

Early Socio-Emotional Skills and Adolescent Offending: Evidence from Administrative Data

Paul Garcia

Institute for Social and Economic Research
University of Essex

No. 2026-07
July 2026



INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH



Economic
and Social
Research Council



University of Essex



ADR UK

Non-Technical Summary

Children with skill deficits are more likely to come into contact with the criminal justice system during adolescence. Using linked education and justice records for pupils in England, this paper examines how early socio-emotional difficulties identified at the start of primary school are associated with later offending, and through which school-specific pathways these effects operate.

Using teacher assessments collected at the end of Reception, the analysis identifies two distinct types of early skill deficits: cognitive (related to language, literacy, numeracy and problem-solving) and socio-emotional (related to emotional regulation, social interaction and communication). Both are associated with a higher probability of being cautioned or convicted between ages 11 and 17, but they operate through different pathways. Cognitive difficulties increase offending risk primarily through lower academic attainment. Socio-emotional difficulties operate largely through school exclusions and retain a direct association with offending that persists even after accounting for differences in school and family environments. They are also associated with the type and frequency of offending, including violent and property-related crimes.

School exclusion emerges as the strongest predictor of adolescent offending, more than doubling the likelihood of a caution or conviction and associated with around 0.7 additional offences on average. This finding highlights the potential harms of removing children from school and suggests that exclusionary discipline may play an important role in the pathway from early disadvantage to crime.

Children with identified Special Educational Needs and Disabilities or known to Children in Need services are less likely to offend than children with similar difficulties who have not been formally identified. This suggests that timely identification and support can be protective, and that unmet need may be a key driver of the link between early skill deficits and later offending.

The findings point to the need for differentiated policy responses. Academic support is particularly important for children with cognitive difficulties, while behavioural and community-based approaches, including alternatives to exclusionary discipline such as restorative and relational practices, may be more relevant for children with socio-emotional difficulties. Early identification of need is critical in both cases.

Early Socio-Emotional Skills and Adolescent Offending: Evidence from Administrative Data[‡]

Paul Garcia[§]

14th July 2026

Abstract

This paper examines how early socio-emotional skills shape adolescent offending in England and the school pathways through which they operate. I use administrative data linking teacher-reported child development assessments at age five to school and criminal justice records through age 17. Factor analysis of these assessments reveals two underlying domains of early difficulty, cognitive and socio-emotional. Both raise the probability of a caution or conviction between ages 11 and 17 by around 15 %, but through distinct channels. Cognitive difficulties act primarily through academic attainment, while socio-emotional difficulties operate largely through school exclusions yet retain a substantial direct effect, accounting for around 40% of the total effect. The socio-emotional effect is robust to unobserved school and family heterogeneity and shapes both the type and frequency of offending. Statutory support for learning difficulties and social care services attenuates the associations between early skill deficits and offending. These findings suggest that differentiated support is needed based on the type of difficulty.

JEL Classification: I21, I28, J13, J24, K42

Keywords: skill deficits, socio-emotional difficulties, school exclusions, adolescent offending

[‡] I would like to thank Emilia Del Bono, Ben Etheridge, Alice Wickersham for helpful discussions. I am grateful for comments and suggestions received at the following conferences: ONS RES (2026), ADR UK (Cardiff, 2025), EUROCRIM (Athens, 2025), and MiSoC public event on *Schools, Behaviours, Skills and Sanctions* (London, 2025), particularly from representatives of Transforming Conflict, the Centre for Emotional Health, and the data owners. This work was funded by ADR UK, award number ES/S012486/1. This work was undertaken in the Office for National Statistics Secure Research Service using Data First linked administrative data. The use of the data in this work does not imply the endorsement of the ONS or data owners (e.g. MoJ and DfE) in relation to the interpretation or analyses of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. National statistics follow consistent statistical conventions over time and cannot be compared to Data First linked datasets. All mistakes and omissions are the authors' responsibility.

[§]University of Essex. E-mail: pagarcia@essex.ac.uk

1 Introduction

Juvenile crime has long been a policy priority, given the high social and economic costs associated with offending.¹ While juvenile crime declined substantially in the UK between 2014 and 2021, this decline has slowed in the years following the pandemic, while offences committed by younger cohorts appear to involve more serious and violent behaviour. Recent figures from the Metropolitan Police show rising violent crime among children aged 10–14 and a widening gap between children’s and adults’ reoffending rates (32% versus 25%).² Together, these patterns suggest that children entering the justice system today may be engaging in more serious violence at younger ages and are more likely to persist in offending, highlighting the urgency of identifying early-life factors that place children at risk.

Early determinants that shape children’s life and offending trajectories, including family environment, socio-economic status, and education, are well documented in the literature (Del Bono et al., 2016; Dickson et al., 2016; Francesconi & Heckman, 2016). From a policy perspective, education is particularly important because it can be more directly influenced than other determinants. Economists have extensively documented the crime-reducing effects of keeping young people occupied in school and off the streets (Hjalmarsson et al., 2015; Landersø et al., 2017; Machin & Sandi, 2025; Machin et al., 2011, 2023; Meghir et al., 2023). But education reduces crime through more than just incapacitation. Policies that raise school academic standards and student performance reduce pupils’ long-run adult crime, an effect attributed to gains in human capital rather than time spent in the classroom or changes in school resourcing (Eren et al., 2025). However, human capital formation begins well before children enter formal education, and the skills children bring with them to the classroom go on to shape both their educational trajectories and their risk of offending.

Existing studies using administrative data find strong associations between school outcomes (e.g., attainment, exclusions) and adolescent crime, but overlook the underlying skills that shape these outcomes. Studies examining the role of early skills typically rely on survey data and face well-known limitations, including small and unrepresentative samples, reliance on self-reported crime, subjective rating bias in skill assessments, and the absence of objective educational outcomes (Bindler et al., 2020; Kremser, 2014).³ Moreover, prior work has not established the school-specific

¹In England and Wales, an estimated £17 billion is spent annually on late intervention, covering areas such as mental health, youth crime, and child neglect (Chowdry & Fitzsimons, 2016).

²See [Youth Justice Statistics: 2023 to 2024](#) for further details.

³Much of the empirical evidence draws on the developmental psychology literature, which focuses on severe antisocial traits such as low self-control and psychopathy, typically measured in late adolescence when these characteristics are most readily observed, leaving limited scope for early intervention (Armstrong et al., 2020; Keating, 2023).

pathways through which early skill deficits translate into later offending, leaving open the question of whether interventions should target low academic attainment or behavioural difficulties.

This paper addresses these gaps using linked administrative records from the Ministry of Justice and the Department for Education. I apply Exploratory Structural Equation Modelling to teacher-reported indicators of child development at the end of the Reception year (ages four or five), identifying two distinct latent dimensions of early skill deficits: cognitive (language, literacy, numeracy, and problem-solving) and socio-emotional (emotional regulation, social interaction, and communication). I then trace how each dimension relates to adolescent offending between ages 11 and 17, including offence type and frequency, through school-specific pathways such as attainment, absenteeism, and exclusions.

Identification combines a rich set of individual controls with school and sibling fixed effects, allowing me to address the most plausible sources of time-invariant unobserved heterogeneity arising from shared environments. School fixed effects control for unobserved institutional factors, while sibling fixed effects, estimated using a within-family proxy identifier, net out shared parental and household influences.

This study makes three main contributions to the literature. First, a bidimensional representation of early skills reveals distinct pathways into offending. While being one standard deviation higher in either cognitive or socio-emotional difficulties is associated with a 15% increase in the probability of adolescent offending (baseline of 2%), cognitive difficulties influence offending primarily through lower academic attainment, while socio-emotional difficulties operate largely through school exclusions. The latter reflects the nature of the socio-emotional construct, capturing poor emotional regulation, difficulties in forming relationships, and in communicating effectively, skills critical for managing conflict and preventing behavioural incidents from escalating into disciplinary actions. Socio-emotional difficulties also retain a meaningful direct association with offending, accounting for approximately 40% of the total association, even after accounting for several school-mediated pathways and conditioning on school and shared family environments. They are further associated with a broad range of offences, including both violent and property-related crimes, as well as overall offending frequency.

Because socio-emotional difficulties operate primarily through exclusions, quantifying exclusion's own association with offending is a natural second contribution. While prior studies have documented the over-representation of excluded pupils in the offending population (Department for Education & Ministry of Justice [DfE & MoJ], 2022; Fuller & McNally, 2023), evidence on its magnitude is limited. School exclusion emerges as the strongest and most consistent predictor of adolescent offending. Beyond acting as a mediating pathway between early socio-emotional diffi-

culties and offending, it has a substantial association, more than doubling the baseline probability of offending and exceeding the magnitude of socio-economic disadvantage (e.g., eligibility for free school meals). School exclusion is also associated with an increase in the number of offences, with an average rise of approximately 0.7 additional offences. These patterns highlight the harmful consequences of removing children from school and speak to a broader literature on the role of school engagement in shaping offending outcomes (Bell et al., 2022; Hjalmarsson et al., 2015; Meghir et al., 2023) and the costs of harsh school discipline (Bacher-Hicks et al., 2024).

As a third contribution, I provide evidence that formal institutional support, through identification for Special Educational Needs and Disabilities (SEND) provision or Children in Need (CIN) services, may be protective for children with early skill deficits. Children with poor cognitive or socio-emotional skills who receive such support are less likely to offend than those with similar difficulties who do not. This is particularly relevant for socio-emotional difficulties, given their direct association with offending that cannot be fully explained by school-based mechanisms. For example, among children without SEND status, a one standard deviation increase in socio-emotional difficulties is associated with 21% higher odds of offending; for those with SEND, the effect falls to approximately 2%. The results suggest the skill–offending relationship is concentrated among children with poor skills but without identified needs, pointing to formal support as a potential mechanism associated with reduced offending among vulnerable children. More broadly, the findings point to the importance of early identification of needs for later-life outcomes (Carneiro et al., 2024).

Across these findings, socio-emotional skill deficits retain a direct effect on offending that is not accounted for by school pathways, and this effect survives rigorous specifications controlling for unobserved school and family heterogeneity in a way cognitive difficulties do not. This asymmetry motivates the case for tailored interventions, since addressing socio-emotional deficits requires more than school-based responses that may effectively tackle cognitive difficulties. Nevertheless, schools play an important role in the pathway between early socio-emotional difficulties and offending. Socio-emotional difficulties are a key predictor of school exclusions, which carry an independent association with offending. Consequently, improving these skills alone will not address the risk that exclusionary practice itself adds. This is particularly relevant given recent changes in school organisation in England, where the expansion of autonomous academies has been associated with more intensive disciplinary practices and higher exclusion rates (Machin & Sandi, 2020). These developments have also intensified policy discussion around relational and restorative practices as alternatives, although evidence on their effectiveness in the English context remains limited.

Taken together, these findings highlight the importance of intervening early while also considering the institutional settings in which children develop. Socio-emotional skills are particularly

malleable in early childhood (Alan et al., 2021; Sorrenti et al., 2025), developing cumulatively through experience and interacting dynamically with cognitive abilities (Heckman et al., 2006; OECD, 2015), which makes early investment valuable not only for reducing offending but also for breaking the intergenerational persistence of disadvantage linked to poor educational outcomes (Blanden et al., 2007; Bolt et al., 2024).

After reviewing the literature in Section 2, the paper proceeds as follows. Section 3 describes the data, Section 4 outlines the empirical methodology, and Section 5 presents the results from the factor analysis and examines the relationships between early skills, school outcomes, and offending. Section 6 concludes and discusses directions for future research.

2 Literature Review

The literature has well established the importance of childhood development and skill formation for shaping later life outcomes. In England and Wales, a primary source of information on child development is the Early Years Foundation Stage Profile (EYFSP). This and other similar frameworks have been used as screening tools to identify children with special educational needs and disabilities (SEND) (Hughes et al., 2018; Wood et al., 2024), predict school outcomes (Atkinson et al., 2022; Bonetti & Blanden, 2020; Department for Education [DfE], 2010a), and flag early neurodevelopmental conditions such as autism spectrum disorder (ASD) (Wright et al., 2019) and attention-deficit/hyperactivity disorder (ADHD) (Ter-Minassian et al., 2022). Evidence on the EYFSP as a predictor of behavioural problems is more limited. Using data from the Born in Bradford cohort, Mooney et al. (2025) find that a one-point increase in the EYFSP total score is associated with a 0.2-point decrease in the Strengths and Difficulties Questionnaire (SDQ) score.

The SDQ is widely used to assess problematic behaviours associated with attention deficits, emotional regulation and social relationships. Such behaviours are known to affect the accumulation of human capital (Cunha & Heckman, 2008; Currie & Almond, 2011), shape labour market outcomes, health, and well-being (Del Bono et al., 2024; Goodman et al., 2015; Heckman et al., 2013), and predict offending (Kautz & Zanoni, 2024; Murray et al., 2010, 2015). However, a large strand of the literature on youth offending, particularly in developmental psychology, has focused on more pronounced self-centred and antisocial behaviours, such as low self-control and psychopathy (Armstrong et al., 2020; Keating, 2023). These are typically assessed in late adolescence, when such behaviours are more evident but opportunities for intervention are more limited.

In the UK, evidence on early behaviours and offending has largely come from cohort studies. Using the BCS70, Murray et al. (2010) show that hyperactivity and internalising behaviours, such as emotional and peer problems, at age five predict criminal convictions from late adolescence to

adulthood among boys. Using a wider set of teacher-reported behavioural items from the BCS70 at age 10, Del Bono et al. (2024) identify additional behavioural domains, including attention and conduct problems, associated with adult offending. Using ALSPAC data, Murray et al. (2015) also show that conduct problems and hyperactivity predict adolescent crime, with comparable patterns observed in Brazil, pointing to the general relevance of early behaviours as predictors of offending across different contexts.

Offending is not driven by a single domain but by a combination of behavioural difficulties that often intersect with poor cognitive abilities. Sallis et al. (2019) highlight the role of the interaction between cognitive difficulties and poor emotional regulation in predicting offending, as limited problem-solving skills prevent effective coping with difficulties, increasing stress and the risk of antisocial personality development. Consistent with this perspective, research drawing on the Cambridge Study in Delinquent Development (CSDD) identifies low cognitive ability, poor concentration, high impulsivity, and high daring measured at age eight as key predictors of criminal behaviour (Basto-Pereira et al., 2024).

Evidence on early behaviours and offending drawn from administrative data remains relatively limited. Kent et al. (2025) show that the relationship between child development and offending is moderated by socio-economic status, with poverty amplifying the effects of lower functional development on the probability of being cautioned or convicted by age 15. It is not surprising that both early development and offending are strongly shaped by the intergenerational transmission of disadvantage and inadequate family support (Basto-Pereira et al., 2024; Farrington et al., 2016).⁴

Beyond family background, shared environments such as communities and schools also shape early behaviours and offending risk. The literature has extensively examined policies that disproportionately affect deprived areas, particularly austerity-driven cuts to local government spending between 2010 and 2015. These cuts reduced funding for a range of youth and community services, such as youth clubs (Villa, 2024; Webb & Bywaters, 2018), green spaces, leisure centres, and crime reduction initiatives (Fahy et al., 2023).

Time spent in school and school quality, including teacher inputs and peer influences, are closely associated with offending. Raising the compulsory school-leaving age reduces crime by keeping young people engaged in education (Bell et al., 2022; Hjalmarsson et al., 2015; Machin et al., 2011; Meghir et al., 2023), with similar effects found for policies encouraging post-compulsory education (Sabates & Feinstein, 2008). In contrast, exposure to high-delinquency peers or attending schools with higher suspension rates is associated with a higher likelihood of offending (Bacher-Hicks

⁴In England and Wales, pupils eligible for free school meals (FSM) are substantially over-represented among the offending population, and around 30% of those cautioned or convicted have been in contact with social care services (DfE & MoJ, 2022).

et al., 2024; Deming, 2011; Farrington et al., 2016). This partly reflects the skill and behavioural composition that pupils bring to the classroom, characteristics that also shape educational and offending outcomes.⁵

Early behaviours can also provide insights into whether offending is limited to adolescence or persists into adulthood.⁶ Life-course-persistent offending is characterised by neurocognitive difficulties, adverse family environments, and temperamental risks emerging in childhood, while adolescence-limited offending reflects a temporary mismatch between biological maturation and social perceptions of maturity, with antisocial behaviour typically declining once adolescents gain access to adult roles (Moffitt, 2018).

Childhood behaviours are also linked to the severity of offending. Around 20% of recorded offences committed by children are classified as serious violent offences, defined as those involving the causing, or intent to cause, serious physical harm (DfE & MoJ, 2022). Children involved in such offences are more likely to have identified educational needs and to experience more disrupted schooling, including school changes, exclusions, and persistent absence, than those involved in other offence types (Fuller & McNally, 2023).

Finally, there is broad consensus that early interventions reduce offending risk among children, including involvement in violent crime (Heller et al., 2017), though the optimal timing remains an open question. Behavioural and cognitive abilities shaping life outcomes can be identified as early as 22 months (Cunha & Heckman, 2008; Feinstein, 2003). Pre-school programmes supporting child development in the UK, such as Sure Start, have been shown to improve attainment and promote earlier identification of SEND (Carneiro et al., 2024). More broadly, programmes enhancing early skills generate lasting benefits across educational, health, and labour market outcomes (Heckman et al., 2013; Heckman & Karapakula, 2019; OECD, 2015). Non-cognitive skills, in particular, remain highly malleable throughout childhood and adolescence (Alan et al., 2021; Del Bono et al., 2016; Sorrenti et al., 2025) and are, in some cases, more predictive of offending than cognitive skills (Heckman et al., 2006; Kautz & Zanoni, 2024).

⁵A high proportion of children who are cautioned or convicted have previously been suspended or excluded, or failed to meet expected standards in English and mathematics (Cornish & Brennan, 2025; DfE & MoJ, 2022; Fuller & McNally, 2023). The latter is consistent with the idea that pupils with poor academic abilities face lower opportunity costs of leaving education, increasing their exposure to offending (Dorsett et al., 2022).

⁶According to DfE & MoJ (2022), 16% of children cautioned or convicted for an offence are classified as prolific offenders, defined as having four or more encounters with the criminal justice system between ages 10 and 17.

3 Data

This paper uses the Ministry of Justice and Department for Education (MoJ-DfE) linked dataset, which combines pupils' educational records attending state-maintained schools in England from the National Pupil Database (NPD) with offending records from the Police National Computer (PNC).⁷ The sample includes all children who entered Reception, the first year of compulsory schooling, in academic years 2006/07, 2007/08, or 2008/09.

Early Years Foundation Stage Profile (EYFSP): I use the original version of the EYFSP, which was in place up to 2012, a teacher-reported assessment of children's development at the end of the Reception year (age five).⁸ Development within each EYFSP scale is assessed using nine binary learning goals (e.g., achieved vs. not achieved) yielding a scale score from 1 to 9. Learning goals 1–3, referred to as *stepping stone points*, are hierarchical and capture foundational developmental milestones. Subsequent learning goals (4–8) are non-hierarchical and exhibit the largest variation in attainment across children (DfE, 2010a). The highest learning goal (point 9) reflects performance beyond the expected level and is awarded only if the pupil has achieved all eight preceding goals; 5–13% of pupils reach this level, depending on the EYFSP scale.

For factor analysis, I focus on learning goals from the core scales used to define a *good level of development* (GLD): *Disposition and attitudes*, *Social development*, *Emotional development*, *Language and thinking*, *Linking sounds and letters*, *Reading*, and *Writing* (DfE, 2010a). I additionally include two numeracy-related scales, *Numbers as labels and counting* and *Calculating*, which capture aspects of cognitive ability. Further information on the EYFSP assessment and moderation process is provided in Appendix A.1.

I also use variation exclusively from learning goals 4–8 for two reasons. First, these are non-hierarchical and can therefore be treated as independent indicators in the factor analysis. Second, they exhibit the largest variation in attainment when comparing pupils who have been cautioned or convicted for an offence with those who have not. Table 1 reports percentage-point (pp) differences in attainment across learning goals within the selected EYFSP scales.

For example, the difference between children who later engage in offending and those who do not in achieving scale point 6 in *Reading*, corresponding to "Reads familiar words independently", is 24 pp. A similar gap is observed at scale point 7 in *Emotional Development*, corresponding to "Considers the consequences of words and actions". Detailed descriptions of the EYFSP learning goals of Table 1 are provided in Table A.1 in the Appendix.

⁷The linkage methodology and data-quality checks are described in Dillon et al. (2022a, 2022b).

⁸The revised EYFSP simplified the original framework by reducing the number of early learning goals from 69 to 17, capturing a child's level of development with fewer measures while reducing administrative burden (Tickell, 2011).

