

# From school absences to crime involvement

Authors: Dr David Buil-Gil and Prof Ken Pease

Date: August 2025

**This Data Insight explores the relationship between school absenteeism and subsequent patterns of criminal activity. It addresses the question ‘*What are the long-term effects of school absenteeism on crime at different stages of life?*’ Specifically, it examines whether authorised and unauthorised absences are associated with future cautions and convictions.**

To investigate this, we used linked data from the Department for Education (DfE) and the Ministry of Justice (MoJ), covering 4,926,067 pupils born between 1990/91 and 1997/98 in England. School absences data (2005–2013) and cautions and convictions (2005–2022) are analysed using descriptive statistics and survival analysis.

Our findings reveal that pupils with higher levels of recorded absences are more likely to be involved in cautions and convictions later in life. The association is particularly strong for unauthorised absences, which increase the risk of first-time offending by 0.5% per session missed. This corresponds to a 9% higher risk of cautions and convictions for pupils with 20 unauthorised absences and a 26% higher risk for those with 50. Authorised absences also show a slightly smaller but still substantial and significant effect. These patterns persist over the life course, with risk rising consistently at higher levels of absence.

## Background

School absenteeism—whether authorised (e.g. due to illness or appointments) or unauthorised (e.g. truancy)—has long been considered a sign of educational disengagement and a potential early indicator of later life challenges. Previous research has linked absenteeism to [lower educational attainment](#), [increased emotional difficulties](#), and [poorer labour market outcomes](#). There is also evidence suggesting a link between persistent absenteeism and [risky behaviours](#), including [criminal activity](#).

However, the existing literature has often relied on smaller-scale longitudinal studies or regional administrative datasets with limited follow-up periods. These studies, while valuable, have generally lacked statistical power and national representativeness to support strong conclusions about long-term trends in crime following school absences.

This study draws on one of the most comprehensive linked administrative datasets available in the UK: the [Ministry of Justice–Department for Education \(MoJ–DfE\) linked dataset](#). By following nearly five million pupils from school age throughout early adulthood, this research provides an unprecedented opportunity to examine the long-term relationship between school absences and future cautions and convictions.

---

## What we did

We used the MoJ-DfE linked dataset, which combines an extensive set of variables from the largest cohort of its kind in England. Records from the MoJ and DfE are linked via unique identifiers, allowing us to track all 4,926,067 individuals born between 1990/91 and 1997/98 in England and examine cautions and convictions up to age 31. This linkage allows for an in-depth investigation of how absenteeism relates to crime across the life course.

In addition to school absences and cautions and convictions, we also extracted data on age, sex, ethnicity, the Income Deprivation Affecting Children Index (IDACI), area of residence, and prison records. These variables will be examined in future phases of the project.

## Measuring absences

Absences data was drawn from the National Pupil Database (NPD), which contains over 15 million records for children born after August 1985 in England. This study focuses on pupils born between 1990/91 and 1997/98. For each pupil, we obtained detailed information on authorised and unauthorised absences by academic year. Pupils were grouped into quintiles based on total absences, with the bottom 20% (lowest absence rates) in the first quintile and the top 20% in the fifth.

We then examined whether pupils in higher quintiles were more likely to have at least one caution or conviction at various ages.

## Measuring crime

Criminal involvement was measured using the Police National Computer (PNC), which contains 1.65 million offence records from 2000–2022. These include offences that resulted in a caution or conviction. We merged PNC data with NPD records using unique identifiers and examined the probability of each pupil having at least one caution or conviction at different ages.

At this stage, our analysis includes all offence types. Disaggregation by offence category will be explored in future work.

Overall, 717,495 of 4,926,067 pupils in this dataset (14.6%) had at least one caution or conviction.



---

## Survival analysis

We began with [descriptive statistics](#) to examine the relationship between school absenteeism (authorised and unauthorised) and later cautions and convictions. We calculated the total number of authorised and unauthorised absences for each pupil across all academic years, and grouped them into quintiles based on their accumulated number of absences. For each group, we calculated the proportion of pupils with at least one caution or conviction during the observed period. These proportions are analysed at various age points to investigate whether absenteeism levels are associated with differences in the likelihood and timing of future criminal activity.

