Discussion Paper Series

IZA DP No. 18599

April 2026

The Timing of School Exclusions and Its Consequences for Peers' Outcomes

Richard Dorsett
University of Westminster

Veruska Oppedisano
University of Westminster, IFS and IZA@LISER

Dave Thomson
FFT Education Datalab

Min Zhang
University of Westminster

IZA@LISER NETWORK
ISSN: 2365-9793

The Timing of School Exclusions and Its Consequences for Peers' Outcomes*

Abstract

This paper studies how the timing of excluding disruptive pupils affects peer outcomes. While removal can generate positive spillovers, delays may harm peers. We interpret exclusions as determining the timing of removal, and estimate the effect of earlier versus later exclusion. Using an instrumental-variables design based on exogenous changes in local capacity to accommodate excluded pupils, we show that exclusions in Year 9 generate the largest negative spillovers: an increase of one excluded pupil per 1,000 in Year 9 pupils child reduces KS4 Maths and English scores by 0.024 and 0.044 standard deviations, lowers Level-2 and Level-3 attainment by around 0.65–0.63 percentage points, it raises the probability of being NEET at 21 by 0.62 percentage points. Peer exclusion effects are highly heterogeneous and timing-dependent, with early exclusions improving labour-market outcomes, mid-secondary exclusions raising attainment but weakening employment trajectories, and late exclusions generating broad losses for disadvantaged pupils. To shed light on the underlying mechanisms, we show that the number of suspension days accumulated by excluded pupils prior to exclusion, a proxy for their level of disruptiveness, explains these negative effects. Together, the results highlight that the timing and nature of exclusions critically shape peer spillovers and provide policymakers with new evidence on the broader social costs of exclusionary discipline.

JEL classification

C36, I2, I20, J24

Keywords

disruptive behaviours, exclusions, schooling, peer effects

Corresponding author

Veruska Oppedisano
v.oppedisano@westminster.ac.uk

* *Acknowledgements*: 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. The authors thank the Department for Education for providing access to the data. The project has been funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org.

Introduction

Managing student behaviour is a central component of the education production process. Schools face a key policy trade-off: removing persistently disruptive students imposes significant long-term costs on the excluded students themselves, yet it may also produce significant benefits for peers and teachers.[1] The literature indicates that the costs for the removed pupils are high, and they accumulate over time, leading to lower educational attainment, worsened employment prospects and earlier involvement in the criminal justice system (McAra and McVie, 2010, Sutherland and Eisner, 2014, Madia et al., 2022 and Buchmueller et al., 2025 in the UK; Cuellar and Markowitz, 2015 in the USA and Sanders et al., 2020 in New Zealand).

In contrast, relatively little is known about the consequences of exclusion for non-excluded peers. While existing research has documented that classrooms with high levels of disruptive behaviour tend to have poorer overall academic outcomes (e.g. Carrell et al., 2018), we lack causal evidence on the spillovers generated when disruptive students are removed and, importantly, when in the school career these removals occur.

Rising levels of classroom misbehaviour have made disruptive peer effects an urgent topic for empirical investigation. In May 2023, 76 percent of teachers in England reported that misbehaviour stopped or interrupted teaching in at least some lessons during the past week, up from 64 percent in June 2022 (DfE, 2024). Such trends raise concerns about the broader costs of inaction, and the potential gains from interventions that modify peer behaviour or reallocate students.

This paper makes a novel contribution by shifting the focus from the disruptive pupil to the implications of their removal, and the point in time at which the removal occurs, through school exclusion on non-excluded peers' short- and longer-term outcomes. Rather than estimating the marginal impact of disruptive students, we quantify the educational consequences for their peers when such students are removed from the classroom at different stages of secondary schooling. In doing so, we provide new evidence on the systemic spillover effects of disciplinary exclusions, evidence that is central to understanding the trade-offs policymakers face when designing school discipline and inclusion policies.

There are two main empirical challenges of credibly estimating causal effects arising from estimating the effects of the removal of disruptive peers. First, measures of disruptive behaviour are rare and often indirectly linked to aspects of the family environment (Carrell and Hoekstra, 2010; Kristoffersenet al, 2015; Carrell et al., 2018; Zhao and Zhao, 2021), making it difficult to separate personality or behavioural traits from family conditions. Second, students self-select into schools based on various factors. Since these factors can be correlated with disruptive behaviour and disciplinary policies, schools with more exclusions, likely have other unobservable characteristics, such as behavioural norms, peer dynamics, or family backgrounds as well as school culture, that could also influence student outcomes, making it difficult to isolate the true effect of exclusion.

[1] We also examine whether peer exclusions affect teacher turnover at the school level, but find no statistically significant effects (results available upon request).

Our study addresses the measurement challenge in the literature by employing a direct and policy-relevant definition of disruptiveness: we identify disruptive peers as those who are permanently excluded from school for behavioural reasons, a clear, institutional signal of severe and/or persistent disruptive behaviour. Permanent exclusions are the most severe disciplinary action for rule violations. They are used to address significant behavioural issues such as persistent disruptive behaviour, bullying, or violence.

To address the endogeneity in both the decision to exclude and the timing of exclusion, we employ an instrumental variable strategy (IV). Our instrument exploits exogenous variation in the local supply of Alternative Provision (AP), the specialized educational settings where excluded pupils tend to be sent after being removed from school. The instrument measures the change in the number of filled AP places in each year group within a 10-kilometre radius of a school 'as the crow flies' from the start of the current academic year to the start of the following academic year, excluding the focal school's own exclusions. Since overall local AP capacity is fixed in the short term, the number of filled places is inversely related to remaining AP capacity, which constrains the ability of schools to exclude pupils. These shocks to filled AP places are plausibly exogenous to the characteristics of the school under study and reflect aggregate exclusion decisions within a local catchment area.

The validity of this strategy hinges on two key features. First, the institutional setup supports exogeneity: because each AP serves multiple mainstream schools within a catchment area, changes in filled capacity over time are driven by the aggregate behavioural environment and exclusion decisions across schools, not solely by the unobservables of a focal school, whose exclusions are removed from the instrument. Second, AP settings are designed to educate pupils who are no longer able to attend mainstream schools, commonly following permanent exclusion; non-excluded peers, whose outcomes we are evaluating, remain enrolled in mainstream secondary schools throughout and do not attend AP settings. Therefore, they are unaffected by change in filled AP supply except indirectly through the removal of disruptive peers. This design reduces the likelihood that our instrument is correlated with unobserved determinants of non-excluded students' outcomes, thereby satisfying the exclusion restriction.

Using data from over two million English secondary school leavers, our analysis leverages the Longitudinal Education Outcomes (LEO) dataset, which is linked to the National Pupil Database. This rich administrative data allows us to estimate our results for four entire student cohorts completing secondary school from the summers of 2014 to 2017.

Our findings show that Year 9 exclusions generate negative spillovers: a one-per-thousand increase in the Year-9 exclusion rate reduces KS4 Maths and English scores by 0.024 and 0.044 standard deviations, lowers the likelihood of achieving a Level 2 qualification by 0.65 percentage points and Level 3 by 0.63 percentage points, raises the probability of being NEET at age 21 by 0.62 percentage points. A one-per-thousand increase in Year 10 exclusions also worsens outcomes, reducing KS4 English by 0.015 standard deviations, lowering the probability of achieving a Level 2 qualification by 0.28 percentage points, and reducing employment at age 21 by 0.36 percentage points. However, these effects are smaller than those associated with Year 9 exclusions. This pattern suggests that much of the cumulative disruption to peers occurs earlier, so that by Year 10 a substantial share of the spillover costs has already materialised.

Our heterogeneity analysis reveals that the effects of peer exclusions vary sharply by pupil background and by the timing of exclusion. Disadvantaged pupils, including those eligible for free school meals, with special educational needs, and with low prior attainment, do not respond uniformly to peer removal. Early exclusions (Year 8) primarily improve labour-market attachment for these groups, with higher employment and lower NEET rates. In contrast, exclusions in Year 9 are particularly damaging for low-ability pupils across all outcomes, even when short-run attainment improves for some groups. By Year 10, exclusions depress attainment and labour-market outcomes for disadvantaged pupils, while low-ability and SEN pupils experience gains in KS4 performance alongside weaker employment outcomes. Taken together, these results show that peer exclusions reshape educational inequality through distinct mechanisms at different stages of secondary schooling, alternately disrupting trajectories, or affecting high-stakes certification margins.

To shed light on the mechanisms behind peer exclusion effects, we use the number of suspension days accumulated by each excluded pupil prior to exclusion, a proxy for their level of disruptiveness. Evaluated at the mean level of suspension intensity (11 days), the interaction accounts for approximately 18–26% of the overall negative effect of Year-9 exclusions. For sufficiently high values of the normalised suspension-days measure (around 70 or above, corresponding to the 95th percentile of the distribution), the interaction fully offsets the baseline negative exclusion effect and may even turn it positive. This pattern suggests that these effects are driven primarily by the degree of disruptiveness rather than exclusion status alone. Disaggregating exclusions by reason shows that the negative spillovers are almost entirely driven by physical-assault and persistent-disruption cases.

Our robustness and falsification exercises confirm that the estimated peer effects are not driven by spurious correlations or unobserved trends in exclusions or AP filled capacity. First, using an alternative instrument based on AP places filled within a 15 km radius, we replicate all main findings with slightly larger but less precise coefficients, suggesting that the results are not sensitive to the specific spatial definition of local AP supply. Restricting the sample to schools with smaller cohort sizes, where peer interactions are stronger but exclusion rates are measured more discretely, produces qualitatively identical patterns, albeit with somewhat attenuated magnitudes, consistent with measurement-induced attenuation rather than weaker underlying spillovers. A joint specification including Year 8, Year 9, and Year 10 exclusions simultaneously yields effects that remain similar in sign and general magnitude to the baseline IV results, though with lower first-stage strength once exclusion rates are jointly instrumented.