Table 1: Percentage-Point Differences in Attainment of EYFSP Learning Goals between Offenders and Non-Offenders

Scales	1	2	3	4	5	6	7	8	9	
Dispositions & attitudes	*	0.1	-0.6	-0.3	-1.7	-12.8	-8.8	-22.5	-7.9	
Social development	*	0.1	-0.2	-9.1	-7.4	-17.5	-21.0	-16.6	-6.9	
Emotional development	*	-0.2	-3.0	-4.8	-15.3	-19.6	-22.8	-20.5	-6.6	
Language & thinking	-0.1	0.1	-1.7	-5.7	-5.9	-11.1	-17.3	-12.4	-5.6	
Linking sounds, letters	-0.1	-3.4	-1.6	-15.6	-13.9	-18.7	-22.1	-18.5	-8.5	[5, 10) pp
Reading	0.0	-0.2	-7.0	-6.7	-13.5	-24.3	-18.8	-16.0	-5.9	[10, 15) pp
Writing	*	-2.3	-7.7	-18.2	-18.8	-21.8	-22.0	-16.2	-4.0	≥ 15 pp
Numbers & counting	0.1	-0.1	-1.5	-2.9	-10.8	-8.4	-15.4	-16.2	-9.3	
Calculating	0.1	-0.5	-4.8	-8.4	-13.7	-17.1	-18.0	-12.8	-4.0	
Number of observations	650,820									

Note: MoJ–DfE data. The table shows the percentage-point difference in attainment of EYFSP learning goals between children who have ever been cautioned or sentenced for an offence and those with no recorded offending between ages 11 and 17. Differences are estimated for pupils with complete information on EYFSP learning goals across three cohorts of pupils (reception years 2006/07, 2007/08 and 2008/09). Total number of observations is rounded to the nearest 10 and cells marked with an asterisk (*) are suppressed for confidentiality.

Altogether, the nine EYFSP scales and five learning goals per scale (scale points 4–8) yield 45 indicators, available for approximately 40% of pupils. Unlike total EYFSP scores, which are observed for the full population, disaggregated point-scale data were not fully collected across all years of analysis. Between 2007 and 2009, schools were under a statutory obligation to submit total EYFSP scores, while the submission of individual learning goal data remained voluntary (DfE, 2010a). As a result, complete data are available for around 36% of pupils in the 2006/07 academic year, increasing by approximately 6 pp per year to reach 48% in the 2008/09 cohort. Despite partial coverage, these data are considered nationally representative, with ongoing evaluations consistently confirming their reliability (DfE, 2010b).

Key Stage 2 (KS2): Pupils undertake standardised national tests in English and Mathematics at the end of primary education, at age 11. For the cohorts of interest, these assessments correspond to the 2012/13, 2013/14, and 2014/15 academic years, respectively.

I construct a continuous standardised KS2 score from test results in Mathematics, Grammar, Punctuation and Spelling, and Reading, following conventions in the literature (Dorsett et al., 2022). As these tests are scored on different scales, each subject score is first standardised, then summed, and the total is standardised again. I also construct a binary indicator for whether the pupil achieved Level 4, the expected standard prior to 2016, or above in the Key Stage 2 assessments. The indicator takes the value 1 if the pupil met this threshold in at least one subject, and 0 if they failed to achieve Level 4 in all subjects.

Absenteeism and School Exclusions: I construct a composite measure of school disengagement capturing absenteeism and disruptions during primary school. This standardised score is derived from the first principal component of a PCA on authorised and unauthorised absences (as a percentage of all possible sessions), number of fixed-term exclusions, and number of establishments attended from Reception to KS2 assessment. Table A.2 in the Appendix presents the component loadings for each variable.

School exclusions are measured using a binary indicator equal to one if a pupil experienced either a temporary or permanent exclusion over the same period.⁹

Special Educational Needs and Disabilities (SEND) and Children in Need (CIN): I identify whether a pupil was ever recorded as having SEND or as a CIN at any point between Reception and the KS2 assessment year. Although some children are identified prior to Reception, often through early years settings or health and social services, many are identified later as learning difficulties, disabilities, or safeguarding concerns become more visible during primary education.

The SEND indicator includes pupils identified as having SEND both with and without specific provision, including those receiving additional support through an Education, Health and Care Plan (EHCP).¹⁰ Analogously, the CIN indicator includes pupils with and without a child protection plan (CPP), including looked-after children (CLA).¹¹

Offending: I identify whether a pupil received a caution or conviction for an offence during adolescence, between ages 11 and 17.¹² For each year, I retain the primary offence with the highest disposal rank, which orders disposals by severity. I also retain cases with outcomes recorded as "Guilty" or "Caution/Warning/Reprimand".

I further use the Home Office (HO) Offence Code to construct four broad offence categories: *Violent Offences*, indicating cautions or convictions for assault, sexual offences, or robbery; *Drug and Weapon Offences*, covering substance-related offences and possession of weapons; *Property and Societal Offences*, including theft, fraud, criminal damage and arson, public-order offences, and other offences against society; and *Summary Offences*, covering minor and motoring offences. Finally, the number of offences is used as a proxy for prolific offending, covering both serious and non-serious violent offences accumulated during adolescence.

⁹Permanent exclusions are rare during primary education; exclusions of any type have a prevalence of around 2%.

¹⁰See the [SEND Code of Practice](#) for further information.

¹¹Since December 2012, children convicted of an offence are classified as CLA under the Legal Aid, Sentencing and Punishment of Offenders Act 2012 (DfE & MoJ, 2022). In this paper, CLA status is restricted to the primary school period, prior to any recorded offending.

¹²Around 1% of pupils with offending records are observed before age 11 and are excluded to ensure offending does not precede the measurement of key explanatory variables. Their exclusion does not meaningfully affect the estimates.

Socio-economic and demographic characteristics: Most socio-economic and demographic characteristics are recorded at the time of the EYFSP assessment.¹³ These include gender (indicator for female), month and year of birth, ethnic background (indicator for non-White ethnicity, including Asian, Black, Chinese, Mixed, and other ethnic groups), first language (indicator for first language other than English, with few cases labelled as "not known but believed to be other than English" or "unclassified" being treated as missing), part-time school attendance, eligibility for free school meals (FSM)¹⁴, deciles of the Income Deprivation Affecting Children Index (IDACI), and identifiers for the school's Local Authority and unique reference number.

Further details on data cleaning and missingness are provided in Appendix A.2 and summary statistics for the main variables are reported in Table A.3.

4 Methodology

4.1 Factorization

I begin by analysing the EYFSP data as detailed in Section 3. It is important to note that the EYFSP is not a psychometric instrument designed to capture well-defined cognitive and non-cognitive domains; rather, it is a curriculum-based assessment aimed at identifying early developmental difficulties. Moreover, non-cognitive domains at such a young age (4–5 years at assessment) are likely to be less clearly differentiated than they are later in childhood. With this in mind, I use Exploratory Structural Equation Modelling (ESEM), which combines the flexibility of Exploratory Factor Analysis (EFA) with the confirmatory structure of Confirmatory Factor Analysis (CFA) within a structural equation modelling framework.

First, EFA is used to uncover the underlying dimensions of early skills, addressing the absence of an established theoretical structure for the EYFSP.¹⁵ The analysis is implemented through an iterative procedure that begins with all items from the selected scales and successively removes items with low factor loadings within each dimension ($|\lambda| < 0.5$) or high cross-loading ratios between dimensions (exceeding 0.75). The number of factors retained at each iteration is determined using a combination of the Kaiser criterion (retaining factors with eigenvalues greater than one) and

¹³Depending on the variable, between 90% and 97% of observations are available at Reception; for the remaining cases, the first available value observed in later years is used.

¹⁴Eligibility is confirmed by the local authority or school and typically applies in cases where one or both parents receive Universal Credit (with income below £7,400), Income Support (including Jobseeker's Allowance), support under Part VI of the Immigration and Asylum Act 1999, and similar benefits.

¹⁵I begin by estimating a tetrachoric correlation matrix, as the EYFSP indicators are binary (achieved vs. not achieved), and then apply an oblique rotation to the factor solution, as the underlying constructs are expected to be correlated.

inspection of the scree plot (retaining factors up to the elbow of the eigenvalue curve, where the sharp initial decline levels off). The process concludes when no further items are removed and a stable factor structure emerges.

The final factor structure identified through EFA is then estimated using CFA within ESEM, where all retained loadings are freely estimated. Unlike traditional CFA, this approach allows indicators to load on multiple domains while maintaining a formal structure for inference on latent relationships, better reflecting the overlapping nature of EYFSP items across scales.¹⁶ Formally, let Z_{ij}^* denote the unobserved continuous response underlying the binary indicator Z_{ij} for individual i and item j :

$$Z_{ij}^* = \sum_{k=1}^K \lambda_{jk} \omega_{ik} + \varepsilon_{ij}, \quad (1)$$

where λ_{jk} denotes the loading of item j on factor k , ω_{ik} is the latent score of factor k for individual i , and ε_{ij} is the measurement error term.

The observed binary outcome is defined through the following threshold model:

$$Z_{ij} = \begin{cases} 1 & \text{if } Z_{ij}^* > \tau_j, \\ 0 & \text{otherwise,} \end{cases}$$

where item-specific thresholds τ_j are estimated jointly with the model.

For identification, I impose standard assumptions from the factor analysis literature: local independence of item residuals, $\text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = 0$ for all $j \neq j'$, and independence between the latent factors and measurement errors, $\text{Cov}(\omega_{ik}, \varepsilon_{ij}) = 0$ (Bolt et al., 2024; Del Bono et al., 2024; Heckman et al., 2013). Factor variances are standardised to 1, $\text{Var}(\omega_k) = 1$, for scale identification.¹⁷

Estimation is conducted using the weighted least squares mean- and variance-adjusted (WLSMV) estimator, which relies on tetrachoric correlations. Factor scores are then obtained

¹⁶In traditional CFA, fixing cross-loadings to zero can artificially inflate latent factor correlations and bias structural relationships (Asparouhov & Muthén, 2009). By allowing small cross-loadings, ESEM allocates shared variance more naturally, yielding factor scores that are less correlated and more clearly differentiated across domains.

¹⁷I implement ESEM using CFA with freely estimated factor loadings and starting values derived from an EFA with oblique geomin rotation. Identification is achieved by designating one anchor item per factor, where anchor items are those with a high primary loading on their respective factor and negligible cross-loadings, and fixing their loadings to the EFA values to preserve factor orientation. Imposing a target rotation structure provides a stronger a priori model and facilitates interpretation of the factor structure (Marsh et al., 2014).

using the Empirical Bayes Modal (EBM) method, which is particularly well suited to binary data, as posterior-based scoring methods outperform linear alternatives.¹⁸

4.2 Offending and School Outcomes

A general specification for modelling the key outcomes is given by

$$g(\mathbb{E}[y_i \mid \Omega_i, X_i]) = \Omega_i' \theta + X_i' \beta + \mu_i, \quad (2)$$

where y_i denotes the outcome of interest for individual i (e.g., exclusions, absenteeism, attainment, or offending). The link function $g(\cdot)$ depends on the nature of the outcome: OLS is used for continuous outcomes, while logit or a linear probability model (LPM) is used for binary outcomes. Ω_i denotes the vector of early skills obtained from the EYFSP data, and X_i captures demographic characteristics, geographical location, and indicators of family background.¹⁹

To assess the contribution of school-specific pathways, I extend equation (2) by including school outcomes. Let M_i denote a vector of school-related outcomes, including absenteeism, exclusions, and attainment. The composition of M_i depends on the outcome y_i : when y_i is an offending outcome, M_i includes all three; when y_i is KS2 attainment, the mediators are restricted to absenteeism and exclusions. The augmented specification is as follows:

$$g(\mathbb{E}[y_i \mid \Omega_i, M_i, X_i]) = M_i' \alpha + \Omega_i' \theta + X_i' \beta + v_i, \quad (3)$$

For each outcome, I report estimates from both the baseline model in equation (2) and the augmented model in equation (3). Comparing the corresponding coefficients provides an initial indication of the extent to which schooling pathways account for the association between early skills and later outcomes. A formal mediation analysis, which quantifies these pathways within a path-decomposition framework, is presented next. For inference, standard errors are clustered at the school level to account for unobserved shocks and persistent features that are common to pupils attending the same school.

¹⁸Although the latent domains are constructed as continuous measures of the observed indicators, the binary nature of the EYFSP indicators is more appropriately captured by this scoring method, which uses the full likelihood of categorical responses rather than assuming continuous and normally distributed indicators.

¹⁹I approximate Ω_i using factor scores obtained from specification 1. These scores are noisy but information-rich proxies for the underlying latent domains. Any remaining measurement error in Ω_i is therefore likely to attenuate the estimated coefficients, implying that the reported associations can be interpreted as conservative estimates of the effects of the latent domains.

4.3 Mediation Framework

To examine the extent to which early skills influence offending through schooling pathways, I implement a mediation decomposition. Although several outcomes are binary, I estimate linear probability models (LPMs) throughout. Preliminary logistic models indicate that average marginal effects closely approximate LPM coefficients. Since the analysis focuses on the share of total effects transmitted through each pathway rather than absolute magnitudes, any approximation bias is likely to affect both total and mediated effects similarly, leaving mediation shares largely unaffected. The linear specification also avoids the path-dependence and interpretational complications inherent in non-linear mediation.

Level 1 Mediation: Early Skills and Offending via School Outcomes

The analysis begins by estimating the direct effects of socio-emotional difficulties, S , and poor cognitive skills, C , on the probability that pupil i has ever been cautioned or convicted for an offence, O_i , conditional on school outcomes. Formally:

$$O_i = \alpha_A A_i + \alpha_E E_i + \alpha_{KS} KS_i + \alpha_S S_i + \alpha_C C_i + \mathbf{X}_i' \boldsymbol{\kappa}^O + \mu_i^O, \quad (4)$$

where A_i represents absenteeism, E_i indicates school exclusion, KS_i denotes Key Stage 2 attainment, $\mathbf{X}_i$ denotes a vector of control variables including pupil demographics, socio-economic characteristics, and school fixed effects, and μ_i^O is an idiosyncratic error term.²⁰

The objective is to obtain the share of the total effect of early skills on offending that is explained by school outcomes. Totally differentiating equation (4) with respect to S_i and C_i then yields:

$$\begin{aligned} \frac{dO_i}{dS_i} &= \alpha_A \frac{dA_i}{dS_i} + \alpha_E \frac{dE_i}{dS_i} + \alpha_{KS} \frac{dKS_i}{dS_i} + \alpha_S, \\ \frac{dO_i}{dC_i} &= \alpha_A \frac{dA_i}{dC_i} + \alpha_E \frac{dE_i}{dC_i} + \alpha_{KS} \frac{dKS_i}{dC_i} + \alpha_C. \end{aligned}$$

I impose theoretically motivated restrictions on the mediation pathways: socio-emotional difficulties operate through absenteeism and exclusions, while cognitive difficulties operate through Key Stage 2 attainment and absenteeism, such that $\frac{dKS_i}{dS_i} = 0$ and $\frac{dE_i}{dC_i} = 0$. This specification is supported empirically. First, preliminary estimates of omitted pathways (e.g., cognitive difficulties affecting exclusions, or socio-emotional difficulties affecting attainment) are substantively small and appear to reflect suppression effects rather than meaningful mediation channels, as shown in

²⁰I rely on school data up to KS2 assessments to preserve a clear sequence from early skills to offending, limiting bias from the potential reverse relationship between offending and school outcomes after age 11. In fact, many pupils entering youth custody lack consistent school records; between 30 and 45% have missing records for their national examinations at age 16 (Machin et al., 2023).

the next section. Second, estimates from the restricted specification closely approximate those from the unrestricted model, indicating minimal loss of information.

Under these restrictions, the relationship between each school outcome and its corresponding skill is given by:

$$A_i = \beta_S^A S_i + \beta_C^A C_i + \mathbf{X}_i' \boldsymbol{\delta}^A + \eta_i^A, \quad (5)$$

$$E_i = \beta_S^E S_i + \mathbf{X}_i' \boldsymbol{\delta}^E + \eta_i^E, \quad (6)$$

$$KS_i = \beta_C^{KS} C_i + \mathbf{X}_i' \boldsymbol{\delta}^{KS} + \eta_i^{KS}, \quad (7)$$

where η_i^A , η_i^E , and η_i^{KS} are error terms. From these equations, it follows that

$$\frac{dA_i}{dS_i} = \beta_S^A, \quad \frac{dA_i}{dC_i} = \beta_C^A, \quad \frac{dE_i}{dS_i} = \beta_S^E, \quad \frac{dKS_i}{dC_i} = \beta_C^{KS},$$

Using equations (4)–(7), the components of the total effects of socio-emotional and cognitive difficulties on offending that operate through school outcomes are:

$$\frac{\partial O_i}{\partial A_i} \frac{dA_i}{dS_i} = \alpha_A \beta_S^A, \quad \frac{\partial O_i}{\partial A_i} \frac{dA_i}{dC_i} = \alpha_A \beta_C^A, \quad \frac{\partial O_i}{\partial E_i} \frac{dE_i}{dS_i} = \alpha_E \beta_S^E, \quad \frac{\partial O_i}{\partial KS_i} \frac{dKS_i}{dC_i} = \alpha_{KS} \beta_C^{KS}.$$

Figure 1 presents the Level 1 mediation structure. Socio-emotional difficulties operate through exclusions and absenteeism (solid orange arrows), while cognitive difficulties operate through absenteeism and Key Stage 2 attainment (solid teal arrows). The dashed arrows represent direct effects on offending that are not mediated by school outcomes.

To estimate the share of the skill effect on offending that operates through each school outcome, I first obtain the total effects of S_i and C_i from the reduced-form regression:

$$O_i = \rho_S S_i + \rho_C C_i + \mathbf{X}_i' \boldsymbol{\phi} + \epsilon_i, \quad (8)$$

where $\rho_S = \frac{dO_i}{dS_i}$ and $\rho_C = \frac{dO_i}{dC_i}$ represent the total effects of socio-emotional and cognitive difficulties on offending, respectively. The share of the socio-emotional effect transmitted through absenteeism is $\alpha_A \beta_S^A / \rho_S$, through exclusions is $\alpha_E \beta_S^E / \rho_S$, and the non-mediated share is α_S / ρ_S . Similarly, for cognitive difficulties, the shares transmitted through absenteeism and Key Stage 2 attainment are $\alpha_A \beta_C^A / \rho_C$ and $\alpha_{KS} \beta_C^{KS} / \rho_C$, with non-mediated share α_C / ρ_C .

This framework assumes that early skills affect offending through non-sequential school outcomes. However, absenteeism and exclusions may first influence attainment, which then constitutes the primary pathway to offending. To account for these sequential pathways, I extend the analysis

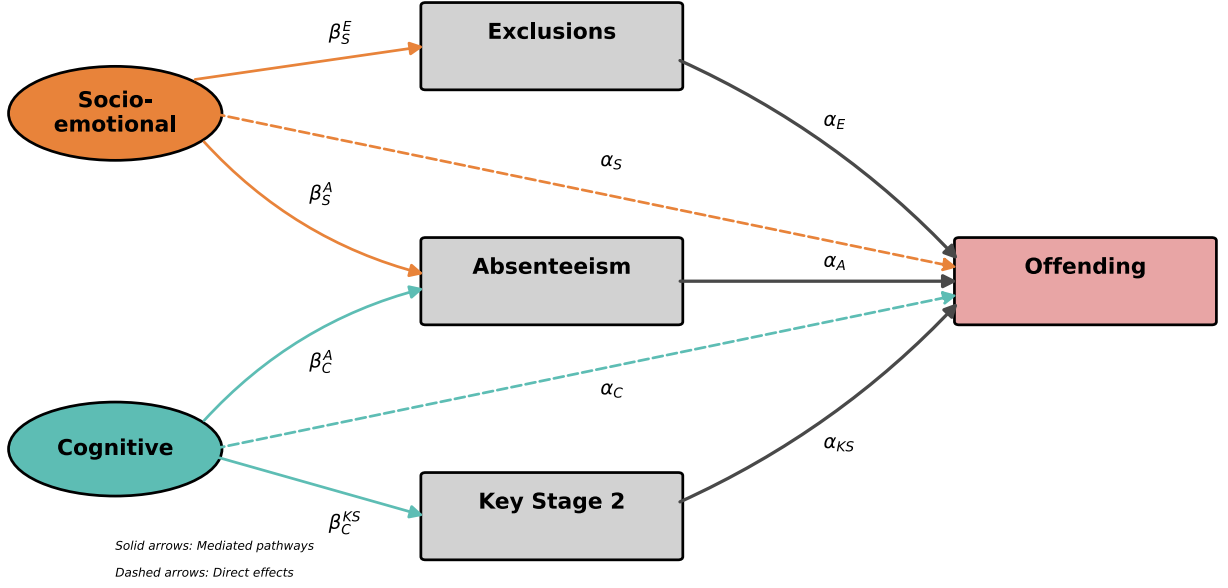


Figure 1: Early Skills and Offending via School Outcomes

Note: Figure shows the parameters of the Level 1 mediation structure, illustrating the non-sequential pathways through which socio-emotional and cognitive difficulties affect offending.

with a second-level mediation model allowing indirect effects of absenteeism and exclusions via Key Stage 2 attainment. The formal derivation is provided in Appendix A.3.

