a given year.	KS4	School-level variables
School-level absence rates	NPD	Quartiles of the percent of sessions missed in an academic year.	KS4	School-level variables

Appendix C: Characteristics of the ALPSAC baseline sample, those in any analysis sample and those who were not in any analyses in our report

		In any analysis sample?		*Baseline*
		No	**Yes**	
Sex	Female	3,701	3,521	7,222
		46.9%	51.4%	49.0%
	Male	4,197	3,329	7,526
		53.1%	48.6%	51.0%
Exposure period (data availability)	Year 10	2,622	5,397	8,019
		65.4%	78.8%	73.8%
	Year 11	1,390	1,453	2,843
		34.7%	21.2%	26.2%
Age of mother at delivery	Mean	28	28	28
	SD	5.2	4.7	5.0
Housing tenure	Owned/mortgaged	4,068	5,386	9,454
		67.2%	78.6%	73.3%
	Other	1,983	1,464	3,447
		32.8%	21.4%	26.7%
Mum smoker	No	4,396	5,244	9,640
		71.1%	76.6%	74.0%
	Yes	1,784	1,606	3,390
		28.9%	23.5%	26.0%
Mother highest education	Degree/A Level	2,142	2,198	4,340
		39.3%	32.1%	35.3%
	None/Vocational/O Level	3,304	4,652	7,956
		60.7%	67.9%	64.7%
Birth parity	0	2,661	3,062	5,723
		44.7%	44.7%	44.7%
	1+	3,293	3,788	7,081
		55.3%	55.3%	55.3%
FSM	0	3,066	6,568	9,634
		89.5%	95.9%	93.7%
	1	361	282	643
		10.5%	4.1%	6.3%
SEN	0	2,675	5,463	8,138
		66.2%	79.8%	74.7%
	1	1,364	1,387	2,751
		33.8%	20.3%	25.3%
Reverse scored quartiles of KS3 total score	High	517	1,886	2,403
		22.5%	27.5%	26.3%

	Upper Middle	540	1,859	2,399
		23.5%	27.1%	26.2%
	Lower Middle	570	1,662	2,232
		24.8%	24.3%	24.4%
	Low	675	1,443	2,118
		29.3%	21.1%	23.1%
IMD Score 2004, quintiles	Least deprived	1,965	2,151	4,116
		26.7%	31.4%	29.0%
	2	1,617	1,592	3,209
		22.0%	23.2%	22.6%
	3	1,388	1,434	2,822
		18.9%	20.9%	19.9%
	4	1,069	957	2,026
		14.5%	14.0%	14.3%
	Most deprived	1,323	716	2,039
		18.0%	10.5%	14.4%
Self-reported violent behaviour at 17.5 or 18.5	No	1,752	2,978	4,730
		91.0%	90.7%	90.8%
	Yes	173	307	480
		9.0%	9.4%	9.2%
Police record any offence	No	2,901	6,165	9,066
		90.8%	95.8%	94.1%
	Yes	293	271	564
		9.2%	4.2%	5.9%
Police record serious violence	No	3,119	6,368	9,487
		97.7%	98.9%	98.5%
	Yes	75	68	143
		2.4%	1.1%	1.5%
KS4 any suspensions	No	3,688	6,442	10,130
		91.9%	94.0%	93.3%
	Yes	324	408	732
		8.1%	6.0%	6.7%
KS4 persistent absence	No	2,947	6,526	9,473
		86.9%	95.4%	92.6%
	Yes	444	314	758
		13.1%	4.6%	7.4%

Appendix D: Clustering within the ALSPAC sample

Most statistical methods assume that measures on individuals in a sample are independent of one another. However, if data are clustered, measures may not be independent. Clustered data can arise from samples being grouped, for example children clustered in schools. If the observations within a cluster are more similar to each other than they are to observations in the wider sample, then the assumption of independence between observations is violated and using standard regression techniques inappropriate. However, clustering should not be viewed solely as a 'statistical nuisance' - it can be of intrinsic interest. For example, the more that the outcomes for individuals within a group are alike, the more likely it is that determinants of that outcome are directly related to the contextual environment itself (Merlo et al., 2005). Multilevel modelling (MLM) is a statistical technique which accounts for clustering in data and can be used to explore the impact of context (e.g. school environment) on outcomes. Another way of taking account of clustering is to use robust standard errors. This is appropriate when the clustering is of no intrinsic interest, but you want to take account of it in your analysis.

In ALSPAC, participants are clustered within schools. At KS4, the majority attended a secondary school in the Avon and Somerset Area, but some had moved elsewhere. The ALSPAC participants who had been linked to the NPD were spread across 900 different schools in KS4. An additional source of clustering within ALSPAC is siblings within a family. Siblings will share many characteristics related to the outcomes, exposures, and covariates being studied.

