



Ministry
of Justice

Data First Research Bulletin

**Harnessing the potential of linked
administrative data for the justice system**

July 2024



Contents

1. What is Data First?	3
2. Ethnicity and the justice system	4
An investigation into racial bias in court case outcomes in England and Wales	
Ethnic inequalities in the Criminal Justice System	
Understanding the intersections between care experiences and ethnicity in criminal justice involvement	
3. Youth Justice	7
An examination of the education and social care background of children cautioned or sentenced for an offence	
School funding, pupil performance and crime: a quasi-experimental study	
The longitudinal association between school performance trajectories and offending behaviour	
4. Repeat Offenders	10
An exploratory analysis of returning defendants and the potential of linked criminal courts data from 2011 to 2019 in England and Wales	
5. Serious and Organised Crime	11
Understanding the nature, extent and outcomes of serious and organised crime (SOC) cases heard before the Crown Courts in England and Wales (2013-2019)	
6. Looking forwards	12
Annex A Examples of ongoing Data First research projects	13
Annex B Additional analytical outputs from the Data First MoJ-DfE data-share	14

Disclaimer

The views expressed are those of the authors and are not necessarily shared by the Ministry of Justice (nor do they represent Government policy).

1. What is Data First?

A pioneering research, data-linking and academic engagement programme delivering powerful new policy insights.

Some of the most vulnerable people in our society have multiple and complex needs. They are likely to interact with a range of public services, from health and social care to housing and benefits, through to the justice system. While the government collects a significant amount of data as the public interacts with these services, this rich source of insights has until now, been largely underused. Data can be fragmented, siloed, and not shared between organisations. This can make it challenging to establish whether government services are effective.

The investment of ADR UK (Administrative Data Research UK) in the Data First programme has enabled the Ministry of Justice (MoJ) to tackle some of these issues and unlock the potential of data generated when people interact with justice services.

Since the inception of Data First in late 2019, using award-winning data-linking software [Splink](#), MoJ has linked and shared eight justice datasets joining cases and people across the criminal, civil and family courts, and prison and probation services. Data First has also released a flagship data-share with the Department for Education (DfE). This linked data is creating opportunities for novel analysis to improve our understanding of justice system users and ‘what works’ to improve their outcomes that has not been possible before.

All datasets created as part of Data First are shared with accredited researchers in an ethical and responsible way, via secure platforms including the ONS Secure Research Server and the SAIL Databank. Making this data accessible to researchers across academia and government is enabling us to enhance our evidence base, providing a more holistic understanding of end-to-end user journeys and outcomes across the justice system and beyond.

Data First reflects the commitment of MoJ to working in partnership with independent expert academics, including via ADR UK funded Fellowships, to deliver innovative new research for policy and practice. To date, the programme has facilitated around 40 projects that align with the department’s strategic evidence priorities as set out in the [MoJ Areas of Research Interest](#) (ARI; 2020). Our partnerships across and beyond government boundaries were recognised through the receipt of the ‘Collaboration’ award at the Analysis in Government Awards 2024.

This bulletin summarises the key findings published to date from research facilitated under Data First, across priority policy themes, such as ethnic disparities in the justice system, youth offending, and repeat offenders. It concludes with a forward look of ongoing projects. To find out more, including data catalogues for the data shared, please see the Data First project pages on [GOV.UK](#) and [ADR UK](#).

2. Ethnicity and the justice system

An investigation into racial bias in court case outcomes in England and Wales

Dr Angela Sorsby – ADR UK Data First Academic Research Fellow

- This research uses multiple regression to explore ethnic disparities in court case outcomes after controlling for factors such as age and offence, using the **Data First linked magistrates’ and Crown Court datasets**.
- For both men and women, defendants in the White ethnic group had significantly more previous convictions before receiving a short sentence of immediate custody than those in any other ethnic group except for the mixed ethnic group. The differences were particularly pronounced for women.
- This indicates people from minority ethnic groups tend to receive a short sentence of immediate custody, as opposed to an alternative sentence such as a community order or suspended sentence order, with fewer previous convictions than those from the White ethnic group.
- Further research explored the impact of a defendant electing to have a Crown Court trial, as compared to having their case heard in the magistrates’ courts, after pleading not guilty to an either-way theft offence. Either-way offences are offences that can be tried in the magistrates’ courts or in the Crown Court. The Data Insight also examines this decision in relation to ethnicity.
- This found that compared to White defendants, Black defendants are significantly more likely (41% more likely) to elect to have their case heard in the Crown Court, when controlling for age, sex, and number of previous convictions. Cases which elect to be heard in the Crown Court are more likely to result in immediate custody, and average sentence lengths are longer, also when controlling for these factors.

