School exclusions and youth custody

Prepared for the Nuffield Foundation

Claire Cathro, Giulia Tagliaferri, Alex Sutherland

January 2023

THE BEHAVIOURAL INSIGHTS TEAM

This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Contents

Tables and Figures

Tables

Figures

Acknowledgements

We are grateful to the following individuals and organisations for their roles in supporting this research:

The Nuffield Foundation funded this research as part of a broader project titled “Youth custody: educational influences and labour market consequences”.

CLOSER, the home of longitudinal research in the UK, along with Ellen Greaves and Bill Nasim, who respectively funded and produced the Consistent Schools Database (*grant reference: ES/K000357/1,* www.closer.ac.uk), which was made available to us on this project in order to track academisation of schools over time.

Professor Sandra McNally and Dr Jenifer Ruiz-Valenzuela provided a dataset from their own research on headteachers at schools with sixth forms which we used in our analysis of headteacher exclusion rates.

Lal Chadeesingh and James Lawrence (BIT), Raj Chande (Nesta), and Richard Dorsett (University of Westminster) reviewed drafts of the report/results and provided insightful comments and suggestions.

Executive summary

What is the gap this work fills?

This project looks at the overlap between two difficult areas of public policy — youth custody and school exclusions. School exclusion means both permanent removal from a school and temporary removal from a school (suspensions, sometimes referred to as 'fixed term exclusions'). There is a strong relationship between being excluded from school and experiences of criminality and custody later in life; however, it is unclear the extent to which this relationship is causal: that is, whether being excluded actually *causes* more people to commit crime and end up in custody. We are therefore trying to answer the question: *what is the causal impact of being excluded on the likelihood of being in custody in the years following exclusion?*

Why is the question important?

This area of research is controversial because there is an ongoing, and often heated, argument about the relationship between exclusions and *crime*. This debate is partly driven by the absence of sufficient causal evidence about the outcomes following exclusion on those excluded, their classmates and their teachers. We are not looking directly at crime as an outcome in terms of 'all offences committed' but instead at custody, or whether a crime of sufficient severity warranting imprisonment has been committed. This approach means we are not able to perfectly investigate the relationship between exclusion and crime as we only see the end of the "funnel" of criminal justice. However, by knowing whether exclusion leads to an increased risk of custody, we know that the intermediate step must include the commission of a criminal offence.

What did we do?

We use linked data from the National Pupil Database and the National Client Caseload Information System for four cohorts of pupils attending state-funded schools in England, particularly looking at their experiences when they are in Year 10, when pupils are 14-15 years old, which occurs in the years between 2009/10 and 2013/14. We analyse the effect of being permanently excluded or temporarily suspended in Year 10 on the probability of experiencing custody at age 15-17 (inclusive). By combining all of this data we ended up with a sample of more than a million pupils for our analysis, covering several cohorts across different years.

Our analysis uses academisation, a change in school status that moves schools out of direct local council control to greater self-governance, to understand the impact of exclusions on custody. Effectively we treat academisation as a 'shock' that increases the likelihood of exclusion then use a statistical approach called instrumental variables that enables us to leverage this change to understand the impact on later custody.

What did we find?

We find that attending **a school that converts to an academy in Year 10, the year when pupils are most likely to be excluded, increases the probability of receiving a suspension or permanent exclusion by 3 percentage points** (statistically significant at 0.1%). This compares to an average exclusion rate of 6% for the whole sample, or 5% among only students that attend schools that have not yet converted to academies. The finding, consistent with other research, that academisation increased the likelihood of exclusion is important, given that academisation has been a priority for government. This alone does not mean that academisation should not be used, but that policymakers can be clearer about the consequences of the policy and possible trade-offs that are being made.

We find evidence that a Year 10 pupil attending a school that academised resulted in a statistically significant increase in the probability of custody age 15-17, with impacts varying depending on the type of exclusion. Specifically, receiving a *permanent* exclusion increases the probability of custody by 33 percentage points (statistically significant at 5%); for *suspension*, the increase is 1.3 percentage points (statistically significant at 5%). Against the custody rate of 0.1% among pupils as a whole (Table 5, col (1)), this constitutes a sizeable increase; however, the estimated impact relates to compliers – those who are excluded as a result of academisation – whose expected custody rate in the absence of academisation is not observed but is likely to be higher than that seen among pupils as a whole.

We checked these results by restricting our analysis to the year immediately following academisation, which reduced our sample size by 40%, and re-running our analysis. This was a much stricter test of the immediate impact of academisation on exclusion. Those robustness checks were not statistically significant regarding the relationship between exclusion and custody, but the coefficients on exclusion were positive in all the models we ran.

What are the main strengths and limitations of the analysis?

Our approach uses multiple cohorts of pupils, we are able to assess our approach empirically (through checking how strong our instrumental variable is), and our findings have support from other evidence. By employing a large and novel dataset and combining this with an approach that allows us to distinguish causal relationships, we are able to make stronger claims about the impact of exclusion on custody. The robustness of our analysis relies on a series of assumptions. A crucial one is that academisation affects the likelihood of custody *only through its impact on exclusion*. That is: being in a school that academises does not increase or decrease the probability of experiencing custody later in life, *except through* the higher probability of being excluded. We cannot know or test whether this is occurring; however, we conduct additional analysis to lend credibility to this assumption.

What are the implications for policy and practice?

Our interpretation of our results is that *exclusion presents a small but non-ignorable risk of increases in custody, warranting, at the very least, further exploration*. Given the consequences (and cost) of both exclusion and custody, the possibility of a negative effect should be taken seriously by researchers and policy practitioners alike. Greater sensitivity to this possibility may also be needed given the long-standing and widespread use of exclusion

in England. Put differently, if these results were reversed – that exclusion *reduced* the risk of custody by up to thirty percentage points – there would be significant interest.

Our results highlight an onward negative criminal justice outcome of education policy that, to the best of our knowledge, has not previously been quantified. This evidence does not mean that exclusion as a policy should cease – although some may advocate that. What it does mean is that headteachers, teachers and policymakers should take seriously the idea of exclusion *as an intervention* that can and should be better understood in terms of its short and long-run impacts. This can happen through better use of data and evidence that is already available to government to look at the impacts of exclusion on excludees but also on their peers (for whom there may be unquantified benefits). This can also happen through the willingness of government, multi-academy trusts and schools to rigorously *test out* approaches to behaviour management - including approaches that are 'known to work' or are promoted but do not have evidence that they work in terms of improving behaviour or other outcomes.

What are the opportunities for future research?

We suggest that future research can build on this work to better understand the relationship between exclusion and crime by: 1) using data for more cohorts of pupils; 2) linking to additional administrative datasets, particularly to better observe contacts with criminal justice system beyond imprisonment, such as convictions and arrest, and to capture episodes of custody later in life; and 3) exploring alternative instruments in order to replicate this result.

Beyond further assessment of the relationship between exclusion and crime, there is a pressing need to quantify basic facts about exclusion and other interventions, specifically costs and benefits of the policy, which would then allow for more informed decision-making.

1. Policy context and research aims

1.1 Youth custody

In any given year in the past decade there have been between 500 and 1,000 children aged 10–17 in custody.[1] That said, the youth custody population in England and Wales has declined by 82% from 2,821 in March 2001, when England and Wales had one of the highest youth custody rates in Europe (Hazel, 2008, Table 8.1), to 516 in March 2021.[2]

Placing children in prison is a last resort punishment, although not one that every country uses (Aebi et al., 2021).[3] In England and Wales, prison for those under 18 is now largely used for the most serious *offences* and repeat *offenders*. There are numerous statistics relating to children in custody and what is happening in the secure estate. For example, while most children in prison are there for more serious crimes, typically violence, one-third are there for non-violent crimes.[4] Of those children in the secure estate, nearly half (48%) were from a minority ethnic background, compared to one-quarter ten years ago. Self-harm and assaults remain high or are increasing (Prison Reform Trust, 2019). Prisoners are much more likely than the general population to have mental health problems or serious psychiatric disorders (Fazel et al., 2016), and children in prison often have reading ages several years lower than their chronological age (Taylor, 2016).

Custody has uncertain impacts on reconviction and rehabilitation in England and Wales. We can see that, for example, there are strong differences in the likelihood of reconviction depending on the type of offence (Howard et al., 2011), but beyond its incapacitative impact, the deterrent effect of prison as *punishment* is unproven (Loeffler & Nagin, 2022; Petrich et al., 2021). Furthermore, the rehabilitative effect of prison, particularly for under-18s, is widely questioned and the policy direction is away from use of custody *apart from* as a last resort (House of Commons Justice Committee, 2018; Taylor, 2016).[5]

[1] Where custody is defined as being in the Secure Estate: Secure Children's Homes (SCH), Secure Training Centres (STC) or Young Offender Institution (YOI) https://www.gov.uk/government/statistics/youth-custody-data (August 2021 release)

[2] Data from https://www.gov.uk/government/statistics/youth-custody-data (August 2021 release). See also: https://wp.unil.ch/europeansourcebook. The reducing use of custody has been a long-standing objective of the Youth Justice Board of England and Wales (Hazel, 2008) and was accompanied by changes 'upstream' in the justice process, specifically the number of first-time entrants to the youth justice system role and fell precipitously during the same period (Sutherland et al., 2017).

[3] See Appendix E for an overview of the costs of imprisonment.

[4] Although that depends on how come offences are categorised. Robbery is inherently violent - it relies on the use or threat of violence - but is sometimes counted as an acquisitive crime rather than a violent one. If one includes robbery as violence then nearly three-quarters (72%) of children in custody in March 2021 were there for violent offences (authors' own calculations based on Table 2.12 of Youth Custody Report June 2021, available from: Youth custody data).

[5] Notwithstanding that prisons' main function is punishment through loss of liberty (Taylor, 2016).

1.2 School exclusion

Throughout this report we use the word exclusion generally to refer to both permanent exclusions and suspensions, except where we indicate otherwise.

A good deal is known about the risks for youth custody and risk factors for youth custody as well. School exclusion is one of the most persistent and well-established risk factors for youth offending (Valdebenito et al., 2018), but, with some exceptions (HM Government, 2019) rarely features in policy relating to crime reduction.

Within the school system in England, sending a pupil away from school either temporarily ("suspensions" or fixed-term exclusions), or permanently removing from a school, are the most severe punishments available in the education system.[6] For each of the last 25 years in England, around three to four hundred thousand suspensions were issued, affecting around 3% of the school-age population, and 5-12,000 pupils were permanently excluded from school each year.[7] So as a 'last resort' (Timpson, 2019), exclusions are resorted to relatively frequently.[8] In spite of this, we know very little about the impacts of exclusion as a sanction on those excluded, classmates of excludees remaining in school, or teachers, although we do know that teachers spend substantial amounts of time, roughly 10-15%, managing behaviour (OECD, 2018, p. 62).