We also conduct three placebo tests to assess the validity of the instrument. Using future changes in AP filled capacity, we show that future AP filled capacity shocks do not predict current exclusion rates, consistent with the instrument capturing contemporaneous, not trending, variation in local AP supply. A second placebo test replaces the key regressors with exclusions occurring in other year groups within the same school. These cross-cohort exclusions have no effect on outcomes, confirming that the peer spillovers operate within the relevant cohort and are not artifacts of school-wide shocks. Finally, using predetermined KS2 attainment and authorised absences as placebo outcomes, we find no effect of exclusion rates on either measure, demonstrating that exclusion rates are orthogonal to pupils' prior academic achievement and unrelated aspects of school attendance. Taken together, these checks provide strong evidence that the main results are not driven by reverse causality, omitted trends, cohort spillovers, or violations of the exclusion restriction.

This study contributes to the literature on peer effects and the persistence of disruptive peer influences on long-term outcomes (Carrell et al., 2018; Bacher-Hicks et al., 2024; Goulas et al., 2024). Our key innovation lies in shifting the empirical focus from exposure to disruptive peers to the systemic effects of their removal and the timing of that removal, a perspective that has received little attention in the existing literature. First, we use permanent exclusion from school as a direct and policy-relevant measure of disruptiveness. This provides a cleaner identification of the type of peer envisaged in Lazear's (2001) model, where classroom disruption imposes negative externalities on others' learning. By observing the consequences for peers when such students are formally removed, and when the removal occurs, we estimate spillovers in a way that is both theoretically grounded and institutionally salient. Second, the richness of our data allows us to trace the outcomes of non-excluded students up to age 21, enabling us to study not only short-term academic impacts, but also longer-run effects on post-school trajectories, an important gap in both research and education policy.

While permanent exclusions are relatively rare, affecting just 0.4 percent of pupils in 2024/25, corresponding to 3,715 pupils (DfE, 2025), their potential reach is far greater. Given that the average secondary school year group contains around 180 pupils, we estimate that approximately half of the of pupils are likely to experience at least one removal of a disruptive pupil within their cohort. This highlights the relevance of understanding not only the costs borne by excluded pupils themselves, but also the broader spillover effects, and their timing, on their peers.[2]

By quantifying how the removal of disruptive pupils at different stages of secondary school affects the outcomes of their non-excluded peers, this paper speaks directly to ongoing policy debates on school discipline and inclusion. Our findings highlight that the timing of intervention is central to understanding the broader social costs of disruptive behaviour: removals that occur only after prolonged exposure to disruption are associated with sizeable negative spillovers for classmates, whereas earlier removals are not. Importantly, these results do not imply that exclusion is an optimal or desirable policy response per se. Rather, they suggest that unresolved behavioural difficulties generate cumulative externalities within classrooms, particularly for disadvantaged pupils, and that delays in addressing such difficulties can exacerbate these costs. By shifting attention from individual-level consequences of exclusion to cohort-level spillovers and their timing, the analysis provides an evidence-based framework for evaluating discipline policies that must balance inclusion, equity, and the efficient provision of learning environments.

Previous Literature

Lazear (2001) presents a model in which classroom learning is disrupted by the least disciplined or least academically engaged students, who impose negative externalities on peers by diverting teacher attention and undermining instructional time. These disruptions reduce

[2] Based on a Poisson approximation using an annual exclusion probability of 0.4% and an average cohort size of 180 pupils.

the efficiency of the classroom as a production unit and may lead to spillovers that affect the learning and behaviour of all students.

Empirical studies have documented these peer effects using various proxies for disruptive behaviour. Carrell and Hoekstra (2010) show that children exposed to domestic violence reduce their primary school peers' test scores and increase disciplinary infractions, with long-lasting consequences for behaviour and labour market outcomes (Carrell et al., 2018). Kristoffersen et al. (2015) find that children with psychiatric diagnoses or from criminally involved households reduce reading achievement among classmates in Danish schools. Similarly, Zhao and Zhao (2021) report that students from families with alcoholic fathers negatively affect peer academic performance. More recently, Lavy and Yancu (2025) show that exposure to peers with a record of violent behaviour depresses academic outcomes, especially for girls and low-SES students, by disrupting the classroom environment, eroding peer and teacher relationships, and reducing study effort. Goulas et al. (2024), using administrative data from Greece, find that disruptive peers, measured by suspension hours, reduce the likelihood of completing high school on time and pursuing competitive academic tracks. These studies highlight the costs of exposure to disruptive behaviour, but they do not directly assess the implications of removing disruptive peers through school exclusion.

A related strand of literature, based on USA data, examines the effects of school-level discipline, often measured by average suspension rates, on peer outcomes. Findings are mixed. Bacher-Hicks et al. (2024) exploit variation in school assignment following district boundary changes and find that higher suspension rates have modest short-term benefits for some student subgroups, particularly white male students, but these gains do not persist in educational attainment or crime reduction. Steinberg and Lacoe (2018) evaluate a Philadelphia policy that limited suspensions for non-violent infractions. Their findings are mixed: partial implementation lowered peer math scores and attendance, while full compliance had small negative effects on English scores. Sorensen et al. (2025) use principal turnover to capture shifts in disciplinary policy and find no significant short-term peer effects.

A key concern in these studies is potential correlation between school disciplinary policies and unobserved school characteristics. Pope and Zuo (2023) address potential endogeneity in disciplinary policy by using year-to-year changes in district-level suspension rates in Los Angeles. They find small, positive spillover effects from higher suspension rates on peers, but large, concentrated negative effects on the suspended students themselves, highlighting the trade-off between equity and efficiency. In contrast, Craig and Martin (2023) study a ban on low-level suspensions in New York City and find positive effects on student achievement, likely driven by mandated non-exclusionary responses that improved school climate.

While these studies provide valuable insights, most examine suspensions rather than permanent exclusions, and often at the school or district level, making it difficult to isolate the effects of peer removal within cohorts. Suspensions are also heterogeneous in nature, encompassing a wide range of behaviours and durations. As a result, the existing literature has limited capacity to directly assess whether removing students who are formally identified as highly disruptive improves outcomes for their peers.

This paper contributes to the literature by shifting the focus from peer exposure to peer removal and timing of removal. Specifically, we examine the causal effects of removing students through permanent exclusion, the most severe disciplinary sanction available to English schools, at different times during secondary school, on the outcomes of their non-excluded peers. To address endogeneity in exclusion decisions, we exploit within-school, within-cohort variation in permanent exclusions instrumented by changes in the local supply of Alternative Provision (AP) places. In short, increases and decreases in AP filled capacity, driven by the local school demand for AP places, present a 'floor' and 'cap' on the total number of permanent exclusions that schools can hand out.

Because fewer than half of the AP intake consists of formally excluded pupils, this diminishes the concern that the supply of AP settings is determined primarily by school-level disruptiveness variation in local capacity. This design allows us to estimate local average treatment effects associated with marginal changes in local capacity to accommodate excluded pupils, providing novel and policy-relevant evidence on how the exclusion and timing of exclusion of disruptive peers affects the broader student body.

Institutional setting

Educational system

In England, compulsory education is structured into four Key Stages: KS1 (ages 5-6) and KS2 (ages 7–10) in primary school, followed by KS3 (ages 11–13) and KS4 (ages 13–16) in secondary school, which concludes with General Certificate of Secondary Education (GCSE) and equivalent qualifications. Our analysis focuses on pupils attending secondary state schools, which account for approximately 93 percent of enrolments and 90 percent of schools nationally. School admissions are coordinated by Local Authorities, who allocate places based on parental preferences and, when schools are oversubscribed, according to published criteria that typically prioritise children in care, those with special educational needs, siblings of current pupils, feeder school attendance, catchment area residence, proximity, or, in some cases, religious affiliation (Burgess et al., 2023). We restrict attention to comprehensive (mixed-ability) schools, which are prohibited from selecting pupils on the basis of prior academic achievement. The study sample comprises 2,609 secondary schools and 2,035,656 pupils.

The probability that a child is successful in gaining a place at her/his preferred school depends on how sought-after the preferred school is but also on the cohort size (since this affects demand for places), the prevalence locally of children with one or more of the priority criteria described above and the presence of siblings in the school. This size and composition of the cohort changes year-on-year, and this affects each school's annual intake. This introduces a near-random component to changes in the patterns of enrolment in secondary schools from year to year. This provides a source of exogenous variation in peer group composition, and we use this in our empirical analysis to address school self-selection by exploiting within-school cross-cohort variation over four cohorts of secondary school leavers.

Within secondary schools, students are grouped with different peers for different subjects. This makes it likely that students in secondary schools will encounter many of the other students in their cohort at some point, justifying our definition of peer group as the group of students in the same cohort who attend the same secondary school. Defining peers as those in the same school and cohort rather than in the same class avoid biases from within-school sorting and selection. However, individuals are likely to be more affected by peers they interact with most frequently, such as classmates, and less by more distant peers. As noted by Ammermueller and Pischke (2009) and Angrist (2014), including irrelevant peers in the school-cohort definition introduces measurement error that may understate the true peer effect. Therefore, our estimated impact should be considered a lower-bound estimate of the peer effect.

Permanent exclusions and AP attendance

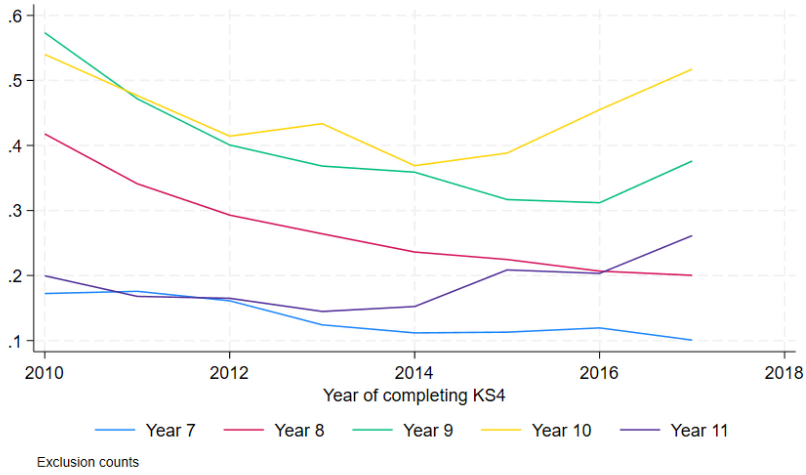
In English secondary schools, disciplinary actions range from minor sanctions to permanent exclusions, which represent the most severe response to student misbehaviour. Different sanctions are designed to be proportionate to the severity of the misconduct and aim to maintain a safe and orderly learning environment. Exclusions are the most severe disciplinary measure. They are used to address significant behavioural issues and are governed by strict guidelines to ensure fairness. According to the relevant Department for Education (2015) guidance, (which was in use during the time-frame of the study hence its relevance) decisions to exclude must be "lawful, rational, reasonable, fair, and proportionate". Exclusion is permitted only for disciplinary reasons; it is unlawful to exclude a pupil for poor academic performance. Common reasons for exclusion include persistent disruptive behaviour, bullying, theft, physical assault, verbal abuse, vandalism, and drug- or alcohol-related offences.[3]

Permanent exclusion rates in England have been relatively stable in recent years, but the longer-term trend shows a gradual rise over the past decade. After declining from 0.23 percent in 2006/07 to 0.12 percent in 2012/13, the rate has increased steadily to 0.25 percent in 2023/24. Figure 1 illustrates this trend by year and grade, with exclusion rates calculated as the share of enrolled students in each grade who are permanently excluded. Exclusions are notably higher in Years 9 and 10 compared to other year groups, with a marked increase in Years 9–11 beginning in 2015/16. Given the low overall incidence of permanent exclusions and their concentration in these year groups, our analysis focuses on Years 8, 9 and 10. This targeted approach reduces dilution and allows for a more precise estimation of the causal impact of permanent exclusion in the population where it is most prevalent. Furthermore, since disruptive pupils removed in Year 9 or 10 had already exposed their peers to up to three or four years of disruptive behaviour, the duration of this initial exposure significantly outweighs the period of exposure to the benefits of their removal. This long-term exposure history justifies the validity of our measure of disruptive behaviour.