5 Results

5.1 Results of factorization

I consider 45 indicators for the factorization analysis as detailed in Section 3. After several iterations, 32 items remain in the final factor structure.²¹ Two factors with eigenvalues greater than one emerge. The scree plot is presented in Figure B.1 in the Appendix, and Table 2 reports the factor loadings for retained items, ordered from highest to lowest within each factor.

The retained latent constructs are labelled as *cognitive* and *socio-emotional* domains. As expected, numeracy abilities and early learning skills, such as reading and writing, load onto the cognitive domain, while items reflecting interpersonal relationships and collaborative work load onto the socio-emotional one. Interestingly, learning goals from *Communication and Thinking*, which are often associated with cognitive abilities, load more strongly onto the socio-emotional

²¹Nine of the 45 indicators are highly correlated ($\rho \geq 0.85$). In these cases, I retain a single indicator and drop the redundant measure. Four additional indicators are dropped due to factor loadings below 0.5 or high cross-loading ratios (secondary loading/primary loading) exceeding 0.75, following conventional thresholds in the literature (Bolt et al., 2024; Heckman et al., 2013).

Table 2: Factor Loadings from EFA using EYFSP data

EYFSP Scales	Point Scales	Cognition	Socio-Emotional
<i>Panel A. Cognition</i>			
Writing	Writes simple regular words	0.965	-0.104
Reading	Reads familiar words independently	0.940	-0.078
Calculating	Finds one more or one less than numbers 1–10	0.885	-0.046
Numbers and counting	Recognizes numerals 1 to 9	0.871	-0.029
Linking sounds and letters	Blends sounds in words	0.854	0.009
Numbers and counting	Uses math ideas to solve problems	0.834	-0.034
Writing	Writes own name and words from memory	0.830	0.045
Calculating	Uses addition/subtraction vocabulary	0.805	0.081
Calculating	Relates addition by combining groups	0.762	0.165
Reading	Understands information from non-fiction texts	0.758	0.011
Reading	Knows print directionality	0.726	0.217
Linking sounds and letters	Links sounds to letters	0.716	0.020
Reading	Retells narratives in sequence	0.705	0.147
Writing	Attempts writing for various purposes	0.702	0.163
Numbers and counting	Says number names in order	0.676	0.208
Reading	Understands story elements	0.662	0.270
Writing	Forms recognizable letters	0.615	0.154
<i>Panel B. Socio-Emotional</i>			
Social development	Forms good relationships with others	-0.139	0.956
Emotional development	Sensitive to the feelings of others	-0.068	0.937
Social development	Works as part of a group	-0.135	0.897
Emotional development	Responds with appropriate feelings	0.000	0.850
Social development	Understands agreed codes of behaviour	0.090	0.813
Emotional development	Considers consequences of words, actions	0.097	0.735
Dispositions and attitudes	Selects activities independently	0.142	0.685
Communication and thinking	Interacts in various contexts	0.256	0.656
Dispositions and attitudes	Interested, motivated to learn	0.263	0.609
Communication and thinking	Uses language to imagine and recreate	0.256	0.596
Emotional development	Respect for own and others' beliefs	0.242	0.585
Dispositions and attitudes	Dresses and undresses independently	0.125	0.584
Emotional development	Understands right and wrong	0.181	0.574
Social development	Respects different needs and cultures	0.284	0.559
Communication and thinking	Listens, responds to stories, rhymes	0.330	0.554
Primary items per construct		17	15
Cronbach's α		0.92	0.86
Number of observations			652,070

Notes: MoJ–DfE data. The table reports factor loadings from the final iteration of an Exploratory Factor Analysis (EFA) based on tetrachoric correlation matrices with an oblique quartimin rotation, estimated for all pupils with complete information on EYFSP learning goals and other relevant variables. Items with low loadings (< 0.5) or high cross-factor loadings (> 0.75) were sequentially excluded across iterations. Cronbach's α coefficients indicate the internal consistency of the primary items within each factor. Total number of observations are rounded to the nearest 10 for confidentiality.

domain, particularly items overlapping with social interactions, such as "uses language to imagine and recreate", "interacts in various contexts", and "listens and responds to stories". Communication skills are often linked to behavioural outcomes through their role in social interactions, supporting pupils' ability to negotiate, mediate, and respond constructively to challenges rather than resorting to aggression. The role of the socio-emotional domain in relation to exclusionary outcomes and offending is discussed later in the paper. Cronbach's alpha for the primary loadings in each factor, reported at the foot of Table 2, confirms high internal consistency for both domains, exceeding the conventional 0.7 threshold.

Model parameters are estimated via CFA within the ESEM framework, as described in Section 4.1. The loadings for this model are reported in Table B.1 in the Appendix, and the fit indices suggest a satisfactory overall model fit. I then extract factor scores using empirical Bayes modal (EBM) scoring.²² Finally, I reverse-code the resulting scores so that higher values indicate greater cognitive and socio-emotional difficulties; I then use these scores to examine associations with later school outcomes and adolescent offending. Pearson correlations between the latent domains and the outcomes of interest are presented in Table B.2 in the Appendix.

Before discussing the results, I briefly summarise how the socio-emotional domain relates to established measures of behavioural problems. Table B.3 in the Appendix shows that some indicators of the socio-emotional construct resemble elements of the SDQ domains (Goodman, 1997), including hyperactivity (e.g., interested, motivated to learn), conduct (e.g., understands agreed codes of behaviour), emotional (e.g., responds with appropriate feelings), peer (e.g., works as part of a group), and prosocial behaviour (e.g., sensitive to the feelings of others). As discussed in Section 2, these domains have been shown to shape later outcomes in distinctive ways, including the probability of engaging in offending. Hyperactivity and conduct problems have been linked to offending and risky behaviours, while emotional and peer problems are less directly associated with offending but negatively relate to welfare outcomes including life satisfaction and mental health (Aebi et al., 2014; Del Bono et al., 2024; Murray et al., 2015).²³

In the following sections, I examine how early cognitive and socio-emotional difficulties relate to school outcomes and offending behaviour. Before turning to the main results, I consider the role of early support identification through Special Educational Needs and Disabilities (SEND) provision and Children in Need (CIN) services. SEND and CIN may act as proxies for under-

²²EBM scores, which are approximately normally distributed in this sample, represent the mode of the posterior distribution of each individual's latent trait given their observed responses and estimated model parameters.

²³A limitation of the socio-emotional domain is that it cannot be disentangled into distinct subdomains. Nevertheless, the EYFSP is less susceptible to rating bias than survey-based inventories such as the SDQ. The EYFSP is externally regulated and requires teachers to discuss behavioural concerns with parents to establish whether similar patterns occur in other settings, helping mitigate implicit bias in assessments.

lying developmental difficulties and socio-economic disadvantage, moderators of the relationship between early skills and later outcomes, or mediators if skill deficits increase the likelihood of support identification. Their role is not easily determined a priori. I therefore include them under alternative specifications and interpret them with caution throughout.

5.2 Special Education Needs and Disabilities (SEND) and Children in Need (CIN)

It is well established that children identified with SEND and those in social care are at higher risk of poorer school outcomes and are over-represented in the offending population (DfE & MoJ, 2022; Fuller & McNally, 2023; Machin et al., 2023). Severe mental health problems, physical impairments or learning disabilities are more likely to trigger early support, while pupils with milder difficulties are often overlooked, and the resulting delay in provision can itself damage their educational and behavioural trajectories. This issue is particularly relevant at a time when the SEND system is under significant strain. Rapid rises in demand and spending have left many local authorities with substantial deficits, while securing support can take months or even years and often fails to meet children's needs in time (Sibieta & Snape, 2025). Children in need of social care face similar delays, with support often arriving only after poor outcomes have already emerged. Around 50% of CIN who have been cautioned or convicted for an offence are first identified as CIN in the same term as their first offence or at a later point (DfE & MoJ, 2022).

Early identification of children who may require support is therefore central to the analysis. Previous work shows that children who do not reach a *good level of development* are substantially more likely to be identified as having SEND, but without distinguishing between the underlying domains driving this relationship (Atkinson et al., 2022; Wood et al., 2024).

Table 3 reports estimates of the associations between early cognitive and socio-emotional difficulties and the probability of first being identified as SEND or CIN throughout primary education.²⁴ Column 1 shows that cognitive difficulties are the strongest predictor of SEND identification, with a one-standard deviation (SD) increase raising the probability by 23 percentage points (pp), far exceeding the 8-pp association with free school meals (FSM) eligibility. By contrast, Column 4 shows that CIN identification is driven primarily by socioeconomic disadvantage, as proxied by FSM eligibility.

Columns 2 and 5 further show that the associations between early skill deficits and SEND/CIN identification are non-linear. Higher levels of cognitive and socio-emotional difficulties are associ-

²⁴For the SEND and CIN regressions, I report linear probability model (LPM) estimates, as these closely approximate logit marginal effects in this setting while offering a simpler interpretation.

Table 3: SEND and CIN on Early Skill Deficits

	SEND			CIN		
	[1]	[2]	[3]	[4]	[5]	[6]
<i>Early Skill Deficits</i>						
Socio-emotional	0.024*** [0.001]	0.019*** [0.001]	0.029*** [0.001]	0.024*** [0.001]	0.019*** [0.001]	0.021*** [0.001]
Cognitive	0.226*** [0.001]	0.218*** [0.001]	0.225*** [0.001]	0.028*** [0.001]	0.026*** [0.001]	0.023*** [0.001]
Socio-emotional sq		0.007*** [0.001]	0.006*** [0.001]		0.007*** [0.001]	0.003*** [0.001]
Cognitive sq		0.025*** [0.001]	0.025*** [0.001]		0.009*** [0.001]	0.007*** [0.001]
<i>Socio-demographic controls</i>						
Female	-0.086*** [0.001]	-0.087*** [0.001]	-0.080*** [0.001]	0.010*** [0.001]	0.010*** [0.001]	0.011*** [0.001]
Non-white	-0.026*** [0.002]	-0.026*** [0.002]	-0.030*** [0.002]	0.007*** [0.002]	0.008*** [0.002]	0.010*** [0.002]
Free school meal	0.080*** [0.001]	0.080*** [0.001]	0.066*** [0.001]	0.220*** [0.002]	0.220*** [0.002]	0.207*** [0.002]
Other backg. controls	X	X	X	X	X	X
Local Authority FE	X	X		X	X	
School FE			X			X
Mean of Dep. var	0.334	0.334	0.334	0.136	0.136	0.136
Observations	638,620	638,620	637,010	647,810	647,810	646,230
R-squared	0.315	0.320	0.370	0.158	0.160	0.198

Note: MoJ–DfE data. Table reports linear probability model (LPM) estimates for SEND, a binary variable indicating whether a pupil was identified as having Special Educational Needs and Disabilities, and CIN, a binary variable indicating whether a pupil was identified as a Child in Need (including looked-after children) between Reception and Key Stage 2 assessments. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, and fixed effects for Local Authorities (LAs) and schools, as indicated. Pupils who were already identified as having SEND or CIN prior to the EYFSP assessment are excluded from the estimation samples, respectively. Standard errors in brackets are clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

ated with a higher probability of SEND or CIN identification. If formal support acts as a protective factor against offending, this raises a natural question: are children with severe skill deficits, who are more likely to receive support, less likely to offend than children with moderate deficits who remain unidentified? I return to this question when examining offending outcomes.

Formal support is, of course, closely tied to the resources available to local authorities and schools. SEND identification rates, in particular, vary substantially across schools not only because of differences in resources, but also because of variation in expertise in identifying early needs

and institutional practices. Moreover, these school-level differences may also influence both the formation of early skills and their assessment.²⁵ Columns 3 and 6 address this concern by including school fixed effects, effectively comparing pupils attending the same school. The results are robust, suggesting that the associations between early skill deficits and SEND/CIN identification are driven by individual-level differences rather than school-level variation in assessment or identification practices.

I further examine the association between early skill deficits and more intensive forms of support: SEND plans among the SEND population, and Child Protection Plans (CPP) and Children Looked After (CLA) status among the CIN population. Results can be found in Table C.1 in the Appendix. Both skill deficits are associated with a higher probability of receiving a SEND plan, with socio-emotional difficulties showing the strongest relationship. A one-SD increase in socio-emotional difficulties raises the likelihood of receiving any SEND plan by about 5 pp, a substantial effect relative to the baseline rate of 7%. Different types of skill deficits are also channelled into different forms of SEND provision. Socio-emotional difficulties are strongly associated with provision for behavioural, emotional, or social difficulties (BESD), while cognitive difficulties are associated with provision for learning needs. Finally, both types are associated with more intensive social-care interventions, such as being looked after or receiving a CPP, with cognitive difficulties more strongly associated with cases involving neglect and socio-emotional difficulties more closely linked to physical and emotional abuse.

5.3 School Outcomes: Absenteeism, Exclusions and Attainment

I now examine the role of early skill deficits in predicting school outcomes, including exclusions, absenteeism, and attainment. Disciplinary incidents such as exclusions are commonly used as indicators of misbehaviour in studies using administrative data, providing an observable alternative to self-reported measures that may be subject to measurement error and social desirability bias (Kautz & Zanoni, 2024). Exclusions are expected to relate more closely to the socio-emotional domain, as misbehaviour often reflects difficulties in social and emotional regulation, while academic attainment should align more strongly with the cognitive domain. Absenteeism, which combines absences, incidents of fixed-term exclusion, and school mobility, captures not only learning disruption from time away from school but also the intensity of disruptive behaviour, likely reflecting difficulties across both domains.

²⁵School quality, particularly teacher inputs, has been shown to influence pupils' early attainment. For example, Bonetti and Blanden (2020) find that pupils exposed to qualified teachers with an academic degree score around 0.3 points higher on the EYFSP scale, particularly on cognitive indicators. The MoJ–DfE linked dataset does not identify classrooms, but differences in practitioner quality are likely to vary more across schools than within them.

Table 4 reports estimates of the associations between skill deficits, absenteeism and school exclusions. Column 1 shows that socio-emotional difficulties are strongly associated with absenteeism, although the association is larger for cognitive difficulties. Column 3 shows the opposite pattern for exclusions, where the socio-emotional domain is the stronger predictor. A one-SD increase in socio-emotional difficulties increases the probability of exclusion by 1.8 pp, an 82% increase relative to the baseline rate of 2.2%. The coefficient on cognitive difficulties is negative for exclusions conditional on socio-emotional difficulties, but its small magnitude likely reflects adjustment for the strong association between socio-emotional difficulties and exclusions rather than a substantive effect.

Table 4: Absenteeism and Exclusions on Early Skill Deficits

	Absenteeism Std. score		Exclusions	
	[1]	[2]	[3]	[4]
<i>Early Skill Deficits</i>				
Socio-emotional	0.014*** [0.004]	0.020*** [0.002]	0.018*** [0.001]	0.017*** [0.000]
Cognitive	0.159*** [0.003]	0.155*** [0.002]	-0.005*** [0.000]	-0.005*** [0.000]
<i>Socio-demographic controls</i>				
Female	0.001 [0.003]	0.013*** [0.002]	-0.027*** [0.001]	-0.025*** [0.000]
Non-white	-0.006 [0.007]	0.010** [0.005]	0.002** [0.001]	0.002*** [0.001]
Free school meal	0.555*** [0.005]	0.495*** [0.004]	0.027*** [0.001]	0.021*** [0.001]
Other backg. controls	X	X	X	X
Local Authority FE	X		X	
School FE		X		X
Mean of Dep. Var.	0.00	0.00	0.022	0.022
Observations	652,070	650,500	652,070	650,500
R-squared	0.185	0.390	0.042	0.138

Note: MoJ-DfE data. The table reports ordinary least squares (OLS) for *Absenteeism score* and linear probability model (LPM) estimates for *Exclusions*. *Absenteeism score* is a standardized predicted score from a PCA on authorised and unauthorised absences, number of fixed-term exclusions, and number of schools attended; *Exclusions* is a binary variable indicating whether a pupil was ever excluded, either permanently or temporarily, both outcomes measured between Reception and Key Stage 2 assessments. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and fixed effects for Local Authorities (LAs) and schools, as indicated. Standard errors in brackets are clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

The literature on early determinants of absenteeism and exclusions has identified systematic differences associated with poverty, race and ethnicity, and other socio-demographic characteristics (Graham et al., 2019). Consistent with these findings, the associations between skill deficits and exclusions, and particularly absenteeism, remain modest compared with indicators of socio-economic disadvantage (e.g., FSM eligibility). School-level characteristics and practices also matter, particularly for exclusions, which are strongly influenced by school-specific disciplinary approaches. To account for differences in school disciplinary practices, Columns 2 and 4 include school fixed effects. The estimates for absenteeism and exclusions remain robust, with the coefficient on socio-emotional difficulties becoming slightly larger for absenteeism once between-school differences are accounted for.

The relationship between socio-emotional difficulties and exclusions is further examined in Table C.2 in the Appendix, which restricts the sample to pupils excluded at least once during primary school and examines how early skills predict the main reasons for exclusion. The most common reasons include physical assault, verbal abuse or threatening behaviour, and serious misconduct such as sexual misconduct, drug- or alcohol-related incidents, damage, theft, and persistent disruptive behaviour. These indicators are not mutually exclusive, as pupils may have been excluded for multiple reasons during primary school. Consequently, the coefficients capture the association between early skills and the likelihood of ever being excluded for a given reason.

Socio-emotional difficulties show strong associations with exclusions for physical assault and serious misconduct, but not for verbal abuse. This is consistent with the idea that the lack of language and communication skills captured by our socio-emotional construct limits children's ability to manage conflict and prevent disagreements from escalating into physical confrontation rather than remaining verbal. As a result, children lacking these skills may face harsher disciplinary sanctions, with implications for later outcomes, a point I return to when examining the relationship between exclusions and offending. By contrast, cognitive difficulties show a negative association with physical assault and no meaningful association with other reasons for exclusion, conditional on socio-emotional difficulties. Cognitive difficulties are generally expected to be more closely associated with learning-related impairments than with aggressive behaviour. Children whose difficulties are primarily cognitive rather than socio-emotional may therefore be less prone to the impulsive or confrontational behaviours that often lead to exclusions for physical assault.

I now turn to the relationship between early skill deficits and attainment, measured using standardised KS2 scores in Grammar, Spelling and Punctuation, Reading, and Maths at age 11.

As expected, cognitive difficulties are a major driver of poor attainment.²⁶ Column 1 of Table 5 shows that moving from -1 SD to $+1$ SD in cognitive difficulties, a typical contrast between high- and low-performing pupils, is associated with a decline of about 1.3 SD in KS2 performance. Although both skill deficits are negatively correlated with KS2 performance (see Table B.2 in the Appendix), the coefficient on socio-emotional difficulties becomes positive, though very small, once cognitive difficulties are controlled for. This is not surprising and likely reflects the fact that cognitive difficulties already account for most of the variation in KS2 scores, leaving little residual variation for socio-emotional difficulties to explain.²⁷

Column 3 examines an alternative attainment measure: achieving Level 4 in at least one core KS2 subject, the expected standard prior to 2016. Here, socio-emotional difficulties show a negative association with attainment. This reflects the nature of Level 4 as a benchmark that separates lower-performing pupils from the rest of the distribution. At this margin, cognitive difficulties remain the dominant predictor, with a magnitude exceeding that of socio-demographic controls. Making the same comparison as before, a pupil 1 SD below the mean in cognitive difficulties is around 17 pp more likely to achieve Level 4 in at least one core subject than a pupil 1 SD above the mean. The equivalent comparison for socio-emotional difficulties adds a further 1.6 pp gap in the probability of achieving this benchmark.

As with previous outcomes, including school fixed effects in Columns 2 and 4 leads to only small changes in the estimates, confirming that within-school differences among pupils rather than school-level characteristics drive most of the observed patterns. I next examine the extent to which absenteeism and exclusions throughout primary school are associated with KS2 performance.

Figure 2 shows that both absenteeism and exclusions have a significant negative association with KS2 score, conditional on school fixed effects. School exclusions are associated with a 0.2 SD decrease in attainment, and because absenteeism also captures the intensity of disruption, each SD increase in absenteeism is associated with a further 0.07 SD decrease in KS2 score. Kautz and Zanoni (2024) similarly show that absences and disciplinary infractions have substantial effects on educational attainment, including longer-term outcomes such as enrolment in post-compulsory education. Notice that the coefficients on cognitive and socio-emotional difficulties remain largely unchanged after controlling for absenteeism and exclusions, suggesting that only a small share of their association with attainment operates through school-level disruptions.