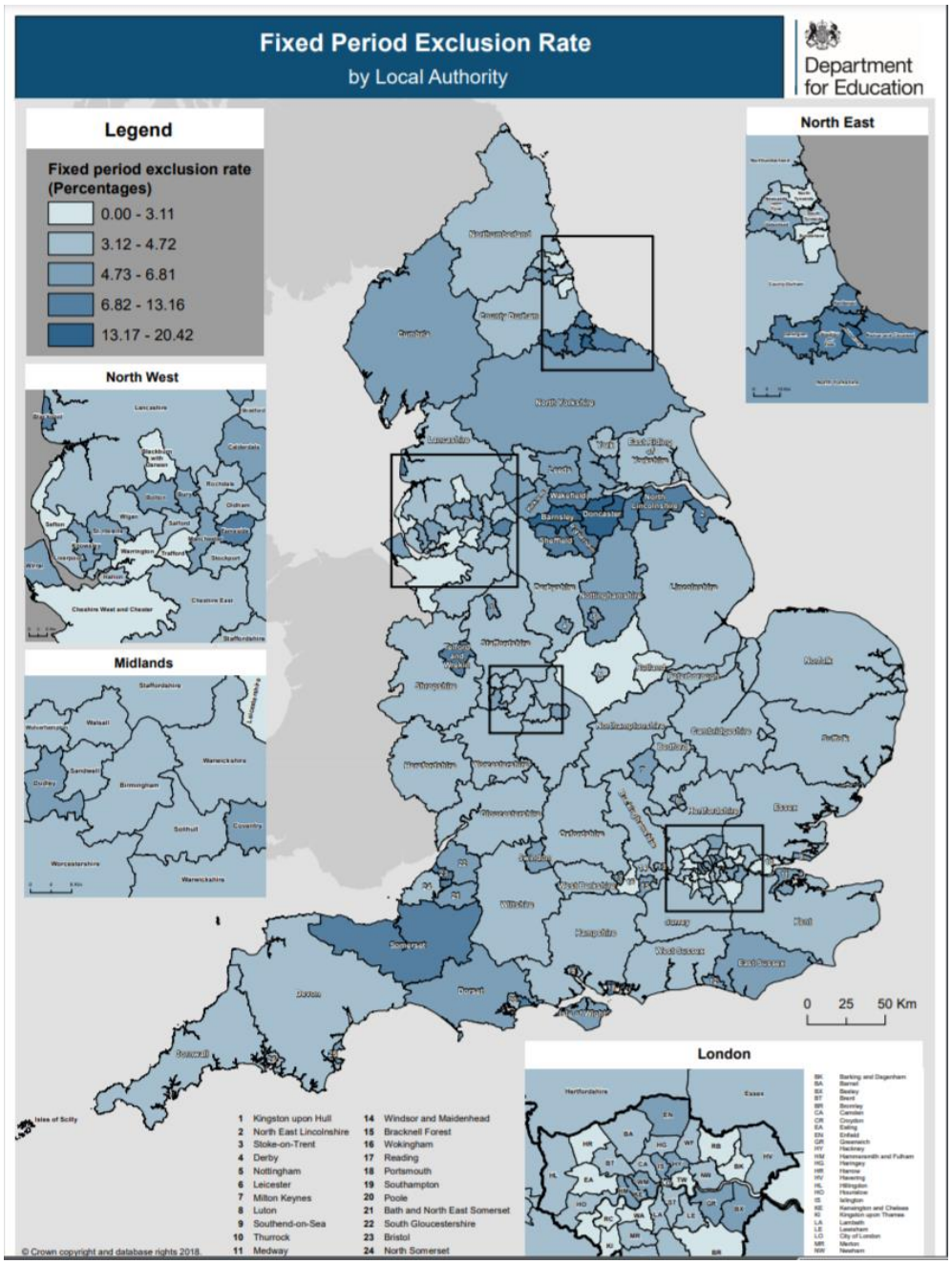
What we do know is that there is a striking consistency in the reasons reported for *why* exclusions happen. Each year, between one-quarter and one-third of suspensions are issued because of what is termed 'persistent disruptive behaviour', nearly one-fifth for assaulting pupils or teachers, and around 18% for verbal abuse of peers or teachers (HM Government, 2021). Overall, this means two-thirds to three-quarters of suspensions are for externalising aggressive behaviour. Again these figures have been similar for some years (Valdebenito et al., 2018). In England, we also know that there is substantial variation in *where* exclusions occur (Figure 1), raising the suggestion that the same behaviour might lead to different outcomes, depending on where a pupil lives.

With regard to *who* is excluded, males and those from some ethnic minorities are consistently more likely to be permanently excluded or suspended (HM Government, 2021). Males are three times more likely than females, and for Black-Caribbean pupils the risk is at least twice as high as for white pupils (HM Government, 2021). The increased likelihood of permanent exclusion for black pupils is still observed when pupil characteristics are accounted for (Timpson, 2019) and some groups, such as Black Caribbean pupils, start secondary school with a higher risk of permanent exclusion than their white peers.[9] This risk

[6] There are also 'managed moves' and off-rolling but these are out of scope for this work. Similarly, 'informal' suspensions, such as sending a child home to 'cool off', which is illegal in England (Department for Education, 2017) are also out of scope as these would not be captured in the data.
[7] With regard to permanent exclusions, the trend over time had been a decrease, certainly from the high point in the late 1990s where 12,000 pupils were permanently excluded annually (see DfE, 2010).
[8] There is often debate about whether it is acceptable to exclude pupils for uniform infractions - published guidance states that schools can exclude pupils for *repeated* uniform infractions (HM Government, n.d.)
[9] Please note that Black Caribbean and Black are not the same groups, hence rates of exclusion can differ between them. Comparing white British and black pupils, rates of exclusion and suspension are very similar (HM Government, 2021; 'Permanent exclusions and suspensions - by characteristic (2006-07 to 2019-20)').

increases term-by-term so by the Autumn term of year 10, Black Caribbean pupils are three times more likely to have been permanently excluded than white pupils (HM Government, 2021, own analysis).

Figure 1: fixed period exclusion (suspension) rate by local authority



Source: Department for Education (2018)

The persistence of gaps over time, particularly by ethnic group, raises questions about where there are disparities in the application of sanctions or genuine differences in behaviour that warrant harsher punishment. This is, understandably, a difficult topic to research but there is some evidence suggesting that teachers may treat pupils differently according to their gender or ethnicity. One novel study from the United States found that following involvement in the same fight, white pupils were treated more leniently than Black pupils (Barrett et al., 2017). Similarly, an experimental US study demonstrated that varying pupil names alone in

vignettes led to differences in teachers' willingness to punish Black pupils (Okonofua & Eberhardt, 2015). Coming at this problem from an intervention perspective, Okonofua et al. (2020) found that through framing avenues for improvements in pupil-teacher relationships it was possible to mitigate the labelling of pupils' behaviour as 'trouble' and teachers' willingness to punish harshly. That study indicates the roles that psychological processes have in disciplinary decision-making, but while promising, that approach has yet to be tested in a field experiment.

1.3 School exclusion, crime, and prison

Around forty percent of serving prisoners had been permanently excluded, and nearly two-thirds (63%) have been suspended (Williams et al., 2012). The relationship between exclusion and later involvement in the justice system is so marked it has been termed 'the school-to-prison pipeline' (Arnez & Condry, 2021; McCarter, 2017; Sanders et al., 2020). This pithy shorthand, although useful rhetorically, lacks precision and does not tally with fact. For example, the 'pipeline' metaphor implies an inevitable and one-way process - but if it were the case that exclusion *deterministically* leads to prison then 100% of excludees would become prisoners. This points to a key gap in knowledge regarding exclusion and the relationship to prison: we do not know whether the observed correlation is *causal*. That is, whether being excluded increases the risk of later prison over and above selection processes or pre-existing behavioural problems (Butler et al., 2021). It might be that pre-existing behavioural problems result in both exclusion and prison, or that exclusion exacerbates existing behaviour problems by providing 'time for crime' (Cohen & Felson, 1979), which then leads to criminal justice involvement. As Sutherland and Eisner (2014, p. 5) put it:

> It is currently unclear whether the disciplinary action itself has a causal effect over and beyond the social, familial and behavioural characteristics of the affected children. To date, studies have used analytical approaches that are unable to reliably establish a robust link between exclusion and outcomes such as criminal behaviour.

What we do know is that exclusion undoubtedly disrupts the education of excludees, leads to labelling of pupils, and that exclusion is a risk factor for juvenile delinquency (Okonofua et al., 2020; Pyle et al., 2020). Exclusion on its own may be neutral in its impact on behaviour, but unsupervised time with peers, a likely *consequence* of exclusion, is a key modifiable risk factor for juvenile delinquency (Wikström et al., 2012) and so may be an intermediary link between exclusion, subsequent offending and later prison. A fundamental problem for making causal claims about the life-course effects of exclusion is that there are no existing randomised trials of exclusion as an intervention, in spite of the high frequency of use and uncertainty about impacts (Rosenbaum, 2020; Sutherland & Eisner, 2014).[10] As such, and for the time-being, we rely on quasi-experimental methods to help identify causal effects. To our knowledge, this is the first such project to attempt to assess whether there is a causal

[10] Put another way, the evidence suggesting a negative impact of exclusion on life outcomes is far greater. What is significantly under-studied is the potential positive societal impact of exclusion for pupils who remain at school without their disruptive peers. Similarly, we do not know what the societal costs (and benefits) of exclusion are, meaning that debates about disciplinary options may be driven more by beliefs, as evidenced by polarised debates regarding 'zero tolerance' and 'restorative practices' (Morton, 2014).

relationship between school exclusion and later custody in the UK using a design able to estimate causal effects.

1.4 Research aims

We aim to provide evidence on the causal effect of school exclusions on being imprisoned as a young person. We focus on the effect of being permanently excluded or suspended in year 10, when pupils are 14-15 years old, on the probability of custody age 15-17 (inclusive). We also aim to look at how the effect of exclusion varies according to the type of exclusion a pupil experiences: permanent exclusion or suspension (also known as a fixed-term exclusion).

We focus on custody in year 10 for two reasons. First, roughly a quarter to a third of all exclusions, both permanent and fixed-term, happen in year 10 (HM Government, 2020; Machin & Sandi, 2020). At 14 years old, puberty peaks, pupils are particularly susceptible to peers' influences (Steinberg & Monahan, 2009) and are more likely to take risks (Blakemore, 2018). Second, as we are unable to observe custody for pupils prior to age 15, focusing on exclusion in year 10 reduces the gap between when we observe a pupil receive an exclusion and when they may start a custodial sentence.

To account for unobserved differences in the characteristics of pupils that receive exclusions, we use an instrumental variable approach. We exploit changes in a pupil's probability of being excluded in year 10 as a result of a change in their school's academy status. This allows us - conditional on some assumptions - to make causal inferences about the impact of exclusion on custody.

We had planned to look at how the probability of custody was impacted by both the timing of exclusions and whether permanently excluded pupils were placed in Pupil Referral Units or mainstream schools. However, given the relative rarity of both exclusion and custody events, the data is fairly sparse, and we were not able to identify the impact on the probability of custody with sufficient precision. Given this, a discussion of directions for future research is included in Section 4. Implications.

2. Analysis strategy

2.1 Data

Our analysis uses a unique set of data made available to UK researchers in 2021. Thanks to the investment in administrative data linkages by the Office for National Statistics and Administrative Data Research England, we have been able to observe pupils' trajectories throughout their primary and secondary education (including any exclusion episodes) and their experiences in the juvenile justice system.

2.1.1 The National Pupil Database (NPD)

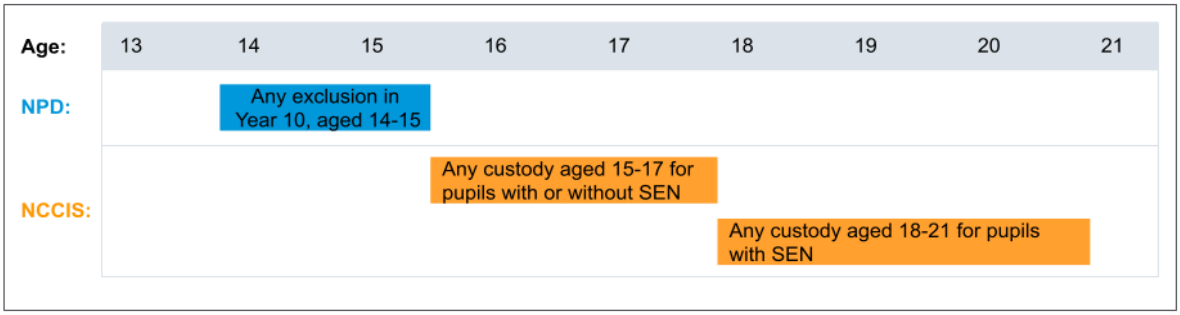
For this project, we focus on pupils born between 1-Sep-1994 and 31-Aug-1998 attending state-funded schools in England. For these pupils, we use data from the Department for Education's NPD from when they are in year 10. This allows us to know what school they attended, their gender, whether they had special education needs (SEN) and whether they were eligible for free school meals in year 10. Even more importantly, the NPD data we access gives us information about whether they received exclusions and of what type (temporary suspensions or permanent exclusions).