To assess the extent of clustering in our outcomes at school-level we compared the fit of multilevel logistic regression models (with individuals clustered within schools) to a standard logistic regression model using likelihood-ratio (LR) tests. The LR tests indicated that a multilevel structure with schools did not improve the model fit for any of the suspension or absence models (χ^2 (1) = 1.20, p = 0.14 for suspension and self-reported violence model; χ^2 (1) = 0.07, p = 0.39 for the suspension and police record model).

In concordance with this, in all models the proportion of variance in the outcomes at school-level was low. For example, in the unadjusted suspension model, the variance in self-reported violence accounted for by schools (number of schools = 329, with a mean of 9.7 pupils, and a range of 1-143 pupils per school) was approximately 1.4% (standard error = 1.5%). Once covariates were included, this reduced to less than 1% of the variance. For the unadjusted suspension model with having a police record as the outcome, 3.1% of the variance was at school-level (SE 2.0%) (number of schools = 172, with a mean of 36.2 pupils, and a range of 1-271 pupils per school). However, once covariates were included, this reduced to less than 1% of the variance (see below).

School-level variance

School-level variance (with standard error) and intraclass correlation (ICC) (with standard error) at school-level and model fit statistics for both self-reported violence and police record for any offence outcomes.

	Self-reported violence			**Police record**		
	Empty model (random effects only)	Unadjusted model (absence + random effects)	Fully adjusted model (absence + covariates + random effects)	Empty model (random effects only)	Unadjusted model (absence + random effects)	Fully adjusted model (absence + covariates + random effects)
School-level variance (SE)	0.037 (0.047)	0.039 (0.048)	0.024 (0.047)	0.095 (0.067)	0.090 (0.065)	0.020 (0.049)
ICC at school-level (SE)	0.011 (0.01)	0.012 (0.01)	0.007 (0.01)	0.028 (0.02)	0.027 (0.019)	0.006 (0.01)
AIC	1950.96	1874.11	1870.96	2135.60	2085.38	1970.24
Log Likelihood	-973.48	-968.09	-913.48	-1065.80	-1039.69	-966.12

The variance at the school-level is relatively low across all models, indicating minimal between-school variability in the outcomes. The ICC for self-reported violence is low in the empty model (1.1%), indicating minimal variance explained by school-level factors. In the fully adjusted model, this drops further to under 1%, suggesting that individual factors (absence and covariates) account for most of the variance. For police-recorded violence, the empty model ICC is slightly higher (2.8%), showing more school-level influence, but this also decreases significantly to less than 1% in the fully adjusted model, with minimal remaining school-level effects. *results were similar for models with suspension as a predictor.

Appendix E: Reason for suspension / exclusion (per suspension period) in the self-report sample (for all school years)

Reason		
Physical assault against a pupil	289	12.7%
Physical assault against an adult	37	1.6%
Verbal abuse/ threatening behaviour against a pupil	54	2.4%
Verbal abuse/ threatening behaviour against an adult	647	28.4%
Bullying	29	1.3%
Racist Abuse	22	1.0%
Sexual Misconduct	11	0.5%
Drug and alcohol related	96	4.2%
Damage	67	2.9%
Theft	57	2.5%
Persistent disruptive behaviour	528	23.2%
Other	443	19.4%

Appendix F: Analysis 1. Suspensions/self-reported violence model

		Unadj. Model SRV		*Adj. Model 2 SRV*		*Adj. Model 3 SRV*	
KS4 any suspensions	*Yes*	6.07	**	4.40	**	4.06	**
		3.98-9.25		2.80-6.92		2.57-6.41	
Sex of child	*Male*			2.57	**	2.48	**
				2.01-3.30		1.93-3.19	
Exposure period	*Year 11*			1.04		1.04	
				0.77-1.40		0.77-1.40	
Age of mother at delivery	*Per increase of 1 year*			1.003		1.005	
				0.97-1.03		0.98-1.04	
Housing tenure	*Other*			1.001		0.96	
				0.71-1.41		0.68-1.36	
Birth parity (0 or 1+)	*1+*			0.94		0.93	
				0.73-1.21		0.72-1.20	
Mum smoker	*Yes*			1.57	**	1.56	**
				1.16-2.12		1.15-2.12	
Mother highest education	*None/Vocational/O Level*			1.39	*	1.31	*
				1.07-1.81		1.00-1.72	
IMD Score 2004, quintiles	*2*			1.03		1.04	
				0.74-1.42		0.75-1.44	
	3			1.05		1.05	
				0.74-1.49		0.74-1.50	
	4			1.19		1.19	
				0.80-1.76		0.80-1.77	
	Most deprived			1.47		1.47	
				0.90-2.39		0.90-2.41	
Is pupil eligible for FSM?	*Yes*					1.37	
						0.67-2.79	
SEN indicated in NPD	*Yes*					1.48	*
						1.07-2.05	
Quartiles of KS3 total score (reverse scored)	*Q2*					1.35	
						0.99-1.84	
	Q3					1.28	
						0.90-1.81	
	Lowest (Q4)					0.98	
						0.64-1.50	
Intercept		0.09	**	0.04	**	0.03	**
		0.08-0.11		0.02-0.10		0.01-0.08	
Number of observations		3284		3284		3284	
AIC		1982.08		1928.33		1926.76	
Log pseudolikelihood		-989.04		-951.17		-945.38	
Pseudo R-squared		0.03		0.07		0.07	

results

Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no suspension, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score.