Following a permanent exclusion, a pupil's educational trajectory in England is largely shaped by placement into Alternative Provision (AP). AP refers to any educational setting for pupils who can no longer attend mainstream or special schools (DCSF, 2008). It encompasses a variety of arrangements for students who have been permanently excluded, lack a school place, refuse to attend school, or have medical needs. For our analysis, we focus on state-funded Pupil Referral Units (PRUs), AP Academies, and AP Free Schools, structured educational settings for pupils outside mainstream schools, whose enrolments are captured in the National Pupil Database (NPD).

[3] The Department for Education statistics show that persistent disruptive behaviour is the main reason for permanent exclusions in England, accounting for around 38–39% of cases in recent years (2023/24–2024/25). Other common reasons include physical assault against a pupil (around 16%) and physical assault against an adult (around 12–13%). The overall permanent exclusion rate was approximately 0.04% in autumn 2024/25. Source: DfE, 2024/25.

Figure 1. Permanent exclusions over time and by year group,
Average number of pupils excluded



Source: NPD Data

While AP is the primary destination for permanently excluded pupils, many enrolled students have not been excluded. Thomson (2019) finds that as of 2017, 45 percent of pupils in state-funded AP schools had a history of permanent exclusion; the remainder are placed due to school refusal, lack of school places, or medical reasons. Because fewer than half of the AP intake consists of formally excluded pupils, this diminishes the concern that the supply of AP settings is determined primarily by school-level disruptiveness.

AP supply over time

Department for Education statistics indicate that the number of state-funded AP schools has declined since 2010, even as pupil enrolment in these schools has increased since 2014 (Thomson, 2019). PRUs have historically been the most common form of AP, growing steadily through the early 2000s before declining after 2010. This decline reflects two main factors: (i) the 2012 Education Regulations Amendment, which allowed high-performing PRUs to convert into independent AP Academies, and (ii) the growing diversity of AP settings provided by local authorities since 2012 (House of Commons Education Committee, 2018).

Since 2012, some institutions have been permitted to operate as APs without official school registration, leading to the expansion of unregistered AP providers not routinely inspected by Ofsted. Although these pupils are recorded in the National Pupil Database (NPD), specific placement information is not tracked, restricting our analysis to mainstream, registered AP settings.

Evidence suggests that increased local AP availability may influence exclusion decisions. A report by the Institute of Education (IoE) and the National Foundation for Educational Research (NFER, 2014) indicates that before permanently excluding a pupil, schools increasingly assess the suitability of AP placements and, following exclusion, monitor progress post-placement. Statutory guidance further indicates that exclusion decisions involve joint consideration by schools and local authorities of the appropriateness, safety, and quality of available AP settings (DfE, 2025). Local commissioning frameworks require providers, including unregistered ones, to meet predefined standards, which must be satisfied before placements can proceed. These procedures imply that fluctuations in local AP filled capacity may affect schools' ability to exclude pupils.

Theoretical framework

We present a stylised theoretical framework to guide the empirical analysis of how the timing of removing disruptive students impact their non-disruptive peers' outcomes. Let $\alpha > 0$ denote the ability of students in the same school and cohort in the absence of disruptive behaviour at the beginning of secondary school. Suppose that disruptive pupils remain present until time τ, where τ indexes how late in schooling removal occurs. Under the assumption that disruptive behaviour continues over time, their continued presence reduces the effective peer learning environment by a cumulative damage function $D(\tau)$, with $D(0) = 0$, $D(\tau) \geq 0$, and $D'(\tau) > 0$. The effective peer ability at the time outcomes is therefore:

$$\tilde{\alpha}(\tau) = \alpha - D(\tau)$$

which declines with later removal.

Student i's outcome is denoted by $y_i(\tau)$ (e.g., test scores) and is produced according to:

$$y_i(\tau) = a_i + \beta e_i(\tau)\tilde{\alpha}(\tau) + \lambda\tilde{\alpha}(\tau)$$

where a_i a fixed parameter that captures the student's baseline ability, $e_i(\tau)$ represents the study effort chosen by student i, and $\tilde{\alpha}(\tau)$ is the average baseline ability of pupils in the same school and cohort excluding i. The parameter $\beta > 0$ measures how strongly effort translates into achievement, while $\lambda > 0$ captures a direct peer composition effect. In this framework, achievement depends both on own effort and on peer quality. The term $\beta e_i(\tau)\tilde{\alpha}(\tau)$ implies that effort is more productive in higher-ability classroom, for example because there are fewer disruptions, more academically engaged peers, and a smoother instructional process, so that each unit of effort yields a larger return when $\tilde{\alpha}(\tau)$ is higher. The term $\lambda\tilde{\alpha}(\tau)$ captures a direct peer effect: even holding effort fixed, a higher average peer ability raises achievement through mechanisms such as exposure to a faster instructional tempo.

Each period, student i chooses effort e_{it} to maximise a simple per-period utility function

$$U_{it} = y_i(\tau) - \frac{c}{2}e_i^2$$

where $c > 0$ governs the marginal disutility of effort. The quadratic (convex) cost of effort is standard in the education production function literature, capturing increasing marginal effort costs and ensuring an interior solution with well-behaved first-order conditions. Substitute the production function into utility:

$$\max_{e_i}\ a_i + \beta e_i\tilde{\alpha}(\tau) + \lambda\tilde{\alpha}(\tau) - \frac{c}{2}e_i^2.$$

Effort maximization yields:

$$e_i^*(\tau) = \frac{\beta}{c}\tilde{\alpha}(\tau)$$

Thus, optimal effort is increasing in the average ability of classmates: $\frac{\partial e_i^*(\tau)}{\partial \tilde{\alpha}} = \frac{\beta}{c} > 0$. When disruptive peers are removed, students endogenously choose to exert more effort, and earlier removal (smaller τ) leads to higher $\tilde{\alpha}(\tau)$, inducing higher effort.

Substituting $e_i^*(\tau)$ into the production function gives the equilibrium learning outcome:

$$y_i(\tau) = a_i + \frac{\beta^2}{c}\tilde{\alpha}(\tau)^2 + \lambda\tilde{\alpha}(\tau)$$

Differentiate with respect to τ:

$$\frac{dy_i(\tau)}{d\tau} = \left(2\frac{\beta^2}{c}\tilde{\alpha}(\tau) + \lambda\right)\tilde{\alpha}'(\tau) < 0$$

because $\tilde{\alpha}'(\tau) = -D'(\tau) < 0$. Thus, later removal reduces final achievement, while earlier removal improves it. Intuitively, disruptive peers depress learning and effort in every period they remain in the school, so removing them in early secondary school prevents multiple years of negative peer effects, whereas removing them only near the end of schooling translates into higher cumulative damage. The mechanism operates through (i) a direct peer-quality effect and (ii) an effort channel, since effort becomes more productive in a higher-quality classroom. The model therefore predicts that exposure to disruptive peers imposes damage that is cumulative and only partially reversible, making early removal more beneficial, and late removal potentially harmful relative to earlier removal.

Empirical strategy

Exclusions are unlikely to be randomly assigned. Schools make exclusion decisions based on a wide range of factors, including school climate, behavioural shocks, and prior performance, which raises concerns that observed exclusion rates are endogenous. The direction of bias in ordinary least squares (OLS) estimates is, however, ambiguous a priori. On the one hand, schools with higher exclusion rates are likely to face more severe underlying behavioural problems, which are themselves negatively correlated with peer outcomes. This would tend to bias OLS estimates downward, overstating the adverse effect of exclusions. On the other hand, exclusion removes disruptive pupils from the classroom, which may improve the learning environment for remaining students and bias OLS estimates upward. The net bias therefore depends on the relative strength of these opposing forces.

To address this challenge and identify the causal effect of exclusions on non-excluded peers, we implement an instrumental variable (IV) strategy.

Our instrument exploits supply-side variation in local Alternative Provision (AP) filled capacity. Specifically, we use the change in the number of places in each year group filled in AP settings within a 10 km radius of a mainstream school from the start of the current academic year to the start of the following academic year, excluding the focal school's own exclusions in that cohort. This leave-one-out measure captures shifts in AP availability driven by neighbouring schools' exclusion decisions, rather than by contemporaneous behavioural shocks within the focal school. In this sense, the instrument provides exogenous variation in a mainstream school's propensity to exclude pupils, consistent with established approaches that rely on local supply shocks to instrument endogenous policy choices (Angrist and Pischke, 2009).

The key identifying assumption is that AP filled capacity influences the outcomes of non-excluded pupils only through its effect on exclusion rates. Several features of the AP market lend credibility to this exclusion restriction. First, AP settings serve multiple mainstream schools within their catchment areas. In 2016, 348 AP schools served approximately 1,372 mainstream schools that issued exclusions, implying that each AP drew pupils from around four nearby mainstream institutions. Consequently, year-on-year changes in filled places at a given AP reflect shifts in aggregate behavioural conditions and exclusion practices in the wider local area, rather than unobserved factors specific to any single mainstream school. Because the instrument excludes the focal school's own exclusions, it is insulated from within-year behavioural shocks originating in that school.