2.1.2 The National Client Caseload Information System Database (NCCIS)

Being the register dataset of all pupils in state schools up until the end of their secondary education, we cannot rely on the NPD alone to assess the impact of exclusions on custody. This is why we then follow pupils into the NCCIS, also provided by the Department for Education. The NCCIS includes the universe of custodial sentences for pupils in the year prior to the compulsory school leaving age and for two years afterwards (DfE, 2014).[11] Therefore, we are able to observe custodial sentences for pupils from age 15 up to their 18th birthday (Figure 2). Local authorities are required to report on activities of pupils with *current* special education needs (SEN) until they are 21, so we have additional years of custody data for some pupils; however, we restrict our analysis to custody before 18 for the sake of consistency. For this reason, we restrict our analysis to custody before 18. We observe some pupils in the NCCIS that begin custody sentences during or before the academic year they are registered in year 10 (21 pupils); we remove them from the sample to reduce potential for reverse causality between exclusions and custody.

[11] The compulsory school leaving age in England is 16; pupils can leave school in June when they are 16 or if they will turn 16 before the end of the school holidays. The Education and Skills Act 2008 introduced changes to educational requirements post-16. From 2013, pupils were required to stay in education or vocational training until 17, and in 2015, this was further raised to 18.

Figure 2: NPD and NCCIS data used in analysis with approximate ages of pupils



2.1.3 The Consistent Schools Database

We also rely on the Consistent Schools Database (CSD) to identify characteristics of schools and how they change over time.[12] Crucially, this dataset allows us to identify the academic year in which a school becomes an academy, though not the exact date of academisation. Therefore, we make the assumption that all pupils attending a school that academises in the academic year when they are in year 10 were affected by academisation from the start of the academic year.

2.1.4 Limitations of combining the different data sources

The combination of these datasets allows us to connect pupils' educational history, including where they go to school and if they are excluded, with their experiences of custody as a young person. After merging the datasets, we have a complete overview of year 10 exclusions and subsequent custodies that start aged 15-17 years old for **2,269,489** pupils across four cohorts, who were in year 10 across six academic years.[13]

However, not all pupils could be linked across all datasets. A detailed explanation of the data merging tasks and the results are included in Appendix A, but this included:[14]

- 4,740 young people from our cohorts who are in custody between 2010/11 to 2017/18 but are not recorded in year 10 between 2009/10 to 2013/14.
- 7,292 pupils who are recorded as having exclusions at schools in which they were not registered.
- 145,059 pupils who are registered in schools we did not have records of in the Consistent Schools Database.
- 33,629 pupils who are observed in year 10 but not in year 9, and for whom we cannot tell if they switch schools.
- 3,344 pupils who are observed in year 9 and year 10 in the same academic year.

[12] This project used the Consistent Schools Database, produced by Ellen Greaves and Bilal Nasim, funded by CLOSER, the home of longitudinal research (grant reference: ES/K000357/1). www.closer.ac.uk

[13] We have data on four cohorts of pupils; however, a small number of students may be behind their cohorts and therefore attending year 10 in a later academic year than the rest of their cohort. We therefore have data from 6 academic years in which we have year 10 students from these four cohorts.

[14] If for some reason these pupils are more likely to experience custody as a direct result of exclusion, our results will be biased downwards; however, based on our understanding of the data matching process, we have no reason to believe so.

2.2 Identifying the causal effect of exclusions on custody

Nearly 50% of the young people we observe in custody aged 15-17 received at least one exclusion when they were in year 10 (Table 1). Pupils receiving an exclusion of any type in year 10 are 1% more likely to also experience custody before age 18, compared with pupils that received no exclusions.[15] This does not directly imply that exclusions lead to custody, as other factors may play a role. For example, it may be that children with family problems are more likely to be excluded from school (DfE, 2019) *and also* to engage in behaviours that lead to custody (Smith, 2017; Local Government Association, 2018).

Table 1: rate of exclusions for pupils that do / do not have custody experiences starting age 15-17

Sample: all pupils				
Pupil receives any exclusion in year 10	***No***		***Yes***	
Pupil experiences custody starting age 15-17	***Frequency***	***Percentage***	***Frequency***	***Percentage***
No	*2,129,846*	*93.98%*	*136,403*	*6.02%*
Yes	*1,737*	*53.61%*	*1,503*	*46.39%*

In this paper, we aim to isolate the causal impact of exclusions on custody. This would help us answer questions like *"if schools relaxed their exclusion policies and excluded half the pupils they are currently excluding, how many pupils could we prevent from ending up in custody?".*

To do so, we rely on a strategy called instrumental variable estimation (Angrist & Krueger, 2001). In a nutshell, we looked for a factor that could influence the chance that a pupil is excluded but does not have an effect on the pupil's likelihood to experience custody later in life - for example, a change in schools' exclusion policies. Unfortunately, a database of schools' policies that could directly allow us to do so does not exist (yet); however, previous research has suggested that after converting to academy status, schools adopt strict disciplinary policies, resulting in a significant increase in the number of pupils excluded post-academisation (Machin & Sandi, 2020).

[15] This is the result from an analysis that accounts for pupils' gender, free school meal status, academic year, and which school they attended in year 10. Results are available in Appendix B, Table 7.

2.2.1 Academisation as an instrument

Until the early 2000s, most state-funded schools in England were run by local authorities; however, starting in the 1990s, there was wide acknowledgement that some of these schools were failing to provide a high quality education to their pupils and suffered from endemic discipline issues (Machin & Sandi, 2020). Under the 1997-2010 Labour government, a policy was adopted to enable state-funded schools to run autonomously from the local authority, leading to the creation of "academy schools", similar to charter schools in the US.

This programme was expanded by the Conservative-led coalition government in 2010 to encompass far more schools, including primary schools and better performing schools, than during the first wave (Eyles, Machin, & McNally, 2017). As of 2021, 79% of all secondary schools and 38% of all primary schools are academies (ONS, 2021), over 85% of which are run as part of a chain of schools, called multi-academy trusts. Like the local authority-run schools that preceded them, academies are state-funded, non-fee charging and non-selective.

Initially the purpose of academisation was to improve the quality of education, particularly in underperforming schools that served primarily disadvantaged pupils, by enabling a board of governors for the school to have far more autonomy over staffing, pay, curriculum and school policies. Critically, autonomy over school policies meant that academy schools could set-up and enforce their own disciplinary procedures. While focusing on whether pupil performance gains in autonomous schools in England could be attributed to the strategic exclusion of poorly performing pupils, Machin and Sandi (2020) find that rather than a means of test score manipulation, the higher exclusion rate in academies reflects more rigorous discipline being enforced. For this reason, we use academisation of local authority schools as a factor which creates a "shock" in exclusion rates for pupils that attend these schools.

A crucial assumption of our analysis strategy is that academisation affects the likelihood of custody *only through its impact on exclusion*. That is: being in a school that academises does not increase or decrease the probability of experiencing custody later in life, *except through* the higher probability of being excluded. We cannot know or test whether this is occurring; however, to address this concern, in our main robustness check we look only at exclusions in the academic year that a school converts to an academy, as we hypothesise that new discipline policies may be implemented and enforced quickly in newly formed academies, compared with other changes that may take longer to be implemented and would also impact probability of custody. This is consistent with anecdotal evidence from some converted academies, in which the newly appointed principal introduced strict discipline policies.[16]

To explore further the robustness of these results we also originally investigated using an alternative instrument: the number of exclusions made by a pupil's headteacher in a previous year. This is because we hypothesised that some headteachers may be more likely to exclude pupils than other headteachers - critically, in a way that does not depend on pupils' characteristics. However, this instrument was not found to have an impact on exclusion

[16] One example comes from schools under the Inspiration Trust academy sponsors' umbrella. In 2017, the BBC covered the sudden change in discipline policy in newly converted Inspiration Trust academies in a series of articles (examples available here and here).

rates, so we have not explored this analysis further (see Appendix D for a discussion and findings).

2.2.2 Sample

Figure 3 provides a visual representation of our sample selection.

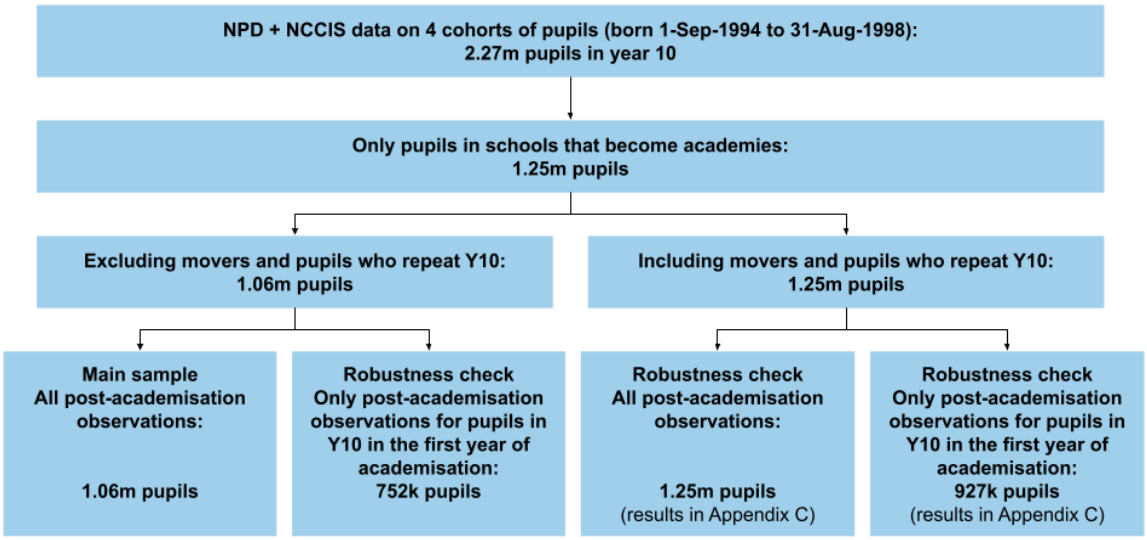
We start from an overall sample of 2,269,489 pupils in year 10 across 4 cohorts (roughly 567k pupil observations in each year). As the process of academisation is voluntary, schools that academise are unlikely to be comparable to those that do not. For this reason, we restrict our analysis to only pupils that attend schools that become academies by academic year 2013/2014 (roughly 44% of the schools attended by pupils in our dataset), enabling us to compare pupils attending similar schools (Eyles, Machin & McNally, 2017). Of the 2,269,489 pupils we have data on, a total of 1,248,037 attended schools that became academies in the years that we study. We focus on pupils that we observed in the same school in both year 9 and year 10 and didn't switch schools while in year 10, in order to reduce the issue of pupils "selecting in" to attending an academy school (Abdulkadiroğlu et al., 2014; Fryer, 2014; Eyles & Machin 2019; Machin & Sandi, 2020). Furthermore, we only include pupils that did not repeat year 10 for the same reason.[17]

As we explained in the previous section, in the robustness analysis we restrict our post-academisation sample to pupils in year 10 in the academic year in which academy conversion takes place only (926,502 pupils overall, 752,675 pupils if excluding those that repeat year 10 or move schools). This means that for this analysis we exclude all pupils who happen to be in year 10 in the academic year *following* conversion, or later. This does reduce the sample size; by applying this restriction we can only include 752,675 pupils instead of 1,064,555 in the analysis. These results are always reported in the last three columns of the results tables.