We then applied [survival analysis](#) to model the timing of first-time cautions and convictions. This method accounts for censored cases (i.e. pupils not known to have offended during the observation period) and estimates the probability of first cautions and convictions over time. We used a Cox proportional hazards model – a statistical method for examining the relationship between the timing of an event and one or more predictor variables – to assess how cumulative absences relate to the hazard (probability and timing) of first-time cautions and convictions across different levels of absenteeism.



## What we found

Authorised absences were nearly universal, recorded for 99.2% of pupils (4,887,321 out of 4,926,067) (Table 1). Unauthorised absences were also widespread, with 66.4% (3,271,283 pupils) having at least one unauthorised session. On average, pupils had 97.5 authorised absences (median = 61) and 19.3 unauthorised absences (median = 3). The maximum recorded for a single pupil exceeded 1,500 authorised and 1,400 unauthorised absences. Overall, 717,495 pupils (14.6%) had at least one recorded caution or conviction by age 31. The mean number of cautions and convictions was 1.6; the median was 0; the maximum was over 350.

**Table 1** – Summary of school absences and cautions and convictions per person

	Min	1 <sup>st</sup> Q	Mean	Prevalence at least 1 (%)	Median	3 <sup>rd</sup> Q	Max*
Authorised absence	0	32	79.5	99.2	61	107	Over 1,500
Unauthorised absence	0	0	19.9	66.4	3	14	Over 1,400
Overall absence	0	37	99.4	99.4	70	127	Over 1,700
Cautions and convictions	0	0	1.6	14.6	0	0	Over 350

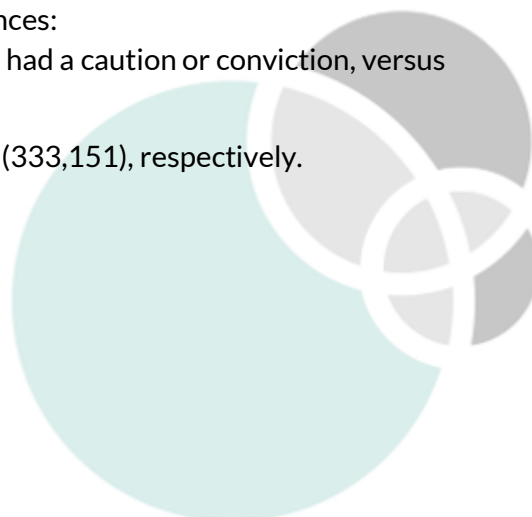
\*Maximum values have been rounded up to the nearest hundred (or nearest ten, for cautions and convictions) at which at least 10 unweighted pupils are represented. This approach minimises the risk of reidentification and follows best practices in statistical disclosure control.

Figure 1 shows the proportion of pupils with cautions and convictions by absence quintile.