Appendix G: Analysis 1. Absence/self-reported violence model results

		Unadj. Model SRV		*Adj. Model 2 SRV*		*Adj. Model 3 SRV*	
KS4 persistent absence	*Yes*	3.243	**	3.185	**	2.854	**
		1.85-5.68		1.76-5.77	**	1.57-5.19	
Sex of child	*Male*			2.861		2.714	**
				2.24-3.66		2.11-3.49	
Exposure period	*Year 11*			1.050		1.052	
				0.78-1.41		0.78-1.41	
Age of mother at delivery	*Per increase of 1 year*			1.000		1.002	
				0.97-1.03		0.97-1.03	
Housing tenure	*Other*			1.049		0.994	
				0.75-1.47		0.70-1.41	
Birth parity (0 or 1+)	*1+*			0.955		0.940	
				0.74-1.23		0.73-1.21	
Mum smoker	*Yes*			1.580	**	1.580	**
				1.18-2.12		1.17-2.13	
Mother highest education	*None/Vocational/O Level*			1.452	**	1.360	*
				1.12-1.89		1.04-1.78	
IMD Score 2004, quintiles	*2*			1.034		1.047	
				0.75-1.43		0.76-1.45	
	3			1.060		1.070	
				0.75-1.50		0.76-1.51	
	4			1.132		1.143	
				0.77-1.67		0.77-1.69	
	Most deprived			1.432		1.441	
				0.89-2.31		0.89-2.35	
Is pupil eligible for FSM?	*Yes*					1.313	
						0.65-2.64	
SEN indicated in NPD	*Yes*					1.587	**
						1.16-2.18	
Quartiles of KS3 total score (reverse scored)	*Q2*					1.327	
						0.97-1.81	
	Q3					1.310	
						0.93-1.85	
	Lowest (Q4)					1.014	
						0.67-1.54	
Intercept		0.099	**	0.040	**	0.032	**
		0.09-0.11		0.02-0.10		0.01-0.08	
Number of observations		3281		3281		3281	
AIC		2024.55		1952.39		1948.64	
Log pseudolikelihood		-1010.28		-963.19		-956.32	
Pseudo R-squared		0.01		0.05		0.06	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no persistent absence, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score.*

Appendix H: Analysis 1. Suspensions/police record model results

		Police record Unadj. Model		Police record Adj. Model 2		Police record Adj. Model 3	
KS4 any suspensions	*Yes*	8.282	**	6.041	**	5.164	**
		6.24-10.98		4.46-8.18		3.77-7.07	
Sex of child	*Male*			2.155		2.052	**
				1.64-2.83		1.56-2.70	
Exposure period	*Year 11*			1.210		1.180	
				0.89-1.63		0.87-1.59	
Age of mother at delivery	*Per increase of 1 year*			0.981		0.987	
				0.95-1.01		0.96-1.01	
Housing tenure	*Other*			1.759	**	1.575	**
				1.32-2.34		1.18-2.10	
Birth parity (0 or 1+)	*1+*			1.422	*	1.333	*
				1.08-1.87		1.01-1.76	
Mum smoker	*Yes*			1.352	*	1.297	
				1.02-1.80		0.97-1.73	
Mother highest education	*None/Vocational/O Level*			1.154		1.019	
				0.84-1.58		0.74-1.40	
IMD Score 2004, quintiles	*2*			1.316		1.270	
				0.90-1.91		0.87-1.85	
	3			0.999		0.953	
				0.68-1.47		0.64-1.41	
	4			1.391		1.254	
				0.93-2.08		0.83-1.88	
	Most deprived			1.342		1.135	
				0.86-2.08		0.72-1.79	
Is pupil eligible for FSM?	*Yes*					1.554	
						0.97-2.48	
SEN indicated in NPD	*Yes*					1.345	*
						1.01-1.79	
Quartiles of KS3 total score (reverse scored)	*Q2*					1.428	
						0.91-2.24	
	Q3					1.780	*
						1.14-2.77	
	Lowest (Q4)					2.116	**
						1.35-3.32	
Intercept		0.032	**	0.017	**	0.011	**
		0.03-0.03		0.01-0.04		0.004-0.02	
Number of observations		6436		6436		6436	
AIC		2084.92		2020.74		2003.62	
Log pseudolikelihood		-1040.46		-997.37		-983.81	
Pseudo R-squared		0.07		0.11		0.12	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no suspension, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score.*