The exact definitions of our sample are outlined in Table 5 in Appendix C.1.

[17] For completeness, we also provide results for the whole sample of pupils (including those who moved schools or who repeated year 10) in Table 9 and Table 10 in Appendix C.

Figure 3: visual representation of the sample selection



3. Findings

Key Takeaways

- **Attending an academy school in year 10 increases the probability of receiving a suspension or permanent exclusion by 2.6 percentage points (statistically significant at 0.1%)**
  - Probability of receiving a suspension increase by 2.6 percentage points (statistically significant at 0.1%). This is against an average exclusion rate of 5.3% among pupils that attend schools that have not yet academised.
  - Probability of receiving a permanent exclusion increases by 0.1 percentage points (statistically significant at 0.1%). This is against an average of 0.2% among pupils that attend schools that have not yet academised.
- **Attending an academy school in year 10 results in an increase in the probability of custody ages 15-17 (statistically significant at 5%)**, suggesting that school exclusions at ages 14-15 *cause* a detectable change in the probability of a pupil entering custody at ages 15-17.
  - Receiving a permanent exclusion increases the probability of custody by 33 percentage points (statistically significant at 5%)
  - For suspensions, the increase is 1.3 percentage points (statistically significant at 5%)
  - This compares to the average custody rate of 0.1% for all pupils in the sample; however, the estimated impacts outlined above relates to compliers – those who are excluded as a result of academisation – whose expected custody rate in the absence of academisation is not observed but is likely to be higher 0.1%.

3.1 Impact of academisation on exclusions

We provide results of the impact of a school's academisation on the probability of being excluded while in year 10 in Table 2.

We find that being a year 10 pupil in a school after it converted to an academy increases the probability of receiving a suspension or permanent exclusion by 2.6 percentage points on average, which is statistically significant at 0.1% ((b = 0.026; t = 12.71; Table 2, column (1)). This compares to a raw exclusion rate of 5.6% of all pupils in our sample, or 5.3% among only pupils in our sample that attend schools that have not yet converted to academies (Table 3, col (1)). 2.6 percentage points is therefore a sizable increase; it means moving from 3 to 4 pupils excluded in every two classes of pupils (52 pupils total).[18] Academisation

[18] To find the number of pupils that receive exclusions in absence of academisation, we multiply the average exclusion rate among pupils attending schools that have not yet converted to academies in our sample (0.053) by the approximate number of students in two classes (26*2=52). This is 0.053*52=2.8 pupils. This compares to the number of pupils in two classes that would receive exclusions given the change in exclusion rate as a result of academisation (the "treatment effect") – we multiply the exclusion rate post-academisation (0.053 + 0.026 = 0.079) by the approximate number of number of pupils in two classes (26*2=52). This is 0.079*52=4.1 pupils.

increases both the probability of permanent exclusion and suspension, but the effect seems to be mainly concentrated on suspensions (2.6 percentage points increase versus a 0.1 percentage points increase for permanent exclusions, Table 2, columns (3) and (2)). This may be because permanent exclusions are much rarer than suspensions; among all pupils in our main sample, only 0.2% receive permanent exclusions (Table 3, col (2)) compared to 5.6% receiving fixed term exclusions (Table 3, col(1)). There are approximately 28 times more pupils with any suspensions than pupils with any permanent exclusions in each academic year in our main sample.

When looking at the effect of academisation in the year of academisation exclusively (that is, restricting the analysis to the year of academy conversion), we find a similar picture. Academisation increases overall exclusion rates by 2.9 percentage points overall (b = 0.029; t = 12.61; Table 2, column (4)), with the probability of experiencing a permanent exclusion increasing by 0.1 percentage points (Table 2, column (5)) and the probability of experiencing a suspension increasing by 2.8 percentage points (Table 2, column (6)) for a pupil in year 10. Again, given that only 6% of pupils in this sample receive suspensions and 0.2% of pupils receive permanent exclusions, these are sizable increases.

When repeating the analysis including pupils who have switched schools during year 10 or between year 9 and year 10 and pupils who repeated year 10, results are remarkably similar (see Appendix C.3). This indicates that there may be little movement of pupils between schools in anticipation of a school's academisation; that is, there is little risk of selection bias.

These results, together with the high F-statistics reported in the table, reassure us that academisation is a strong predictor (strong instrument) for the probability of receiving exclusions, regardless of the type of exclusion and whether this is in the longer or shorter term.

Table 2: effect of attending an academy on probability of receiving exclusion *(first stage, 2SLS)*

Analysis:	***Main analysis:* Attendance at academy in year of or after conversion, excluding students that switch schools or repeat y10**			***Robustness check*: Attendance at academy in year of conversion, excluding students that switch schools or repeat y10**		
Dependent variable: receiving exclusion or suspension in year 10 of following types:	(1) Any	(2) Permanent exclusion	(3) Suspension	(4) Any	(5) Permanent exclusion	(6) Suspension
Explanatory variable: Attendance at an academy (b/t)	0.026** (12.71)	0.001** (4.30)	0.026** (12.60)	0.029** (12.61)	0.001** (3.54)	0.028** (12.52)
F	161.4	18.5	158.7	159.0	12.5	156.7
N	1,064,555	1,064,555	1,064,555	752,675	752,675	752,675
+ p<0.05, * p<0.01, ** p<0.001. Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school						

fixed effects.
1 singleton observation dropped from regressions in columns (4)-(6).

Table 3: average exclusion rates by sample and academy status

	***Main analysis:* Attendance at academy in year of or after conversion, excluding students that switch schools or repeat y10**			***Robustness check*: Attendance at academy in year of conversion, excluding students that switch schools or repeat y10**		
Type of exclusion:	(1) Any	(2) Permanent exclusion	(3) Suspension	(4) Any	(5) Permanent exclusion	(6) Suspension
Exclusion rate among all pupils in sample	0.057 (N=1,064,555)	0.002 (N=1,064,555)	0.056 (N=1,064,555)	0.054 (N=752,676)	0.002 (N=752,676)	0.053 (N=752,676)
Exclusion rate among only pupils that attend schools that have not yet converted to academies	0.053 (N=561,249)	0.002 (N=561,249)	0.053 (N=561,249)	0.053 (N=561,249)	0.002 (N=561,249)	0.053 (N=561,249)

3.2 Impact of exclusions on custody

When turning to the impact on custody, we find that ultimately the higher probability of receiving an exclusion in the academic year of academisation (because of the academisation) affects the probability of custody at ages 15-17. Receiving any exclusion in year 10 while attending a school that has academised causes a 1.3 percentage point increase in the probability of custody, statistically significant at 5% (b = 0.013 *;* t = 2.20; Table 4, column (1)); receiving a suspension has a similar effect size and is also statistically significant at 5% (b = 0.013 *;* t = 2.20*;* Table 4, column (3)). The overall effect masks an even larger impact of permanent exclusion on custody: a permanent exclusion in year 10 while attending a school that has academised causes a 33% increase in probability of custody, statistically significant at 5% (b = 0.328*;* t = 2.04; Table 4, column (2)). Against the custody rate of 0.1% among pupils as a whole (Table 5, col (1)), this constitutes a sizeable increase; however, the estimated impact relates to compliers – those who are excluded as a result of academisation – whose expected custody rate in the absence of academisation is not observed but is likely to be higher than that seen among pupils as a whole.

When looking at the impact of exclusions on custody including cohorts attending academy schools in the year of conversion only, we find results that are consistent with the main analysis in that they are directionally positive but, these are non-significant (Table 4, columns (4), (5), (6)). Specifically, experiencing any exclusion in year 10 (age 14-15) while attending a school that academises affects the probability of custody when 15-17 by 1.1 percentage points ((b = 0.011; t = 1.85; Table 4, column (4)); receiving a suspension has a similar effect, and also is not statistically significant at conventional levels ((b = 0.011; t = 1.84; Table 4,

column (6)). Last, experiencing a permanent exclusion increases the probability of custody by 31 percentage points ((*b* = 0.313; *t* = 1.73; Table 4, column (5)), not statistically significant at the 5% level.

Both analyses, with pupils in only schools converting in that academic year and with pupils attending longer established academies, give similar estimates of the effect size. The main analysis using multiple years *is* significant thanks to its larger sample size, but it is possible that the effect is somewhat overstated if academies introduced other measures in subsequent years that impact on custody through means other than exclusion.

When repeating the analysis including pupils who have switched schools during year 10 or between year 9 and year 10 and pupils who repeated year 10, point estimates are directionally positive, but smaller and no longer statistically significant at 5% (see Appendix C.3).

Finally, Table 6 includes the results from the reduced form regression, looking at the impact of the instrument, academisation, on the ultimate outcome measure, custody age 15-17. For the main analysis, this shows that academisation is correlated with 0.03 percentage point increase in the probability of custody, statistically significant at 5% (Table 6, column (1)). The result is reassuring, given that the reduced form provides an estimate that is unaffected by bias and still shows a positive, and statistically significant result. We note that the point estimates for the robustness check (in Table 6, column (2) and in Appendix C.3) are all similar, though no longer statistically significant. As we discuss in the limitations section below, because the events are rare, the lack of statistical significance for the other samples could be the result of smaller samples sizes or greater variation in the outcome data due to less strict sample criteria.

Table 4: effect of receiving exclusion on probability of custody age 15-17 (*second stage, 2SLS*)

Analysis:	***Main analysis:* Attendance at academy in year of or after conversion, excluding students that switch schools or repeat y10**			***Robustness check*: Attendance at academy in year of conversion, excluding students that switch schools or repeat y10**		
Dependent variable:	Experience of custody, age 15-17					
Explanatory variable: predicted exclusion or suspension of following types: (b/t)	(1) Any	(2) Permanent exclusion	(3) Suspension	(4) Any	(5) Permanent exclusion	(6) Suspension
	0.013+ (2.20)	0.328+ (2.04)	0.013+ (2.20)	0.011 (1.85)	0.313 (1.73)	0.011 (1.84)
N	1,064,555	1,064,555	1,064,555	752,675	752,675	752,675
+ p<0.05, * p<0.01, ** p<0.001. Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects. 1 singleton observation dropped from regressions in columns (4)-(6).						

Table 5: average custody rate by sample & academy status

Analysis:	*Main analysis:* Attendance at academy in year of or after conversion, excluding students that switch schools or repeat y10	*Robustness check*: Attendance at academy in year of conversion, excluding students that switch schools or repeat y10
Custody rate among all pupils in sample	0.001 (N=1,064,555)	0.001 (N=752,676)
Custody rate among only pupils that attend schools that have not yet converted to academies	0.001 (N=561,249)	0.001 (N=561,249)

Table 6: effect of attending an academy on probability of custody age 15-17 (reduced form)

Analysis:	Main analysis: Attendance at academy in year of or after conversion, excluding students that switch schools or repeat y10	Robustness check: Attendance at academy in year of conversion, excluding students that switch schools or repeat y10
Dependent variable:	Experience of custody, age 15-17	
Explanatory variable: Attendance at an academy (b/t)	0.0003+ (2.21)	0.0003 (1.86)
N	1,064,555	752,675
+ p<0.05, * p<0.01, ** p<0.001. Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects. 1 singleton observation dropped from first year academisation regressions (column (2)).		

4. Implications

4.1 Implications for policy

In this analysis, we provide additional evidence that **academisation increased the number of pupils excluded from school, both permanently and for temporary suspensions**. We estimate that for every 1,000 year 10 pupils attending schools that converted to academies in a given academic year, approximately 26 more pupils received permanent exclusions or suspensions than if their school had not academised.[19] This works out to approximately **13,100 additional pupils receiving exclusions** among the 503,306 pupils in our sample who attended year 10 in the years after the school academised (out of total N =1,064,555).[20] This is an important finding given that it has been a government priority to see "every school in the country" join multi-academy trusts (DfE, 2021).