Find out more:

- **Blog:** [An investigation into racial bias in court case outcomes in England and Wales](#)
Data Explained: [An investigation into racial bias in court case outcomes in England and Wales](#)
- **Data Insight:** [The number of convictions before the imposition of a short sentence of immediate custody: ethnicity and gender differences](#)
- **Blog:** [Number of convictions before receiving a short, immediate prison sentence: ethnicity and gender differences](#)

- **Data Insight 2:** [Deciding to have a Crown Court jury trial for a theft offence: Consequences, and relationships with ethnicity](#)

Ethnic inequalities in the Criminal Justice System

Dr Kitty Lymperopoulou – ADR UK Data First Academic Research Fellow

- This research, developed in response to the call for evidence by the Lammy Review to explain ethnic disparities in the Criminal Justice System (CJS), examined the extent of ethnic disparities in remand and sentencing based on defendant appearances in the **Data First linked magistrates’ and Crown court datasets** in England and Wales.
- Both remand and sentencing outcomes were shown to vary substantially between ethnic groups after controlling for defendant, case and court characteristics.
- Ethnic minority defendants from Black, Mixed, Chinese, and Other White groups were more likely to be remanded in custody by the police than White British defendants.
- Those from ethnic minority groups appearing in magistrates’ courts were more likely to be committed to the Crown Court for trial compared to white British defendants.
- Among defendants appearing in the Crown Court who entered a plea, ethnic minority defendants had higher rates of a Non-Guilty (NG) plea than white British defendants.
- Among those convicted in the Crown Court, defendants from the Black Caribbean, Chinese and Other White group were more likely to receive a custodial sentence.
- Bangladeshi, Pakistani and Black Caribbean groups received significantly longer sentences than the White British group.

Find out more:

- **Blog:** [Ethnic Inequalities in the Criminal Justice System](#)
- **Data Explained:** [Ethnic inequalities in the criminal justice system](#)
- **Data Insight:** [Ethnic Inequalities in Sentencing in the Crown Court - Evidence from the MoJ Data First Criminal Justice datasets](#)
- **Data comic:** [Ethnic Inequalities in the Criminal Justice System](#)
- **Policy briefing:** [Ethnic Inequalities in the Criminal Justice System: a briefing.](#)
- **Journal article:** [Ethnic Inequalities in Sentencing: Evidence from the Crown Court in England and Wales](#)

Understanding the intersections between care experiences and ethnicity in criminal justice involvement

Dr Katie Hunter – ADR UK Data First Academic Research Fellow

- This research examines the relationship between ethnicity, care experience and youth justice involvement using the **Data First MoJ-DfE data-share**.
- Care-experienced children were disproportionately likely to have youth justice involvement compared to those without care experience. 1 in 3 care-experienced children (33%) born between 1996 and 1999 received a youth justice caution or conviction between the ages of 10 and 17 compared to 4% of those without care experience.
- Some groups of ethnic minority children were even more likely to have youth justice involvement. The analysis showed higher levels of youth justice involvement among nearly all ethnic minority groups compared to White British children. Depending on their ethnicity, care-experienced children were between 2 and 16 times more likely to have youth justice involvement than those who had not been in care.
- The gap in youth justice involvement between care-experienced children and non-care-experienced children widened over time. This gap widened further still for some groups of ethnic minority children.
- The extent of youth justice involvement also varied by care experience. Typically, care-experienced children had more youth justice involvement than non-care-experienced children. care-experienced children received around four cautions or convictions, while those without received approximately one. Some groups of ethnic minority care-experienced children had even higher levels of youth justice involvement.
- A significantly higher proportion of care-experienced children received a custodial sentence compared non-care-experienced children. Custodial sentences were twice as common among Black and Mixed ethnicity care-experienced children compared to White care-experienced children.

Find out more:

- **Blog:** [Exploring ethnicity, care experience and justice systems involvement](#)
- [Data Explained](#)
- **Policy Briefing:** [Care experience, ethnicity and youth justice involvement - key trends and policy implications](#)
- **Report:** [Double Discrimination - Black care-experienced young adults navigating the criminal justice system](#)
- **Animation:** [Challenging \(In\)Justice](#)

3. Youth Justice

An examination of the education and social care background of children cautioned or sentenced for an offence