Appendix I: Analysis 1. Absence/police record model results

		Police record Unadj. Model		Police record Adj. Model 2		Police record Adj. Model 3	
KS4 persistent absence	*Yes*	4.847	**	3.792	**	3.097	**
		3.46-6.80		2.65-5.43		2.14-4.47	
Sex of child	*Male*			2.680	**	2.427	**
				2.04-3.52		1.84-3.20	
Exposure period	*Year 11*			1.085		1.075	
				0.80-1.46		0.79-1.45	
Age of mother at delivery	*Per increase of 1 year*			0.981		0.988	
				0.95-1.01		0.96-1.01	
Housing tenure	*Other*			1.811	**	1.590	**
				1.35-2.42		1.18-2.14	
Birth parity (0 or 1+)	*1+*			1.432	*	1.331	*
				1.09-1.88		1.01-1.75	
Mum smoker	*Yes*			1.395	*	1.341	*
				1.05-1.85		1.01-1.79	
Mother highest education	*None/Vocational/O Level*			1.194		1.027	
				0.87-1.63		0.75-1.40	
IMD Score 2004, quintiles	*2*			1.276		1.224	
				0.88-1.85		0.84-1.78	
	3			1.017		0.964	
				0.69-1.50		0.65-1.43	
	4			1.213		1.101	
				0.81-1.81		0.73-1.65	
	Most deprived			1.294		1.115	
				0.84-2.00		0.71-1.75	
Is pupil eligible for FSM?	*Yes*					1.403	
						0.90-2.19	
SEN indicated in NPD	*Yes*					1.566	**
						1.19-2.05	
Quartiles of KS3 total score (reverse scored)	*Q2*					1.515	
						0.96-2.38	
	Q3					1.944	**
						1.25-3.02	
	Lowest (Q4)					2.329	**
						1.49-3.63	
Intercept		0.037	**	0.017	**	0.010	**
		0.03-0.04		0.01-0.04		0.004-0.03	
Number of observations		6426		6426		6426	
AIC		2162.04		2073.38		2044.70	
Log pseudolikelihood		-1079.02		-1023.69		-1004.35	
Pseudo R-squared		0.03		0.08		0.10	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no persistent absence, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score.*

Appendix J: Analysis 1. Suspensions/serious violence record model results

		SV record Unadj. Model		SV record Adj. Model 2		SV record Adj. Model 3	
KS4 any suspensions	*Yes*	15.643	**	10.148	**	8.047	**
		9.61-25.47		6.00-17.17		4.66-13.91	
Sex of child	*Male*			2.146	**	1.926	*
				1.22-3.77		1.09-3.41	
Exposure period	*Year 11*			1.029		1.026	
				0.55-1.94		0.55-1.93	
Age of mother at delivery	*Per increase of 1 year*			0.958		0.963	
				0.91-1.01		0.91-1.01	
Housing tenure	*Other*			2.317	**	2.002	*
				1.34-4.02		1.14-3.51	
Birth parity (0 or 1+)	*1+*			2.166	*	1.961	*
				1.19-3.93		1.07-3.60	
Mum smoker	*Yes*			1.144		1.091	
				0.68-1.92		0.64-1.85	
Mother highest education	*None/Vocational/O Level*			2.394		2.064	
				1.00-5.75		0.86-4.97	
IMD Score 2004, quintiles	*2*			1.550		1.520	
				0.63-3.80		0.61-3.78	
	3			1.530		1.465	
				0.68-3.47		0.63-3.43	
	4			2.631	*	2.363	*
				1.17-5.90		1.02-5.50	
	Most deprived			1.380		1.160	
				0.55-3.48		0.43-3.11	
Is pupil eligible for FSM?	*Yes*					1.173	
						0.51-2.72	
SEN indicated in NPD	*Yes*					1.701	
						0.99-2.91	
Quartiles of KS3 total score (reverse scored)	*Q2*					1.852	
						0.60-5.71	
	Q3					1.474	
						0.47-4.61	
	Lowest (Q4)					2.923	
						0.98-8.76	
Intercept		0.006	**	0.002	**	0.001	**
		0.004-0.008		0.0003-0.0099		0.0001-0.0072	
Number of observations		6436		6436		6436	
AIC		661.37		631.56		627.18	
Log pseudolikelihood		-328.68		-302.78		-295.59	
Pseudo R-squared		0.13		0.20		0.22	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no suspension, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score.*

