

# Childhood educational predictors of re-offending trajectories

**Author:** Dr Hannah Dickson

**Date:** August 2024

## Summary

This Data Insight examines the feasibility of using childhood educational information to develop predictive models of re-offending trajectories. A re-offending trajectory is the pattern by which a person commits crimes throughout their lifetime. [Read more about what this means](#). This research aimed to explore whether we could use de-identified administrative educational and social care data to potentially identify the characteristics of individuals who may go on to become persistent offenders.

We used the Ministry of Justice and Department for Education linked dataset – England. This included de-identified data from the Police National Computer (PNC) and National Pupil Database (NPD)<sup>1</sup>.

## Background

Criminal behaviour is a global public health problem, associated with a wide range of poor health and social outcomes for both victims and perpetrators. Such behaviour typically follows distinct pathways or trajectories, with some individuals behaving antisocially throughout their life, and others for only short periods of time such during adolescence. This is known as the taxonomy of antisocial behaviour<sup>2</sup> – taxonomy is the practice of categorisation.

In the first part of my ADR UK Research Fellowship, using information taken from the PNC, I found five different re-offending trajectories:

1. **Non-violent\* adolescent-limited prolific (AdoLim-prolific):** this group committed a higher than average number of non-violent offences during adolescence only.
2. **Non-violent adolescent-limited low density (AdoLim-Low):** this group committed a low number of non-violent offences in adolescence only.

---

\* Violent offences were considered to be all 'Violence against the person' offences and contact sexual offences. Non-violent offending was all other offence codes.

3. **Mixed offence type life course prolific (LCP):** this group committed different offence types at a higher-than-average rate throughout their lives.
4. **Mixed offence type 'later' onset with escalating offence history (Adol-Late Onset):** this group committed different offence types, with the number of offences rising more steeply in adulthood than in adolescence.
5. **Non-violent adult-limited low density (Adult-low):** this group committed a low number of non-violent offences in adulthood only.

Efforts to prospectively identify the different re-offending trajectories using longitudinal cohort data (which follows a group of people over a long period of time) have reaped limited rewards. This is because of small sample sizes for disadvantaged groups, selection bias and infrequency of data collection<sup>3,4</sup>. Larger child-to-adult longitudinal datasets are therefore required to build better prospective models. These could help inform early targeted interventions in education and criminal justice settings which could reduce the risk of later offending<sup>2</sup>.

Administrative education and social care data in England represents a rich but under-utilised longitudinal source of de-identified information on every child in state education. This includes demographic, education and socioeconomic information. The aim of the second part of my ADR UK-funded fellowship was to examine whether it was possible to use administrative education and social care data to identify the characteristics of children and adolescents that indicate they are more likely to become persistent and prolific offenders, before involvement with the criminal justice system begins. This insight could support better targeting of possible therapeutic interventions.

## What we did

We used de-identified data from the PNC and the NPD for individuals born between 1 September 1990 and 31 August 1999, with at least one record in the PNC between January 2000 and December 2017. To develop our prospective statistical models, we used NPD variables known to be associated with juvenile prolific offending reported by the Ministry of Justice and Department for Education in 2022<sup>5</sup>. In developing the final model variables listed below, we only used information where it preceded the first offence or caution.

- Eligible for free school meals: Yes or no
- Receiving special educational needs support\*: Yes or no
- Ethnicity: White, Black, Asian, Other and Mixed ethnic categories
- Gender: Male or female
- Exclusions (permanent and suspensions): Yes or no
- Income Deprivation Index score
- Key stage 1 (KS1) attainment based on reading, writing, maths and science at 6/7 years: Below versus at/above the national KS1 attainment thresholds
- Gender and KS1 attainment interaction

We used modelling to see how well these educational variables predicted whether an individual was in the AdolLim-prolific, LCP, Adol-Late Onset, or Adult-Low trajectory, compared to the

---

\* SEN support also includes all those with an Education, Health and Care (EHC) plan, where applicable.

AdolLim-Low comparison trajectory – a small number of non-violent offences in adolescence only. This strategy shows us which combination of educational variables are important for each re-offending trajectory.

We made our analysis more robust by repeating the whole modelling process in 100 samples drawn with replacements from the original sample. We then tested how well each model was able to differentiate between each trajectory relative to our comparison trajectory, as well as how accurate the predictions were.

## What we found

The findings are based on 694,192 individuals born between 1 September 1990 and 31 August 1999, who offended between January 2000 and December 2017 and had information on month and year of birth in the NPD. Figure 1 shows how some of the model educational variables differed according to the different re-offending trajectories.

From Figure 1, we can see that that there were larger proportions of individuals who:

- were male
- were eligible for free school meals
- performed below KS1 attainment thresholds
- were receiving special educational needs support

in the LCP compared to other re-offending trajectories.

We also found higher proportions of individuals from Black ethnic backgrounds in the LCP and Adol-Late Onset trajectories. There were proportionally more females in the AdolLim-Low trajectory compared to the others. Approximately one-third of all individuals in this analysis had been excluded from school prior to their first involvement with the criminal justice system. However, for individuals in the AdolLim-Low trajectory, this was only around 25%.