Ministry of Justice (MoJ) and Department for Education (DfE) joint publication

- Cross-government MoJ and DfE publications investigate the educational and social care backgrounds of children who had been cautioned or sentenced for an offence. This analysis is based on the **Data First MoJ-DfE data-share**.
- A high proportion of children cautioned or sentenced for an offence had been permanently excluded (10%), suspended (71%) or persistently absent from school (81%). Whilst the majority (59%) of children who had been excluded had been cautioned or sentenced for an offence, most children who had been suspended (23%) or persistently absent (9%) were not in the offending groups.
- High proportions of offending groups were recorded as having Special Educational Needs. 80% of those who had been cautioned or sentenced for an offence, and 87% of those cautioned or sentenced for a serious violence offence had been recorded as ever having SEN up to the end of Key Stage 4. This compared with 45% of the overall pupil population.
- 69% of children who were cautioned or sentenced for an offence were eligible for Free School Meals, compared to 34% of the overall pupil population. However, only 10% of children who had ever been eligible for Free School Meals had been cautioned or sentenced for an offence.
- Children in the offending groups were more likely to have been known to children's social care than the overall population. 32% of children who had been cautioned or sentenced for an offence were or had been a Child in Need, and 11% were or had been a Child Looked After.

Find out more in the [publication](#).

School funding, pupil performance and crime: a quasi-experimental study

Dr William Cook – ADR UK Data First Academic Research Fellow

- This research uses regression discontinuity models to examine the impact of two area-based education policies on academic attainment and offending, using the **Data First MoJ-DfE data-share**.
- The Leadership Incentive Grant¹ did not seem to affect crime outcomes. While the research replicated earlier published work on the positive impact of the grant on GCSE attainment, implementing the same methods on crime outcomes indicated that the policy had no effect on crime outcomes.
- The results from the analysis of the Pupil Learning Credits Programme² were mixed. The analysis replicated the results of the official evaluation, showing an increase in maths attainment at age 14 by around five percentage points. It also extended these findings, to show that the programme increased subsequent maths performance at GCSE, particularly for children eligible for free school meals. However, the results for crime outcomes were less certain. Whilst crime reducing effects were found, these were not statistically significant.

Find out more:

- **Blog:** [Area-based education policy - what can we learn from past efforts?](#)
- [Data Explained](#)
- **Data Insight:** [School funding, pupil performance and crime](#)

The longitudinal association between school performance trajectories and offending behaviour

Dr Alice Wickersham – ADR UK Data First Academic Research Fellow

- This research uses growth mixture modelling and multilevel regression modelling to analyse changing school performance as a key predictor of offending behaviour, using the **Data First MoJ-DfE data-share**.

¹ A programme from 2003-2006 that provided additional funding (approximately £0.5 million) to secondary schools with low GCSE pass rates and/or high proportions of pupils eligible for free school meals.

² An early forerunner of the pupil premium policy where, in 2001, secondary schools in mostly urban areas with high proportions of pupils eligible for free school meals were allocated additional per pupil funding. This was to support additional provision for disadvantaged pupils aged 11-14.

- Pupils could generally be described as following one of five possible school performance trajectories: Average Consistent³, Average Increasing⁴, Average Declining⁵, Low Increasing⁶ and Low Consistent⁷.
- Pupils with Average Declining and Low Consistent school performance trajectories were at higher risk of offence conviction or caution during young adulthood. Analysis showed that pupils in these groups were just over twice as likely to be convicted or cautioned for a first offence after Year 11 and up to age 21 than those in the Average Consistent group.
- The association between school performance and offence conviction or caution was similar across different offence types. The risk of subsequent offending was consistently highest in either the Average Declining or Low Consistent trajectory groups, irrespective of what type of offence it was.
- Other sociodemographic characteristics played an important role in these relationships. After adjusting for school performance trajectory, there was an overall higher risk for subsequent offence conviction or caution among:
 - males (as compared to females)
 - Black and mixed ethnic groups (as compared to white ethnic groups)
 - pupils eligible for free school meals
 - pupils who are looked after children in Year 11
 - pupils receiving low levels of Special Educational Needs provision.

Find out more:

- **Blog:** [Exploring educational attainment patterns and criminal offending](#)
- **Blog:** [Using linked data to evaluate special educational needs provision and offending risk](#)
- **Public engagement case study:** [Consulting a pre-existing advisory group about a research project on ethnicity data](#)
- **Data Insight:** [Changes in school performance and involvement in the criminal justice system](#)
- **Data Explained:** [Discrepancies in gender/sex and ethnicity data between the National Pupil Database and the Police National Computer](#). See also: [detailed methods supplement](#).

³ Average Consistent – Pupils generally perform at average or above average levels throughout school.

⁴ Average Increasing – Pupils generally perform around average levels in Year 2, and show a relative increase in their school performance, such that by Year 11 they perform at above average levels.