Turning to our focal causal analysis, using pupil attendance in the years after a school academises as an instrument for exclusions, we find evidence that **receiving an exclusion of any type leads to a higher probability of starting a custodial sentence between the ages of 15-17**. Permanent exclusion in particular has a large effect on the later probability of custody, raising this by approximately 33 percentage points. Although there is some uncertainty around this estimate given the marginal statistical significance, the uncertainty is also a function of the low base-rate for permanent exclusions; only 2,128 pupils of the 1,064,555 (0.2%) in our sample experienced a permanent exclusion. (A low base-rate means that it is much more difficult to find statistical differences between groups.) Suspensions were found to have a smaller impact, on average increasing the probability of custody by 1 percentage point. However, given that only 1,065 of the 1,064,555 pupils (0.1%) in our sample start custody age 15-17 at all (Table 5), a 1 percentage point increase is still substantial. Furthermore, more than 60,000 year 10 pupils in our main sample (approximately 6% of all pupils) received fixed-term exclusions, meaning even a small increase in the probability they end up in custody is substantial when you take into account the size of the population it affects.

We caveat that additional analysis (robustness checks) were not statistically significant at conventional levels. However, the coefficients on exclusion were positive in all the models we ran; this provides further reassurance on the likely direction of the effect.

We also recognise that using attendance at an academy school in any year as an instrument for exclusions may be problematic as i) in time, schools may implement changes beyond their discipline policies that could affect the chances of custody (not via exclusions) ii) there

[19] We calculate this by multiplying the increase in probability of receiving any exclusion or suspension as a result of attending an academy, 2.6% (see Table 2, column 1) by 1,000 students. 1,000*0.026=26

[20] We calculate this by multiplying the number of Year 10 students in our sample that attended academies (503,306 of 1,064,555) by the increase in probability of receiving any exclusion or suspension as a result of attending an academy, 2.6% (see Table 2, column 1). 503,306*0.026=13,086

are likely more significant selection problems (pupils moving schools to go to academies).[21] However, using this instrument allows us to use a sample size that includes 41% more pupils, substantially increasing statistical power, compared to the analysis reported as robustness check.

Our interpretation of our results, even with the caveats given above, is that exclusion presents a small but non-ignorable increase in the risk of criminality and thus custody. We think it is a reasonable assertion if one considers the very small numbers of pupils experiencing our outcome, custody (<1% of the sample) and the consistency of results in terms of the direction (i.e., all showing a positive relationship across different model specifications).

Furthermore, custody is an end-point of often lengthy criminal justice processes; it is reserved for more serious offences, and its use for under-18s has been steadily decreasing in recent years (Bowyer, Dorsett & Thomson, 2021). There is also a substantial amount of attrition between crimes committed and final criminal justice outcomes, *particularly* custody. This means we are looking at the end of the 'funnel' of criminal justice, with tiny numbers of young people actually going to prison, making it very difficult to find statistical relationships. That we find anything at all is surprising; that we find changes as great as a 30 percentage point increase in the risk of custody following permanent exclusion suggests this requires greater attention from both policy and research.

Put another way, if this were another 'intervention' and it looked like there was a *beneficial* relationship that was as great as a 30pp improvement in an outcome, we think it reasonable to assume that there would be significant interest in demonstrating effectiveness further. Given the consequences here, we also think it reasonable that one should be much more sensitive to the possibility of harmful outcomes than positive ones.

4.2 Limitation and directions for future research

This research used a novel dataset comprised of both NPD and NCCIS data which had not been available to researchers prior to this year. As a result of linkages in administrative datasets, we were able to look at the trajectories of pupils from education through to young adulthood. Despite this, there are some limitations of our research strategy that can be addressed and improved upon in future research.

4.2.1 Using data for more cohorts of pupils

Custody and school exclusions are fairly rare events which make them difficult to study. On average, only 6% of year 10 pupils in our main sample receive any type of exclusion each year (Table 3) and only 0.1% start a custody sentence between 15-17 (Table 5). In this project, we had access to data on four cohorts of pupils. Looking at a greater number of pupil cohorts may allow for more consistent estimates of the impact of exclusion on custody as well as the ability to break down how the impact varies based on timing of exclusion, reason

[21] We observe only whether students switch schools in year 10, or between year 9 and 10. If academisation happens earlier than when a student is in year 10, we don't observe whether they switched schools in an earlier year to attend the academy school.

for exclusion or other factors. It would also allow a more detailed look into what happens to pupils that are permanently excluded, for example whether they are placed in Pupil Referral Units or mainstream schools and what this means for their trajectories later in life.

4.2.2 Improving matching across datasets

We lose more than 190,000 observations from the original dataset due to merging issues: that is, pupils and schools whose data could not be matched across the many datasets used in the analysis. Losing these observations harms the statistical power of our analyses, but more detrimentally introduces room for selection in the sample as we are ultimately forced to exclude these observations for our analysis. Additional work could be done to improve matching between datasets.

4.2.3 Linking to additional data sources on criminality and custody

We have limited data available on young people's experiences of custody. NCCIS data covers most young people only when they are aged 15 to 17, meaning that any custody experiences before age 15 or at age 18 or later are not captured in our analysis. Further, we do not observe any experience with the justice system outside of young adulthood; much more could be learned if further data linkages were made which allowed researchers to follow pupils later in life, for example by linking with Ministry of Justice data.

As discussed above, custody is one end of a long funnel and in the data we have access to, we are unable to observe whether pupils have other contacts with the justice system that do not end in custody. Additionally, our data on custody does not allow us to observe when pupils commit an offence that leads to custody, opening the potential that the custody sentences we observe were from offences that actually occurred prior to exclusion or academisation.

Moving further upstream from custody to focus on criminality would allow researchers to assess the relationship between exclusion and crime more generally and with greater accuracy. This could be done by using the same methodology as in here with linked NPD and Police National Computer (PNC) data available in the ONS SRS. Reliably and consistently assessing the relationship between exclusion and crime would provide a more solid basis for policy and practice.

4.2.4 Exploring alternative instruments

Much more could also be learned from using different instruments. One instrument we explored was the prior exclusion rate of a pupil's headteacher, though this was not found to have a statistically significant impact on exclusion rates (see Appendix D for a full discussion of this analysis and findings). This analysis was ultimately constrained by the number of cohorts we had data on and the quality of the data on headteachers. With better data on headteachers and their movement across schools, as well as more cohorts of pupils on which to develop baseline exclusion rates, there is much that could potentially be learned about the link between exclusion and custody.

Authors

Claire Cathro – Senior Research Advisor

Claire is a Senior Advisor in the Consumer and Business team based in BIT's London office. Prior to joining BIT, Claire worked as a consultant with Deloitte, supporting Canadian public sector organisations on transformation programmes and technology implementations.
She holds a Master of Public Administration from the London School of Economics and a Bachelor of Arts in Global Development Studies from Queen's University, Canada.

Dr Giulia Tagliaferri – Head of Quantitative Research

Giulia is Head of Quantitative Research at BIT. She leads BIT's analytical work in complex quantitative evaluations (observational data, quasi-experimental methods, thorny RCTs).
Prior to joining BIT, Giulia worked for the Centre for Vocational Education Research (CVER) at the London School of Economics, conducting research in the field of education.
She completed her PhD in Economics at Queen Mary University of London in 2018, where her research focused on the economics of identity. She holds an MSc and a BA in Economics and Social Sciences from Universita' Bocconi in Milan.

Dr Alex Sutherland – Former Chief Scientist and Director of Research and Evaluation

Dr Alex Sutherland worked at BIT as Chief Scientist and Director of Research and Evaluation from 2019 to 2022. His recent published work has been on data linkage in government, police body-worn cameras and he has led a
number of large-scale randomised-controlled trials in education.

References

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., & Pathak, P. A. (2016). Charters without lotteries: Testing takeovers in New Orleans and Boston. *American Economic Review*, *106*(7), 1878-1920.

Aebi, M.F, Caneppele, S., Harrendorf, S., Hashimoto, Y. Z., Jehle, J.-M., Khan, T.S., Kühn, O., Lewis, C., Molnar, L., Smit, P., Þórisdóttir, R. (2021). European Sourcebook of Crime and Criminal Justice Statistics 2021 (6th ed.). Series UNILCRIM, (1)2021. Retrieved from: https://wp.unil.ch/europeansourcebook/printed-editions-2/

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives*, *15*(4), 69-85.

Arnez, J., & Condry, R. (2021). Criminological perspectives on school exclusion and youth offending. *Emotional and Behavioural Difficulties*, *26*(1), 87-100. https://doi.org/10.1080/13632752.2021.1905233

Barrett, N., McEachin, A., Mills, J. N., & Valant, J. (2017). What are the sources of school discipline disparities by student race and family income. *Education Research Alliance for New Orleans.* Retrieved from: https://educationresearchalliancenola.org/files/publications/111417-Barrett-McEachin-Mills-Valant-What-Are-the-Sources-of-School-Discipline-Disparities-by-Student-Race-and-Family-Income.pdf

BBC News. (2017). Great Yarmouth Charter Academy: Head defends strict rules. Retrieved from: https://www.bbc.co.uk/news/uk-england-norfolk-41265485

Blakemore, S. J. (2018). *Inventing ourselves: The secret life of the teenage brain*. Penguin.

Bowyer, A., Dorsett, R., & Thomson, D. (2021). The risk of experiencing custody aged 16-18. FFT *Education Data Lab.* Retrieved from: https://ffteducationdatalab.org.uk/2021/03/the-risk-of-experiencing-custody-aged-16-18/

Butler, H. D., Caudill, J. W., Craig, J. M., DeLisi, M., & Trulson, C. R. (2021). 99 Percenters: An Examination of the Misconduct Careers of the Most Violent and Disruptive Incarcerated Delinquents. *Aggression and Violent Behavior*, 101520. https://doi.org/10.1016/j.avb.2020.101520

Closer. (2018). New resource to improve research into school performance. Available here: https://www.closer.ac.uk/news-opinion/news/resource-improve-research-school-performance/

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American sociological review*, *44*(4), 588-608. https://doi.org/10.2307/2094589

Department for Education. (2014). *NCCIS management information requirement.* Retrieved from: https://www.gov.uk/government/publications/nccis-management-information-

requirement

Department for Education. (2017). *Exclusion from maintained schools, academies and pupil referral units in England*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/921405/20170831_Exclusion_Stat_guidance_Web_version.pdf

Department for Education. (2018). *Fixed Period Exclusion Map*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/733474/Fixed_Period_Exclusion.pdf

Department for Education (2019). *School exclusion: a literature review on the continued disproportionate exclusion of certain children*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/800028/Timpson_review_of_school_exclusion_literature_review.pdf

Department for Education. (2021). What is an academy and what are the benefits? *The Education Hub.* Retrieved from: https://educationhub.blog.gov.uk/2021/10/14/what-is-an-academy-and-what-are-the-benefits/

Eyles, A., & Machin, S. (2019). The Introduction of Academy Schools to England's Education. *Journal of the European Economic Association, 17(4),* 1107–1146. https://doi.org/10.1093/jeea/jvy021

Eyles, A., Machin, S., & McNally, S. (2017). Unexpected school reform: Academisation of primary schools in England. *Journal of Public Economics*, *155*, 108-121. https://doi.org/10.1016/j.jpubeco.2017.09.004

Fazel, S., Hayes, A. J., Bartellas, K., Clerici, M., & Trestman, R. (2016). Mental health of prisoners: prevalence, adverse outcomes, and interventions. *The Lancet Psychiatry*, *3*(9), 871-881. https://doi.org/10.1016/S2215-0366(16)30142-0