Second, Alternative Provision (AP) settings serve pupils who are unable to attend mainstream schools. While these placements often follow permanent exclusion, only around 45% of AP pupils have been permanently excluded; others enter due to school refusal, insufficient school places, or medical needs. The non-excluded pupils whose outcomes we study remain enrolled in mainstream secondary schools throughout and do not attend AP settings, including PRUs, AP academies, or AP free schools. As a result, changes in local AP filled capacity cannot directly affect their educational experiences or outcomes. Any impact of AP availability on non-excluded pupils must therefore operate indirectly, by altering schools' ability to remove persistently disruptive peers from mainstream classrooms. Taken together, this institutional separation alleviates concerns that variation in AP filled capacity is mechanically correlated with unobserved, school-specific determinants of non-excluded pupils' outcomes. By exploiting plausibly exogenous variation in the local supply of AP placements, the instrument strengthens the credibility of our causal interpretation.

The identifying variation compares cohorts in which disruptive pupils are excluded at a given point in secondary school because sufficient alternative provision capacity exists, to otherwise similar cohorts in which exclusion is delayed due to capacity constraints. These estimates therefore identify a local average treatment effect for pupils whose exclusion status responds to variation in local AP capacity. These "complier" pupils are excluded when AP placements are available but remain in mainstream school when capacity constraints bind. The counterfactual is therefore not the absence of disruptive behaviour, but the continued presence of the same marginally disruptive pupil in the classroom when AP capacity prevents exclusion.

Data

This paper uses data from the Longitudinal Educational Outcomes (LEO), a uniquely rich linked administrative dataset of (nearly) all young people in England, sourced from the records of multiple government departments. The NPD component of LEO provides detailed records of all pupils in state schools across England, tracking them from their early years through to post-16 education. It includes data on academic attainment, pupil backgrounds, absences, and exclusions, enabling us to determine if a pupil has experienced fixed-term or permanent exclusions during year 11. More specifically, it provides information on pupils' school, gender, age, ethnicity, whether English is an additional language (EAL), Special Educational Needs or disabilities (SEN) and entitlement to Free School Meals (FSM). It records whether the child has been "looked after" in state care (CLA), whether "in need" in some other way (CiN) and whether the child has an Education, Health and Care (EHC) plan.

It also provides a measure of local deprivation relevant to children (IDACI - income deprivation affecting children index).
Our analysis focuses on four full cohorts of English maintained secondary school pupils: those completing KS4 in the academic years 2014-2017.[4]

Our estimation sample comprises 2,035,656 observations. Table 1 reports summary statistics for pupil characteristics by whether the school's exclusion rate across Years 8 and 10 is above or below the median (where the median is zero). Schools with positive exclusion rates exclude, on average, about 2.5 student every one-thousand student in Year 8, 3.9 in Year 9; and 4.8 in Year 10 during the sample period. Across schools and cohorts, the leave-one-out change in the number of filled AP places within a 10 km radius ranges from 0 to 0.15 on average. The mean change is slightly larger for schools with zero exclusions than for those with at least one exclusion. The distribution includes both positive and negative values, reflecting year-to-year increases and decreases in local AP filled capacity.

Pupil characteristics vary systematically with exclusion exposure. Schools with above-median exclusion rates have a disproportionate share of pupils who are eligible for free school meals (FSM), of Black ethnic origin, or identified as having special educational needs (SEN). Pupils in these schools also have lower average attainment at the end of primary school, as measured by Key Stage 2 (KS2) test scores, and are more likely to live in urban and more deprived areas, as measured by the Income Deprivation Affecting Children Index (IDACI).

Outcomes also differ across these groups. Pupils in schools with above-median exclusion rates have lower average Key Stage 2 (KS4) scores. Looking beyond secondary school at outcomes by age 21, pupils exposed to positive exclusions are less likely to complete Level 2 and Level 3 qualifications by age 21, as well as less likely to be employed, more likely to not be in education, employment, or training (NEET).

[4] The number of unique schools is 2,609. The figure reported in this table is larger because some schools are counted more than once, depending on whether they experienced an exclusion in a given year.

Table 1 – Descriptive statistics by positive/zero exclusions

	Mean (sd) No exclusion	Mean (sd) Positive exclusions
Exclusion per 1k in Y8	2.509	0
	(4.035)	
Exclusion per 1 in Y9	3.934	0
	(4.904)	
Exclusion per 1 in Y10	4.827	0
	(5.267)	
Change AP filled cap Y8 over acad year, within 10km, leave school out	0.017	0.045
	(0.297)	(0.286)
Change AP filled cap Y9 over acad year, within 10km, leave school out	0.075	0.103
	(0.506)	(0.499)
Change AP filled cap Y10 over acad year, within 10km, leave school out	0.123	0.158
	(0.72)	(0.715)
<u>Controls</u>		
Female	0.489	0.515
Ever FSM	0.407	0.321
Standardised KS2 math scores	-0.067	0.08
	(0.913)	(0.918)
Standardised KS2 English scores	-0.082	0.074
	(0.93)	(0.924)
SEN	0.465	0.4
EAL	0.167	0.151
White British ethnicity	0.714	0.734
White Other ethnicity	0.048	0.042
Asian ethnicity	0.1	0.11
Black ethnicity	0.07	0.052
Mixed ethnicity	0.044	0.041
Unauthorised absences at age 11	0.009	0.007
IDACI score	0.248	0.208
	(0.178)	(0.17)
Urban	0.883	0.83
<u>Outcomes</u>		
Standardised KS4 math scores	-0.059	0.138
	(1.019)	(0.999)
Standardised KS4 English scores	-0.033	0.15
	(0.945)	(0.929)
A-level or more at age 21	0.579	0.656
Level 2 qual at age 21	0.805	0.853
Employed at age 21	0.779	0.812
NEET at age 21	0.189	0.155
Months as NEET age 21	6.173	4.763
Months in employment age 21	8.823	9.156
Number of schools	1,842	2,121
Number of pupils	922,912	1,112,744

Note: Data from National Pupil Database for students in comprehensive (non-selective) state schools. Descriptives for children in care had to be suppressed for statistical disclosure control.

Estimating the Effect of Exclusion on peers' outcome

The aim of our empirical work is to estimate the effect of the removal of disruptive peers on non-excluded peers' outcomes. We estimate the following specification,

$$y_{is} = \beta_0 + \beta_k r_s^k + \beta_X^{'} X_{is} + \sigma_s + \tau_t + u_{ist} \quad (1)$$

where y_{is} represents a measure of academic achievement and labour market outcome for student i in school s. Our outcomes cover both educational and labour market dimensions. For education, we examine academic achievement, measured by standardised Maths and English Key Stage 4 test scores at age 16. To capture longer-term effects, we also consider outcomes by age 21, specifically the likelihood of completing Level 2 qualifications and Level 3 qualifications. A Level 2 qualification is defined as having achieved five or more GCSEs at grades A*–C (or 9–4 under the reformed grading system) or equivalent vocational qualifications, while a Level 3 qualification is defined as attaining 2 or more A-levels, a BTEC National Diploma, or an equivalent post-16 academic or vocational qualification by age 21. We also examine longer-run labour market outcomes measured by age 21, including whether the individual is in paid employment and whether she is classified as not in education, employment, or training (NEET).[5]

r_s^k denote the percentage of pupils in school s, who are permanently excluded in Year k, with $k \in \{8,9,10\}$. Thus, the timing-specific exclusion rates r_{st}^k map directly into the model's $\tilde{\alpha}(\tau)$. The main parameters of interest, β_k represent the effect of increasing exclusion in a given year group by one pupil every 1,000. We scale the exclusion treatment to correspond to a 0.1 percentage point (pp) increase in the exclusion rate, equivalent to one additional excluded pupil per 1,000 peers. This unit reflects realistic policy-relevant variation and facilitates interpretation of effect sizes. In our data, the average cohort size is below 200 pupils, so a single exclusion corresponds to an increase of approximately 0.5 pp in the cohort exclusion rate. Nationally, the permanent exclusion rate is around 0.1 pp, implying that a 0.1 pp increase represents a doubling of the national baseline. We therefore interpret our estimates as the effect of modest but substantively meaningful changes in exclusion intensity, well within the range of observed variation across schools and cohorts.

X_{is} is a vector of controls measured at the pupil level. These characteristics include prior attainment (KS2 scores in English and maths); ethnicity dummies; whether English is an additional language; free school meal eligibility; special educational needs; whether ever looked after by the state, whether ever classified as in need; gender; and unauthorised absences at the age 11 years. We also control for area deprivation (IDACI) measured at the Lower-layer Super Output Area (LSOA) level (neighbourhood) using decile dummies of this measure, and for whether the LSOA is an urban or rural area. σ_s and τ_t are a set of school level and cohort level fixed effects. To address potential dependence among observations, standard errors are clustered at the school-cohort level. This approach accounts for correlations among students who share the same peer environment, which is the natural unit of exposure to excluded peers' disruptive behaviour.

The control variables in Eq. (1) account for certain observable variables that are correlated with exclusions. However, estimates of β_k will remain biased if unobservable time-varying, within-school characteristics are correlated with exclusions and also with y_{is}. This is the motivation for our IV strategy, that addresses the endogeneity of exclusion by using changes

[5] 2019/20 is the most recent year for which Higher Education Student Data (HESA) are available. As a result, university enrolment can only be observed for a single cohort, and we therefore do not analyse higher education outcomes.

in AP filled capacity as a source of exogenous variation in the probability that a school excludes a pupil. The first-stage equations relate the timing-specific exclusion rates to the instrument:

$$r_s^k = \delta_1^k \Delta A_{-s,10km}^k + \delta_2^k X_{is} + \sigma_s + \tau_t + e_{is}^k, \quad k \in \{8,9,10\}. \ (2)$$

Here, r_s^k denotes the fraction of pupils in school s, who are permanently excluded in Year k. The instrument $\Delta A_{-s,10km}^k$ measures the year group specific change in the number of AP places filled within a 10-km radius from the start of the current academic year to the start of the following academic year, excluding the focal school's own exclusions, normalised by the number of pupils in the same age group within the 10 km radius. X_{is} is the vector of student-level controls, and σ_s and τ_t denote school and time fixed effects. Standard errors are clustered at the school-cohort level.