Figure 1. Breakdown of educational variables according to re-offending trajectories



Abbreviations:

- *Adult-low: Non-violent adult-limited low density*
- *Adol-Late Onset: Mixed offence type 'later' onset with escalating offence history*
- *LCP: Mixed offence type life course prolific*
- *AdoLim-Low: Non-violent adolescent-limited low density*
- *AdoLim-Prolific: Non-violent adolescent-limited prolific*

In Table 2 below, we present the basic prediction model statistics although there are others. We present findings for just one comparison; LCP compared to the AdolLim-Low trajectory. Discrimination shows how well the model can discriminate between different outcomes (or trajectories). A good C-statistic has a value over 0.70<sup>6</sup>. Measures of calibration show how accurate the model is at predicting risk. Calibration-in-the-large should be as close to 0 as possible<sup>7</sup> and the calibration slope should ideally be 1<sup>7</sup>. Our findings indicate that the predictive model shows good discrimination between the LCP and AdolLim-Low trajectories. Measures of calibration show that the model does not appear to over- or under-estimate risk.

**Table 2. Discrimination and calibration model performance measures for LCP compared to the AdolLim-Low trajectory**

Performance measures	
Discrimination: C-statistic	0.752
Calibration-in-the-large	0.003
Calibration slope	1.011

## Why it matters

We identified re-offending trajectories that have previously been identified including life course persistent or prolific and adolescent limited trajectories<sup>2,8</sup>. We then went onto develop a prospective statistical model using some early basic education variables to see if it was possible to discriminate between the different re-offending trajectories. This has the potential to help inform early intervention strategies and criminal justice responses to antisocial behaviour, and by extension contribute towards evidence-based policy making.

The one place that most young people have in common is school. Schools are, therefore, the best place in which to identify and support those at risk of offending before such problems escalate. School-based interventions to build students' social and emotional skills or those that support effective problem solving and anger management may help students avoid involvement with the criminal justice system.

Our findings indicate that we can use administrative education information to determine the likelihood of following different re-offending trajectories, prior to first involvement with the criminal justice system. These findings are striking because our statistical model did not include key characteristics that are known to be associated with prolific and persistent offending, such as being known to children's social care, persistent absenteeism, school type and academic attainment during adolescence.<sup>5,9</sup> As part of this research I did engage with young people and professionals in education and criminal justice system via a third sector partner (Synergy, an arts charity). Some of the feedback received was positive "If this dataset could be used to identify the vulnerable groups so early before they start interacting with the legal system, (it could be useful)".

Importantly, our findings show that administrative data is a powerful low-cost alternative to expensive and lengthy prospective longitudinal studies. Administrative data research can produce

findings with high external validity and applicability to UK policy making that can be used to inform public services for young people.

## What next?

In this research fellowship we used a mixed gender sample of individuals who had offended. Although it is widely recognised that women have different drivers of offending behaviours and worse outcomes in the criminal justice system, there is little research examining female offenders.<sup>10,11</sup>

Importantly, there is little consensus about the development of female offending over time. This is because most research on patterns of re-offending consists mostly of male samples. It is important to examine re-offending trajectories in females, because studies of male cohorts suggest that certain re-offending trajectories are associated with more negative outcomes than others. Moreover, as female offenders are more likely to be primary carers than male offenders, there are significant downstream consequences of offending and custodial sentences on their children and families.

In the next phase of our research, we have funding from ADR UK for a PhD student to look at the education and social care drivers of re-offending trajectories in females only. This has the potential to inform criminal justice responses as recommended in the [MoJ's Female Offender Strategy report](#) published in 2018.



---

## Disclaimer

This work was produced using administrative data accessed through the Secure Research Service, part of the Office for National Statistics (ONS). The use of ONS data in this work does not imply the endorsement of the ONS or data owners in relation to the interpretation or analysis of 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.

## References

1. Ministry of Justice and Department of Education, released 17 November 2021, ONS Secure Research Service dataset, Bespoke Police National Computer and National Pupil Database extract.
2. Moffitt, T. E. (2018). Male antisocial behaviour in adolescence and beyond. *Nature Human Behavior*, 2, 177–186.
3. Bosick, S.J., Bersani, B.E. & Farrington, D.P. (2015). Relating Clusters of Adolescent Problems to Adult Criminal Trajectories: A Person-Centered, Prospective Approach. *Journal of Developmental Life Course Criminology*, 1, 169–188.
4. Jolliffe, D., Farrington, D. P., Piquero, A. R. (2017). Prevalence of life-course persistent, adolescence-limited, and late-onset offenders in prospective longitudinal studies. *Aggression and Violent Behavior*, 33, 4–14.
5. Department for Education and Ministry of Justice (2022). *Education, Children's social care and offending*. Descriptive statistics. London.
6. Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J. & Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 21(1), 128-38.
7. Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., Steyerberg, E. W; On behalf of the Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative. (2019). Calibration: the Achilles heel of predictive analytics. *BMC Medicine*. 16, 17(1), 230.
8. Morizot, J. (2019). Trajectories of criminal behaviour across the life course. *The Oxford Handbook of Developmental and life course criminology*. Oxford Online Handbooks. Oxford
9. Department for Education (2023). *Education, children's social care and offending: multi-level modelling*. Research Report. London

10. Odgers, C., Moffit, T. E., Broadbent, J. M, et al (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology* 20, 673-716

11. *Ministry of Justice (2018). Female Offender Strategy. London*

## 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/W002647/1]. The author would like to thank George Vamvakas, Professor Nigel Blackwood, and Dr Roxy Short for their invaluable contribution to this research.

## Contact

Name: Dr Hannah Dickson

Email: [hannah.dickson@kcl.ac.uk](mailto:hannah.dickson@kcl.ac.uk)