Fryer Jr, R. G. (2014). Injecting charter school best practices into traditional public schools: Evidence from field experiments. *The Quarterly Journal of Economics*, *129*(3), 1355-1407. https://doi.org/10.1093/qje/qju011

Hazel, N. (2008). Cross-national comparison of youth justice. London: Youth Justice Board. https://dera.ioe.ac.uk/7996/1/Cross_national_final.pdf

Heeks, M., Reed, S., Tafsiri, M., Great Britain, & Home Office. (2018). *Economic and social costs of crime.* Retrieved from: https://www.webarchive.org.uk/access/resolve/20180730161657/https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/727958/the-economic-and-social-costs-of-crime-horr99.pdf

HM Prison and Probation Service & Youth Custody Service. (2021). Youth custody data. Retrieved from: https://www.gov.uk/government/statistics/youth-custody-data

HM Government. (n.d.) *School uniform.* Retrieved from: https://www.gov.uk/school-uniform

HM Government. (2019). *A whole-system multi-agency approach to serious violence prevention.* Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/862794/multi-agency_approach_to_serious_violence_prevention.pdf

HM Government. (2020). *Data catalogue*. Retrieved from: https://explore-education-statistics.service.gov.uk/data-catalogue/permanent-and-fixed-period-exclusions-in-england

HM Government. (2021). *Permanent exclusions and suspensions in England, Academic Year 2019/20.* Retrieved from: https://explore-education-statistics.service.gov.uk/find-statistics/permanent-and-fixed-period-exclusions-in-england/2019-20#dataBlock-630edb4f-e9de-4276-be0d-71aed81b1300-tables

House of Commons Justice Committee. (2018). *Young adults in the criminal justice system.* Retrieved from: https://publications.parliament.uk/pa/cm201719/cmselect/cmjust/419/419.pdf

Howard, P. (2011). *Hazards of different types of reoffending.* Ministry of Justice. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/217377/research-reoffending-hazards.pdf

Local Government Association. (2018). *The relationship between family violence and youth offending*. Retrieved from: https://www.local.gov.uk/sites/default/files/documents/15%2034%20-%20The%20relationship%20between%20family%20violence%20and%20youth%20offending-V4_1.pdf

Loeffler, C. E., & Nagin, D. S. (2021). The Impact of Incarceration on Recidivism. *Annual Review of Criminology*, *5*(1). https://doi.org/10.1146/annurev-criminol-030920-112506

Machin, S., & Sandi, M. (2020). Autonomous schools and strategic pupil exclusion. *The Economic Journal*, *130*(625), 125-159.

McCarter, S. (2017). The school-to-prison pipeline: A primer for social workers. *Social Work*, *62*(1), 53-61. https://doi.org/10.1093/sw/sww078

Morton, R. (2014). Returning decision to school discipline decisions: An analysis of recent, anti-zero tolerance legislation. *Wash. UL Rev.*, *91*, 757. https://openscholarship.wustl.edu/law_lawreview/vol91/iss3/8/

National Audit Office. (2011). *The cost of a cohort of your offenders.* 36. NAO: London. https://www.nao.org.uk/wp-content/uploads/2010/12/1011663_technical_paper.pdf

OECD. (2019). *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners | READ online.* Oecd-ilibrary.org. Retrieved from: https://read.oecd-ilibrary.org/education/talis-2018-results-volume-i_1d0bc92a-en

Office for National Statistics. (2021). *Open academies, free schools, studio schools and UTCs.* Retrieved from: https://www.gov.uk/government/publications/open-academies-and-academy-projects-in-development

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological science*, *26*(5), 617-624. https://doi.org/10.1177/0956797615570365

Okonofua, J. A., Perez, A. D., & Darling-Hammond, S. (2020). When policy and psychology meet: Mitigating the consequences of bias in schools. *Science advances*, *6*(42), eaba9479. http://doi.org/10.1126/sciadv.aba9479

Petrich, D. M., Pratt, T. C., Jonson, C. L., & Cullen, F. T. (2021). Custodial Sanctions and Reoffending: A Meta-Analytic Review. *Crime and Justice*, *50*(1), 000-000. https://doi.org/10.1086/715100

Prison Reform Trust. (2019). Prison: the facts. Bromley Briefings Summer 2019. Retrieved from:

http://www.prisonreformtrust.org.uk/Portals/0/Documents/Bromley%20Briefings/Prison%20the%20facts%20Summer%202019.pdf

Pyle, N., Flower, A., Williams, J., & Fall, A. M. (2020). Social risk factors of institutionalized juvenile offenders: A systematic review. *Adolescent Research Review*, *5*(2), 173-186. https://doi.org/10.1037/h0034537

Richardson, H. (2017). School discipline: How strict is too strict?. *BBC News.* Retrieved from: https://www.bbc.co.uk/news/education-41441639

Rosenbaum, J. (2020). Educational and criminal justice outcomes 12 years after school suspension. *Youth & Society*, *52*(4), 515-547. http://doi.org/10.1177/0044118X17752208

Sanders, J., Liebenberg, L., & Munford, R. (2020). The impact of school exclusion on later justice system involvement: Investigating the experiences of male and female students. *Educational Review*, *72*(3), 386-403. https://doi.org/10.1080/00131911.2018.1513909

Shapland, J., & Hall, M. (2007). What do we know about the effects of crime on victims?. *International Review of Victimology*, *14*(2), 175-217. https://doi.org/10.1177/026975800701400202

Smith, I. D. (2007). Being tough on the causes of crime: Tackling family breakdown to prevent youth crime. *Social Justice Challenge*, 1-15.

Steinberg, L., & Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental psychology*, *43*(6), 1531.

Sutherland, A., Disley, E., Cattell, J., & Bauchowitz, S. (2017). An analysis of trends in first time entrants to the youth justice system. London: Ministry of Justice. https://www.gov.uk/government/publications/analysis-of-trends-in-first-time-entrants-to-the-youth-justice-system

Sutherland, A., & Eisner, M. (2014). The treatment effect of school exclusion on unemployment. *Available at SSRN 2380956*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2380956

Taylor, C. (2016). Review of the youth justice system in England and Wales. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/577105/youth-justice-review-final-report-print.pdf

Timpson, E. (2019). *Timpson review of school exclusion*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/807862/Timpson_review.pdf

Valdebenito, S., Eisner, M., Farrington, D. P., Ttofi, M. M., & Sutherland, A. (2018). School-based interventions for reducing disciplinary school exclusion: a systematic review. *Campbell Systematic Reviews*, *14*(1), i-216. https://doi.org/10.4073/csr.2018.1

Wikström, P. O. H., Oberwittler, D., Treiber, K., & Hardie, B. (2012). *Breaking rules: The social and situational dynamics of young people's urban crime*. Oxford University Press.

Williams, K., Papadopoulou, V., & Booth, N. (2012). Prisoners' childhood and family backgrounds. *Results from the Surveying Prisoner Crime Reduction (SPCR) longitudinal cohort study of prisoners. Ministry of Justice Research Series*, *4*, 12. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachme

nt_data/file/278837/prisoners-childhood-family-backgrounds.pdf

Written questions and answers - Written questions, answers and statements - UK Parliament. (n.d.). Retrieved 9 September 2021, from https://questions-statements.parliament.uk/written-questions/detail/2018-05-15/144303

Appendices

Appendix A: Data

A.1 Datasets

As outlined above in Section 2.1, we have NPD and NCCIS data on pupils from four cohorts of pupils who are born between 1-Sep-1994 and 31-Aug-1998. We focus on these pupils when they are in year 10 which roughly covers the academic years 2009/10 to 2012/13. Some pupils that are behind their cohort, either because they started school late or were held back to repeat a year, will be in year 10 in a later academic year than this, so in reality we have data on pupils in our cohorts for year 10 up to 2014/15.

We use the following pupil-level data to construct a dataset with information on year 10 pupils and any experiences of custody as a young person later in life:

1. **NPD pupil census data -** pupil censuses are done at three points through an academic year, autumn, spring and summer, and capture which pupils were registered at which schools. We use pupil census data to identify what schools were attended by pupils when they are in year 10 from academic years 2009/10 through to 2014/15. We also use pupil census data for year 9 to determine what schools pupils attended in year 9 and whether this was different from the schools attended in year 10.
2. **NPD year 10 exclusions data** to identify year 10 pupils who were excluded between the academic years 2009/10 to 2014/15, what type of exclusions they experienced (permanent or temporary suspension), and when they occurred
3. **NCCIS data** captures whether young people go into further education, training or employment following the compulsory school leaving age. It covers pupils in the year prior to the compulsory school leaving age, and for two years after (more for pupils with SEN). We use the NCCIS data to identify which pupils end up in custody during the years 2010/11 to 2017/18, which is indicated by activity codes 130 (custodial sentence), 290 (custodial institution, juvenile offender) and 710 (custody, young adult offender).

In addition to data at the pupil level, we also use school level data. This comes from the Consistent Schools Database (CSD)**.**[22] In the Department for Education's School Level Database (SLD), schools are identified by unique reference numbers (URNs); however, these reference numbers change when specific characteristics of a school change, for example, when a school becomes an academy. The CSD enables schools to be tracked

[22] This project used the Consistent Schools Database, produced by Ellen Greaves and Bilal Nasim, funded by CLOSER, the home of longitudinal research (grant reference: ES/K000357/1). www.closer.ac.uk

through time with a “consistent URN” that does not change. This dataset also critically enables us to identify the years in which schools become academies.

A.2 Merging and cleaning

We import year 10 NPD census data from three different terms for all pupils in our cohorts and make it unique on the combination of a pupil identifier number (PMRA), academic year and school reference number (URN). Any pupil that attends more than one school in year 10 or is registered in year 10 in more than one year will have multiple records, while pupils that attend only one school in one academic year will have one record.

This data is merged with our year 10 exclusions data on PMRA, academic year and URN to identify whether pupils in a specific year and school received exclusions and of what type. 7,309 observations from the exclusion data do not merge with the census data, indicating pupils we have recorded as having an exclusion in schools or years they did not attend. We drop pupils whose PMRA we never observe in the census data (543 pupils) and keep the pupils who receive exclusions at schools or in years not in the census data, assuming that these pupils did actually attend the schools they received exclusions at, and this was “missed” in the school census.

We then merge our combined census and exclusions data with the Consistent School Database on the combination of URN and year. In this merge,145,059 observations (pupils in specific schools and years) from 893 different schools in our census data could not be merged to the school level data; these are pupils that were recorded as attending a school that did not exist in our school level data. We drop these pupils.

We next merge in the custody data using the pupil identifier (PMRA). 4,740 observations from the custody data do not merge, meaning that we observe these pupils in custody, but we do not observe them in year 10; these individuals are dropped. We also drop 21 kids who were in custody in the same academic year or earlier than when they were registered in year 10.

Finally, we compare year 10 census data with year 9 census data to identify pupils that switch schools between the two years. If any of the following are true, we assume that the pupil *did not* switch schools between year 9 and year 10:

1) URN in year 9 = URN in year 10
2) Consistent URN in year 9 = Consistent URN in year 10
3) School name in year 9 = school name in year 10
4) URN in year 9 = Consistent URN in year 10
5) Consistent URN in year 9 = URN in year 10

To make these comparisons, we drop all pupils where we either a) observe them in year 9 but not year 10 (36,389 pupils), b) observe them in year 10 but not year 9 (31,790 pupils) or c) observe them registered in year 9 and year 10 in the same academic year (3,343 pupils).