A key empirical choice concerns whether to estimate the effects of removals in Year 8, Year 9 and Year 10 jointly within a single regression or separately across three timing-specific regressions. When we instrument the three exclusion fractions simultaneously, the Kleibergen–Paap statistic falls to around 9, indicating a weaker first stage once the instrument must explain variation in all three timing margins at the same time. In contrast, when each exclusion year is estimated in its own regression, the instrument is substantially stronger. This pattern is expected: the same AP filled capacity shock induces variation in each timing margin, but becomes weaker once the three endogenous variables are partialled out against one another. For this reason, our preferred specifications are the timing-specific IV regressions, which are more strongly identified and offer cleaner causal estimates of the effects of early, mid-, and late-year removals. We also report the joint specification for completeness, as it allows for within-model comparisons across timing categories; however, those results are interpreted cautiously and accompanied by weak-instrument-robust inference.

Results

Fixed effects estimates

This section presents the relationship between pupil exclusions and a range of non-excluded peers' educational and labour market outcomes. We begin by reporting fixed effects estimates of the association between the fraction of exclusions in Year groups 8 to 10 and attainment KS4 outcomes and educational attainment, as well as labour market outcomes by age 21, before turning to our IV strategy to address potential endogeneity concerns.

The OLS estimates in Table 2 provide consistent evidence that higher exclusion rates within a cohort are associated with poorer educational and early-adulthood outcomes for non-excluded peers. Although the magnitude of the coefficients is small, the pattern across outcomes and year groups is systematic. For academic outcomes, exclusions in Year 10 show the most pronounced associations. A one-per-thousand increase, which is double the national average observed rate of exclusion in Year 10, in the Year 10 exclusion rate is linked to a significant decline in KS4 Mathematics scores, significantly lower probabilities of achieving Level 2 and Level 3 qualifications by age 21. For post-16 destinations, higher Year 10 exclusions reduce the probability of being employed at age 21 and increase the likelihood of being NEET.[6] Across outcomes, the emerging pattern is that exclusions occurring later in secondary school,

[6] The last column in Table 2 shows the total number of observations for each outcome. For the labour market outcomes by age 21, the number of observations is lower due to sample attrition over time.

particularly in Year 10, are most consequential for peers, consistent with the idea that prolonged disruptive behaviour imposes greater costs on classmates.

Table 2 – Fixed effects estimate of exclusions on educational and labour market outcomes

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls	School FE	Cohort FE	N
KS4 Maths	-0.0009**	-0.0003	-0.0010***	Yes	Yes	Yes	2,035,656
	[0.000]	[0.000]	[0.000]				
KS4 English	-0.0009*	-0.0008**	-0.0015***	Yes	Yes	Yes	2,035,656
	[0.000]	[0.000]	[0.000]				
Level 2	-0.0004**	-0.0003**	-0.0006***	Yes	Yes	Yes	2,035,656
	[0.000]	[0.000]	[0.000]				
Level 3	-0.0003+	-0.0003**	-0.0003**	Yes	Yes	Yes	2,035,656
	[0.000]	[0.000]	[0.000]				
Employed at 21	-0.0001	-0.0002	-0.0007***	Yes	Yes	Yes	1,630,561
	[0.000]	[0.000]	[0.000]				
NEET at 21	-0.0000	0.0000	0.0006***	Yes	Yes	Yes	1,630,561
	[0.000]	[0.000]	[0.000]				

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Standard errors in brackets, clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

First stage results

The first-stage results in Table 3 show that changes in locally available AP filled capacity in the same academic year, leaving out school own exclusions, are a strong and statistically significant predictor of a school's exclusion rate in each year group. We show results using normalised changes in AP filled capacity within each year group within a 10km and a 15 km radius. Across all specifications, the coefficients on the AP filled capacity instrument are negative and highly significant. This implies that when neighbouring schools fill more AP places, thereby tightening local AP capacity, the focal school excludes fewer pupils. The pattern is consistent with the institutional environment: when AP placements become scarce, schools face stronger constraints on their ability to transfer disruptive pupils and correspondingly reduce their use of permanent exclusion.

In our preferred 10 km specification, an additional AP placement filled in nearby provision is associated with a reduction of 0.95 excluded pupils per 1,000 in Year 8, 0.76 pupils per 1,000 in Year 9, and 0.42 pupils per 1,000 in Year 10 (all $p < 0.001$). Because the outcome is measured as exclusions per 1,000 pupils, these coefficients can be interpreted directly as marginal changes in exclusion intensity induced by shifts in local AP utilisation. Relative to mean exclusion rates of approximately 1.1 per 1,000 in Year 8, 1.78 per 1,000 in Year 9, and around 2.2 per 1,000 in Year 10, these estimates are equivalent to roughly 86%, 43%, and 19% of the corresponding mean exclusion rates, respectively. The magnitudes are therefore economically meaningful, particularly in earlier secondary years, and indicate that local AP capacity operates as a binding institutional margin shaping schools' disciplinary decisions.
Although the number of permanent exclusions is highest in Year 10, the instrumental variation generated by changes in local AP capacity primarily affects earlier disciplinary margins. By Year 10, many exclusions involve severe cases that would likely occur regardless of AP availability, limiting the extent to which AP capacity shocks alter exclusion decisions. As a result, the estimates are more strongly identified for exclusions occurring in Year 8 and 9, where a larger share of exclusions appears to lie close to the margin influenced by AP capacity.

Instrument strength is high: first-stage F-statistics exceed the conventional weak-instrument threshold of 10 by a wide margin (F = 67.97 for Year 8, 31.90 for Year 9, and 49.35 for Year 10). We adopt the 10 km radius as our preferred specification because it most plausibly captures the relevant local placement market; the 15 km results serve as a robustness check.

Table 3 – First stage: the effect of change in AP filled spaces on exclusions

	(1)	(2)	(3)	(4)	(5)	(6)
	Exclusion per 1k in Y8		Exclusion per 1k in Y9		Exclusion per 1k in Y10	
Yearly change AP Y8, 10km	-0.9500***					
	[0.115]					
Yearly change AP Y8, 15km		-0.7625***				
		[0.119]				
Yearly change AP Y9, 10km			-0.4236***			
			[0.075]			
Yearly change AP Y9, 15km				-0.2444**		
				[0.085]		
Yearly change AP Y10, 10km					-0.4169***	
					[0.059]	
Yearly change AP Y10, 15km						-0.2568***
						[0.061]
Constant	1.1681***	1.1648***	1.8031***	1.7912***	2.2283***	2.2084***
	[0.028]	[0.028]	[0.036]	[0.036]	[0.039]	[0.039]
Mean outcome	1.137	1.137	1.784	1.784	2.189	2.189
Observations	2,035,656	2,035,656	2,035,656	2,035,656	2,035,656	2,035,656
F-test	67.97	41.28	31.90	8.252	49.35	17.63
p-value	0.000	0.000	0.000	0.004	0.000	0.000

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

To further probe the mechanism through which our instrument operates, we examine whether local AP filled capacity affects the time it takes for an excluded pupil to be placed into an AP setting. For each school–cohort–year-group cell, we compute the average number of days between a pupil's exclusion and their first recorded AP placement, and regress this waiting time on our normalised AP-filled-capacity-change measure within a 10 km radius, including school, cohort, and year-group fixed effects. Table 4 reports the results. The estimates reveal clear year-group heterogeneity consistent with institutional practice. For Year 8, increases in filled capacity significantly reduce waiting times. For Year 9, the effect is near zero and statistically insignificant. In contrast, for Year 10, where the number of exclusions is the highest and supply constraints are therefore most binding, increased filled capacity significantly increases delays to AP placement, as there are less places available for excluded pupils. A one-unit increase in the normalised Year-10 filled capacity measure raises waiting times by 6.3 days ($p < 0.05$). Given that the mean waiting time in our sample is 81 days, the Year-10 effect is meaningful: a one-unit increase in filled capacity corresponds to roughly an

8% increase in the average delay to placement. This pattern strongly supports the interpretation that AP supply constraints bind more tightly in Year 10 than in earlier year groups, amplifying the effect of local capacity fluctuations on placement speed. This result is consistent with the interpretation that our filled capacity measure reflects the availability of AP slots in the local market: as capacity becomes more fully utilized, excluded pupils in Y10 must wait longer to be placed, suggesting that AP provision in Y10 is operating under binding capacity constraints.

Table 4 – Waiting time for AP Placement and filled AP spaces

	Avg number of days between exclusions to AP placement		
Yearly change AP Y8, 10km	-48.753*		
	[22.917]		
Yearly change AP Y9, 10km		0.173	
		[3.953]	
Yearly change AP Y10, 10km			6.306*
			[2.693]
Constant	120.8774***	68.6529***	38.9558***
	[5.089]	[1.829]	[0.876]
Mean outcome	81.14	81.14	81.14
Observations	1,240	2,394	3,399

Notes: *** p<0.01, ** p<0.05, * p<0.1. School fixed effect linear regression estimated. Standard errors in brackets clustered at school and cohort level. Controls: year-group dummies, cohort dummies.

IV estimates

Table 5 reports 2SLS estimates of the effect of exclusions in Years 8–10 on peers' educational and early-adulthood outcomes, using changes in local AP filled capacity as an instrument. The Kleibergen–Paap Wald F-statistics are comfortably above conventional weak-instrument thresholds (all above 30, and often around 50–70), indicating that the instrument is strongly correlated with cohort-level exclusion rates and that the IV estimates are not driven by weak identification.

Relative to the OLS results in Table 2, the IV coefficients are generally larger in magnitude, especially for exclusions in Year 9, suggesting that OLS estimates were attenuated by endogenous selection into exclusion. For educational outcomes, the IV estimates reinforce the pattern that mid- and late-secondary exclusions are particularly harmful. A one-per-thousand increase in the Year 8 exclusion reduces KS4 Mathematics scores by 0.01 standard deviation (SD), while a one-per-thousand increase in Year 9 exclusion reduces KS4 Maths by 0.024 SD and KS4 English by 0.044 standard deviations, compared with much smaller OLS associations. Year 10 exclusion decreases English attainment by 0.015 SD, though the effects on Maths become imprecisely estimated.