²⁶Consistent with the skill accumulation literature, cognitive ability is the primary driver of educational attainment, with poor academic performance potentially signalling a higher opportunity cost of schooling from an early age (Cunha & Heckman, 2008; Currie & Almond, 2011).

²⁷Robustness checks using a bifactor specification, which imposes orthogonality between the cognitive and socio-emotional domains, yield an insignificant socio-emotional coefficient in the KS2 regression, a point I return to later in this section.

Table 5: Key Stage 2 on Early Skill Deficits

	KS2 Std. score		KS2 Level 4 or higher	
	[1]	[2]	[3]	[4]
<i>Early Skill Deficits</i>				
Socio-emotional	0.045*** [0.003]	0.037*** [0.002]	−0.008*** [0.001]	−0.005*** [0.001]
Cognitive	−0.658*** [0.003]	−0.665*** [0.002]	−0.084*** [0.001]	−0.081*** [0.001]
<i>Socio-demographic controls</i>				
Female	−0.042*** [0.002]	−0.056*** [0.002]	−0.011*** [0.001]	−0.013*** [0.001]
Non-white	0.082*** [0.005]	0.090*** [0.004]	0.007*** [0.001]	0.009*** [0.001]
Free school meal	−0.195*** [0.003]	−0.171*** [0.003]	−0.018*** [0.001]	−0.017*** [0.001]
Other backg. controls	X	X	X	X
Local Authority FE	X		X	
School FE		X		X
Mean of Dep. Var.	0.00	0.00	0.939	0.939
Observations	652,070	650,500	652,070	650,500
R-squared	0.427	0.510	0.154	0.240

Note: MoJ–DfE data. Columns 1–2 show ordinary least squares (OLS) estimates for the standardised KS2 scores in Grammar, Spelling and Punctuation, Reading, and Maths from national curriculum assessments at age 11. Columns 3–4 report linear probability model (LPM) estimates for achieving Level 4 or higher at KS2 in at least one of the assessed subjects. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and fixed effects for Local Authorities (LAs) and schools, as indicated. Standard errors in brackets are clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

Figure 2 also shows how these associations change after controlling for indicators of whether pupils were ever identified as SEND or CIN prior to the KS2 assessment. Both indicators of formal support are strong predictors of poor attainment and slightly moderate the associations between skill deficits, absenteeism, exclusions, and attainment. This is not surprising given that SEND and CIN identification itself partly reflects these early deficits. SEND and CIN likely also capture additional dimensions of vulnerability not fully accounted for in the baseline specification. The SEND coefficient is particularly large, with SEND pupils performing approximately 0.6 SD below their non-SEND peers on KS2 assessments, with similar gaps already evident earlier in the educational

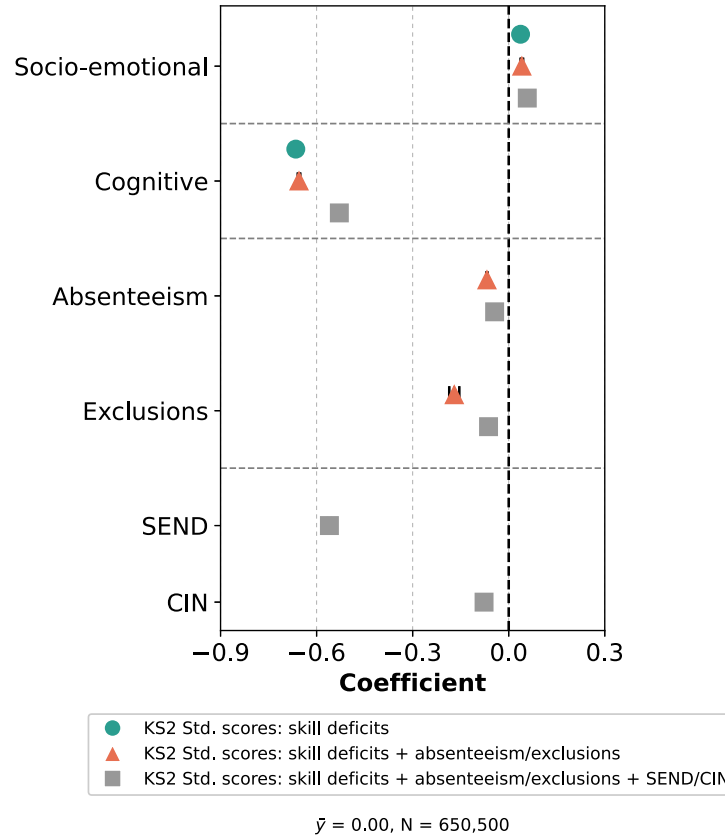


Figure 2: Key Stage 2 on Early Skill Deficits, Absenteeism, and Exclusions

Note: MoJ–DfE data. The figure shows ordinary least squares (OLS) estimates for the standardised KS2 score in Grammar, Spelling and Punctuation, Reading, and Maths scores from national curriculum assessments at age 11. Early Skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. See Table 4 notes for details on the construction of the *Absenteeism score* and *Exclusions* variables. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, prior SEND and CIN status before the EYFSP assessment, and school fixed effects. Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level. Total number of observations are rounded to the nearest 10 for confidentiality.

trajectory.²⁸ This large negative association may partly arise from delays in SEND provision and its limited effectiveness at this early stage of pupils’ educational trajectories. Although the timing of SEND/CIN identification is beyond the scope of this paper, it likely plays an important role in shaping later outcomes.²⁹

²⁸Atkinson et al. (2022) estimate that around 78% of pupils who did not achieve a *good level of development*, as measured by the EYFSP, and were identified as SEND also performed below the expected standard at Key Stage 1 at age 7, a precursor to the Key Stage 2 assessments.

²⁹Carneiro et al. (2024) show that a substantial share of the benefits of early-years interventions such as Sure Start stems from increasing the likelihood that children receive SEND support earlier in childhood.

5.4 Offending

I now examine the relationship between early skill deficits and whether a pupil has ever been cautioned or convicted for an offence between ages 11 and 17. Table 6 reports odds ratios (ORs) and marginal effects (ME) from logistic specifications of offending.³⁰ Both cognitive and socio-emotional difficulties are positively and significantly associated with offending, and the magnitudes are similar when school outcomes are not included. Column 2 shows that a one-SD increase in either skill deficit raises the probability of offending by around 0.30 pp. While modest in absolute terms, this represents a 15% increase relative to the baseline offending rate of 2%.

Behavioural and cognitive problems are well-established predictors of later offending (Basto-Pereira et al., 2024; Farrington et al., 2016). The estimates in this study are broadly consistent with this evidence, though somewhat more conservative. Kent et al. (2025) report ORs of around 0.7 per SD increase in childhood development, equivalent to around 1.45 per SD increase in developmental difficulties ($1/0.7 \approx 1.45$), with the larger magnitude likely reflecting broader measures of development combining aspects of cognitive and socio-emotional domains. Murray et al. (2015) report risk ratios of around 1.4 for conduct problems and 1.2 for hyperactivity predicting adolescent offending based on SDQ measures at age 11. The larger magnitude here likely reflects not only the use of self-reported crime and smaller samples, but also the fact that behavioural difficulties are measured in late childhood.

Most survey-based studies measure behavioural difficulties in mid-to-late childhood, by which point school experiences may already have shaped these behaviours. Consistent with this, behavioural measures collected later in childhood tend to show stronger associations with offending, particularly when based on teacher rather than parent reports (see for example Moffitt, 2018). This may suggest that teacher assessments of older children partly reflect accumulated school-specific responses in addition to underlying skill deficits. Unlike these studies, I observe skill deficits at school entry, before sustained school exposure can shape behaviour, and treat later behavioural outcomes (e.g., exclusions) as part of the pathway to offending. Before turning to this analysis, it is useful to briefly consider other important determinants of adolescent offending.

As expected, socio-economic status and ethnicity are both strongly associated with offending, with FSM eligibility doubling the baseline rate of adolescent offending. Gender differences follow familiar patterns in the criminology literature; boys are approximately 2.7 pp more likely to offend

³⁰I estimate logit models rather than LPMs for the offending analysis. While LPMs yield similar estimates in baseline specifications, these become less appropriate once exclusions are included, as offending rates are approximately ten times higher among excluded pupils, implying non-constant marginal effects.

Table 6: Offending on Early Skill Deficits

	ORs [1]	ME $\times$ 100 [2]	ORs [3]
<i>Early Skill Deficits</i>			
Socio-emotional	1.186*** [1.152, 1.220]	0.317*** [0.028]	1.209*** [1.172, 1.247]
Cognitive	1.168*** [1.133, 1.203]	0.289*** [0.028]	1.180*** [1.143, 1.216]
<i>Socio-demographic controls</i>			
Female	0.240*** [0.229, 0.252]	-2.655*** [0.051]	0.245*** [0.234, 0.258]
Non-white	1.270*** [1.195, 1.400]	0.444*** [0.058]	1.262*** [1.180, 1.350]
Free school meal	3.420*** [3.273, 3.569]	2.287*** [0.045]	3.086*** [2.949, 3.228]
Other backg. controls	X	X	X
Local Authority FE	X	X	
School FE			X
Mean of Dep. Var.	0.020	0.020	0.028
Observations	652,070	652,070	457,560
Pseudo R-squared	0.134	0.134	0.123

Note: MoJ-DfE data. The table reports odds ratios (ORs) and marginal effects (ME) multiplied by 100 from logit models (conditional logit in Column 3) for adolescent offending, a binary variable indicating whether a pupil has ever been cautioned or convicted for an offence between ages 11 and 17. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and fixed effects for Local Authorities (LAs) and schools, as indicated. Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

than girls.³¹ Finally, to account for unobserved school-level confounders, Column 3 reports ORs from a conditional logistic regression. The resulting odds ratios are substantively similar to those from the baseline model without school fixed effects, suggesting that estimates are primarily driven by within-school variation.³²

³¹Gender differences are also evident in other outcomes; for example, girls tend to outperform boys in early learning goals, particularly in communication, language, and literacy (Qualifications and Curriculum Authority [QCA], 2008), which are key indicators of the socio-emotional construct used in this study.

³²With approximately 14,000 schools, a standard logit with fixed effects is computationally infeasible; a conditional logit is used instead, exploiting within-school variation. ME are not reported because fixed effects are conditioned out, preventing comparable predicted probabilities, while ORs remain comparable across specifications.

Figure 3 presents marginal effects from logit models corresponding to the specification in Column 1 of Table 6, sequentially adding controls for absenteeism/exclusions and attainment. These estimates reflect direct associations rather than mediating pathways, which are examined later in the paper.

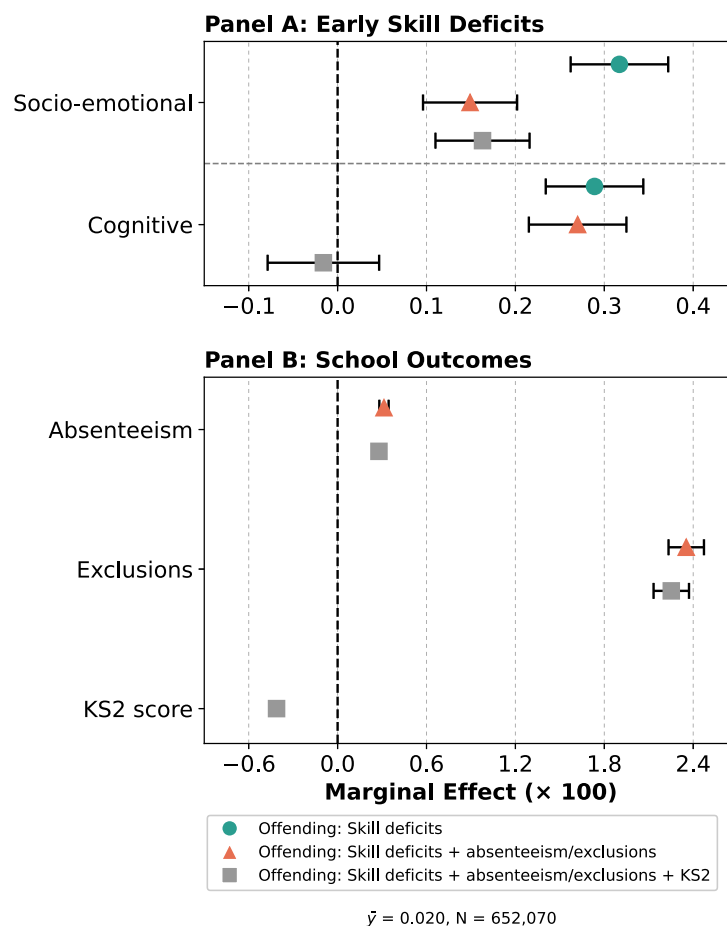


Figure 3: Offending on Early Skill Deficits and School Outcomes

Note: MoJ–DfE data. The figure shows marginal effects (ME) multiplied by 100 from logit models for adolescent offending, a binary variable indicating whether a pupil has ever been cautioned or convicted for an offence between ages 11 and 17. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. See Table 4 and 5 notes for details on the construction of the *Absenteeism score*, *Exclusions* and *KS2 score* variables. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and fixed effects for Local Authorities (LAs). Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level. Total number of observations are rounded to the nearest 10 for confidentiality.

School exclusions and absenteeism are strongly associated with offending. Exclusions show the largest effect, raising the probability of offending by approximately 2.4 pp and more than doubling the baseline rate. The magnitude is consistent with Cornish and Brennan (2025), who estimate a causal effect of exclusion on offending of similar size using a target trial framework applied to linked

national education and justice data. In this study, however, exclusions are neither the treatment of interest nor a proxy for underlying misbehaviour (Kautz & Zanoni, 2024), but rather a potential consequence of early skill deficits. Distinguishing the effect of early socio-emotional difficulties on offending from the consequences of school disciplinary responses is therefore non-trivial, given that such responses are often disproportionately imposed on children displaying poor socio-emotional skills. School exclusions carry independent harmful effects on pupils' outcomes that arise from removing pupils from the educational environment. Exclusions disrupt learning and limit peer interactions; even brief absences can deepen academic and socio-emotional vulnerabilities, both linked to higher offending (Farrington, 2021).³³ The positive association between exclusion and adolescent offending aligns with the literature on the incapacitation effect of education on crime, which shows that time spent in school keeps young people off the streets and reduces offending (Bell et al., 2022; Hjalmarsson et al., 2015; Meghir et al., 2023). This effect arises even over short periods of time, such as during fixed-term exclusions, and in contexts where absence is unrelated to disciplinary practices. For example, Luallen (2006) finds that juvenile crime rises by around 20% on days when teachers strike.

Notice that the marginal effects of both skill deficits decline once exclusions and absenteeism are included, suggesting that part of their association with offending operates through these school outcomes. This reduction is particularly pronounced for the socio-emotional domain, declining from 0.32 pp to 0.15 pp, consistent with socio-emotional difficulties being a key predictor of exclusionary responses. However, the remaining association is still substantial, indicating that exclusions do not fully capture all dimensions of poor socio-emotional skills linked to offending. From a policy perspective, the association between socio-emotional difficulties, school exclusions, and offending may be easier to address through early intervention if better understood, since exclusionary practices are partly policy-driven and socio-emotional difficulties can potentially be identified before disciplinary escalation occurs.

Table 2 showed that, in addition to indicators of social interaction and emotional regulation, the socio-emotional domain also captures elements of language and communication. The latter are increasingly recognised as crucial skills for managing conflict and preventing behavioural incidents from escalating into disciplinary actions. In 2023, London's Violence Reduction Unit invested £3 million in primary schools to support language and communication development, recognising that children with communication difficulties face higher risks of poor long-term outcomes, including violent offending.³⁴ The harmful effects of school exclusions on pupils'

³³Although permanent exclusions are rare in primary school, fixed-term exclusions are not. Around 20% of pupils who later offend have at least one exclusion record during primary school, compared with just 2% of non-offenders (see Table A.3 in the Appendix section).

³⁴See: [Child violent crime suspects 'getting younger'](#) (Accessed: November 2025).

outcomes have also motivated interest in alternative disciplinary approaches such as relational and restorative practices. These approaches replace punitive discipline with dialogue, accountability, and collective responsibility, addressing the underlying causes of misbehaviour while keeping pupils in the classroom. Evidence from the US suggests that these practices improve pupil outcomes (Adukia et al., 2025), though adoption in the UK is more recent and the evidence base remains limited (Bonell et al., 2018).

Consistent with evidence that academic achievement reduces exposure to criminogenic environments by increasing school engagement (Machin et al., 2023), higher Key Stage 2 performance is strongly negatively associated with the probability of offending. A one-SD increase in KS2 attainment reduces offending by 0.4 pp, a 20% decline relative to the offending baseline rate. Once KS2 performance is included in the specification, the association between cognitive difficulties and offending is no longer statistically significant, suggesting that the relationship operates primarily through attainment. In contrast, socio-emotional difficulties remain positively and significantly associated with offending even after controlling for all school-related outcomes.

Table 7 further provides insights into the relationship between SEND/CIN identification and adolescent offending. Column 2 shows that both are positively associated with offending, suggesting that formal identification captures additional vulnerabilities, beyond early skill deficits, that contribute to the over-representation of these groups in the offending population (DfE & MoJ, 2022; Machin et al., 2023). The coefficients on cognitive and socio-emotional difficulties attenuate once SEND/CIN are included, consistent with earlier results showing that children with the most severe difficulties are more likely to receive formal support (see Table 3).

Column 3 examines whether SEND and CIN status moderate the association between skill deficits and offending. The interaction terms indicate meaningful heterogeneity across support groups. SEND support attenuates the association between both skill deficits and offending, while CIN status attenuates only the association between cognitive difficulties and offending. Two distinct mechanisms could drive such effects. First, SEND or CIN support may divert vulnerable pupils from offending pathways. Second, severe skill deficits may themselves reduce the probability of offending once impairments exceed a threshold that makes offending less feasible. For example, pupils with severe social withdrawal may be less likely to engage in offences requiring social interaction or group participation.³⁵

Among pupils receiving SEND support, socio-emotional difficulties exhibit little additional association with offending ($1.21 \times 0.84 \approx 1.02$), consistent with Fuller and McNally (2023), who find that a SEND support plan can operate as a protective factor against offending. Column 4 reports

³⁵Farrington et al. (2016) identify "having few friends" as a promotive factor negatively associated with offending.

Table 7: Offending on Early Skill Deficits, School Outcomes and SEND/CIN

	ORs	ORs	ORs	ME (X 100) (<i>SEND/CIN</i> = 0)	ORs
	[1]	[2]	[3]	[4]	[5]
<i>School outcomes</i>					
Absenteeism score	1.167*** [1.147, 1.188]	1.106*** [1.087, 1.126]	1.107*** [1.088, 1.127]	0.139*** [0.013]	1.119*** [1.097, 1.140]
Exclusions	3.490*** [3.270, 3.725]	3.180*** [2.980, 3.400]	3.206*** [3.001, 3.421]	1.595*** [0.057]	3.248*** [3.017, 3.495]
KS2 score	0.796*** [0.779, 0.813]	0.845*** [0.826, 0.864]	0.835*** [0.817, 0.854]	-0.246*** [0.018]	0.782*** [0.763, 0.801]
<i>Skill Deficits & SEND/CIN</i>					
Socio-emotional	1.095*** [1.063, 1.127]	1.055*** [1.025, 1.086]	1.207*** [1.144, 1.273]	0.258*** [0.038]	1.236*** [1.170, 1.305]
Cognitive	0.991 [0.958, 1.026]	0.948** [0.916, 0.982]	1.042 [0.983, 1.104]	0.056 [0.041]	1.013 [0.954, 1.076]
Socio-emotional × SEND			0.839*** [0.790, 0.893]		0.839*** [0.787, 0.893]
Cognitive × SEND			0.927** [0.868, 0.990]		0.923** [0.863, 0.988]
Socio-emotional × CIN			0.961 [0.908, 1.016]		0.983 [0.927, 1.042]
Cognitive × CIN			0.883*** [0.832, 0.937]		0.877*** [0.825, 0.934]
SEND		1.391*** [1.320, 1.465]	1.441*** [1.371, 1.515]	0.293*** [0.035]	1.383*** [1.312, 1.458]
CIN		2.190*** [2.093, 2.293]	2.433*** [2.314, 2.557]	1.085*** [0.030]	2.404*** [2.282, 2.532]
Other backg. controls	X	X	X	X	
Local Authority FE	X	X	X	X	
School FE					X
Mean of Dep. var	0.020	0.020	0.020	0.020	0.028
Observations	652,070	652,070	652,070	652,070	457,560
Pseudo R-squared	0.166	0.178	0.179	0.179	0.176

Note: MoJ–DfE data. The table report odds ratios (ORs) from logit models (conditional logit in Column 5) and marginal effects (ME) multiplied by 100 for adolescent offending, a binary indicator for whether a pupil has ever been cautioned or sentenced for an offence between ages 11 and 17. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and fixed effects for Local Authorities (LAs) and schools, as indicated. Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

the marginal effect of socio-emotional difficulties on offending among pupils without SEND/CIN support, based on the specification in Column 3. A one-SD increase in socio-emotional difficulties increases the probability of offending by around 0.26 pp among unsupported children, 0.11 pp

larger than the baseline marginal effect of 0.15 pp reported in Figure 3, which averages across supported and unsupported pupils.