Appendix K: Analysis 1. Absence/serious violence record model results

		SV record Unadj. Model		*SV record Adj. Model 2*		*SV record Adj. Model 3*	
KS4 persistent absence	*Yes*	9.200	**	6.264	**	4.794	**
		5.37-15.76		3.53-11.11		2.73-8.43	
Sex of child	*Male*			3.295	**	2.768	**
				1.86-5.83		1.54-4.99	
Exposure period	*Year 11*			0.755		0.775	
				0.40-1.44		0.41-1.48	
Age of mother at delivery	*Per increase of 1 year*			0.955		0.958	
				0.91-1.01		0.91-1.01	
Housing tenure	*Other*			2.534	**	2.128	*
				1.43-4.51		1.19-3.81	
Birth parity (0 or 1+)	*1+*			2.133	*	1.914	*
				1.18-3.84		1.05-3.49	
Mum smoker	*Yes*			1.208		1.174	
				0.71-2.05		0.69-2.01	
Mother highest education	*None/Vocational/O Level*			2.444	*	2.005	
				1.02-5.87		0.82-4.88	
IMD Score 2004, quintiles	*2*			1.311		1.230	
				0.53-3.23		0.49-3.08	
	3			1.511		1.391	
				0.67-3.39		0.60-3.23	
	4			2.018		1.764	
				0.91-4.46		0.78-3.99	
	Most deprived			1.212		1.040	
				0.48-3.04		0.39-2.75	
Is pupil eligible for FSM?	*Yes*					0.987	
						0.45-2.19	
SEN indicated in NPD	*Yes*					2.008	**
						1.19-3.40	
Quartiles of KS3 total score (reverse scored)	*Q2*					1.819	
						0.58-5.70	
	Q3					1.644	
						0.52-5.19	
	Lowest (Q4)					3.254	*
						1.09-9.74	
Intercept		0.008	**	0.002	**	0.001	**
		0.005-0.010		0.0004-0.011		0.0002-0.008	
Number of observations		6426		6426		6426	
AIC		693.49		650.95		641.30	
Log pseudolikelihood		-344.75		-312.47		-302.65	
Pseudo R-squared		0.06		0.15		0.18	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no persistent absence, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score.*

Appendix L: Additional analysis of persistent absence at the 10% threshold

Logistic regression results for self-reported violence and persistent absence at the 10% threshold

		Unadjusted model	M2	M3
Odd ratio for KS4 persistent absence (at 10% threshold)	*Yes*	1.95**	1.97**	1.86**
95% CI		[1.47 – 2.59]	[1.46 – 2.66]	[1.38 – 2.50]

*Note. ** p<.01, * p<.05. Reference category: No persistent absence.*
M2 (adjusted for sex + exposure period + maternal smoking + housing tenure + mother's highest education + birth parity + mother age at delivery + neighbourhood IMD)
M2 (adjusted for Model 2 + school attainment + eligibility for FSM + SEN)

Logistic regression results for having a police record and persistent absence at the 10% threshold

		Unadjusted model	M2	M3
Odds ratio for KS4 persistent absence (at 10% threshold)	*Yes*	2.68**	2.31**	1.97**
95% CI		[2.09 – 3.45]	[1.77 – 3.01]	[1.50 – 2.58]

*Note. ** p<.01, * p<.05. Reference category: No persistent absence.*
M2 (adjusted for sex + exposure period + maternal smoking + housing tenure + mother's highest education + birth parity + mother age at delivery + neighbourhood IMD)
M2 (adjusted for Model 2 + school attainment + eligibility for FSM + SEN)

Appendix M: Analysis 2. School-level variables with suspensions/self-reported violence

		Unadj.		*M2*		*M3*		*M4*		*M5*	
KS4 any suspensions	*No*	Ref		Ref		Ref		Ref		Ref	
	Yes	6.260**	(1.383)	4.083**	(0.954)	4.055**	(0.949)	4.122**	(0.963)	4.071**	(0.955)
Quartiles: % sessions absent @ school-level	*Q1 (Lowest)*			Ref							
	Q2			1.033	(0.189)						
	Q3			0.893	(0.173)						
	Q4 (Highest)			1.032	(0.229)						
Quartiles: % eligible FSM	*Q1 (Lowest)*					Ref					
	Q2					1.181	(0.211)				
	Q3					0.974	(0.189)				
	Q4 (Highest)					1.189	(0.275)				
Quartiles: % achieve L2 threshold	*Q1 (Lowest)*							Ref			
	Q2							1.257	(0.257)		
	Q3							1.367	(0.276)		
	Q4 (Highest)							1.018	(0.232)		
Quartiles: size of the school year starting KS4	*Q1 (Lowest)*									Ref	
	Q2									1.226	(0.269)
	Q3									1.147	(0.260)
	Q4 (Highest)									1.199	(0.276)
Intercept		0.089**	(0.007)	0.025**	(0.013)	0.024**	(0.012)	0.021**	(0.011)	0.021**	(0.012)
var(_cons[ks4sch])		0.046	(0.050)	0.028	(0.048)	0.012	(0.044)	0.006	(0.042)	0.029	(0.049)
ICC		0.014	(0.015)	0.008	(0.014)	0.004	(0.013)	0.0003	(0.013)	0.009	(0.015)
AIC		1901.42		1854.11		1853.16		1851.20		1853.90	
N (individuals)		3189		3189		3189		3189		3189	
N (schools)		329		329		329		329		329	
Log likelihood		-947.71		-905.06		-904.58		-903.60		-904.95	