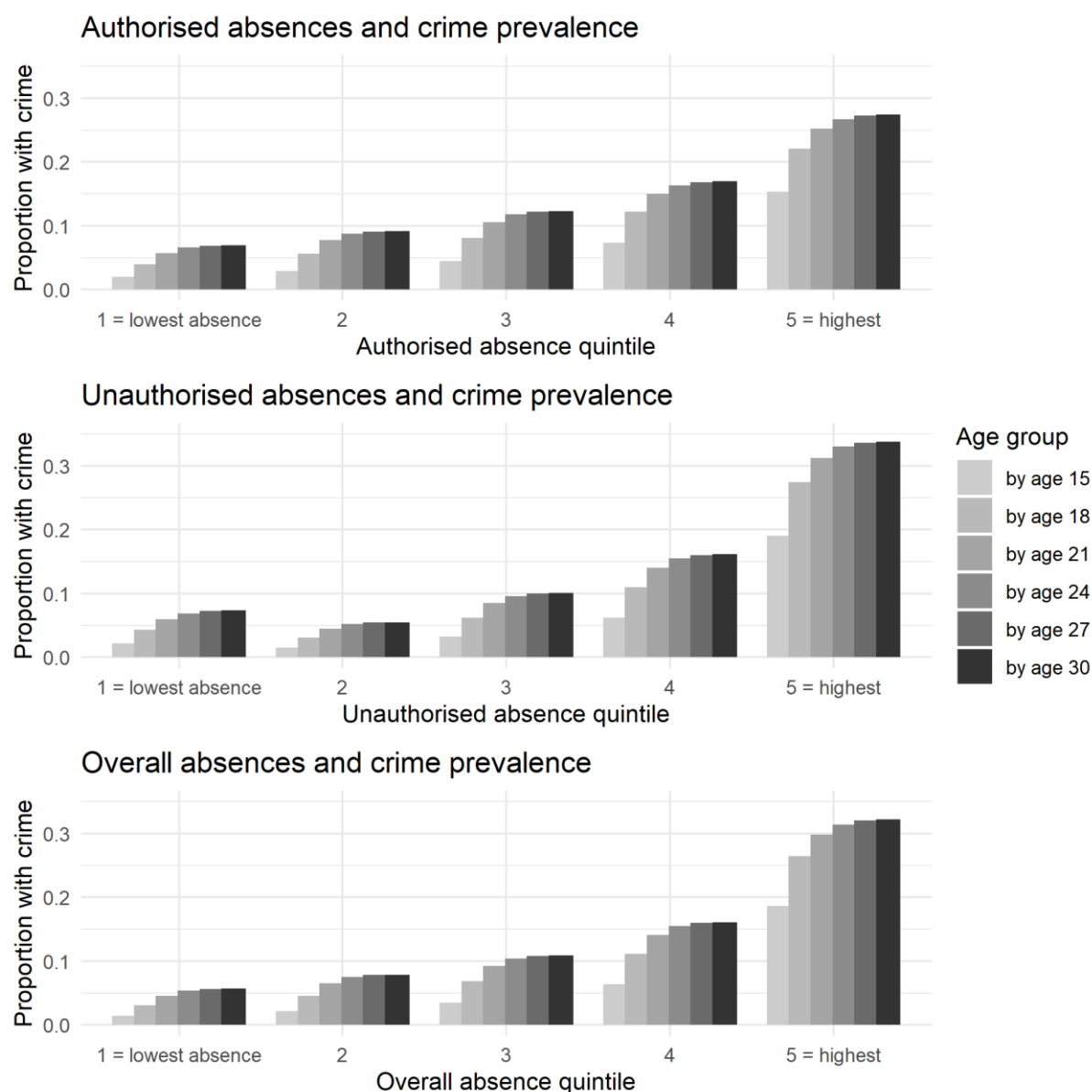
- By age 15, 2% of pupils in the lowest authorised absence quintile had a caution or conviction (19,630 out of 985,214), compared with nearly 15% in the highest (151,361 out of 985,213).
- By age 30, these figures were around 7% (68,560) and 27% (270,682), respectively.

The pattern is even more pronounced for unauthorised absences:

- By age 18, 4% of pupils in the lowest quintile (42,609) had a caution or conviction, versus 27% in the highest (270,590).
- By age 30, these figures rose to 7% (72,401) and 33% (333,151), respectively.