Appendix B: Cross-sectional analysis custody on exclusion

B.1 Analytical strategy

We run the following regression to observe the relationship between custody at age 15-17, and exclusion in year 10.

$$custody\_before18_{isy} = \theta * exclusion\_[type]_{isy}^{Y10} + \nu * covariates_{i}^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $custody\_before18_{isy}$ is a binary variable that is equal to 1 if pupil i was observed starting custody in between 15 and 17 (inclusive). For all analyses, we exclude pupils that are in custody during or before the academic year they are in year 10.
- $exclusion\_[type]_{isy}^{Y10}$ is a binary indicator equal to 1 if a year 10 pupil i attending school (s) in academic year (y) received an exclusion of a specific $[type]$ and 0 otherwise.
- $[type]$ is one of *any*, *permanent*, or *fixed.*
- $covariates_{i}^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including *gender*, *free-school meal eligibility*, and *special education needs.*
- γ_Y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school level.

B.2 Results

Table 7: cross-sectional regression, custody age 15-17 on exclusion in year 10

Sample: All pupils			
Dependent variable	Custody, age 15-17		
Explanatory variable: receiving exclusion or suspension in year 10 of following types: (b/t)	(1) Any	(2) Permanent exclusions	(3) Suspension
	0.008** (26.04)	0.029** (11.76)	0.008** (25.78)
Constant	0.002** (27.22)	0.002** (33.27)	0.002** (25.78)
r^2	0.023	0.022	0.023
N	2,294,420	2,294,420	2,294,420
+ p<0.05, * p<0.01, ** p<0.001 Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects.			

Appendix C: Technical annex for academisation analysis

C.1 Sample

Table 8 outlines the samples that we use in both academisation analyses, and the differences between the main analyses, results of which are presented in the main paper above, and the additional analyses, results from which are included in this appendix.

Table 8: samples used in academisation analysis

	Attendance at academy in year of or after conversion		Attendance at academy in year of conversion	
	Main analysis	Robustness check	Robustness check	Robustness check
Number of observations	1,064,555	1,260,759	752,676	935,090
Number of year 10 pupils[23]	1,064,555	1,248,037	752,676	926,502
Years covered (when pupils are in year 10)	2009/10 to 2013/14	2009/10 to 2013/14	2009/10 to 2013/14	2009/10 to 2013/14
Requirement for pupil to be included in sample:				
Attends a school that has greater than 5 year 10 pupils in their year	✓	✓	✓	✓
Attends a school that is ever an academy	✓	✓	✓	✓
Attends a school that has either not academised or is in year of conversion			✓	✓
Does not switch schools in year 10 or between year 9 and year 10	✓		✓	
Does not repeat year 10	✓		✓	

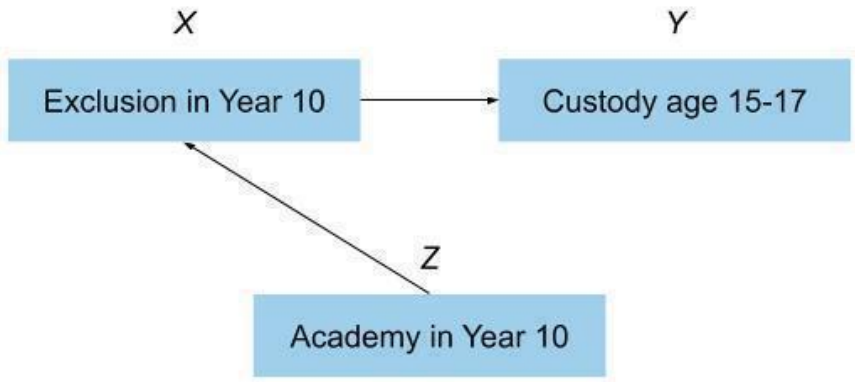
[23] Number of pupils is different than number of observations in some samples as students that repeat year 10 or are registered in more than one school in year 10 will have multiple observations, one for each different school and/or academic year

C.2 Analytical strategy

C.2.1 Academy

We use attendance in an academy in year 10 as an instrument for exclusion for year 10, as outlined in Figure 4. The regression equations, for both the first stage, second stage and reduced form, are outlined below. The results for the main analyses, which exclude pupils that repeat year 10 or switch schools between year 9 and 10 or during year 10, are provided above in Table 2 (first stage) and Table 4 (second stage) in columns (1) through (3) and Table 6 (reduced form) in the main text above. Additional results are provided below in Table 9 in columns (1) through (3) and Table 10 for a larger sample which includes pupils either repeat year 10 or switch schools.

Figure 4: academy as an instrument



First stage regression:

The basic first stage regression of exclusion on academy status, for year 10 pupil (i), attending some school (s) in academic year (y) is as follows:

$$exclusion\_[type]_{isy}^{Y10} = \beta * academy_{sy}^{Y10} + \nu * covariates_{i}^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $exclusion\_[type]_{isy}^{Y10}$ is a binary indicator equal to 1 if a year 10 pupil i attending school (s) in academic year (y) received an exclusion of a specific $[type]$ and 0 otherwise.
- $[type]$ is one of *any*, *permanent* or *fixed.*
- $academy_{sy}^{Y10}$is a binary indicator equal to 1 if pupil i attended year 10 in a school (s) that was an academy and 0 if the pupil's school had not yet converted to an academy
- $covariates_{i}^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_y is a vector of year effects.
- δ_s is a vector of school effects.

- ϵ_{isy} are the residual errors, clustered at the school-level.

Second stage regression

$$custody\_before18_{isy} = \theta * \widehat{exclusion\_[type]}_{isy}^{Y10} + \nu * covariates_i^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $custody\_before18_{isy}$ is a binary variable that is equal to 1 if pupil i was observed starting custody in between 15 and 17 (inclusive). For all analyses, we exclude pupils that are in custody during or before the academic year they are in year 10.
- $\widehat{exclusion\_[type]}_{isy}^{Y10}$ is the predicted probability that a year 10 pupil (i) attending school (s) in academic year (y) received an exclusion of a specific $[type]$ and 0 otherwise.
- $[type]$ is one of *any*, *permanent*, or *fixed.*
- $covariates_i^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_Y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school-level.

Reduced form regression

$$custody\_before18_{isy} = \beta * academy_{sy}^{Y10} + \nu * covariates_i^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

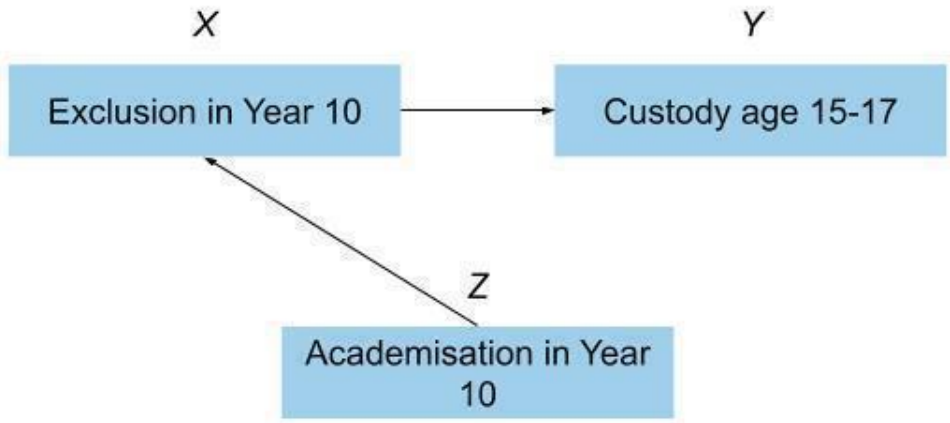
- $custody\_before18_{isy}$ is a binary variable that is equal to 1 if pupil i was observed starting custody in between 15 and 17 (inclusive). For all analyses, we exclude pupils that are in custody during or before the academic year they are in year 10.
- $academy_{sy}^{Y10}$is a binary indicator equal to 1 if pupil i attended year 10 in a school (s) that was an academy and 0 if the pupil's school had not yet converted to an academy.
- $covariates_i^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school-level.

C.2.2 Attendance at a school in the academic year of academy conversion

We use pupil attendance in year 10 at a school that converts to an *academy in that academic year* as an instrument for exclusion for year 10, as outlined in Figure 5. The regression equations, for both the first stage, second stage and reduced form are outlined below. The results for the main analyses, which exclude pupils that repeat year 10 or switch schools between year 9 and 10 or during year 10, are provided above in Table 2 (first stage), and

Table 4 (second stage) in columns (4) through (6) and Table 6 (reduced form) in the main text above. Additional results are provided below in Table 9 columns (4) through (6) and Table 10 (reduced form) for a larger sample which includes pupils either repeat year 10 or switch schools.

Figure 5: first-year academy as an instrument



First stage regression:

The basic first stage regression of exclusion on academy status, for year 10 pupil (i), attending some school (s) in academic year (y) is as follows:

$$exclusion\_[type]_{isy}^{Y10} = \beta * academy\_firstyear_{sy}^{Y10} + \nu * covariates_{i}^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $exclusion\_[type]_{isy}^{Y10}$ is a binary indicator equal to 1 if a year 10 pupil i attending school (s) in academic year (y) received an exclusion of a specific $[type]$ and 0 otherwise.
- $[type]$ is one of *any*, *permanent* or *fixed.*
- $academy\_firstyear_{sy}^{Y10}$ is a binary indicator equal to 1 if a pupil (i) attended year 10 in a school (s) that converted to an academy in academic year (y) (i.e., it was not an academy in year $y-1$) and 0 if the pupil's school has not yet converted to an academy.
- $covariates_{i}^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school-level.

Second stage regression

$$custody\_before18_{isy} = \theta * \widehat{exclusion\_[type]}_{isy}^{Y10} + \nu * covariates_{i}^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $custody\_before18_{isy}$ is a binary variable that is equal to 1 if pupil i was observed starting custody in between 15 and 17 (inclusive). For all analyses, we exclude pupils that are in custody during or before the academic year they are in year 10.
- $\widehat{exclusion\_[type]}_{isy}^{Y10}$ is the predicted probability that a year 10 pupil (i) attending school (s) in academic year (y) received an exclusion of a specific $[type]$ and 0 otherwise.
- $[type]$ is one of *any*, *permanent*, or *fixed*
- $covariates_i^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_Y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school-level.

Reduced form regression

$$custody\_before18_{isy} = \beta * academy\_firstyear_{sy}^{Y10} + \nu * covariates_i^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $custody\_before18_{isy}$ is a binary variable that is equal to 1 if pupil i was observed starting custody in between 15 and 17 (inclusive). For all analyses, we exclude pupils that are in custody during or before the academic year they are in year 10.
- $academy\_firstyear_{sy}^{Y10}$ is a binary indicator equal to 1 if a pupil (i) attended year 10 in a school (s) that converted to an academy in academic year (y) (i.e., it was not an academy in year $y-1$) and 0 if the pupil's school has not yet converted to an academy.
- $covariates_i^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school-level.

C.3 Robustness checks

Table 9: effect of being in an academy on probability of custody aged 15-17 (2SLS) - robustness checks

First stage						
Analysis:	Attendance at academy in year of or after conversion, including students that switch schools or repeat year 10			Attendance at academy in year of conversion, including students that switch schools or repeat year 10		
Dependent variable: receiving exclusion or suspension in year 10 of following types:	(1) Any	(2) Permanent exclusion	(3) Suspension	(4) Any	(5) Permanent exclusion	(6) Suspension
Explanatory variable: Attendance at an academy (b/t)	0.026** (14.42)	0.001** (6.28)	0.026** (14.27)	0.030** (14.91)	0.001** (6.34)	0.029** (14.78)
F	207.8	39.5	203.7	222.2	40.3	218.5
Mean	0.055	0.002	0.055	0.051	0.002	0.050
N	1,260,759	1,260,759	1,260,759	935,090	935,090	935,090
Second stage						
Analysis:	Attendance at academy in year of or after conversion, including students that switch schools or repeat year 10			Attendance at academy in year of conversion, including students that switch schools or repeat year 10		
Dependent variable:	Experience of custody, age 15-17					
Explanatory variable: predicted exclusion or suspension of following types: (b/t)	(1) Any	(2) Permanent exclusion	(3) Suspension	(4) Any	(5) Permanent exclusion	(6) Suspension
	0.009 (1.83)	0.188 (1.80)	0.009 (1.83)	0.006 (1.19)	0.123 (1.20)	0.006 (1.19)
F	77.2	76.3	77.2	58.1	58.5	58.1
Mean	0.001	0.001	0.001	0.001	0.001	0.001
N	1,260,759	1,260,759	1,260,759	935,090	935,090	935,090

+ p<0.05, * p<0.01, ** p<0.001.
Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects.