*Note. ** p<.01, * p<.05. Odds ratios and standard errors presented within the table.*

Appendix N: Analysis 2. School-level variables with absence/self-reported violence

		Unadj.		*M2*		*M3*		*M4*	*M5*	
KS4 persistent absence	*No*	Ref		Ref		Ref			Ref	
	Yes	3.560**	(1.032)	3.208**	(0.994)	3.174**	(0.981)		3.188	(0.993)
Quartiles: % sessions absent @ school-level	*Q1 (Lowest)*			Ref						
	Q2			1.037	(0.187)					
	Q3			0.891	(0.170)					
	Q4 (Highest)			1.007	(0.221)					
Quartiles: % eligible FSM	*Q1 (Lowest)*					Ref				
	Q2					1.179	(0.206)			
	Q3					0.943	(0.180)			
	Q4 (Highest)					1.152	(0.262)			
Quartiles: % achieve L2 threshold	*Q1 (Lowest)*									
	Q2									
	Q3									
	Q4 (Highest)									
Quartiles: size of the school year starting KS4	*Q1 (Lowest)*								Ref	
	Q2								1.244	(0.270)
	Q3								1.118	(0.250)
	Q4 (Highest)								1.154	(0.264)
Intercept		0.096**	(0.007)	0.027**	(0.014)	0.025**	(0.013)		0.023**	(0.012)
var(_cons[ks4sch])		0.039	(0.048)	0.024	(0.047)	0.007	(0.043)		0.026	(0.048)
ICC		0.012	(0.014)	0.007	(0.014)	0.002	(0.013)		0.008	(0.014)
AIC		1942.18		1874.11		1872.95			1873.79	
N (individuals)		3189		3189		3189			3189	
N (schools)		329		329		329			329	
Log likelihood		-968.09		-915.05		-914.48			-914.90	

*Note. Model 4 failed to converge and is therefore not presented.** p<.01, * p<.05. Odds ratios and standard errors presented within the table*

Appendix O: Analysis 2. School-level variables with suspensions/police record for any offence

		Unadj.		*M2*		*M3*		*M4*		*M5*	
KS4 any suspensions	*No*	Ref						Ref			
	Yes	8.182**	(1.238)					5.000**	(0.811)		
Quartiles: % sessions absent @ school-level	*Q1 (Lowest)*										
	Q2										
	Q3										
	Q4 (Highest)										
Quartiles: % eligible FSM	*Q1 (Lowest)*										
	Q2										
	Q3										
	Q4 (Highest)										
Quartiles: % achieve L2 threshold	*Q1 (Lowest)*							Ref			
	Q2							1.020	(0.192)		
	Q3							0.883	(0.178)		
	Q4 (Highest)							0.782	(0.195)		
Quartiles: size of the school year starting KS4	*Q1 (Lowest)*										
	Q2										
	Q3										
	Q4 (Highest)										
Intercept		0.031**	(0.003)					0.010**	(0.005)		
var(_cons[ks4sch])		0.107	(0.075)					0.012	(0.046)		
ICC		0.031	(0.021)					0.004	(0.014)		
AIC		1986.30						1912.55			
N (individuals)		6223						6223			
N (schools)		172						172			
Log likelihood		-990.15						-934.28			

*Note. Models 2,3 and 5 failed to converge and are not presented; results were similar to M4. ** p<.01, * p<.05. Odds ratios and standard errors presented within the table.*

Appendix P: Analysis 2. School-level variables with absence/ police record for any offence

		Unadj.		*M2*		*M3*		*M4*		*M5*	
KS4 persistent absence	*No*	Ref		Ref				Ref		Ref	
	Yes	4.524**	(0.828)	2.863**	(0.566)			2.830**	(0.559)	2.881**	(0.569)
Quartiles: % sessions absent @ school-level	*Q1 (Lowest)*			Ref							
	Q2			1.049	(0.213)						
	Q3			1.079	(0.212)						
	Q4 (Highest)			0.752	(0166)						
Quartiles: % eligible FSM	*Q1 (Lowest)*										
	Q2										
	Q3										
	Q4 (Highest)										
Quartiles: % achieve L2 threshold	*Q1 (Lowest)*							Ref			
	Q2							1.130	(0.210)		
	Q3							0.896	(0.179)		
	Q4 (Highest)							0.819	(0.202)		
Quartiles: size of the school year starting KS4	*Q1 (Lowest)*									Ref	
	Q2									0.791	(0.147)
	Q3									0.800	(0.159)
	Q4 (Highest)									0.657	(0.144)
Intercept		0.037**	(0.003)	0.008**	(0.005)			0.009**	(0.005)	0.012**	(0.007)
var(_cons[ks4sch])		0.090	(0.065)	0.004				0.012		0.001	
ICC		0.027	(0.019)	0.001	(0.014)			0.003	(0.013)	0.0002	(0.013)
AIC		2085.38		1972.49				1973.82		1972.63	
N (individuals)		6223		6223				6223		6223	
N (schools)		172		172				172		172	
Log likelihood		-1039.69		-964.25				-964.91		-964.32	