To facilitate comparison with the peer-effects literature, it is useful to interpret the one-per-thousand increase in the Year 9 exclusion rate as a marginal change in peer composition at the cohort level. In a typical secondary school cohort of around 200 pupils per grade, a one-per-thousand increase corresponds to approximately one additional excluded pupil per five cohorts, or equivalently a small increase in the probability that any given student is exposed to a highly disruptive peer during the school year. While this exposure is more diffuse than the classroom-level shocks studied in much of the literature, the resulting achievement effects are remarkably similar in magnitude.

For example, Carrell and Hoekstra (2010) show that adding one disruptive peer to a classroom of 20 students reduces test scores by around 0.07 standard deviations, while Carrell et al. (2018) find that exposure to one additional disruptive peer in a class of 25 lowers achievement by approximately 0.02 standard deviations. These studies identify disruptive peers using indicators such as exposure to domestic violence in the household, which capture children at higher risk of behavioural problems but who typically remain in the classroom. By contrast, our measure focuses on students who are actually excluded from school, a group likely to represent the most severe cases of disruptive behaviour. At the same time, our estimates reflect a more diffuse cohort-level exposure across multiple classrooms, rather than direct classroom interaction.

Interpreting the similarity in magnitudes therefore requires accounting for the structure of peer interactions in English secondary schools, where pupils rotate across subjects and interact with a broad set of peers over the school day and academic year. In this setting, exposure to a disruptive or excluded peer is not confined to a single classroom but is instead diffuse and repeated across multiple learning environments, with heterogeneous intensity across pupils. As a result, relatively small changes in the exclusion rate can affect a large share of the cohort through cumulative low-intensity interactions, rather than concentrated classroom-level shocks. Consistent with this interpretation, studies examining larger shifts in the composition of disadvantaged peers at the grade level, such as Lavy et al. (2012) and Lavy and Yancu (2025), document achievement losses of around 0.05–0.11 standard deviations associated with substantial increases in the share of very low-achieving or violent peers in Israel. Taken together, this evidence suggests that the spillovers we identify, 0.024 standard deviations in Mathematics and 0.044 standard deviations in English, are quantitatively meaningful and broadly aligned with existing estimates once differences in institutional context and the structure of peer interactions are taken into account.

The IV estimates point to economically meaningful, though moderate, peer effects on qualification outcomes. A one-per-thousand increase in the exclusion rate in Year 9 reduces the probability of achieving a Level 2 qualification by 0.65 percentage points, corresponding to around 0.8 per cent of the baseline mean (0.83), and lowers the likelihood of attaining a Level 3 qualification by 0.63 percentage points. Exclusions occurring in Year 10 are also associated with reductions in Level 2 completion (0.28 percentage points), although the corresponding effects on Level 3 attainment are smaller and imprecisely estimated. By contrast, exclusions in Year 8 generate weaker and less precisely estimated impacts, suggesting that peer spillovers on attainment are most pronounced when exclusions occur closer to key educational examinations.

Turning to post-16 outcomes, the IV results broadly corroborate the OLS pattern that higher exclusion rates in the later years of secondary school shift peers towards poorer labour-market trajectories. A one-per-thousand increase in the Year 9 exclusion rate leads to the largest deterioration in outcomes, raising the probability of being NEET at age 21 by 0.62 percentage points, an increase of approximately 3.6 per cent relative to the sample mean (0.171). Exclusions in Year 10 are associated with a decline in the probability of employment at age 21 of 0.36 percentage points (around 0.5 per cent of the mean), consistent with weaker but still adverse effects on post-school labour-market attachment.

The comparatively limited and less precisely estimated peer effects for Year 10 exclusions should not be interpreted as evidence that disruption at this stage is benign. Rather, they are consistent with a setting in which much of the cumulative peer harm has already materialised by the final year of compulsory education. AP-induced variation in exclusion timing during Year 10 affects only a short remaining period before GCSE completion, potentially leaving insufficient scope for additional disruption, or recovery, to translate into detectable differences in post-school outcomes.

Table 5 – Two Stage Least Square estimates of exclusions

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N	Mean outcome
KS4 Maths						
Coeff	-0.0100**	-0.0244***	-0.0038	Yes	2,035,656	0.103
[se]	[0.004]	[0.006]	[0.003]			
Kleibergen-Paap Wald F stat	67.97	31.90	49.35			
KS4 English						
Coeff	0.0016	-0.0436***	-0.0150**	Yes	2,035,656	0.130
[se]	[0.004]	[0.010]	[0.005]			
Kleibergen-Paap Wald F stat	67.97	31.9	49.35			
Level 2						
Coeff	-0.0024+	-0.0065**	-0.0028*	Yes	2,035,656	0.831
[se]	[0.001]	[0.002]	[0.001]			
Kleibergen-Paap Wald F stat	31.90	49.35	71.38			
Level 3						
Coeff	0.0011	-0.0063**	-0.0007	Yes	2,035,656	0.621
[se]	[0.002]	[0.002]	[0.001]			
Kleibergen-Paap Wald F stat	67.97	31.90	49.35			
Employed at 21 - coeff						
Coeff	0.0025	-0.0034	-0.0036*	Yes	1,630,561	0.797
[se]	[0.002]	[0.002]	[0.002]			
Kleibergen-Paap Wald F stat	71.38	30.49	44.40			
NEET at 21						
Coeff	-0.0021	0.0062**	0.0025	Yes	1,630,561	0.171
[se]	[0.001]	[0.002]	[0.002]			
Kleibergen-Paap Wald F stat	71.38	30.49	44.40			

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

The IV estimates should also be interpreted as local average treatment effects for pupils whose exclusion status responds to variation in local AP capacity. These "marginal" cases are excluded when alternative provision places are available but remain in mainstream school when capacity constraints bind. Importantly, such pupils typically accumulate substantial behavioural incidents prior to exclusion, as reflected in the large number of suspension days observed in the data. The comparison therefore occurs between cohorts exposed to prolonged disruption where the pupil is eventually excluded and similar cohorts where the pupil remains in school because exclusion is constrained. In this setting, much of the peer harm may occur during the extended period of disruptive behaviour prior to exclusion.

A potential concern when comparing IV estimates across year groups is that differences in magnitudes may reflect variation in first-stage strength rather than genuine differences in causal effects, given that the IV estimator is the ratio of the reduced form to the first stage. In our setting, however, the first stage is strong across all year groups, with F-statistics well above conventional thresholds. Moreover, first-stage coefficients are largest in earlier years, while the largest IV effects arise in Year 9. This pattern is inconsistent with differences in first-stage strength driving the results mechanically. Instead, it suggests that the observed variation across year groups primarily reflects differences in the underlying reduced-form relationships, consistent with a timing-based interpretation in which the cumulative exposure to disruptive peers plays a central role.

In the Appendix, we report results for two intensive-margin labour-market outcomes: total months in employment by age 21 and total months spent NEET by age 21. Table A1 shows that exclusions in Year 9 are associated with a substantial increase in time spent NEET, around 0.26 additional months by age 21, which corresponds to approximately 4 per cent of the sample mean of 6.1 months. Exclusions in Year 10 reduce employment by roughly two weeks (about 0.55 per cent of the mean of nine months), consistent with a deterioration in labour-market attachment along the intensive margin. By contrast, exclusions in Year 8 are associated with an increase in months as NEET (0.13 additional months) and a small increase in months employed by age 21 (around 0.08 months), significant but very small in magnitude.

These patterns are already evident at earlier ages, though with smaller magnitudes. Results reported in Table A2 show that a one-per-thousand increase in the Year 8 exclusion rate is associated with a 0.5 percentage point increase in employment by age 19, alongside an increase in months of continuous employment. In contrast, Year 9 exclusions raise months spent NEET by age 19 by around 0.109 months. While the effect is very small in magnitude, it points towards the view that peer spillovers from exclusions may intensify as pupils approach the end of compulsory schooling.

We empirically assess heterogeneity of impacts by conducting heterogeneity analysis, examining differences in outcomes by socio-economic status, and prior ability in Table 6. The results reveal substantial heterogeneity in the spillover effects of peer exclusions across pupil subgroups and across year groups, with both the magnitude and the distributional incidence of effects varying sharply with the timing of exclusion.

In Year 8 (Panel A), exclusions occurring early in secondary school generate similar effects for disadvantaged pupils. FSM and low-ability pupils display a broadly similar pattern: peer exclusions in Year 8 are associated with improved labour-market outcomes (higher employment probabilities and lower NEET rates).

In Year 9 (Panel B), peer effects become more clearly differentiated across groups. FSM pupils show improvements in educational attainment, but weaker labour-market outcomes, with lower employment and higher NEET risk, suggesting that improvements in the classroom environment may come at the cost of less stable post-school pathways. In contrast, low-ability pupils experience large and pervasive negative effects across attainment, qualification completion, and labour-market outcomes, indicating that mid-secondary disruption is particularly harmful for those with weaker academic foundations.

In Year 10 (Panel C), FSM pupils experience negative effects across nearly all outcomes, including KS4 attainment, qualification completion, and labour-market attachment. Low-ability pupils show a different pattern: modest gains in KS4 attainment alongside weaker labour-market outcomes. This suggests that late removal of disruptive peers may improve exam performance at a critical stage, but these gains do not translate into sustained employment advantages. Overall, Year 10 exclusions are associated with broad penalties for disadvantaged pupils, while low-ability pupils face a trade-off between short-run attainment gains and weaker labour-market outcomes, highlighting the importance of timing for both the magnitude and direction of effects.