**Figure 1 – Percentage of pupils with cautions and convictions by absence quintiles**



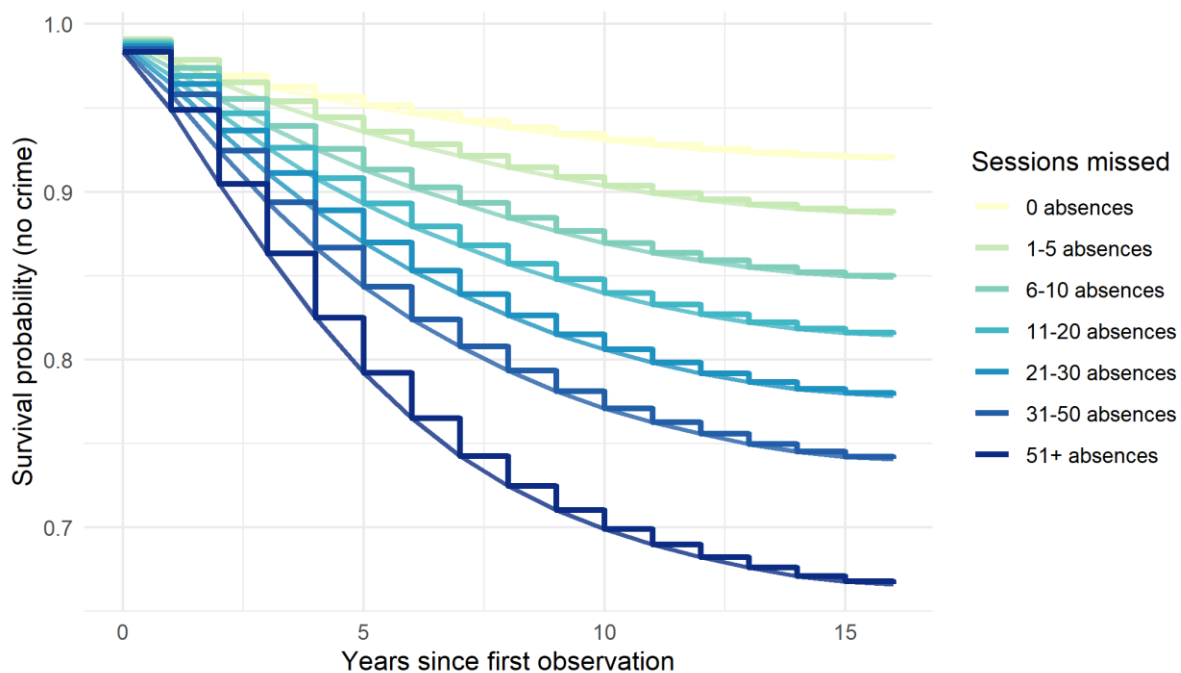
The Cox proportional hazards model shows that each additional unauthorised absence increases the risk of first-time cautions or convictions by 0.5% (hazard ratio [HR] = 1.0046,  $p < 0.001$ ). This equates to a 9% higher risk for pupils with 20 unauthorised absences and 26% for those with 50. Authorised absences also raise the risk, with a 0.4% increase per session ( $p < 0.001$ ). The same 0.4% increase applies when combining both authorised and unauthorised absences.

Figure 2 presents survival curves showing the probability of no recorded crime over time by level of unauthorised absences. After 16 years:

- 92% of pupils with no unauthorised absences remained crime-free (1,547,741 out of 1,681,818)
- 89% for those with 1–5 absences (1,154,695 of 1,300,851)
- 85% for 6–10 (438,576 of 516,613)
- 81% for 11–20 (397,283 of 487,701)
- 78% for 21–30 (195,288 of 251,038)
- 74% for 31–50 (187,438 of 253,472)
- 66% for over 50 absences (288,174 of 434,574)

Similar trends, though less pronounced, are observed for authorised absences.

Figure 2 – Survival probability until first caution or conviction by unauthorised absence groups



## Why it matters

Understanding the long-term implications of school absenteeism is vital for designing more effective early intervention policies. This research offers evidence that higher levels of school absences, particularly unauthorised absences, are strongly associated with increased risk of later criminal involvement, even after many years. These findings underscore the importance of viewing absenteeism not just as an educational issue, but as a cross-sectoral concern with implications for criminal justice, social care, and community safety.

The public value of this research lies in its potential to support early identification of risk and inform targeted interventions. By identifying patterns that link absenteeism with future offending, policymakers and practitioners can better allocate resources to prevent long-term harm. This is especially relevant for supporting vulnerable groups, reducing educational and social inequalities, and informing cross-departmental strategies that span education, justice, and public health.

## What next?

Future stages of this project will explore how the link between absenteeism and crime varies by demographic and socio-economic characteristics. We will assess differences by sex, ethnicity, deprivation level, area characteristics, and type of offence, to determine whether certain groups are disproportionately affected or involved in specific forms of criminal behaviour.

This is vital for developing more effective early intervention strategies. By understanding the differential impact of absenteeism, practitioners can better target resources to reduce inequality, prevent crime, and minimise the long-term harms associated with criminal justice involvement.



## Disclaimer

This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners (e.g. Department for Education, Ministry of Justice) in relation to the interpretation or analysis.

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.

## Acknowledgements

This work is supported by ADR UK (Administrative Data Research UK). ADR UK is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve people's lives. ADR UK is an Economic and Social Research Council (ESRC) investment (part of UK Research and Innovation). [Grant number: ES/Z503204/1].

## Contact

Name: Dr David Buil-Gil  
Email: david.builgil@manchester.ac.uk