*Note. Model 3 failed to converge and is therefore not presented. ** p<.01, * p<.05. Odds ratios and standard errors presented within the table.*

Appendix Q: Analysis 3. Suspensions/self-reported violence model results

		Unadj. Model SRV		*Adj. Model 2 SRV*	
KS4 any suspensions	*Yes*	4.855	**	2.360	**
		2.81-8.38		1.27-4.40	
Sex of child	*Male*			2.875	**
				2.09-3.96	
Exposure period	*Year 11*			1.149	
				0.81-1.64	
Age of mother at delivery	*Per increase of 1 year*			1.002	
				0.97-1.04	
Housing tenure	*Other*			1.119	
				0.72-1.73	
Birth parity (0 or 1+)	*1+*			0.939	
				0.69-1.28	
Mum smoker	*Yes*			1.264	
				0.83-1.92	
Mother highest education	*None/Vocational/O Level*			1.249	
				0.90-1.73	
IMD Score 2004, quintiles	*2*			1.179	
				0.80-1.73	
	3			1.122	
				0.74-1.71	
	4			0.997	
				0.60-1.65	
	Most deprived			1.372	
				0.73-2.60	
Is pupil eligible for FSM?	*Yes*			1.324	
				0.58-3.00	
SEN indicated in NPD	*Yes*			1.092	
				0.72-1.66	
Quartiles of KS3 total score (reverse scored)	*Q2*			1.278	
				0.88-1.86	
	Q3			1.420	
				0.94-2.15	
	Lowest (Q4)			1.132	
				0.67-1.91	
Mum or partner ever in trouble with law	*Yes*			1.268	
				0.90-1.79	
Child smoked cigarettes as teenager	*Yes*			2.074	**

				1.45-2.97	
Any ACEs 5-11yrs	*Yes*			1.450	*
				1.07-1.97	
High (top 10%) TDS score	*Yes*			1.488	
				0.92-2.40	
Mother interested in what child does at school	*Yes, mostly/No, not really*			1.304	
				0.87-1.95	
Intercept		0.090	**	0.020	**
		0.07-0.10		0.01-0.07	
Number of observations		2390		2390	
AIC		1405.35		1350.40	
Log pseudolikelihood		-700.68		-652.20	
Pseudo R-squared		0.02		0.09	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no suspension, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score, mum or her partner never in trouble with law, child did not smoke cigarettes, no ACEs at age 5-11 years, did not have high TDS, mother very interested in what child does at school.*

Appendix R: Analysis 3. Absence/self-reported violence model results

		Unadj. Model SRV	***Adj. Model 2 SRV***
KS4 persistent absence	*Yes*	2.644 **	2.093
		1.20-5.82	0.87-5.01
Sex of child	*Male*		3.026 **
			2.20-4.17
Exposure period	*Year 11*		1.161
			0.82-1.64
Age of mother at delivery	*Per increase of 1 year*		1.001
			0.97-1.04
Housing tenure	*Other*		1.150
			0.74-1.78
Birth parity (0 or 1+)	*1+*		0.939
			0.69-1.28
Mum smoker	*Yes*		1.281
			0.85-1.93
Mother highest education	*None/Vocational/O Level*		1.266
			0.91-1.75
IMD Score 2004, quintiles	*2*		1.178
			0.80-1.73
	3		1.130
			0.75-1.71
	4		0.983
			0.60-1.62
	Most deprived		1.369
			0.72-2.61
Is pupil eligible for FSM?	*Yes*		1.325
			0.60-2.93
SEN indicated in NPD	*Yes*		1.133
			0.75-1.71
Quartiles of KS3 total score (reverse scored)	*Q2*		1.260
			0.87-1.83
	Q3		1.433
			0.95-2.17
	Lowest (Q4)		1.169
			0.70-1.96
Mum or her partner ever in trouble with law	*Yes*		1.273
			0.90-1.80
Child smoked cigarettes as teenager	*Yes*		2.176 **

				1.53-3.09	
Any ACEs 5-11yrs	*Yes*			1.461	*
				1.08-1.98	
High (top 10%) TDS score	*Yes*			1.507	
				0.95-2.41	
Mother interested in what child does at school	*Yes, mostly/No, not really*			1.327	
				0.88-1.99	
Intercept		0.095	**	0.019	**
		0.08-0.11		0.01-0.06	
Number of observations		2390		2390	
AIC		1426.25		1355.20	
Log pseudolikelihood		-711.12		-654.60	
Pseudo R-squared		0.00		0.08	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no suspension, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score, mum or her partner never in trouble with law, child did not smoke cigarettes, no ACEs at age 5-11 years, did not have high TDS, mother very interested in what child does at school.*