Table 6 – Heterogeneous analysis

	KS4 math	KS Eng	Lev 2	Lev 3	Emp 21	Neet 21
Panel A: year 8						
Exclusion per 1K in Y8	-0.0084*	0.0070	-0.0041**	0.0002	-0.0089***	0.0110***
	[0.004]	[0.004]	[0.002]	[0.002]	[0.002]	[0.002]
FSM x Excl Y8	-0.0047	-0.0154*	0.0049	0.0023	0.0301***	-0.0346***
	[0.006]	[0.007]	[0.003]	[0.003]	[0.006]	[0.006]
Exclusion per 1K in Y8	-0.0091+	0.0049	-0.0079***	0.0004	-0.0174***	0.0178***
	[0.005]	[0.006]	[0.002]	[0.003]	[0.004]	[0.004]
Low KS2 X Excl Y8	-0.0014	-0.0056	0.0093*	0.0011	0.0326***	-0.0327***
	[0.007]	[0.007]	[0.004]	[0.003]	[0.006]	[0.006]
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561
Panel B: year 9						
Exclusion per 1K in Y9	-0.0478***	-0.0578***	-0.0108***	-0.0140***	0.0021	-0.0017
	[0.008]	[0.010]	[0.002]	[0.003]	[0.002]	[0.002]
FSM x Excl Y9	0.0811***	0.0496***	0.0149**	0.0266***	-0.0178**	0.0257***
	[0.018]	[0.012]	[0.005]	[0.007]	[0.006]	[0.007]
Exclusion per 1K in Y9	0.0545***	0.0169	0.0085**	0.0095**	0.0147***	-0.0125***
	[0.012]	[0.013]	[0.003]	[0.00]	[0.004]	[0.003]
Low KS2 X Excl Y9	-0.1466***	-0.1122***	-0.0277***	-0.0293***	-0.0320***	0.0332***
	[0.026]	[0.022]	[0.006]	[0.006]	[0.006]	[0.006]
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561
Panel C: year 10						
Exclusion per 1K in Y10	0.0109*	-0.0053	0.0040*	0.0002	0.0116***	-0.0170***
	[0.004]	[0.007]	[0.002]	[0.002]	[0.003]	[0.004]
FSM x Excl Y10	-0.0386***	-0.0255**	-0.0180***	-0.0025	-0.0368***	0.0472***
	[0.008]	[0.008]	[0.004]	[0.003]	[0.007]	[0.009]
Exclusion per 1K in Y10	-0.0535***	-0.0563***	-0.0005	-0.0103***	0.0117***	-0.0158***
	[0.010]	[0.010]	[0.002]	[0.003]	[0.003]	[0.004]
Low KS2 X Excl Y10	0.0874***	0.0726***	-0.0040	0.0169***	-0.0254***	0.0303***
	[0.016]	[0.013]	[0.003]	[0.004]	[0.005]	[0.005]
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Taken together, the three panels reveal a clear timing gradient in the effects of peer exclusions. In Year 8, early exclusions primarily improve labour-market attachment for disadvantaged and low-ability pupils, suggesting gains concentrated along non-academic margins. In Year 9, effects diverge: low-ability pupils experience large and pervasive losses across outcomes, while FSM pupils show attainment gains that do not translate into improved labour-market trajectories. By Year 10, exclusions depress both attainment and employment outcomes for disadvantaged pupils, while low-ability pupils exhibit improved KS4 performance alongside weaker labour-market attachment. Overall, the timing of exclusion shapes both the magnitude and direction of peer effects, generating distinct trade-offs between academic performance and longer-run progression across groups.[7]

To investigate the mechanisms through which exposure to excluded peers affects subsequent educational and labour market outcomes, we exploit heterogeneity in the severity of behavioural problems among excluded pupils. Specifically, we proxy peer disruptiveness using the cumulative number of suspension days accrued by each pupil prior to exclusion. This measure captures repeated and serious behavioural incidents that warranted temporary removal from the classroom and therefore reflects persistent disruptive behaviour rather than isolated infractions. This measure also provides direct evidence that many exclusions occur only after prolonged behavioural escalation, allowing us to test whether the negative peer effects documented above are driven by the cumulative disruption experienced prior to exclusion.

One potential concern is that pupils excluded in different years may differ systematically in the severity or trajectory of their behavioural problems. Our empirical strategy mitigates this concern by exploiting variation in exclusion decisions generated by changes in local AP capacity rather than by differences in pupil behaviour. Moreover, the suspension-based mechanism analysis shows that the magnitude of peer spillovers increases with the cumulative number of suspension days prior to exclusion. This pattern indicates that the negative effects are driven primarily by prolonged exposure to disruptive behaviour rather than by differences in the types of pupils excluded.

We interact this measure with the fraction of excluded pupils in the cohort. The results, reported in Table 7, show that this interaction term explains a non-trivial share of the adverse peer spillovers associated with exclusions in Year 9. Evaluated at the mean level of suspension intensity (11 days), the interaction accounts for approximately 18–26% of the overall negative effect of Year-9 exclusions across the outcomes reported in the first four columns. For sufficiently high values of the normalised suspension-days measure (around 70 or above, corresponding to the 95th percentile of the distribution), the interaction fully offsets the baseline negative exclusion effect and may even turn it positive. This pattern suggests that excluding pupils with the most severe and persistent behavioural problems can generate modest positive spillovers for their peers. Results for Year-10 exclusions are not interpreted. Estimates are imprecise and the associated first-stage diagnostics, including the F-test, indicate weak instrument strength, limiting the reliability of inference in this case.

[7] We report in the Appendix results for females (Table A3). They show that female pupils benefit from exclusions in Years 8 and 9 across most outcomes, including educational progression and labour-market attachment, with the exception of maths attainment. However, in Year 10, they experience declines in KS4 attainment, suggesting that peer disruption in the final pre-GCSE year may be particularly detrimental to their academic performance.

Table 7 – 2 Stage Least Square estimates of exclusion, interacted with days suspended

	KS4 math	KS4 Eng	Lev 2	Lev 3	Emp at 21	Neet at 21
Exclusion per 1K in Y8	-0.00524	0.00757	-0.00090	0.00271	0.00117	0.00008
	[0.005]	[0.006]	[0.002]	[0.002]	[0.002]	[0.002]
Excl Y8*Days suspended	-0.00004	-0.00005+	-0.00001	-0.00001	0.00001	-0.00002+
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen-Paap Wald F stat	13.92	13.92	35.73	35.73	38.54	38.54
Avg days suspended in Y8			6.092			
Exclusion per 1K in Y9	-0.04665***	-0.06580***	-0.00853**	-0.01018**	-0.00033	0.00461
	[0.011]	[0.015]	[0.003]	[0.003]	[0.003]	[0.003]
Excl Y8*Days suspended	0.00011*	0.00011**	0.00001	0.00002*	-0.00002	0.00001
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen-Paap Wald F stat	14.62	14.62	14.02	14.02	16.91	16.91
Coeff int*Days suspended	0.0259	0.0184		0.0217		
Avg days suspended in Y9			11.007			
Exclusion per 1K in Y10	0.00196	-0.01537*	0.00033	0.00148	0.00322	-0.00589
	[0.005]	[0.007]	[0.002]	[0.002]	[0.004]	[0.004]
Excl Y10*Days suspended	-0.00002	0.00000	-0.00001	-0.00001	-0.00002+	0.00003+
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen-Paap Wald F stat	4.315	4.315	1.735	1.735	4.142	4.142
Avg days suspended in Y10			13.547			
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

These findings indicate that a substantial share of the detrimental impact of exposure to excluded peers operates through the degree of disruptiveness of those peers, rather than through exclusion per se. In other words, the negative effects we document are largely driven by exposure to highly disruptive classmates, highlighting a behavioural channel through which peer spillovers from exclusion materialise.

To further shed light on mechanisms, we disaggregate exclusions by their main recorded reasons: persistent disruptive behaviour, physical assault, and verbal abuse, and re-estimate the models by year group (Appendix Table A4). The results show that negative peer effects are driven primarily by exclusions related to persistent disruption and physical assault, with the strongest and most consistent impacts arising in Year 9. By contrast, exclusions for verbal abuse are associated with smaller and less consistent effects, suggesting that more severe and sustained behavioural problems account for the bulk of observed spillovers. Effects in Years 8 and 10 are generally weaker and less stable, particularly where first-stage strength is limited.

Overall, these findings support an interpretation in which peer effects operate through exposure to disruptive behaviour rather than exclusion status per se.

Robustness check

We conduct a range of robustness checks, with detailed results reported in the Appendix. First, to assess sensitivity to the spatial definition of the instrument, we use an alternative measure of AP capacity shocks based on a 15 km radius rather than 10 km. The resulting estimates are broadly consistent with the baseline specification, with similar signs and slightly larger, though less precisely estimated, coefficients (Appendix Table A5).

Second, we examine whether results are driven by cohort size by restricting the sample to schools with smaller year groups, where peer interactions may be more concentrated. The qualitative pattern of effects remains unchanged, with Year 9 exclusions continuing to generate the strongest negative spillovers, while effects in Years 8 and 10 are weaker. Estimates in this subsample are generally smaller and less precise, likely reflecting coarser measurement of exclusion rates in smaller cohorts (Appendix Table A6).

Third, we estimate a joint specification including exclusion rates in Years 8–10 simultaneously, allowing timing effects to be identified conditional on exposure in adjacent years. The results confirm the main patterns, with Year 9 exclusions exhibiting the most consistent negative effects on educational and labour-market outcomes, albeit with reduced precision due to weaker first-stage strength (Appendix Table A7).

Overall, the consistency of signs, timing, and relative magnitudes across these alternative specifications supports the robustness of the main findings and suggests that the estimated peer effects are not driven by the choice of instrument, sample composition, or omitted exposure to exclusions in adjacent years.

Falsification check

To assess the validity of our identification strategy, we conduct a set of falsification tests. Across all specifications, we find no evidence that the instrument captures spurious trends or unrelated cohort dynamics.

First, we test whether *future* changes in local AP capacity predict current exclusion rates. If the instrument reflects exogenous shocks, future capacity should have no predictive power once contemporaneous variation is accounted for. To implement this, we estimate:

$$r_s^k = \delta_1^k \, \Delta\widetilde{A}_{-s,10km,t+j}^{res,k} + \delta_2^k X_{is} + \sigma_s + \tau_t + e_{is}^k, \quad k \in \{8,9,10\}. \quad (3)$$

Here, $\Delta\widetilde{A}_{-s,10km,t+j}^{res,k}$ denotes the change in AP filled capacity within 10 km in year t + j, residualised with respect to the contemporaneous measure used in the main first stage. This step purges all variation in the future instrument that is predictable from the current instrument, ensuring that any remaining association with exclusions reflects improper "reverse timing" and not mechanical correlation across years.

The results show that future AP capacity changes have no systematic relationship with current exclusions, while the contemporaneous instrument remains strong (Table 8). This supports the interpretation that the instrument captures year-specific shocks rather than underlying trends or anticipatory behaviour.