Consistent with earlier findings, the association between cognitive difficulties and offending is not significant once school outcomes, particularly attainment, are accounted for. The interaction with CIN status in Column 3 is nevertheless below one and significant ($1.04 \times 0.88 \approx 0.92$), suggesting that among CIN pupils, cognitive difficulties are more weakly associated with offending than among non-CIN pupils.

Finally, odds ratios change very little when school fixed effects are added in Column 5 compared to Column 3, indicating that unobserved differences across schools account for little of the estimated relationship.

5.5 Number and Types of Offences

There is growing concern about a post-pandemic rise in violent offending among children and young people in the UK. ONS figures show that offences committed by those aged 10 to 18 rose by 4% in 2023/24, marking a second consecutive annual increase. The sharpest increases were in sexual offences, violence against the person, and property-related offences such as theft, robbery, and burglary.³⁶

Figure 4 shows estimates of the relationship between early skill deficits and different types of offences, as well as overall offence counts as described in Section 3.³⁷ Cognitive difficulties show no direct association with offence type or number of offences once school outcomes are accounted for. In contrast, socio-emotional difficulties retain small but statistically significant associations with violent and property-societal offending, as well as with the overall number of offences. However, school exclusions stand out as the strongest predictor of offence types and counts. Being excluded increases the probability of violent and property-societal offending by around 4 pp (15% and 10% of their respective baselines) and of summary offending by around 9 pp (19% of baseline), and is associated with approximately 0.7 additional offences, a 30% increase relative to the mean. The latter is consistent with Kautz and Zandoni (2024), who report that programmes targeting non-cognitive skills, measured using suspensions and absences, reduce arrest rates by around 32%.

An interesting pattern is that school exclusions predict offending broadly across categories, consistent with evidence that time away from school facilitates opportunistic offending (Luallen, 2006), while socio-emotional difficulties show a more selective pattern concentrated in violent and

³⁶See [Youth Justice Statistics: 2023 to 2024](#) for more details.

³⁷Offence categories are not mutually exclusive, and individuals may appear in more than one category. The exception is *Violent offences*: individuals with any caution or conviction for a violent offence are classified exclusively as violent offenders, with all other category indicators set to zero.

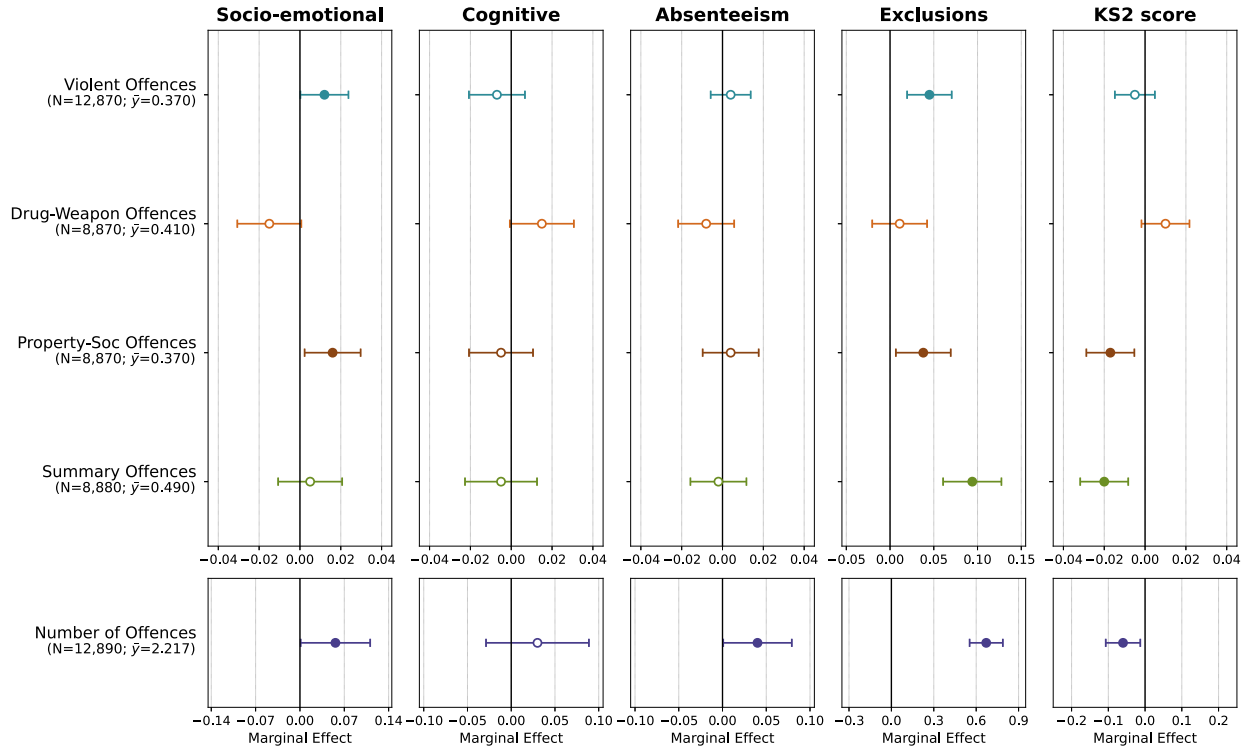


Figure 4: Type and Number of Offences on Early Skill Deficits and School Outcomes

Note: MoJ–DfE data. The figure shows marginal effects (ME) from logit models for type of offences and negative binomial for number of offences. Offending outcomes are shown on the y-axis and regressors on the x-axis, including early skill deficits and school outcomes. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, IDACI deciles, month and year of birth, SEND and CIN status prior to the EYFSP assessment, and fixed effects for Local Authorities (LAs). Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level. Total number of observations are rounded to the nearest 10 for confidentiality.

acquisitive crime. This distinction suggests different underlying mechanisms. Exclusion removes children from a structured and supervised environment that would otherwise constrain offending opportunities, while socio-emotional difficulties may capture behavioural and interpersonal tendencies more closely associated with serious offending. The two channels are therefore complementary rather than competing, and their interaction may be particularly important for children exposed to both.

Higher KS2 attainment acts as a protective factor, reducing the probability of committing property-societal and summary offences as well as overall offence counts, but shows no significant association with violent offending. This pattern is consistent with Sabates (2008), who finds that educational attainment reduces offences such as burglary, theft, and vandalism, but has little association with violent offending in a late-1990s cohort study.

5.6 Mediation Analysis Results

The results in Section 5.4 provided estimates of the direct associations between early skill deficits, school outcomes, and adolescent offending. I now examine the extent to which cognitive and socio-emotional difficulties affect offending through school exclusions, absenteeism, and attainment, as well as an alternative specification that additionally considers SEND and CIN as mediating mechanisms, using the mediation framework described in Section 4.3.

Table 8 reports the share of the total effect attributable to each mediation pathway, where each school outcome serves as a distinct mediator linking skill deficits to offending. Column 1 shows that absenteeism and exclusions account for 42% of the total effect of socio-emotional difficulties on offending, with school exclusions representing the dominant pathway (39%). The remaining 58% reflects a substantial direct effect not mediated by either channel, suggesting that socio-emotional difficulties influence offending through additional mechanisms beyond disciplinary infractions and attendance. These may include broader aspects of school quality, including peer effects and teacher inputs, identified as important pathways shaping later outcomes (Deming, 2011). In this analysis, these dimensions are not directly observed in the data but are instead absorbed by school fixed effects.

For cognitive difficulties (Column 2), the pathway decomposition reveals that attainment (KS2) more than fully accounts for the total effect on offending, implying a negative direct component. Cognitive difficulties thus operate through competing mechanisms: the primary effect is indirect, working through reduced attainment, while the residual direct effect is slightly protective once school outcomes are held constant. One possible explanation, consistent with earlier findings in Table 7, is that children with severe cognitive difficulties are more likely to receive formal support that reduces offending through channels beyond attainment, such as increased supervision or reduced exposure to criminogenic environments.

Columns 3 and 4 examine the role of formal support classifications, SEND and CIN, as pathways linking early skill deficits to later offending. These classifications capture not only underlying vulnerabilities but also institutional responses, including targeted interventions and support services. Table 3 showed that skill deficits strongly predict SEND and CIN identification, which are themselves associated with adolescent offending (Column 2 in Table 7), suggesting that these mechanisms may also operate as mediating pathways.³⁸

For socio-emotional difficulties (Column 3), including SEND and CIN as additional mediating pathways increases the mediated share from 42% to 57%, with CIN accounting for a larger share

³⁸For the mediation analysis, SEND and CIN are modelled as outcomes of cognitive and socio-emotional difficulties, analogous to equation (5) for absenteeism in Section 4.3.

Table 8: Share of Early Skill Deficits Explained by School Outcomes & SEND/CIN (Level 1 - Non-sequential Pathways)

	Only School Outcomes		Including SEND and CIN	
	Socio-emotional	Cognitive	Socio-emotional	Cognitive
Absenteeism	0.029*** [0.019, 0.038]	0.273*** [0.243, 0.304]	0.021*** [0.014, 0.028]	0.202*** [0.168, 0.236]
Exclusions	0.388*** [0.360, 0.416]	—	0.379*** [0.357, 0.402]	—
KS2 score	—	1.289*** [1.185, 1.392]	—	1.044*** [0.955, 1.135]
SEND	—	—	0.031*** [0.020, 0.043]	0.286*** [0.242, 0.330]
CIN	—	—	0.143*** [0.123, 0.162]	0.208*** [0.175, 0.241]
Observations	650,500			

Notes: MoJ–DfE data. Table reports the share of the total effect of early skill deficits on offending explained by each mediating pathway. Columns 1 and 2 include only absenteeism, exclusions, and attainment as mediators. Columns 3 and 4 additionally include SEND and CIN as mediators. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, IDACI deciles, month and year of birth, SEND and CIN status prior to the EYFSP assessment (Columns 1 and 2), and school fixed effects. Confidence intervals are reported at the 95% level and based on standard errors from 250 bootstrap replications clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

than SEND. Nevertheless, the direct effect remains substantial, accounting for around 43% of the total association. For cognitive difficulties (Column 4), both SEND and CIN account for meaningful shares of the total effect, capturing part of the association otherwise mediated through absenteeism and KS2 attainment, though the latter remains the dominant pathway. The positive mediated share suggests that SEND and CIN identification amplify rather than attenuate the pathway from early cognitive and socio-emotional difficulties to offending, consistent with a labelling or cumulative disadvantage mechanism rather than a protective one.

Table C.3 in the Appendix presents analogous estimates from a specification allowing for sequential mediation, acknowledging that absenteeism, exclusions, and SEND/CIN may affect offending partly through their effects on KS2 attainment.³⁹ The sequential specification matters primarily for cognitive difficulties, as KS2 fully captures their indirect pathway to offending. The results point to a substantial SEND → KS2 → offending pathway, with the SEND share rising sharply by 20 pp while the KS2 share declines to 80%. This suggests that part of the association

³⁹In this specification, SEND and CIN may also affect offending sequentially via KS2 attainment, analogous to equation (9) in the Appendix, with SEND and CIN indicators additionally included in the KS2 regression.

previously attributed to KS2 attainment is itself transmitted through earlier SEND identification, highlighting the potential importance of early institutional recognition and support in shaping the pathways through which cognitive difficulties translate into later offending.

Together, these findings point to distinct mediation patterns. For cognitive difficulties, school and institutional channels account for the bulk of the total effect, suggesting that early identification and attainment support could substantially reduce offending risk. For socio-emotional difficulties, a large direct component remains even after accounting for exclusions and care-system responses, implying that school-based interventions may need to be complemented by broader behavioural and family-based support.

5.7 Robustness Checks

Family-level unobserved heterogeneity: Although the specifications in previous sections control for an extensive set of covariates, including several proxies for family socio-economic status, unobserved factors, particularly those related to family environments and parental involvement, are still likely to influence both early skill formation and later offending. To address family-level unobserved heterogeneity, I estimate sibling fixed-effects models that exploit within-family variation using a proxy sibling identifier constructed from address and surname information, available in the MoJ–DfE linked dataset.⁴⁰

Table 9 reports results for the full sibling sample and the brothers-only subsample. Strikingly, the estimated effect of socio-emotional difficulties on offending remains positive and statistically significant when sibling fixed effects are included, while the effect of cognitive difficulties is fully attenuated. This suggests that within-family variation in socio-emotional difficulties continues to predict differential offending risk among siblings, while cognitive difficulties appear more closely tied to family-level characteristics shared across siblings. The estimated effect in Column 2 is even larger when restricting the sample to brothers only, consistent with the large gender gap in offending. Females are around 76% less likely to offend than males (OR = 0.24, Table 6), and the stronger effect among brothers may also reflect the larger dispersion of socio-emotional difficulties among boys. Column 3 shows that this association remains robust even after conditioning on school outcomes, though smaller in magnitude, consistent with earlier evidence that part of the association between socio-emotional difficulties and offending is mediated through exclusions.⁴¹

⁴⁰Two aspects of this analysis are worth noting. First, because the analytical sample combines three pupil cohorts, I can only identify siblings born one or two years apart. Second, because the identifier is not preserved across years, I use the sibling group identifier available in 2015/16, when all three cohorts had completed primary education.

⁴¹These estimates should be interpreted with caution, as they rely on a linear probability model to accommodate sibling fixed effects. Some non-linearity may be present in this relationship, although this is more likely to affect the magnitude of the estimates than their significance or direction.

Table 9: Sibling Fixed Effects Estimates for Offending

	All	Boys sample	
	[1]	[2]	[3]
<i>School outcomes</i>			
Absenteeism score			0.167 [0.484]
Excluded			9.058*** [1.776]
KS2 score			-1.272*** [0.300]
<i>Early Skill Deficits</i>			
Socio-emotional	0.344*** [0.130]	0.778*** [0.291]	0.621** [0.290]
Cognitive	-0.105 [0.145]	-0.311 [0.323]	-0.916** [0.355]
<i>Other backg. controls</i>			
Sibling FE	X	X	X
Mean of Dep. var	0.020	0.032	0.032
Observations	86,660	24,310	24,310
R-squared	0.578	0.608	0.612

Note: MoJ–DfE data. The table reports linear probability model (LPM multiplied by 100) estimates of offending with sibling fixed effects. Columns [2] and [3] report estimates for the boys-only sample. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM. Specification (1) controls for gender, and all specifications control for individual characteristics including birth order, part-time attendance, month and year of birth, and SEND and CIN status prior to the EYFSP assessment. Standard errors in brackets are clustered at the sibling group level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

Sample selection: The analytical sample in this paper comprises pupils with complete point-scale EYFSP data, which are required to construct domain-specific skills. Such data are available for approximately 40% of pupils, as detailed in Section 3. This raises concerns about whether the available data are representative of the full pupil population, or whether missingness is correlated with pupil characteristics that are relevant for later offending outcomes.

To address potential sample selection bias, I implement inverse probability weighting (IPW). Since the EYFSP is assessed and recorded by reception schools, selection into the analytical sample may reflect school-level recording practices rather than individual pupil characteristics. I model selection as a function of socio-economic status, demographic characteristics, geographical location, and school-level averages of pupil characteristics using a logistic regression.⁴² Observations are

⁴²School fixed effects are not included in the selection model due to the incidental parameters problem in logistic regression; school-level averages of pupil characteristics are used as a proxy.

then weighted by the inverse of their predicted selection probability, giving higher weight to under-represented observations. The resulting weighted sample more closely reflects the characteristics of the full population.

Column 1 in Table C.4 in the Appendix shows that, in specifications without school-level controls, pupils with SEND or CIN status prior to the EYFSP assessment are significantly less likely to have complete point-scale data. This may reflect the higher administrative burden associated with recording detailed assessments for pupils with early-life vulnerabilities in less well-resourced schools. Once within-school averages of pupil characteristics are included in Column 2, the odds ratios for SEND and CIN are no longer statistically significant. Consistent with this interpretation, reweighting yields negligible changes in coefficient estimates for key outcomes, particularly when controlling for school fixed effects, as shown in Table C.5 in the Appendix.

Alternative factorization: I examine whether additional EYFSP learning goals, such as *Knowledge and understanding of the world* and *Physical development*, reveal latent dimensions beyond cognitive and socio-emotional domains. Table B.4 in the Appendix reports item loadings from the first iteration of an EFA that includes all available EYFSP learning goals. Loadings for retained items, following iterative refinement (see table note), are shown in bold.

Although there is initial evidence of a potential third dimension capturing *Creativity* (Factor 3), its weak internal consistency becomes apparent during iterative refinement. In fact, most additional items are dropped after a few iterations. The baseline two-domain structure remains stable, with the same key indicators, those with the highest loadings, consistently retained within each domain. In general, additional measures exhibit unstable loading patterns and tend to cluster children into high- and low-achieving groups rather than forming distinct dimensions. This likely reflects that high-performing children tend to score well across all developmental areas, and may indicate the presence of an underlying general dimension.

Motivated by this evidence, I estimate a bifactor specification for the baseline indicators, in which a general dimension captures overall skill deficits, while the remaining variance is explained by the cognitive and socio-emotional domains. Table C.6 in the Appendix reports estimates from offending regressions using the bifactor specification within a conditional logit framework to accommodate school fixed effects. Overall, the findings closely mirror those from the two-domain specification. While the general dimension, capturing overall skill deficits, accounts for a substantial share of the variation in offending, the cognitive and socio-emotional domains retain similar magnitude and remain substantively meaningful in explaining offending.

Column 3 shows that once school outcomes are included, the general domain becomes non-significant, mirroring the pattern observed for cognitive difficulties in the baseline specification.

This suggests that general skill deficits operate primarily through school engagement and performance. Finally, Column 5 adds interaction terms with SEND and CIN indicators, consistent with the attenuation pattern observed in the main specification. This attenuation is particularly strong for the general domain, reinforcing the interpretation that children with general skill deficits are more readily identified and supported within the education system, while those with domain-specific deficits may go unidentified and are therefore at higher risk of offending.

6 Summary and Conclusions

In this paper, I examine how early skills are associated with adolescent offending. Using the MoJ–DfE linked dataset, I identify two domains of early skill deficits: cognitive and socio-emotional. Both are associated with adverse school outcomes, including absenteeism, exclusions, and lower attainment. While cognitive difficulties affect offending primarily through attainment, socio-emotional difficulties do so through school exclusions. However, socio-emotional difficulties retain a substantial direct association with offending that is not fully explained by school-mediated pathways and persists within schools and within families.

The findings contribute to the literature in several ways. First, I show that early skill deficits comprise two distinct domains, cognitive and socio-emotional, with different school-mediated pathways linking them to offending. This distinction matters because treating them as a single construct would provide limited guidance on how interventions should be targeted, which school pathways should be prioritised, and which existing support mechanisms are most effective. Second, school exclusions are not simply markers of underlying socio-emotional difficulties but may independently increase offending risk, potentially through reduced exposure to structured environments and supervision. Third, although SEND and CIN children are overrepresented in the offending population, the findings suggest that SEND and CIN support attenuate offending risk among children with skill deficits. Finally, the paper highlights the value of early childhood developmental data for predicting later outcomes and informing targeted interventions, a contribution particularly relevant given recent simplifications to early years assessments aimed at reducing teacher workload (Tickell, 2011). Whether such changes yield net benefits remains an open question, as reduced information may limit the capacity to identify children at risk.⁴³

The association between early skill deficits and offending is not deterministic, and schools play an important protective role. Academic support may be particularly valuable for children with cognitive difficulties, while relational and restorative disciplinary practices may be more

⁴³Bonetti and Blanden (2020) document increased clustering in scores under the revised EYFSP, with around 19% of children receiving identical scores, potentially reducing the system’s ability to identify specific support needs.

effective than punitive approaches for those with socio-emotional difficulties. This is particularly relevant given the strong association between exclusions and adolescent offending, including violent offending. Children with poor communication skills, a key indicator of socio-emotional difficulties, are especially at risk, as conflict can more easily escalate into aggression and exclusion. Programmes targeting these skills are being implemented in London primary schools, though existing evidence suggests that earlier intervention may be more effective, given that socio-emotional skills are most malleable in early childhood, potentially before school entry.

One data limitation is worth noting. The reliance on administrative records means that only formally recorded offences are observed, which may limit comparability with studies using self-reported offending. Moreover, temporal limitations in the data prevent follow-up beyond adolescence, leaving the longer-term consequences of early socio-emotional difficulties an important area for future research.

Future work could also examine how policy decisions shape the trajectory from early skill deficits to offending. The cohorts in this study attended school during a period of substantial cuts to public spending on youth and community services, including youth clubs, leisure facilities, and crime reduction initiatives, services that provide structured environments supporting children's wellbeing and socio-emotional development. How their availability shapes offending risk remains an open question.