Appendix S: Analysis 3. Suspensions/police record model results

		Unadj. Model SRV		***Adj. Model 2 SRV***	
KS4 any suspensions	*Yes*	9.629	**	4.684	**
		5.65-16.42		2.52-8.70	
Sex of child	*Male*			1.934	**
				1.18-3.17	
Exposure period	*Year 11*			1.468	
				0.86-2.51	
Age of mother at delivery	*Per increase of 1 year*			1.000	
				0.94-1.06	
Housing tenure	*Other*			1.117	
				0.59-2.12	
Birth parity (0 or 1+)	*1+*			1.099	
				0.67-1.81	
Mum smoker	*Yes*			1.065	
				0.53-2.12	
Mother highest education	*None/Vocational/O Level*			0.704	
				0.41-1.20	
IMD Score 2004, quintiles	*2*			1.035	
				0.57-1.88	
	3			0.627	
				0.31-1.27	
	4			1.106	
				0.53-2.30	
	Most deprived			1.322	
				0.55-3.21	
Is pupil eligible for FSM?	*Yes*			1.939	
				0.74-5.11	
SEN indicated in NPD	*Yes*			1.463	
				0.83-2.57	
Quartiles of KS3 total score (reverse scored)	*Q2*			1.168	
				0.60-2.29	
	Q3			1.105	
				0.54-2.26	
	Lowest (Q4)			2.269	*
				1.17-4.40	
Mum or her partner ever in trouble with law	*Yes*			1.572	
				0.95-2.60	
Child smoked cigarettes as teenager	*Yes*			2.029	**

				1.27-3.24	
Any ACEs 5-11yrs	*Yes*			1.131	
				0.72-1.77	
High (top 10%) TDS score	*Yes*			1.336	
				0.72-2.47	
Mother interested in what child does at school	*Yes, mostly/No, not really*			1.803	*
				1.03-3.17	
Intercept		0.021	**	0.007	**
		0.02-0.03		0.00-0.04	
Number of observations		3066		3066	
AIC		710.40		698.09	
Log pseudolikelihood		-353.20		-326.04	
Pseudo R-squared		0.07		0.14	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no suspension, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score, mum or her partner never in trouble with law, child did not smoke cigarettes, no ACEs at age 5-11 years, did not have high TDS, mother very interested in what child does at school.*

Appendix T: Analysis 3. Absence/police record model results

		Unadj. Model SRV		***Adj. Model 2 SRV***	
KS4 persistent absence	*Yes*	6.850	**	3.470	**
		3.48-13.53		1.52-7.92	
Sex of child	*Male*			2.122	**
				1.31-3.44	
Exposure period	*Year 11*			1.461	
				0.87-2.46	
Age of mother at delivery	*Per increase of 1 year*			0.994	
				0.94-1.05	
Housing tenure	*Other*			1.137	
				0.58-2.24	
Birth parity (0 or 1+)	*1+*			1.098	
				0.67-1.80	
Mum smoker	*Yes*			1.108	
				0.57-2.14	
Mother highest education	*None/Vocational/O Level*			0.711	
				0.42-1.21	
IMD Score 2004, quintiles	*2*			0.970	
				0.53-1.76	
	3			0.660	
				0.32-1.35	
	4			1.046	
				0.52-2.12	
	Most deprived			1.247	
				0.52-3.00	
Is pupil eligible for FSM?	*Yes*			1.902	
				0.76-4.80	
SEN indicated in NPD	*Yes*			1.564	
				0.91-2.69	
Quartiles of KS3 total score (reverse scored)	*Q2*			1.225	
				0.63-2.39	
	Q3			1.145	
				0.56-2.34	
	Lowest (Q4)			2.356	*
				1.23-4.53	
Mum or her partner ever in trouble with law	*Yes*			1.568	
				0.96-2.57	
Child smoked cigarettes as teenager	*Yes*			2.224	**

				1.40-3.55	
Any ACEs 5-11yrs	*Yes*			1.183	
				0.76-1.84	
High (top 10%) TDS score	*Yes*			1.595	
				0.89-2.87	
Mother interested in what child does at school	*Yes, mostly/No, not really*			1.935	*
				1.13-3.33	
Intercept		0.024	**	0.007	**
		0.02-0.03		0.00-0.04	
Number of observations		3066		3066	
AIC		738.54		711.03	
Log pseudolikelihood		-367.27		-332.51	
Pseudo R-squared		0.03		0.12	

*Note. ** p<.01, * p<.05. Odds ratios and 95% CI presented within the table. Reference categories: no persistent absence, female, Year 10, homeowner, no previous births, non-smoker, degree/A-level qualifications, least deprived, not eligible for FSM, no SEN indicated, highest KS3 total score, mum or her partner never in trouble with law, child did not smoke cigarettes, no ACEs at age 5-11 years, did not have high TDS, mother very interested in what child does at school.*