⁵ Average Declining – Pupils generally perform around average levels in Year 2, but show a relative decline in their school performance, such that by Year 11 they perform at below average levels.

⁶ Low Increasing – Pupils generally perform at below average levels in Year 2, but show a relative increase in their school performance, such that by Year 11 they perform around average levels.

⁷ Low Consistent – Pupils generally perform at below average levels throughout school.

4. Repeat Offending

An exploratory analysis of returning defendants and the potential of linked criminal courts data from 2011 to 2019 in England and Wales

Tom Jackson, Caris Greyson, Dr Ian Rickard and Professor Andromachi Tseloni, Ministry of Justice

- This research enabled detailed analysis of the nature and scale of repeat offending, based on the **Data First linked magistrates' and Crown court datasets**.
- Upon a defendant's first known case in 2011, over half (56%) returned to the magistrates' courts by 2019. The rate of return varied by offence from 82% for theft to 43% for sexual offences (possibly due to longer custodial sentences).
- Defendants were most likely to return for the same offence group, suggesting a specialisation in offending behaviour. Excluding summary offences this was particularly evident for theft. In addition, theft offences were the most common offence group defendants returned for regardless of initial offence.
- Increased returns were more likely to lead to a case being sent to the Crown Court, suggesting an escalation in offending and sentences imposed.
- Defendants across most offence groups were over-represented in the most deprived areas of England and Wales (43% compared with 20% of the general population). This was highest for robbery and theft (50%) and lowest for sexual offences (36%).
- The over-representation of defendants for acquisitive offences residing in the most deprived areas of England and Wales supports existing evidence on the relationship between neighbourhood deprivation and offending. It suggests that area-based policy tools are likely to be appropriate for rehabilitation; for example, initiatives to improve residents' health, education and employment outcomes.
- This Ministry of Justice [publication](#) won the [Linked Administrative Data Award](#) for the **ONS Research Excellence Awards 2022**.

5. Serious and Organised Crime

Understanding the nature, extent and outcomes of serious and organised crime (SOC) cases heard before the Crown Courts in England and Wales (2013-2019)

Tim McSweeney – ADR UK Data First Academic Research Fellow

- This research aims to better understand the nature, extent and outcomes of SOC cases heard before the Crown Courts in England and Wales, using the **Data First Crown Court dataset**. It provides an estimate of the proportion and number of SOC cases appearing before the higher courts and describes how SOC is distributed across different offence types (for example, violence, sexual assault, drugs, and theft).
- The proportion of defendants meeting the study's definition of SOC remained stable between 2013 and 2018, before falling from 2019 (impact of pandemic). Six percent of defendants and 3% of all cases dealt with by the Crown Court met this definition. Most SOC defendants were male, aged in their early 30s, and self-identified as 'white'.
- 57% defendants considered to be SOC-related were appearing before the higher courts for drugs offences. They were over-represented in offences relating to drugs, fraud and money laundering, but comparatively under-represented in offences involving violence, theft, sexual crimes, and possession of weapons.
- On average, when compared with other cases, SOC-related cases involved more hearings and took longer to conclude; they had higher conviction rates which were more stable over the eight-year period examined.
- SOC appearances generated higher levels of crime-related harm, which were distributed and experienced differently across England and Wales. SOC defendants typically resided in more affluent areas.
- Fewer SOC defendants reappeared at court for further offences within two years. They took longer to reappear, their charges had reduced in severity, and there were fewer reappearances overall.

Find out more:

- **Blog:** [What do we know about serious and organised crime and how effective is our response to it?](#)
- **Data Explained:** [Understanding the nature, extent and outcomes of serious and organised crime cases heard before the Crown Court in England and Wales](#)

- **Data Insight:** [Emerging findings on the nature, extent and outcomes of serious and organised crime cases prosecuted in England and Wales](#)
- **Data Insight 2:** [The outcomes of serious and organised crime cases appearing before the criminal courts in England and Wales](#)
- **Animation:** [Who is prosecuted for serious and organised crimes, and what happens to them?](#)

6. Looking forwards

Data First will continue to work in partnership with academic researchers to maximise the use of linked datasets and create a sustainable body of knowledge on justice system users. As more data is linked and shared, further avenues for research will open.

Ongoing research projects are exploring a range of priority themes in line with the MoJ ARI. These include the impact of social care early interventions on offending; an evaluation of enforced alcohol abstinence orders; a review of domestic abuse in family court proceedings; and the outcomes for children subject to private family law cases.