References

- Adukia, A., Feigenberg, B., & Momeni, F. (2025): From retributive to restorative: An alternative approach to justice in schools. *American Economic Review*, 115(8), 2722–2754.
- Aebi, M., Giger, J., Plattner, B., Metzke, C. W., & Steinhausen, H.-C. (2014): Problem coping skills, psychosocial adversities and mental health problems in children and adolescents as predictors of criminal outcomes in young adulthood. *European child & adolescent psychiatry*, 23, 283–293.
- Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021): Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, 136(4), 2147–2194.
- Armstrong, T. A., Boisvert, D., Wells, J., Lewis, R. H., Cooke, E., & Woeckner, M. (2020): Assessing potential overlap between self-control and psychopathy: A consideration of the grasmick self-control scale and the levenson self-report psychopathy scale. *Journal of Criminal Justice*, 70, 101725.
- Asparouhov, T., & Muthén, B. (2009): Exploratory structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, 16(3), 397–438.

- Atkinson, A. L., Hill, L., Pettinger, K., Wright, J., Hart, A., Dickerson, J., & Mon-Williams, M. (2022): Can holistic school readiness evaluations predict academic achievement and special educational needs status? evidence from the early years foundation stage profile. *Learning and Instruction*, 77, 101537.
- Bacher-Hicks, A., Billings, S. B., & Deming, D. J. (2024): The school-to-prison pipeline: Long-run impacts of school suspensions on adult crime. *American Economic Journal: Economic Policy*, 16(4), 165–193.
- Basto-Pereira, M., Farrington, D. P., & Maciel, L. (2024): Unraveling the sequences of risk factors underlying the development of criminal behavior. *Journal of Developmental and Life-Course Criminology*, 1–23.
- Bell, B., Costa, R., & Machin, S. (2022): Why does education reduce crime? *Journal of political economy*, 130(3), 732–765.
- Bindler, A., Ketel, N., & Hjalmarsson, R. (2020): Costs of victimization. In K. F. Zimmermann (Ed.), *Handbook of labor, human resources and population economics*. Springer. https://doi.org/10.1007/978-3-319-57365-6_140-2
- Blanden, J., Gregg, P., & Macmillan, L. (2007): Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal*, 117(519), C43–C60.
- Bolt, U., French, E., Hentall-MacCuish, J., & O’Dea, C. (2024): *The intergenerational elasticity of earnings: Exploring the mechanisms* (tech. rep.). IFS Working Papers.
- Bonell, C., Allen, E., Warren, E., McGowan, J., Bevilacqua, L., Jamal, F., Legood, R., Wiggins, M., Opondo, C., Mathiot, A., et al. (2018): Effects of the learning together intervention on bullying and aggression in english secondary schools (inclusive): A cluster randomised controlled trial. *The Lancet*, 392(10163), 2452–2464.
- Bonetti, S., & Blanden, J. (2020): Early years workforce qualifications and children’s outcomes: An analysis using administrative data. *Education Policy Institute/Nuffield Foundation*.
- Carneiro, P., Cattan, S., & Ridpath, N. (2024): *The short-and medium-term impacts of sure start on educational outcomes*. Institute for Fiscal Studies.
- Chowdry, H., & Fitzsimons, P. (2016): *The cost of late intervention: Eif analysis 2016*. Early Intervention Foundation London.
- Cornish, R., & Brennan, I. (2025): Exclusion from school and risk of serious violence: A target trial emulation study. *The British Journal of Criminology*, 65(6), 1221–1240.
- Cunha, F., & Heckman, J. J. (2008): Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources*, 43(4), 738–782.
- Currie, J., & Almond, D. (2011): Human capital development before age five. In *Handbook of labor economics* (pp. 1315–1486, Vol. 4). Elsevier.
- Del Bono, E., Etheridge, B., & Garcia, P. (2024): *The economic value of childhood socio-emotional skills* (tech. rep.). Institute for Social and Economic Research.
- Del Bono, E., Francesconi, M., Kelly, Y., & Sacker, A. (2016): Early maternal time investment and early child outcomes. *The Economic Journal*, 126(596), F96–F135.
- Deming, D. J. (2011): Better schools, less crime? *The Quarterly Journal of Economics*, 126(4), 2063–2115.

- Department for Education. (2010a): *Achievement of Children in the Early Years Foundation Stage Profile* [Available at: <https://assets.publishing.service.gov.uk/media/5a7acbb8e5274a319e77ac7e/DFE-RR034.pdf>, Accessed: 06 January 2025]. Research Report.
- Department for Education. (2010b, December): *Early Years Foundation Stage Profile Attainment by Pupil Characteristics, England 2009/10* [Available at: <https://assets.publishing.service.gov.uk/media/5a7b875c40f0b62826a0429c/sfr28-2010.pdf>, Accessed: 17 December 2024]. Statistical Release.
- Department for Education & Ministry of Justice. (2022, March): *Education, children's social care and offending: Descriptive Statistics* [Available at: <https://www.gov.uk/government/publications/education-childrens-social-care-and-offending>, Accessed: August 2024]. Research report.
- Dickson, M., Gregg, P., & Robinson, H. (2016): Early, late or never? when does parental education impact child outcomes? *The Economic Journal*, 126(596), F184–F231.
- Dillon, A., Boyd, A., Mummé, M., Tilling, K., Brennan, I., & Cornish, R. (2022a, July): *Feasibility study 1: Moj-dfe linked dataset – data quality report* (tech. rep.) (Available at: https://www.adruk.org/fileadmin/uploads/adruk/Documents/Feasibility_study_1_MoJ-DfE_linked_dataset_Data_quality_report.pdf, Accessed: September 2024). ADR UK.
- Dillon, A., Boyd, A., Mummé, M., Tilling, K., Brennan, I., & Cornish, R. (2022b, July): *Feasibility study 2: Moj-dfe linked dataset – generating matched controls* (tech. rep.) (Available at: https://www.adruk.org/fileadmin/uploads/adruk/Documents/Feasibility_study_2_MoJ-DfE_linked_dataset_Generating_matched_controls_.pdf, Accessed: September 2024). ADR UK.
- Dorsett, R., Gorman, E., Morando, G., Oppedisano, V., & Zhang, M. (2022, July): *School peers' ability and youth custody* (tech. rep.) (Also associated with IZA, Bonn, Germany). University of Westminster. 35 Marylebone Road, London NW1 5LS.
- Eren, O., Figlio, D., Mocan, N., & Ozturk, O. (2025): School accountability, long-run criminal activity, and self-sufficiency. *Journal of Human Resources*.
- Fahy, K., Alexiou, A., Mason, K., Bennett, D., Egan, M., Taylor-Robinson, D., & Barr, B. (2023): Inequalities in local government spending on cultural, environmental and planning services: a time-trend analysis in England, Scotland, and Wales. *BMC Public Health*, 23(1), 408.
- Farrington, D. P. (2021): Youth crime and antisocial behaviour. In *The social child* (pp. 353–392). Psychology Press.
- Farrington, D. P., Ttofi, M. M., & Piquero, A. R. (2016): Risk, promotive, and protective factors in youth offending: Results from the Cambridge study in delinquent development. *Journal of criminal justice*, 45, 63–70.
- Feinstein, L. (2003): Inequality in the early cognitive development of british children in the 1970 cohort. *Economica*, 70(277), 73–97.
- Francesconi, M., & Heckman, J. J. (2016): Child development and parental investment: Introduction. *The Economic Journal*, 126(596), F1–F27.
- Fuller, M., & McNally, J. (2023, July): *Education, children's social care and offending: multi-level modelling* (Research report) (Available at: <https://www.gov.uk/government/publications/>

- [education - childrens - social - care - and - offending](#), Accessed: August 2024). Department for Education.
- Goodman, A., Joshi, H., Nasim, B., & Tyler, C. (2015): Social and emotional skills in childhood and their long-term effects on adult life [Report commissioned by the Cabinet Office].
- Goodman, R. (1997): The strengths and difficulties questionnaire: A research note. *Journal of child psychology and psychiatry*, 38(5), 581–586.
- Graham, B., White, C., Edwards, A., Potter, S., & Street, C. (2019): School exclusion: A literature review on the continued disproportionate exclusion of certain children. *Department for Education*.
- Heckman, J., Pinto, R., & Savelyev, P. (2013): Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American economic review*, 103(6), 2052–2086.
- Heckman, J. J., & Karapakula, G. (2019): *The perry preschoolers at late midlife: A study in design-specific inference* (tech. rep.). National Bureau of Economic Research.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006): The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3), 411–482.
- Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., & Pollack, H. A. (2017): Thinking, fast and slow? some field experiments to reduce crime and dropout in chicago. *The Quarterly Journal of Economics*, 132(1), 1–54.
- Hjalmarsson, R., Holmlund, H., & Lindquist, M. J. (2015): The effect of education on criminal convictions and incarceration: Causal evidence from micro-data. *The Economic Journal*, 125(587), 1290–1326.
- Hughes, C., Foley, S., White, N., & Devine, R. T. (2018): School readiness in children with special educational needs and disabilities: Psychometric findings from a new screening tool, the brief early skills, and support index. *British Journal of Educational Psychology*, 88(4), 606–627.
- Kautz, T., & Zanon, W. (2024): Measurement and development of noncognitive skills in adolescence: Evidence from chicago public schools and the onegoal program. *Journal of Human Capital*, 18(2), 272–304.
- Keating, B. E. (2023): Disagreeable and out of control: Applicability of the big five and low self-control in predicting offending. *Journal of Criminal Justice*, 86, 102069.
- Kent, H., Hogarth, L., Williams, W. H., Cornish, R., & Leckie, G. (2025): A ‘double disadvantage’: Neurodevelopmental profile and poverty confer synergistic risk of youth justice involvement. *Longitudinal and Life Course Studies*, 1–18.
- Kremser, J. (2014): School crime (problems in measurement). *The Encyclopedia of Criminology and Criminal Justice*, 1–5.
- Landersø, R., Nielsen, H. S., & Simonsen, M. (2017): School starting age and the crime-age profile. *The Economic Journal*, 127(602), 1096–1118.
- Luallen, J. (2006): School’s out. . . forever: A study of juvenile crime, at-risk youths and teacher strikes. *Journal of urban economics*, 59(1), 75–103.
- Machin, S., Marie, O., & Vujić, S. (2011): The crime reducing effect of education. *The Economic Journal*, 121(552), 463–484.

- Machin, S., McNally, S., & Ruiz-Valenzuela, J. (2023): *School qualifications and youth custody* (tech. rep.). Centre for Economic Performance, LSE.
- Machin, S., & Sandi, M. (2020): Autonomous schools and strategic pupil exclusion. *The Economic Journal*, 130(625), 125–159.
- Machin, S., & Sandi, M. (2025): Crime and education. *Annual Review of Economics*, 17.
- Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014): Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual review of clinical psychology*, 10(1), 85–110.
- Meghir, C., Palme, M., & Schnabel, M. (2023): *The effect of education policy on crime: An inter-generational perspective* (Cowles Foundation Discussion Paper No. 2327). Cowles Foundation for Research in Economics, Yale University.
- Moffitt, T. E. (2018): Male antisocial behaviour in adolescence and beyond. *Nature human behaviour*, 2(3), 177–186.
- Mooney, K. E., Welch, C., Palliser, G., Cheung, R. W., Nielsen, D., Eddy, L. H., & Blower, S. L. (2025): An assessment of the teacher-completed early years foundation stage profile as a routine measure of child developmental health. *PLoS ONE*, 20(3), e0302771.
- Murray, J., Irving, B., Farrington, D. P., Colman, I., & Bloxsom, C. A. (2010): Very early predictors of conduct problems and crime: Results from a national cohort study. *Journal of child psychology and psychiatry*, 51(11), 1198–1207.
- Murray, J., Menezes, A. M., Hickman, M., Maughan, B., Gallo, E. A. G., Matijasevich, A., Gonçalves, H., Anselmi, L., Assunção, M. C. F., Barros, F. C., et al. (2015): Childhood behaviour problems predict crime and violence in late adolescence: Brazilian and british birth cohort studies. *Social psychiatry and psychiatric epidemiology*, 50, 579–589.
- OECD. (2015): *Skills for social progress: The power of social and emotional skills*.
- Qualifications and Curriculum Authority. (2008): *Early Years Foundation Stage Profile Handbook*. London.
- Sabates, R. (2008): Educational attainment and juvenile crime: Area-level evidence using three cohorts of young people. *The British Journal of Criminology*, 48(3), 395–409.
- Sabates, R., & Feinstein, L. (2008): Effects of government initiatives on youth crime. *Oxford Economic Papers*, 60(3), 462–483.
- Sallis, H., Szekely, E., Neumann, A., Jolicoeur-Martineau, A., Van IJzendoorn, M., Hillegers, M., Greenwood, C. M., Meaney, M. J., Steiner, M., Tiemeier, H., et al. (2019): General psychopathology, internalising and externalising in children and functional outcomes in late adolescence. *Journal of Child Psychology and Psychiatry*, 60(11), 1183–1190.
- Sibieta, L., & Snape, D. (2025): England’s SEND crisis: Costs, challenges and the case for reform [Accessed: 5 November 2025]. <https://ifs.org.uk/articles/englands-send-crisis-costs-challenges-and-case-reform>
- Sorrenti, G., Zölitz, U., Ribeaud, D., & Eisner, M. (2025): The causal impact of socio-emotional skills training on educational success. *Review of Economic Studies*, 92(1), 506–552.
- Ter-Minassian, L., Viani, N., Wickersham, A., Cross, L., Stewart, R., Velupillai, S., & Downs, J. (2022): Assessing machine learning for fair prediction of ADHD in school pupils using a retrospective cohort study of linked education and healthcare data. *BMJ open*, 12(12), e058058.

- Tickell, C. (2011): The early years: Foundations for life, health and learning. *An independent report on the early years foundation stage to Her Majesty's Government.*
- Villa, C. (2024): *The effects of youth clubs on education and crime* (tech. rep.). Institute for Fiscal Studies.
- Webb, C. J., & Bywaters, P. (2018): Austerity, rationing and inequity: Trends in children's and young peoples' services expenditure in england between 2010 and 2015. *Local Government Studies*, 44(3), 391–415.
- Wood, M. L., Gunning, L., Relins, S., Sohal, K., Wright, J., Mon-Williams, M., & Atkinson, A. L. (2024): Potential for england's statutory school entry assessment to identify special educational needs and reveal structural inequalities: A population-based study. *Archives of Disease in Childhood*, 109(1), 52–57.
- Wright, B., Mon-Williams, M., Kelly, B., Williams, S., Sims, D., Mushtaq, F., Sohal, K., Blackwell, J. E., & Wright, J. (2019): Investigating the association between early years foundation stage profile scores and subsequent diagnosis of an autism spectrum disorder: a retrospective study of linked healthcare and education data. *BMJ Paediatrics Open*, 3(1).

Appendix for:

Early Socio-Emotional Skills and Adolescent Offending: Evidence from Administrative Data

Paul Garcia

A Further Details on the Data and Methodology	1
A.1 EYFSP assessment and moderation process	1
A.2 Further details on data cleaning and missingness	2
Table A.1: Descriptions of EYFSP Learning Goals	3
Table A.2: Loadings for the first component of a PCA on School Disengagement	4
Table A.3: Summary Statistics of Main Variables	5
A.3 Level 2 Mediation: Sequential Pathways via Key Stage 2 Attainment	6
Figure A.1: Sequential Pathways via Key Stage 2 Attainment	7
B Additional Factorization Results	9
Figure B.1: Eigenvalues from EFA of EYFSP data: Final Iteration	9
Table B.1: CFA Loadings from a Two-factor ESEM	10
Table B.2: Correlation Matrix for Early Skill Deficits and Key Outcomes	11
Table B.3: Conceptual correspondence between SDQ items and EYFSP socio-emotional indicators	12
Table B.4: Factor Loadings from EFA using EYFSP Point Scales	14
C Additional Results Quoted in the Main Paper	16

Table C.1: SEND (Primary Needs) and CIN (CPP and CLA) status on Early Skill Deficits . . .	16
Table C.2: Main Reasons for Exclusions on Early Skill Deficits	17
Table C.3: Share of Early Skill Deficits Explained by School Outcomes & SEND/CIN (Level 2 - Sequential Mediation via KS2)	18
Table C.4: Selection into the Main Analytical Sample	19
Table C.5: Key Outcomes Re-weighted Using IPW for Sample Selection	20
Table C.6: Offending on Early Skill Deficits, School Outcomes and SEND/CIN (Bifactor Specification)	21

A Further Details on the Data and Methodology

A.1 EYFSP assessment and moderation process

The Foundation Stage Profile, introduced in 2002/03 and later known as the EYFSP from 2008, provides teacher-reported assessments of child development at the end of the Reception year (age five). This version remained in place until 2012 and covers six learning areas: personal, social and emotional development; communication, language and literacy; problem solving, reasoning and numeracy; knowledge and understanding of the world; physical development; and creative development. These are measured through 13 scales, each scored from 1 to 9, with 9 indicating the highest level of proficiency (QCA, 2008).

A child was classified as having achieved a *good level of development* (GLD) if they obtained a total score of at least 78 points across all learning goals, with a minimum of 6 points in each of the following key scales: *Disposition and attitudes*, *Social development*, *Emotional development*, *Language and thinking*, *Linking sounds and letters*, *Reading*, and *Writing* (DfE, 2010a).⁴⁴ The main analysis draws on these seven scales, together with two numeracy-related scales (*Numbers as labels and counting* and *Calculating*), yielding nine scales in total as described in Section 3.⁴⁵

The EYFSP is a nationally standardised framework implemented by trained teachers. Judgements are made based on observed attainment, using evidence gathered from both planned and spontaneous activities, whether child-initiated or practitioner-guided. The Department for Children, Schools and Families defines a self-initiated activity as: "...an activity wholly decided on by the child and is the result of an intrinsic motivation to explore a project or express an idea". Where activities are initiated by the practitioner, evaluation is based on the child's willingness to engage and extend their involvement. In this case, practitioners assess the extent to which children can complete the activity independently or with appropriate support.

Moderation of the EYFSP is overseen by local authorities. Moderation processes became more rigorous after 2009 following reforms led by the Qualifications and Curriculum Development Agency (QCDA), which reviewed local authorities' strategies and tools for collecting EYFSP data.

⁴⁴Children with SEND status may be assessed using alternative approaches depending on their individual needs. When a learning goal cannot be assessed due to SEND, it is recorded as "not applicable" (DfE, 2010b). These cases represent a small share of the sample, varying by scale (e.g., 0.13% for personal, social and emotional development and 0.27% for problem solving, reasoning and numeracy).

⁴⁵I exclude scales that lack a clear distinction between cognitive and non-cognitive domains. Evidence from Mooney et al. (2025) suggests that a latent domain associated with behavioural difficulties is largely captured by the learning goals from the core EYFSP scales, while others, such as *Knowledge and Understanding of the World*, contribute little to distinguishing between these domains. Including all learning goals in the factor analysis may therefore add noise rather than help identify the relevant domains. However, as shown in Table B.4, this does not alter the overall factor structure.

Moderation involves discussions among practitioners, headteachers, and external bodies, including meetings to review children's attainment, consideration of supporting evidence, consultation with parents, and support for practitioners in assessing specific scales. One notable change documented by DfE (2010a) is a marked decline in the proportion of children awarded the highest scale points (8 and 9) in 2009, consistent with the introduction of more stringent assessment practices. These limitations should be taken into account when comparing EYFSP outcomes over time or pooling information across cohorts.

A.2 Further details on data cleaning and missingness

EYFSP: Some pupils appear more than once in the EYFSP and School Census data. To ensure that each child is represented only once, I retain a single EYFSP record per pupil, using the earliest available observation. In the School Census, duplicate records can arise through school transfers or dual registration; removing these affects fewer than 1% of observations.

KS2 assessments: Approximately 3% of KS2 marks are missing. However, for these cases the data still record the achievement level attained by the pupil in each subject. I use this information by imputing missing marks using the population mean mark among all pupils who achieved the same level in the same subject. These levels range from Level 2 ("Below expected standard") to Level 6 ("Exceptionally high performance"), so the imputed values are consistent with the national grading structure and reflect meaningful differences in attainment.

Absenteeism score: The standardised score is derived from the first principal component of a PCA on authorised and unauthorised absences (as a percentage of possible sessions), number of fixed-term exclusions, and number of establishments attended from Reception to KS2 assessment. For pupils attending multiple establishments within a year (approximately 3% of the sample), absences are summed across establishments. Where data on total possible sessions are missing (around 2%), I impute using the population mean.

Offending and Number of Offences: A small share of offence records (around 6%) are classified as "Not Guilty", "Non-Conviction", or "Pending", and are excluded to ensure that the analysis focuses on confirmed offending outcomes.