Table 10: effect of attending an academy on probability of custody aged 15-17 (reduced form) – robustness checks

Analysis:	Attendance at academy in year of or after conversion, including students that switch schools or repeat year 10	Attendance at academy in year of conversion, including students that switch schools or repeat year 10
Academy	0.0002 (1.84)	0.0002 (1.19)
N	1,260,759	935,090
+ p<0.05, * p<0.01, ** p<0.001. Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects.		

Table 11: average exclusion rates by sample and academy status

Analysis:	Attendance at academy in year of or after conversion, including students that switch schools or repeat year 10			Attendance at academy in year of conversion, including students that switch schools or repeat year 10		
Type of exclusion:	(1) Any	(2) Permanent exclusion	(3) Suspension	(4) Any	(5) Permanent exclusion	(6) Suspension
Exclusion rate among all pupils in sample	0.054 (N=1,248,271)	0.002 (N=1,248,271)	0.054 (N=1,248,271)	0.050 (N=926,527)	0.002 (N=926,527)	0.049 (N=926,527)
Exclusion rate among only pupils that attend schools that have not yet converted to academies	0.050 (N=668,258)	0.002 (N=668,258)	0.049 (N=668,258)	0.050 (N=668,258)	0.002 (N=668,258)	0.050 (N=668,258)

Table 12: average custody rate by sample & academy status

Analysis:	Attendance at academy in year of or after conversion, including students that switch schools or repeat year 10	Attendance at academy in year of conversion, including students that switch schools or repeat year 10
Custody rate among all pupils in sample	0.00108 (N=1,248,271)	0.00105 (N=926,527)
Custody rate among only pupils that attend schools that have not yet converted to academies	0.00112 (N=668,258)	0.00112 (N=668,258)

Appendix D: Alternative instrument, headteacher exclusion rates

We explored, using an alternative instrument, headteacher exclusion rates; however, this was found to be a weak instrument. A headteacher's baseline exclusion rate did not have a statistically significant impact on whether a year 10 pupil they teach in a future year will experience an exclusion. As a result, this instrument was not pursued further. This section contains details on how the instrument was set-up and the results of the first stage regressions.

D.1 Datasets

In addition to the datasets described in Appendix A, we had access to a dataset of headteacher names at all schools with a sixth form between the years 2009-2014.[24] We create unique headteacher IDs using the combination of first name, last name, and imputed gender based on title and name.

Using the pupil census data described above, we create a dataset that is at the school level, showing how many year 10 pupils there were at a school in a given academic year, and how many exclusions of a specific type (any, permanent or fixed) there were in that year and school. From this we calculate the exclusion rate for each school and year per 100 pupils:

$$exclusion\ rate\ per\ 100\ pupils\ = \frac{number\ of\ year\ 10\ exclusions\_[type]}{number\ of\ year\ 10\ pupils} * 100$$

We merge the school-level dataset to our headteacher data. 9,512 observations (schools in particular years) from the principals data do not match with schools in the schools data. These are schools that we have principal data on but do not have any year 10 pupils attending; these are dropped. 8,134 observations (schools in particular years) from the census data do not match with schools in the headteacher data; these are schools that we have records of year 10 pupils attending but we do not have data on the headteacher. 8,646 observations match between the two datasets; these are schools where we observe both year 10 pupils and the headteacher.

We use 2009/2010 as a "baseline" year and create a variable for each headteacher that is their exclusion rate in the baseline year. We assume that for each headteacher, their preference for using exclusions as a disciplinary strategy is roughly represented by the number of exclusions per 100 pupils that happen in their school in that year. That is, a headteacher that has a high exclusion rate in 2009/2010 is likely to have a high exclusion rate in future years. We drop all observations for headteachers that we do not observe in 2009/2010 (2,146 observations). We also drop observations for the year 2009/2010. This

[24] Principals Longitudinal Dataset, provided by Dr J. Ruiz-Valenzuela (Centre for Vocational Education Research (CVER) at the London School of Economics), previously used in the CVER publication *Effectiveness of CEOs in the Public Sector: Evidence from Further Education Institutions (J. Ruiz-Valenzuela, C. Terrier, C. Van Effenterre, 2017)*, available at https://cver.lse.ac.uk/textonly/cver/pubs/cverbrf005.pdf.

leaves us with data on what schools headteachers taught at for the academic years 2010/2011 to 2013/14 and their baseline exclusion rate.

We merge this teacher data back to the individual pupil level census data. We drop all pupils where we do not observe their headteacher (1,486,434 observations), leaving us with 807,986 pupils. Table 13 outlines the two samples we create from this dataset.

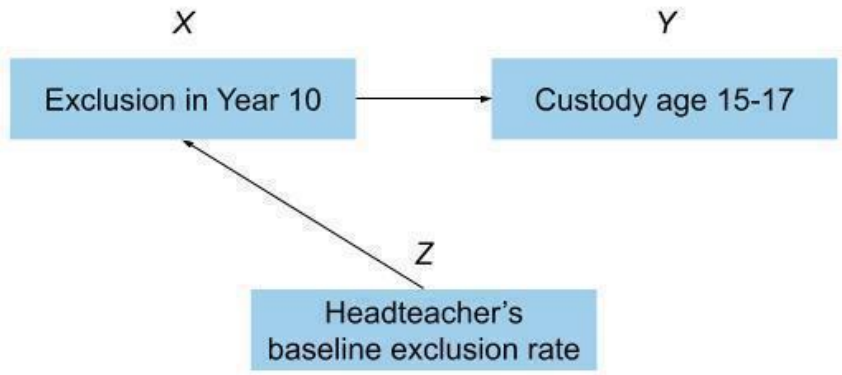
Table 13: samples used for headteacher instrument

	Headteacher exclusion rates	
	Primary sample	Secondary sample
Number of observations	722,074	807,986
Number of year 10 pupils	722,074	800,130
Years covered (when pupils are in year 10)	2010/11 to 2013/14	2010/11 to 2013/14
Requirement for pupil to be included in sample:		
Attends a school that has greater than 5 year 10 pupils in their year	✓	✓
Attends a school with a 6th form when in year 10	✓	✓
Attends a school in year 10 with a headteacher whose school record we observe in 2009/10	✓	✓
Does not switch schools in year 10 or between year 9 and year 10	✓	
Does not repeat year 10	✓	

D.2 Analytical strategy

Figure 6 depicts how we use a headteacher's baseline exclusion rate as an instrument for exclusion in Year 10. The regression equations, for both the first stage and second stage of the regression are outlined below. The results for the main analyses, which exclude pupils that repeat year 10 or switch schools between year 9 and 10 or during year 10, are provided in Table 14 (first stage only). Additional results are provided below in Table 15 (first stage only) for a larger sample which includes pupils either repeat year 10 or switch schools. Since the first stage effect is extremely small (and not significant in the case of permanent exclusions), we do not proceed to the second stage.

Figure 6: headteacher baseline exclusion rate as an instrument



First stage regression:

The basic first stage regression of exclusion on headteacher's baseline exclusion rate, for year 10 pupil (i), attending some school (s) in academic year (y) is as follows:

$$exclusion_{[type]_{isy}}^{Y10}$$

$$= \beta * ht\_exclusionrate\_[type]_{sy}^{baseline} + \nu * covariates_i^{Y10} + \gamma_y + \delta_s + \epsilon_{isy}$$

Where:

- $exclusion\_[type]_{isy}^{Y10}$ is a binary indicator equal to 1 if a year 10 pupil i attending school (s) in academic year (y) received an exclusion of a specific $[type]$ and 0 otherwise.
- $[type]$ is one of *any*, *permanent* or *fixed.*
- $ht_exclusionrate_[type]_{sy}^{baseline}$is the baseline exclusion rate for a specific $[type]$ of exclusion of the headteacher that pupil i has in year 10 in school (s) in academic year (y).
- $covariates_i^{Y10}$ are pupil-level covariates from when the pupil was in year 10, including gender, free-school meal eligibility, and special education needs.
- γ_y is a vector of year effects.
- δ_s is a vector of school effects.
- ϵ_{isy} are the residual errors, clustered at the school-level.

D.3 Results

Table 14: effect of headteacher's baseline exclusion rate on probability of exclusion in year 10 (first stage, 2SLS) - primary sample

Analysis: First stage regression with pupils that attend year 10 in a school with a sixth form college, excluding pupils that switch schools or repeat year 10			
Dependent variable: receiving exclusion or suspension in year 10 of following types:	(1) Any	(2) Permanent exclusion	(3) Suspension
Explanatory variable: Headteacher's baseline exclusion rate (b/t)	0.001 (1.13)	<0.001 (0.49)	0.001 (1.06)
F	1.3	0.2	1.1
Mean exclusion rate	0.053	0.002	0.053
N	722,064	722,064	722,064
+ p<0.05, * p<0.01, ** p<0.001 Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects 10 singleton observations dropped from regressions (columns (1)-(3))			

Table 15: effect of headteacher's baseline exclusion rate on probability of exclusion in year 10 (first stage, 2SLS) - secondary sample

Analysis: First stage regression with pupils that attend year 10 in a school with a sixth form college, including pupils that switch schools or repeat year 10			
Dependent variable: receiving exclusion or suspension in year 10 of following types:	(1) Any	(2) Permanent exclusion	(3) Suspension
Explanatory variable: Headteacher's baseline exclusion rate (b/t)	0.001+ (2.12)	<0.001 (0.27)	0.001+ (2.03)
F	4.5	0.07	4.1
Mean exclusion rate	0.053	0.002	0.053
N	807,982	807,982	807,982
+ p<0.05, * p<0.01, ** p<0.001 Regressions control for: gender, FSM eligibility in year 10, SEN in year 10, year fixed effects, school fixed effects 4 singleton observations dropped from regressions (columns (1)-(3))			

Appendix E: Costs of crime and custody

Crime, in particular violence, carries significant societal costs, first and foremost to victims, many of whom suffer long-lasting effects (HM Government, 2019; Shapland & Hall, 2007). The financial costs of crime to the public purse are also significant. In 2015/16 the total cost of crime was estimated to be nearly £60bn (Heeks et al., 2018). The costs of crime arising from a single cohort of 83,000 young offenders who entered the justice system in 2000 – first time offenders – is £664,000,000, an average of £8,000 per offender but with substantial variation across the cohort in terms of persistence and seriousness of offending (National Audit Office, 2011). We can also quantify the overt costs of punishment – specifically custody. Table 16 below sets out costs for different types of establishments in the under-18 secure estate, which also provides a benchmark for the potential savings possible through not using the secure estate (notwithstanding that costs reflect estate and personnel).

Table 16: average price per place per year of secure Children's Homes, Secure Training Centres and Under-18 Young Offender Institutions

Secure Children's Homes (SCHs)	*£210,000*
Secure Training Centres (STCs)	*£160,000*
Under-18 Young Offender Institutions (YOIs)	*£76,000*

Source: (Written Questions and Answers - Written Questions, Answers and Statements - UK Parliament, n.d.)