The complete list of projects is published on GOV.UK and can be found on this [link](#) under 'Approved External Data Request Log'. Annex A provides some examples of research projects in the pipeline. The MoJ-DfE datashare has also been used in a range of additional analytical outputs. See Annex B for some key examples.

Please get in touch with the [Data First team](#) if you would like further information.

Annex A Examples of ongoing Data First research projects

Family Justice

Born into care: Newborns in care proceedings in Wales

To provide the first estimate, and profile, of cases of newborns subject to care proceedings in Wales using linked Data First family court and Cafcass data.

Private family law cases in England and Wales: profiles, pathways and outcomes

To provide a detailed exploration of the characteristics of the families involved in private law legal disputes, how they compare over time and to the broader population and other court users. The research will look at how cases are handled by the family justice system, and how children and adults are doing after court intervention.

Reducing Reoffending

Enforced alcohol abstinence: does it reduce reoffending?

To perform analysis of criminal court and probation data to develop a better understanding of how alcohol treatment and abstinence orders are being used and with what impact on reoffending (using a proxy measure). See this [blog](#) for further information.

Impact of the Universal Credit welfare reform on recidivism

To explore the impact of the Universal Credit welfare reform on recidivism. The project will investigate whether any reported impact varies across several dimensions – for instance, by type of offence; the age, gender, or ethnicity of an offender; or the type or severity of an offender’s previous sentence.

The links between social care, education and offending

An investigation into the impact of education policies on crime

Econometric analysis will be used to estimate the effects of recent UK education policies on youth crime. Three separate projects will explore how the introduction of the Literacy and Numeracy Hour programmes affected youth crime and school exclusion; how the UK school-leaving age reforms in 2013 and 2015 affected youth crime; and how school-closures for teacher-training (INSET) days affected youth crime. See more [project details](#).

Evaluation of the impact of Sure Start Local Programmes and Children's Centres on youth offending and contact with the CJS

The MoJ-DfE dataset will be used to evaluate the impact of Sure Start Local Programmes and Children’s Centres on youth offending and contacts with the CJS. This will shed light on whether one of the most important early years initiatives from the last 20 years benefitted

children (and when and how any benefits arose) and to inform the future design of early years policy.

Education and social care predictors of offending trajectories: An administrative data linkage study

Using a retrospective cohort design and latent class analysis, the MoJ-DfE dataset will be used to develop trajectories of (re)offending. The identification of children and young people at higher risk for persistent offending has the potential to inform early intervention approaches and criminal justice responses to reduce offending and contribute to evidence-based policy making. See this [2022 blog](#), [2023 blog](#) and [learnings from the dataset](#).

Annex B Additional analytical outputs from the Data First MoJ-DfE data-share

The MoJ-DfE data-share has been used in a range of other analytical and statistical outputs. For example (this is a non-exhaustive list):

Women and the Criminal Justice System (Ministry of Justice)

This publication compiles statistics from data sources across the CJS, to provide a combined perspective on the typical experiences of females who come into contact with it. It includes a section on the educational background of young people who had been cautioned or sentenced for an offence. See the [2021 publication](#).

Ethnicity and the Criminal Justice System (Ministry of Justice)

This publication compiles statistics from data sources across the CJS, to provide a combined perspective on the typical experiences of different ethnic groups. It includes a summary of Key Stage 4 educational background, covering attainment, Free School Meals, Special Educational Needs, and persistent absence and exclusions. See the [2022 publication](#).

The education and social care background of young people who interact with the criminal justice system (Office for National Statistics)

This [publication](#) examines the educational attainment and provision, social care provision and demographics of young people educated in England who subsequently received a custodial sentence, and comparing with their peers who did not.

Reanalysing education experiments with linked administrative data (Behavioural Insights Team)

This [publication](#) combines administrative data with randomised control trials to understand whether educational interventions led to additional benefits or costs – beyond those initially evaluated.

Education, children’s social care and offending: Local Authority level dashboard (Department for Education)

This [dashboard](#) was created to complement the published [Education, children’s social care and offending](#) with the aim of providing local authority area specific breakdowns, for each local authority in England, where possible.

Education, children’s social care and offending: multi-level modelling (Department for Education)

This [report](#) looks in detail at the demographic, educational, children’s social care and local authority area characteristics of children who have been cautioned or sentenced for any offence and/or a serious violence offence.

Feasibility study to explore the use of the MoJ-DfE linked dataset to aid evaluations (University of Bristol)

This [report](#) explores whether the MoJ-DfE linked dataset can be used for evaluation of early interventions for violent crime prevention. More specifically, it explores how well the dataset can be used to create robust matched control groups.