Despite prior data cleaning, some duplicate records may persist where offences are similarly described but recorded under different codes. To reduce the influence of such cases, particularly on the offence count measure, the upper tail of the distribution is winsorised at the top 3%.

Table A.1: Descriptions of EYFSP Learning Goals

Scale Point	Score Desc.	Dispositions & Attitudes	Social Development	Emotional Development	Language for Comm. & Thinking	Linking Sounds & Letters	Reading	Writing	Numbers as Labels & Counting	Calculating
1	Progressing towards early learning goals	Interested in classroom activities	Plays alongside others	Separates from carer with support	Listens & responds	Joins in rhyming	Interested in books	Experiments with mark making	Says some numbers in familiar cont.	Responds to maths vocabulary
2		Dresses & undresses with support	Builds relationships	Communicates about home, community	Initiates communication	Aware of rhyme & alliteration	Knows print has meaning	Uses letters to communicate meaning	Counts up to three	Recognizes quantity differences
3		Highly involved in activities	Takes turns, shares with support	Expresses needs, feelings appropriately	Talks, reflects, modifies actions	Links some sounds to letters	Recognizes a few familiar words	Represents some sounds in writing	Counts up to six	Finds one more (less) from 5 objects
4	Attainment of early learning goals (non-hierarchical)	Dresses & undresses independently	Works as part of a group	Responds with appropriate feelings	Listens, responds to stories, rhymes	Links sounds to letters	Knows print directionality	Writes own name & words from memory	Says number names in order	Relates addition by comb. groups
5		Selects activities independently	Forms good relationships with others	Sensitive to the feelings of others	Uses language to imagine & recreate	Hears & says sounds in words	Understands story elements	Forms recognizable letters	Recognizes numerals 1 to 9	Relates subtraction to taking away
6		Interested, motivated to learn	Understands agreed codes of behaviour	Respect for own & others' beliefs	Interacts in various contexts	Blends sounds in words	Reads familiar words independently	Attempts writing, various purposes	Counts up to 10 objects	Uses addition/subtraction vocabulary
7		Confident to try new activities	Respects different needs & cultures	Considers conseq. of words, actions	Organises & clarifies thinking	Uses phonic knowledge to read words	Retells narratives in sequence	Writes simple regular words	Orders numbers up to 10	Finds one more (less), numbers 1-10
8		Maintains attention & concentrates	Expects others to respect needs, views	Understands right & wrong	Speaks confident, listener awareness	Attempts to read more complex words	Understands non-fiction texts	Forms captions & simple sentences	Uses math ideas to solve problems	Uses math ideas to solve problems
9	Beyond goals	Perseveres, solves problems independently	Values others' ideas	Displays positive self-id. & emotions	Talks & shows awareness by including details	Uses sounds, words when read & write	Reads books of own choice with accuracy	Communicates meaning via phrases	Counts, orders, uses numbers up to 20	Uses strategies for addition, subtraction

Notes: Information from QCA (2008). Early Years Foundation Stage Profile (EYFSP) scale point descriptors for developmental areas. Scale points 1–3 indicate children progressing towards early learning goals, points 4–8 represent attainment of goals (non-hierarchical) and are used in the factorization analysis, and point 9 indicates advancement beyond the expected goals.

Table A.2: Loadings for the first component of a PCA on School Disengagement

Absenteeism Variable	Loadings
Share of absences	0.639
Number of schools attended	0.572
Number of fixed-term exclusions	0.514
Variance	0.412
Number of observations	652,070

Notes: MoJ–DfE data. Table reports loadings for the first component from a principal component analysis on: share of absences (authorized and unauthorized) from all possible sessions, number of fixed-term exclusions, and number of schools attended, all measured from school reception to KS2 assessment year. Total number of observations are rounded to the nearest 10 for confidentiality.

Table A.3: Summary Statistics of Main Variables

Variables	Offending				Non-offending				Mean / Pr.
	Mean	SD	Min	Max	Mean	SD	Min	Max	p-value
<i>Demographics</i>									
Female	0.18	0.38	0	1	0.50	0.50	0	1	0.00
Non-white	0.22	0.42	0	1	0.20	0.40	0	1	0.00
English not first language	0.11	0.31	0	1	0.14	0.35	0	1	0.00
<i>Socio-economic characteristics</i>									
Eligible for FSM	0.66	0.47	0	1	0.29	0.45	0	1	0.00
CIN [§] (including CLA)	0.40	0.50	0	1	0.10	0.30	0	1	0.00
IMD deciles	3.62	2.59	1	10	5.17	2.95	1	10	0.00
Number of siblings	2.22	1.13	1	*	2.15	0.94	1	*	0.00
First born	0.56	0.50	0	1	0.56	0.50	0	1	0.20
<i>School</i>									
KS2 score	-0.67	1.04	-3.80	1.60	0.01	0.99	-4.10	1.90	0.00
KS2 Level 4 +	0.50	0.50	0	1	0.75	0.43	0	1	0.00
Excluded	0.20	0.40	0	1	0.02	0.13	0	1	0.00
SEND	0.63	0.48	0	1	0.32	0.47	0	1	0.00
Schools attended	0.82	0.88	0	*	0.58	0.72	0	5	0.00
Part-time	0.02	0.14	0	1	0.03	0.16	0	1	0.00
Absenteeism score	0.83	1.48	-1.40	3.90	-0.02	0.98	-1.40	3.90	0.00
<i>EYFSP factor scores[†]</i>									
<i>2-factor solution</i>									
Cognitive difficulties	0.55	0.96	-1.40	2.90	-0.01	1.00	-1.40	2.90	0.00
Socio-emotional difficulties	0.57	0.97	-1.30	3.30	-0.01	1.00	-1.30	3.30	0.00
<i>Bifactor solution</i>									
General difficulties	0.54	0.94	-1.40	3.30	-0.01	1.00	-1.40	3.30	0.00
Cognitive difficulties	0.30	1.08	-3.60	3.80	-0.01	1.00	-4.30	3.90	0.00
Socio-emotional difficulties	0.31	1.19	-3.40	4.60	-0.01	0.99	-3.70	4.80	0.00
Number of obs.	12,890				639,180				

Note: MoJ–DfE data. The table reports summary statistics for the main variables used in this paper, separately for pupils with at least one recorded offence and those with no recorded offences between ages 11 and 17. The last column reports p-values from two-sided tests of differences in means (for continuous variables) or proportions (for categorical variables) between the two groups. Values marked with an asterisk (*) are suppressed and total number of observations are rounded to the nearest 10 for confidentiality.

[†] Early skill deficits are standardized (reverse-coded) factor scores derived from ESEM (2-factor solution) and confirmatory bifactor model (bifactor solution) using EYFSP data (see Section 4.1 for further details).

A.3 Level 2 Mediation: Sequential Pathways via Key Stage 2 Attainment

I next allow for indirect effects of absenteeism and exclusion on offending operating through Key Stage 2 attainment. Specifically, I estimate:

$$KS_i = \gamma_A A_i + \gamma_E E_i + \gamma_C C_i + \mathbf{X}_i' \boldsymbol{\kappa}^{KS} + \mu_i^{KS}, \quad (9)$$

where γ_A and γ_E capture the effects of absenteeism and exclusions on KS2 attainment, and γ_C captures the direct effect of cognitive difficulties on KS2 attainment net of these school outcomes. Consistent with the Level 1 restrictions, I continue to impose that socio-emotional difficulties affect KS2 attainment only indirectly, through absenteeism and exclusions.

Combining equations (4) and (9), and the Level 1 restriction that cognitive difficulties do not affect exclusions ($dE_i/dC_i = 0$), the total derivatives of offending with respect to S_i and C_i are:

$$\begin{aligned} \frac{dO_i}{dS_i} &= \alpha_{KS} \left[\frac{\partial KS_i}{\partial A_i} \frac{dA_i}{dS_i} + \frac{\partial KS_i}{\partial E_i} \frac{dE_i}{dS_i} \right] + \alpha_A \frac{dA_i}{dS_i} + \alpha_E \frac{dE_i}{dS_i} + \alpha_S \\ &= \alpha_{KS} [\gamma_A \beta_S^A + \gamma_E \beta_S^E] + \alpha_A \beta_S^A + \alpha_E \beta_S^E + \alpha_S, \\ \frac{dO_i}{dC_i} &= \alpha_{KS} \left[\frac{\partial KS_i}{\partial A_i} \frac{dA_i}{dC_i} + \frac{\partial KS_i}{\partial C_i} \right] + \alpha_A \frac{dA_i}{dC_i} + \alpha_C \\ &= \alpha_{KS} [\gamma_A \beta_C^A + \gamma_C] + \alpha_A \beta_C^A + \alpha_C, \end{aligned}$$

The total effect of socio-emotional difficulties on offending operating through absenteeism can now be decomposed into a non-sequential component and a sequential component via KS2:

$$\underbrace{\alpha_A \beta_S^A}_{\text{non-sequential via } A} + \underbrace{\alpha_{KS} \gamma_A \beta_S^A}_{\text{sequential: } S \rightarrow A \rightarrow KS \rightarrow O}. \quad (10)$$

Similarly, the total effect of socio-emotional difficulties operating through exclusions is

$$\underbrace{\alpha_E \beta_S^E}_{\text{non-sequential via } E} + \underbrace{\alpha_{KS} \gamma_E \beta_S^E}_{\text{sequential: } S \rightarrow E \rightarrow KS \rightarrow O}. \quad (11)$$

For cognitive difficulties, the total effect operating through absenteeism is:

$$\underbrace{\alpha_A \beta_C^A}_{\text{non-sequential via } A} + \underbrace{\alpha_{KS} \gamma_A \beta_C^A}_{\text{sequential: } C \rightarrow A \rightarrow KS \rightarrow O}. \quad (12)$$

Figure A.1 shows the Level 2 mediation structure, where exclusions and absenteeism can also affect offending indirectly through Key Stage 2 attainment. The solid arrows indicate sequential mediation pathways (e.g., $S \rightarrow A \rightarrow KS \rightarrow O$), while dashed arrows represent non-sequential paths from absenteeism and exclusions to offending. The dotted arrows show the direct effects from early skills to offending, operating independently of all school mediators.

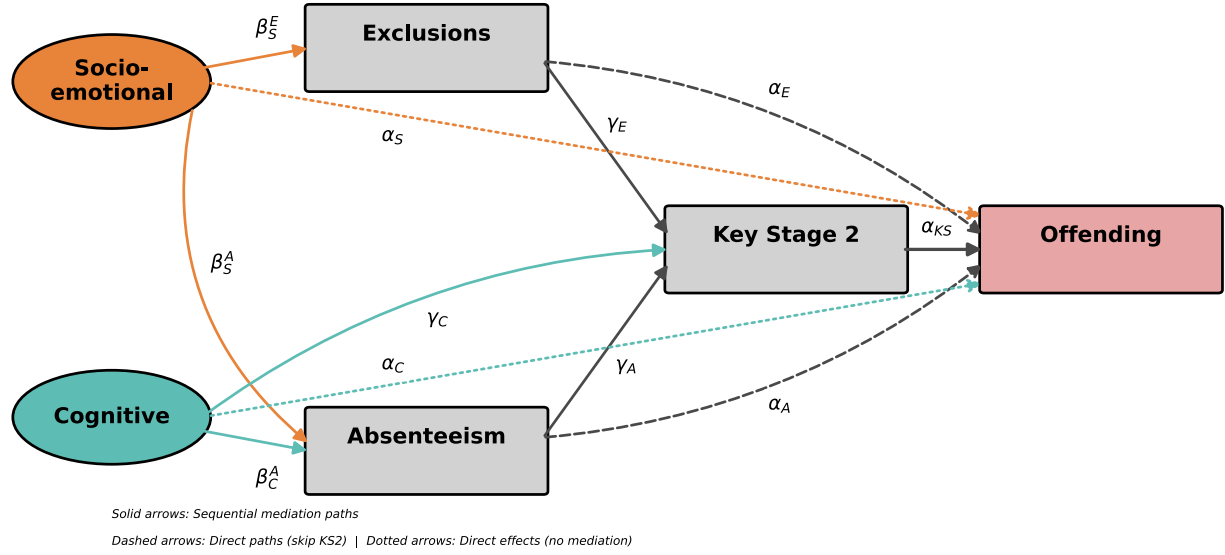


Figure A.1: Sequential Pathways via Key Stage 2 Attainment

Note: Figure shows the parameters of the Level 2 mediation structure, illustrating the non-sequential and the sequential pathways (via Key Stage 2) through which socio-emotional and cognitive difficulties affect offending.

The contribution of each pathway can be expressed as a share of the total effect by dividing equations (10) and (11) by ρ_S , and equation (12) by ρ_C , coefficients obtained from equation (8). For example, the share of the socio-emotional effect on offending that operates through absenteeism (including its sequential component via KS2) is:

$$\frac{\alpha_A \beta_S^A + \alpha_{KS} \gamma_A \beta_S^A}{\rho_S},$$

Finally, note that the cognitive effect on offending transmitted through KS2 attainment, $\alpha_{KS} \beta_C^{KS} / \rho_C$, can be further decomposed. Equation (7) gives the total effect of cognitive difficulties on KS2, β_C^{KS} , while equation (9) gives the direct effect, γ_C , net of the pathway operating through absenteeism. The share of the cognitive effect on KS2 that operates directly (independently

of other school outcomes) is therefore γ_C/β_C^{KS} . This allows the following decomposition:

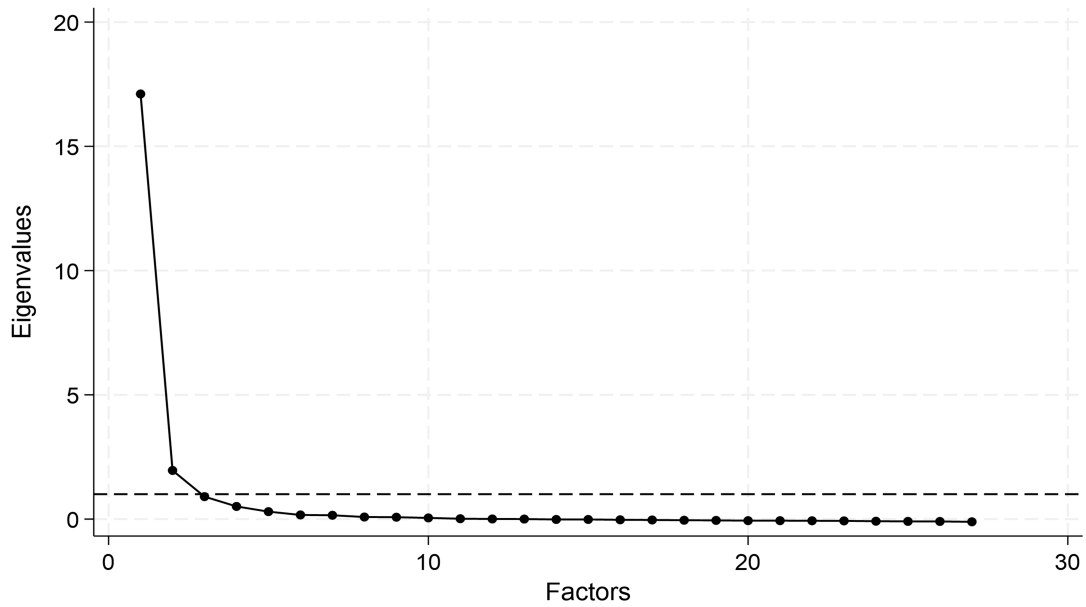
$$\frac{\alpha_{KS}\beta_C^{KS}}{\rho_C} = \underbrace{\left(1 - \frac{\gamma_C}{\beta_C^{KS}}\right) \frac{\alpha_{KS}\beta_C^{KS}}{\rho_C}}_{\text{sequential via A}} + \underbrace{\frac{\alpha_{KS}\gamma_C}{\rho_C}}_{\text{direct } C \rightarrow KS \rightarrow O},$$

where the sequential term can be expressed in terms of structural parameters, $\alpha_{KS}\gamma_A\beta_C^A/\rho_C$, corresponding to the pathway $C \rightarrow A \rightarrow KS \rightarrow O$ derived in equation (12).⁴⁶ This separates the KS2-mediated cognitive effect on offending into a sequential component operating through absenteeism and a direct component from cognitive difficulties to attainment. In contrast, for socio-emotional difficulties, the entire effect on offending through KS2, $\alpha_{KS}(\gamma_A\beta_S^A + \gamma_E\beta_S^E)$, is sequential by construction, as socio-emotional difficulties do not directly affect attainment under the identifying restrictions.

⁴⁶From equation (9), taking dKS_i/dC_i and imposing $dE_i/dC_i = 0$ gives $dKS_i/dC_i = \gamma_A\beta_C^A + \gamma_C$. Since equation (7) defines $dKS_i/dC_i = \beta_C^{KS}$, it follows that $\beta_C^{KS} - \gamma_C = \gamma_A\beta_C^A$, and therefore $\left(1 - \frac{\gamma_C}{\beta_C^{KS}}\right) \frac{\alpha_{KS}\beta_C^{KS}}{\rho_C} = \frac{\alpha_{KS}(\beta_C^{KS} - \gamma_C)}{\rho_C} = \frac{\alpha_{KS}\gamma_A\beta_C^A}{\rho_C}$.

B Additional Factorization Results

Figure B.1: Eigenvalues from EFA of EYFSP data: Final Iteration



Notes: MoJ–DfE data. The figure shows eigenvalues from the last iteration of an Exploratory Factor Analysis (EFA) using tetrachoric correlation matrices and oblique quartimin rotation (see Section 4.1 and Table 2 for more details). Eigenvalues are estimated for all pupils with complete information on EYFSP point scales ($N = 652,070$).

Table B.1: CFA Loadings from a Two-factor ESEM

EYFSP Point Scales	Cognition	Socio-Emotional
Writes simple regular words	0.981	-0.104
Reads familiar words independently	0.962	-0.098
Recognizes numerals 1 to 9	0.909	-0.072
Writes own name and words from memory	0.895	-0.025
Blends sounds in words	0.885	-0.024
Finds one more or one less than numbers 1–10	0.853	-0.009
Relates addition by combining groups	0.789	0.132
Uses math ideas to solve problems	0.782	0.045
Uses addition/subtraction vocabulary	0.780	0.108
Knows print directionality	0.764	0.178
Says number names in order	0.762	0.126
Links sounds to letters	0.748	-0.022
Attempts writing for various purposes	0.702	0.164
Understands information from non-fiction texts	0.701	0.109
Forms recognizable letters	0.672	0.087
Retells narratives in sequence	0.669	0.201
Understands story elements	0.653	0.275
Forms good relationships with others	-0.139	0.940
Sensitive to the feelings of others	-0.078	0.930
Works as part of a group	-0.140	0.886
Considers consequences of words, actions	-0.025	0.879
Understands agreed codes of behaviour	0.034	0.863
Responds with appropriate feelings	0.004	0.835
Respects different needs and cultures	0.128	0.763
Understands right and wrong	0.049	0.737
Respect for own and others' beliefs	0.125	0.731
Selects activities independently	0.156	0.664
Interacts in various contexts	0.255	0.643
Interested, motivated to learn	0.260	0.599
Uses language to imagine and recreate	0.262	0.587
Dresses and undresses independently	0.154	0.545
Listens, responds to stories, rhymes	0.344	0.533
Tucker-Lewis index (TLI)		0.973
Comparative fit index (CFI)		0.976
Std. RMSR		0.045
RMSEA		0.042
Number of observations		652,070

Notes: MoJ–DfE data. Table reports factor loadings from Exploratory Structural Equation Modelling (ESEM) using EYFSP indicators (see Section 4.1 for more details). Loadings in **bold** indicate primary items within each factor. Goodness-of-fit statistics include TLI (Tucker and Lewis, 1973) and CFI (Bentler, 1990) have a scale from 0 to 1, with values close to 1 indicating a good fit. RMSR (Joreskog and Sorbom, 1986) and RMSEA (Browne and Cudeck, 1992; Steiger, 1990) also range from 0 to 1, with values close to 0 indicating a good fit. Total number of observations are rounded to the nearest 10 for confidentiality.

Table B.2: Correlation Matrix for Early Skill Deficits and Key Outcomes

	Offend	KS2	Absent	Excl	CIN	SEND	Cogn	Semo
Offending	1							
KS2 score	-0.095	1						
Absenteeism	0.118	-0.260	1					
Excluded	0.173	-0.137	0.332	1				
CIN	0.120	-0.222	0.261	0.148	1			
SEND	0.093	-0.518	0.228	0.149	0.187	1		
Cognitive	0.079	-0.622	0.241	0.102	0.193	0.383	1	
Socio-emotional	0.080	-0.479	0.204	0.133	0.182	0.285	0.719	1
N. observations	652,070							

Notes: MoJ–DfE data. The table reports correlations between the outcomes of interest, including indicators of whether a pupil was identified as a Child in Need (CIN) or as having Special Educational Needs and Disabilities (SEND), KS2 assessments, and factor scores for early skills (see Table B.1 for further details). The correlation between cognitive and socio-emotional difficulties corresponds to the correlation between their latent constructs estimated in the ESEM model. Correlations are estimated for all pupils with complete information on EYFSP point scales. Total number of observations is rounded to the nearest 10 for confidentiality.

Table B.3: Conceptual correspondence between SDQ items and EYFSP socio-emotional indicators

SDQ Item	Closest EYFSP Indicator(s)
<i>Hyperactivity</i>	
(i) Restless, overactive, cannot stay still for long	
(ii) Constantly fidgeting or squirming	
(iii) Easily distracted, concentration wanders	
(iv) Thinks things out before acting (reversed)	Considers consequences of words, actions
(v) Sees tasks through to the end, good attention span (reversed)	Interested, motivated to learn
<i>Conduct problems</i>	
(i) Often has temper tantrums or hot tempers	
(ii) Generally obedient, usually does what adults request (reversed)	Understands agreed codes of behaviour
(iii) Often fights with other children or bullies them	Forms good relationships with others
(iv) Often lies or cheats	
(v) Steals from home, school or elsewhere	
<i>Emotional problems</i>	
(i) Often complains of headaches, stomach-aches or sickness	
(ii) Many worries, often seems worried	
(iii) Often unhappy, down-hearted or tearful	
(iv) Nervous or clingy in new situations, easily loses confidence	Interacts in various contexts
(v) Many fears, easily scared	Responds with appropriate feelings
<i>Peer problems</i>	
(i) Rather solitary, tends to play alone	Works as part of a group
(ii) Has at least one good friend (reversed)	
(iii) Generally liked by other children (reversed)	
(iv) Picked on or bullied by other children	
(v) Gets on better with adults than with other children	
<i>Prosocial behaviour</i>	
(i) Considerate of other people's feelings	Sensitive to the feelings of others; Respects different needs and cultures
(ii) Shares readily with other children	

Continued on next page

Table B.3 – continued from previous page

SDQ Item	Closest EYFSP Indicator(s)
(iii) Helpful if someone is hurt, upset or feeling ill	
(iv) Kind to younger children	
(v) Often volunteers to help others (parents, teachers, other children)	

Notes: The table aligns items with similar definitions from the 25-item SDQ and the 32 EYFSP learning goals retained in the factorization analysis. The SDQ comprises five scales: hyperactivity, conduct problems, emotional problems, peer problems, and prosocial behaviour, each containing five items (Goodman, 1997). These are compared with the 15 items underlying the socio-emotional domain derived in this paper. Note that not all 15 EYFSP items have close SDQ matches.

Table B.4: Factor Loadings from EFA using EYFSP Point Scales

EYFSP Scales	Point Scales	Factor 1	Factor 2	Factor 3	Factor 4
Numbers & counting	Recognizes numerals 1 to 9	0.955	-0.010	-0.117	-0.008
Reading	Reads familiar words independently	0.929	0.011	-0.022	-0.101
Writing	Writes own name & words from memory	0.920	0.045	-0.139	0.054
Writing	Writes simple regular words	0.862	0.046	0.092	-0.198
Linking sounds & letters	Blends sounds in words	0.851	0.047	-0.007	-0.045
Calculating	Relates addition by combining groups	0.797	0.039	0.026	0.143
Linking sounds & letters	Links sounds to letters	0.792	-0.007	-0.101	0.064
Calculating	Finds one more (less) (numbers 1-10)	0.782	0.007	0.162	-0.125
Reading	Knows print directionality	0.776	0.113	-0.018	0.141
Numbers & counting	Says number names in order	0.743	0.073	-0.021	0.181
Calculating	Uses addition/subtraction vocabulary	0.704	-0.008	0.227	0.034
Writing	Forms recognizable letters	0.703	0.113	-0.120	0.129
Shape, space & measures	Recreates simple patterns	0.615	0.073	0.072	0.269
Reading	Understands story elements	0.614	0.083	0.180	0.178
Numbers & counting	Uses math ideas to solve problems	0.613	0.009	0.382	-0.196
Writing	Attempts writing for various purposes	0.600	0.113	0.172	0.051
Shape, space & measures	Describes shape & size	0.549	0.022	0.284	0.150
Shape, space & measures	Uses words to describe position	0.524	0.050	0.224	0.247
Shape, space & measures	Compares quantities	0.470	0.033	0.430	0.000
Reading	Retells narratives in sequence	0.464	0.080	0.427	-0.031
Reading	Understands info. from non-fiction texts	0.471	0.033	0.484	-0.188
Physical development	Demonstrates fine motor control	0.355	0.231	0.025	0.307
Communication & thinking	Listens, responds to stories, rhymes	0.346	0.293	0.115	0.310
Social Development	Works as part of a group	0.084	0.869	-0.290	0.132
Emotional Development	Considers conseq. of words, actions	0.009	0.860	0.130	-0.170
Social Development	Understands agreed codes of behaviour	0.108	0.809	0.033	-0.011
Emotional Development	Sensitive to the feelings of others	0.055	0.781	-0.052	0.180
Social Development	Forms good relationships with others	-0.028	0.758	-0.026	0.237
Emotional Development	Understands right & wrong	-0.028	0.753	0.262	-0.256
Dispositions & attitudes	Maintains attention & concentrates	0.184	0.614	0.227	-0.242
Social development	Respects different needs & cultures	0.064	0.557	0.377	-0.077
Emotional development	Responds with appropriate feelings	0.074	0.533	0.069	0.375
Dispositions & attitudes	Interested, motivated to learn	0.178	0.488	0.206	0.128
Emotional development	Respect for own & others' beliefs	0.115	0.443	0.290	0.116
Communication & thinking	Interacts in various contexts	0.165	0.433	0.277	0.211
Creativity development	Expresses ideas with materials	-0.048	0.091	0.782	0.109
Knowledge of the world	Builds & constructs with objects	-0.007	0.084	0.687	0.000
Creativity development	Uses imagination in creative forms	-0.011	0.126	0.664	0.243
Shape, space & measures	Uses math ideas to solve problems	0.386	-0.007	0.596	-0.142
Physical development	Recognizes importance of health	0.157	0.173	0.591	-0.019
Communication & thinking	Speaks confidently, shows awareness	0.137	0.216	0.578	0.036
Communication & thinking	Organises & clarifies thinking	0.247	0.166	0.563	0.043

Continued on next page

Table B.4 – continued from previous page

EYFSP Scales	Point Scales	Factor 1	Factor 2	Factor 3	Factor 4
Dispositions & attitudes	Confident to try new activities	0.056	0.257	0.529	0.148
Knowledge of the world	Identifies uses of technology	0.200	0.071	0.506	0.074
Physical development	Handles tools, materials safely	0.055	0.298	0.457	0.138
Knowledge of the world	Knows about past & present events	0.247	0.134	0.421	0.207
Knowledge of the world	Asks questions about how things work	0.335	-0.001	0.368	0.361
Creativity development	Recognizes sounds, matches movements	0.239	0.134	0.350	0.293
Creativity development	Explores color, texture, & shape	0.172	0.150	0.299	0.452
Knowledge of the world	Investigates using all senses	0.305	0.106	0.279	0.451
Communication & thinking	Uses language to imagine & recreate	0.169	0.198	0.328	0.432
Creativity development	Sings simple songs from memory	0.280	0.192	0.145	0.432
Physical development	Moves with confidence & imagination	0.122	0.362	0.024	0.431
Dispositions & attitudes	Selects activities independently	0.118	0.377	0.176	0.394
Dispositions & attitudes	Dresses & undresses independently	0.199	0.343	-0.024	0.390
Physical development	Uses small & large equipment	0.154	0.193	0.268	0.383
Eigenvalues		32.85	2.77	2.17	1.17
Number of observations			652,070		

Notes: MoJ–DfE data. Table reports factor loadings from the first iteration of an Exploratory Factor Analysis (EFA) using tetrachoric correlation matrices and oblique quartimin rotation. Eigenvalues for each factor are also reported. The analysis uses 65 indicators from all EYFSP scales for pupils with complete information on EYFSP point scales and variables of interest. Items then are removed first due to high correlations (≥ 0.85) and subsequently at each iteration due to low loadings (< 0.5) or high cross-factor loadings (> 0.75). Loadings shown in **bold** correspond to the primary items for each factor that are retained in the final iteration and satisfy these criteria. Total number of observations are rounded to the nearest 10 for confidentiality.

C Additional Results Quoted in the Main Paper

Table C.1: SEND (Primary Needs) and CIN (CPP and CLA) status on Early Skill Deficits

Panel A: SEND	Any plan	BESD	Learn.Disab.	Phys.Disab.
<i>Early Skill Deficits</i>				
Socio-emotional	0.051*** [0.001]	0.088*** [0.001]	-0.028*** [0.001]	0.027*** [0.001]
Cognitive	0.016*** [0.001]	-0.079*** [0.001]	0.151*** [0.002]	0.025*** [0.001]
Mean of Dep. var	0.070	0.118	0.251	0.152
Observations	210,930	210,930	210,930	210,930
R-squared	0.334	0.181	0.210	0.123
Panel B: CIN	CPP	CPP Neglect	CPP Abuse	CLA
<i>Early Skill Deficits</i>				
Socio-emotional	0.005*** [0.002]	0.001 [0.001]	0.002* [0.001]	0.010*** [0.001]
Cognitive	0.014*** [0.002]	0.008*** [0.001]	0.001 [0.001]	0.009*** [0.001]
Mean of Dep. var	0.095	0.030	0.045	0.06
Observations	85,650	85,650	85,650	85,650
R-squared	0.151	0.138	0.132	0.163
<i>Background controls</i>	X	X	X	X
School FE	X	X	X	X

Note: MoJ-DfE data. The table reports linear probability model (LPM) estimates for the following outcomes: *Any plan* indicates whether a pupil has been assigned any SEND support plan. *BESD* refers to behaviour, emotional, and social difficulties as the primary SEND need. *Learn.Disab.* indicates whether the primary need is a learning disability, including moderate, severe, specific, or multiple learning difficulties. *Phys.Disab.* indicates whether the primary need is a physical disability, including speech, language and communication needs, hearing, visual, or multi-sensory impairments. *CPP* indicates that a child has been issued a child protection plan (e.g. neglect and physical/emotional abuse), and *CLA* indicates whether a pupil has ever been identified as a Child Looked After. All outcomes are measured from school reception to the KS2 assessment year. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, and school fixed effects. Pupils who were already identified as having SEND or CIN prior to the EYFSP assessment are excluded from the estimation samples, respectively. Standard errors in brackets are clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. The total number of observations is rounded to the nearest 10, and sample means (%) for CLA are rounded to zero decimal places for confidentiality.

Table C.2: Main Reasons for Exclusions on Early Skill Deficits

	Physical Assault	Verbal Abuse	Serious Misbehaviour
<i>Early Skill Deficits</i>			
Socio-emotional	0.038*** [0.007]	0.002 [0.007]	0.027*** [0.007]
Cognitive	-0.013** [0.007]	0.001 [0.007]	-0.01* [0.007]
Background controls	X	X	X
School FE	X	X	X
Mean of Dep. var	0.563	0.312	0.387
Observations	11,790	11,790	11,790
R-squared	0.329	0.302	0.330

Note: MoJ–DfE data. The table reports linear probability model (LPM) estimates for the main reasons for exclusion, whether fixed-term or permanent. *Physical Assault* refers to exclusions for physical assault against another pupil or an adult. *Verbal Abuse* refers to verbal abuse or threatening behaviour towards a pupil or an adult. *Serious Misbehaviour* captures exclusions related to sexual misconduct, drug- or alcohol-related incidents, damage or theft, and persistent disruptive behaviour. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM, as shown in equation 1. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and school fixed effects. Standard errors in brackets are clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

Table C.3: Share of Early Skill Deficits Explained by School Outcomes & SEND/CIN (Level 2 - Sequential Mediation via KS2)

	Only School Outcomes		Including SEND and CIN	
	Socio-emotional	Cognitive	Socio-emotional	Cognitive
Absenteeism	0.031*** [0.022, 0.040]	0.295*** [0.265, 0.325]	0.021*** [0.014, 0.028]	0.214*** [0.180, 0.248]
Exclusions	0.390*** [0.362, 0.418]	—	0.379*** [0.357, 0.402]	—
KS2 score	—	1.261*** [1.157, 1.364]	—	0.798*** [0.716, 0.880]
SEND	—	—	0.052*** [0.036, 0.068]	0.491*** [0.434, 0.548]
CIN	—	—	0.143*** [0.123, 0.162]	0.211*** [0.177, 0.245]
Observations	650,500			

Notes: MoJ–DfE data. Table reports the share of the total effect of early skill deficits on offending explained by each mediation pathway, accounting for sequential mediation via KS2 attainment. Columns 1 and 2 include only absenteeism, exclusions, and KS2 attainment as mediators. Columns 3 and 4 additionally include SEND and CIN as mediators. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, IDACI deciles, month and year of birth, SEND and CIN status prior to the EYFSP assessment (Columns 1 and 2), and school fixed effects. Confidence intervals are reported at the 95% level and based on standard errors from 250 bootstrap replications clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

Table C.4: Selection into the Main Analytical Sample

	[1]	[2]
<i>Early Childhood Characteristics</i>		
SEND (before EYFSP assessment)	0.879*** [0.817, 0.946]	0.989 [0.950, 1.030]
CIN (before EYFSP assessment)	0.792*** [0.734, 0.855]	0.967 [0.922, 1.014]
Female	1.007 [0.997, 1.018]	1.005 [0.998, 1.012]
Number of siblings	0.998 [0.990, 1.007]	1.001 [0.997, 1.005]
First_born	0.997 [0.984, 1.009]	0.999 [0.991, 1.008]
Non-white	0.997 [0.957, 1.037]	0.989 [0.976, 1.002]
Non-English first lang.	0.954* [0.904, 1.007]	0.992 [0.977, 1.008]
Free school meals	0.993 [0.972, 1.015]	0.999 [0.989, 1.008]
Most deprived area	1.036 [0.931, 1.151]	1.053 [0.989, 1.120]
School Level Vars		X
Local Authorities FE	X	X
Mean of Dep. var	0.430	0.430
Observations	1,513,470	1,513,470
Pseudo R-squared	0.490	0.490

Note: MoJ–DfE data. Table reports odds ratios (ORs) from logit models for the probability that a pupil is included in the main analytical sample, defined as having complete EYFSP point-scale information. All specifications control for gender, SEND and CIN prior to the EYFSP assessment, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, first language other than English, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, Local Authority (LA's) fixed effects and school-level variables (within-school averages of early childhood characteristics), as indicated. Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

Table C.5: Key Outcomes Re-weighted Using IPW for Sample Selection

	Absenteeism		Excluded		KS2 Score		Offending	
	OLS	OLS (IPW)	LPM	LPM (IPW)	OLS	OLS (IPW)	ME	ME (IPW)
<i>School outcomes</i>								
Absenteeism score					-0.068*** [0.001]	-0.069*** [0.002]	0.280*** [0.016]	0.268*** [0.022]
Exclusions					-0.167*** [0.008]	-0.169*** [0.011]	2.256*** [0.062]	2.431*** [0.083]
KS2 score							-0.411*** [0.021]	-0.418*** [0.030]
<i>Early Skill Deficits</i>								
Socio-emotional	0.020*** [0.002]	0.020*** [0.003]	0.017*** [0.000]	0.017*** [0.001]	0.040*** [0.002]	0.035*** [0.003]	0.168*** [0.027]	0.161*** [0.039]
Cognitive	0.155*** [0.002]	0.154*** [0.003]	-0.005*** [0.000]	-0.005*** [0.000]	-0.660*** [0.002]	-0.660*** [0.003]	-0.017 [0.032]	0.034 [0.047]
<i>Socio-demog. controls</i>								
Female	0.013*** [0.002]	0.013*** [0.003]	-0.025*** [0.000]	-0.023*** [0.001]	-0.061*** [0.002]	-0.063*** [0.003]	-2.366*** [0.050]	-2.391*** [0.070]
No. siblings	-0.037*** [0.001]	-0.035*** [0.002]	0.001*** [0.000]	0.001*** [0.000]	0.003*** [0.001]	0.006*** [0.002]	0.009 [0.020]	0.015 [0.028]
First born	0.070*** [0.002]	0.068*** [0.003]	-0.001** [0.000]	-0.001 [0.001]	0.055*** [0.002]	0.056*** [0.003]	0.001 [0.040]	0.042 [0.056]
Non-white	0.012*** [0.005]	0.011* [0.006]	0.002*** [0.001]	0.003*** [0.001]	0.091*** [0.004]	0.086*** [0.005]	0.492*** [0.058]	0.551*** [0.080]
Free school meals	0.495*** [0.004]	0.476*** [0.005]	0.021*** [0.001]	0.021*** [0.001]	-0.131*** [0.003]	-0.130*** [0.003]	1.794*** [0.044]	1.836*** [0.060]
Most deprived area	0.234*** [0.008]	0.223*** [0.009]	0.009*** [0.001]	0.008*** [0.001]	-0.137*** [0.006]	-0.137*** [0.007]	1.340*** [0.107]	1.218*** [0.147]
Local Authority FE							X	X
School FE	X	X	X	X	X	X		
Mean of Dep. Var.								
Observations	644,890	644,890	644,890	644,890	644,890	644,890	646,480	646,480
R-squared	0.390	0.415	0.138	0.158	0.517	0.522		
Pseudo R-squared							0.166	0.175

Note: MoJ-DfE data. Table reports OLS, linear probability model (LPM), and marginal effects (ME multiplied by 100) estimates for absenteeism, school exclusions, KS2 score, and offending, from a baseline model and a model using inverse probability weighting (IPW) to account for sample selection. Early skill deficits are standardised (reverse-coded) factor scores derived from ESEM using information from EYFSP point scales. All specifications control for gender, SEND and CIN status prior to the EYFSP assessment, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, and fixed effects for Local Authorities (LAs) and schools, as indicated. Standard errors in brackets are clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.

Table C.6: Offending on Early Skill Deficits, School Outcomes and SEND/CIN (Bifactor Specification)

	[1]	[2]	[3]	[4]	[5]
<i>School outcomes</i>					
Absenteeism score			1.181*** [1.158, 1.203]	1.116*** [1.095, 1.138]	1.117*** [1.096, 1.139]
Exclusions			3.355*** [3.118, 3.611]	3.124*** [2.901, 3.365]	3.150*** [2.923, 3.391]
KS2 score			0.751*** [0.733, 0.769]	0.791*** [0.773, 0.811]	0.785*** [0.766, 0.804]
<i>Early Skill Deficits</i>					
Socio-emotional		1.250*** [1.225, 1.274]	1.175*** [1.152, 1.199]	1.150*** [1.127, 1.174]	1.230*** [1.187, 1.275]
Cognitive		1.244*** [1.217, 1.269]	1.100*** [1.075, 1.125]	1.073*** [1.049, 1.098]	1.138*** [1.095, 1.183]
General	1.327*** [1.301, 1.354]	1.208*** [1.181, 1.236]	0.994 [0.968, 1.021]	0.944*** [0.919, 0.970]	1.116*** [1.070, 1.163]
Socio-emotional × SEND					0.924*** [0.886, 0.964]
Cognitive × SEND					0.955** [0.914, 0.999]
General × SEND					0.816*** [0.779, 0.855]
Socio-emotional × CIN					0.960** [0.923, 0.999]
Cognitive × CIN					0.919*** [0.881, 0.959]
General × CIN					0.905*** [0.867, 0.945]
<i>Socio-demographic controls</i>					
SEND				1.294*** [1.225, 1.366]	1.359*** [1.289, 1.433]
CIN				2.178*** [2.078, 2.283]	2.408*** [2.286, 2.539]
Other backg. controls	X	X	X	X	X
School FE	X	X	X	X	X
Mean of Dep. var	0.028	0.028	0.028	0.028	0.028
Observations	457,560	457,560	457,560	457,560	457,560
Pseudo R-squared	0.120	0.127	0.162	0.176	0.178

Note: MoJ–DfE data. Table reports odds ratios (ORs) from conditional logit models for adolescent offending, a binary indicator for whether a pupil has ever been cautioned or sentenced for an offence between ages 11 and 17. Early skill deficits are standardised (reverse-coded) factor scores obtained from a dedicated bifactor measurement system. All specifications control for gender, birth order, part-time attendance, eligibility for free school meals, non-white ethnic background, English as an additional language, number of siblings, deciles of the Income Deprivation Affecting Children Index (IDACI), month and year of birth, SEND and CIN status prior to the EYFSP assessment, and school fixed effects. Confidence intervals are reported at the 95% level and based on standard errors clustered at the school level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Total number of observations are rounded to the nearest 10 for confidentiality.