Table 8 - Falsification Test: Effects of Future instrument on Exclusion

	Excl Y8	Excl Y9	Excl Y10
Instrument year t	-0.9500***	-0.4236***	-0.4169***
	[0.115]	[0.075]	[0.059]
	2,035,656	2,035,656	2,035,656
Future instrument year t+1	-0.3083*	-0.0951	0.0147
	[0.125]	[0.089]	[0.082]
	1,528,724	1,528,724	1,528,724
Future instrument year t+2	-0.0722	0.0494	-0.0013
	[0.104]	[0.077]	[0.078]
Observations	1,025,277	1,025,277	1,025,277

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Second, we examine whether exclusions in other year groups affect outcomes within a given cohort. Since peer disruption operates primarily within cohorts, we expect no cross-cohort effects. Consistent with this, we find no significant effects of exclusions in adjacent year groups on pupil outcomes (Appendix Table A8), indicating that spillovers are localised within cohorts.

Third, we estimate placebo regressions using predetermined outcomes, KS2 attainment, and an unrelated behavioural measure (authorised absences). As these outcomes are fixed prior to secondary school or unrelated to disruptive behaviour, any effect would indicate spurious correlation. We find no evidence of significant effects across these placebo outcomes (Table A9), suggesting that the instrumented variation in exclusions is orthogonal to pre-existing pupil characteristics and unrelated behavioural factors.

Taken together, these tests provide strong support for the exogeneity of the instrument and the interpretation of our estimates as capturing causal peer spillover effects.

Conclusion

This paper provides the first large-scale causal evidence on how the timing of permanent exclusions shapes the educational and early-adulthood trajectories of pupils who remain in mainstream schools. Leveraging year-to-year fluctuations in local Alternative Provision (AP) filled capacity as a strong and plausibly exogenous instrument, we document substantial peer spillovers generated by the removal of disruptive classmates. Our findings show that mid-to-late secondary exclusions, particularly in Year 9, impose the largest academic and labour-market costs on peers, while earlier exclusions in Year 8 generate small but positive longer-run effects.

Across a wide range of 2SLS estimates, a one-per-thousand increase in the Year 9 exclusion rate reduces GCSE Maths and English scores by 0.024 and 0.044 standard deviations respectively, lowers Level 2 and Level 3 qualification attainment by 0.65 and 0.63 percentage points, raises the probability of being NEET at age 20 by 0.62 percentage points. Year 10 exclusions also worsen post-16 outcomes, particularly reducing employment probability, though their academic effects are lower. The pattern suggests that exclusions in Years 9 and 10 may reflect behavioural problems that intensify over time, generating larger spillovers before removal occurs. This interpretation is supported by the mechanism analysis, which indicates that most of the negative effects of exclusion are explained by the number of days pupils are suspended, a proxy for the severity of their disruptiveness, and that peer effects are driven predominantly by exclusions for physical assault and persistent disruptive behaviour.

It is important to clarify the interpretation of these estimates. Our empirical design identifies the impact of marginal changes in exclusion rates within each year group rather than the causal effect of explicitly delaying exclusion until that point in secondary school. However, because pupils excluded later typically have longer histories of disruptive behaviour prior to removal, the timing pattern we document likely reflect the peer consequences of excluding pupils whose disruptive behaviour has typically persisted for longer prior to removal. Because these marginal exclusions are induced by variation in AP capacity, the estimates capture the effect of excluding pupils whose removal depends on the availability of alternative placements, rather than the effect of excluding the most severely disruptive pupils.

Heterogeneity analysis reveals strong timing and distributional effects: early exclusions improve labour-market attachment for disadvantaged pupils, mid-secondary exclusions are most harmful for low-ability students, while late exclusions reshape outcomes around high-stakes exams, generating attainment gains for some marginal students but weaker employment prospects overall. This implies that decisions about exclusion timing matter. Interventions that stabilise classrooms earlier may support vulnerable pupils' longer-term outcomes, while mid-secondary disruption appears particularly damaging for those already at academic risk.

To explore the mechanisms behind peer exclusion effects, we proxy disruptiveness using the number of suspension days accumulated prior to exclusion. At the mean level of suspension intensity (11 days), the interaction explains roughly 18–26% of the overall negative effect of Year 9 exclusions. At very high levels of prior suspension (around the 95th percentile), the interaction fully offsets, and may even reverse, the baseline negative effect. This indicates that peer spillovers are driven less by exclusion per se and more by the severity of disruptive behaviour. Consistent with this interpretation, disaggregating by exclusion reason shows that negative spillovers are concentrated among cases involving physical assault and persistent disruptive behaviour, rather than verbal abuse.

A series of robustness and placebo tests confirms the validity of our empirical strategy. Using an alternative (15 km) instrument, restricting the sample to smaller schools, and jointly estimating all exclusion years yields results consistent with the main specification. Placebo tests using future AP filled capacity shocks, cross-cohort exclusions, pre-treatment KS2 scores, and authorised absences all show null effects, supporting both the exogeneity of the instrument and the interpretation that peer spillovers are confined to disruptions within the same cohort.

Taken together, the results point to a key policy insight: the peer costs of disruptive behaviour are real, persistent, and highly sensitive to the timing of intervention. Exclusions that occur after several years of unresolved disruption are associated with substantial losses in attainment and early labour-market outcomes for classmates, with these harms falling disproportionately on pupils who are already academically or socially vulnerable. At the same time, a large body of evidence documents the significant long-run costs borne by excluded pupils themselves, implying that exclusion involves a fundamental trade-off between competing social costs. Our analysis does not assess alternative behavioural interventions, nor does it advocate exclusion as a preferred response (Valbenito et al., 2025). Instead, it shows that delaying action in the presence of persistent disruption magnifies peer spillovers, suggesting that policies which identify and address unmet behavioural and educational needs earlier, through targeted support, inclusive practices, or alternative pathways, may be more effective in limiting cumulative harm within cohorts while avoiding the substantial long-term costs associated with exclusion.

Appendix

Table A1- IV for months as NEET and as employed

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N	Mean outcome
Months NEET (by age 21)						
Coeff	0.1327**	0.2598***	0.0558	Yes	1,630,509	6.105
[se]	[0.044]	[0.073]	[0.043]			
Kleibergen-Paap Wald F stat	71.37	30.50	44.40			
Months employed (by age 21)						
Coeff	0.0809***	-0.0460	-0.0501*	Yes	1,630,561	9.004
[se]	[0.022]	[0.029]	[0.022]			
Kleibergen-Paap Wald F stat	71.38	30.49	44.40			

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A2- IV for post-16 outcomes by age 19

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N	Mean outcome
Employment at age 19						
Coeff	0.0049**	0.0039+	-0.0012	Yes	1,833,230	0.752
[se]	[0.002]	[0.002]	[0.001]			
Kleibergen-Paap Wald F stat	70.12	31.97	47.89			
Months employed (by age 19)						
Coeff	0.0402*	0.0468+	-0.0086	Yes	1,833,230	8.04
[se]	[0.017]	[0.024]	[0.016]			
Kleibergen-Paap Wald F stat	70.12	31.97	47.89			
NEET at 19						
Coeff	0.0014	-0.0027	0.0013	Yes	1,833,230	0.152
[se]	[0.001]	[0.002]	[0.001]			
Kleibergen-Paap Wald F stat	70.12	31.97	47.89			
Months NEET (by age 19)						
Coeff	0.0349+	0.1086***	0.0158	Yes	1,832,864	2.9286
[se]	[0.020]	[0.032]	[0.019]			
Kleibergen-Paap Wald F stat	70.12	31.96	47.87			

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A3- Heterogeneity analysis for females

	KS4 math	KS Eng	Lev 2	Lev 3	Emp 21	Neet 21
Panel A: year 8						
Exclusion per 1K in Y8	-0.0095*	-0.0196***	-0.0050**	-0.0031	-0.0019	-0.0005
	[0.004]	[0.006]	[0.002]	[0.002]	[0.002]	[0.002]
Female X Excl Y8	-0.0012	0.0458***	0.0056*	0.0089**	0.0091**	-0.0034
	[0.005]	[0.009]	[0.003]	[0.003]	[0.003]	[0.003]
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561
Panel B: year 9						
Exclusion per 1K in Y9	-0.0334***	-0.0359**	-0.0051*	-0.0074*	-0.0014	0.0116***
	[0.008]	[0.011]	[0.003]	[0.003]	[0.003]	[0.003]
Female X Excl Y9	0.0179**	-0.0154+	-0.0027	0.0022	-0.0039	-0.0105**
	[0.006]	[0.009]	[0.003]	[0.003]	[0.003]	[0.003]
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561
Panel C: year 10						
Exclusion per 1K in Y10	0.0038	-0.0246**	-0.0100***	-0.0059**	-0.0057*	0.0077***
	[0.000]	[0.008]	[0.002]	[0.002]	[0.002]	[0.002]
Female X Excl Y10	-0.0156**	0.0197**	0.0147***	0.0105**	0.0042	-0.0105***
	[0.005]	[0.008]	[0.003]	[0.003]	[0.003]	[0.003]
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A4 – 2SLS of exclusion, by reason for exclusion

	KS4 math	KS4 Eng	Lev 2	Lev 3	Emp at 21	Neet at 21	Univ at 21
Excl Y8, persistent distr beh	-0.3008*	0.0559	-0.0776+	0.0290	0.0698	-0.0611	0.0255
	[0.131]	[0.127]	[0.044]	[0.050]	[0.048]	[0.046]	[0.043]
Kleibergen-Paap Wald F stat	20.43	20.43	20.43	20.43	21.19	21.19	21.19
Excl Y9, persistent distr beh	-0.6026**	-1.0894**	-0.1725**	-0.1656*	-0.1019+	0.1727*	-0.1161*
	[0.201]	[0.335]	[0.064]	[0.067]	[0.059]	[0.069]	[0.057]
Kleibergen-Paap Wald F stat	14.24	14.24	14.24	14.24	13.24	13.24	13.24
Excl Y10, persistent distr beh	-0.1293	-0.5115*	-0.0948+	-0.0252	-0.1182*	0.0812	0.0776
	[0.118]	[0.221]	[0.049]	[0.049]	[0.059]	[0.053]	[0.061]
Kleibergen-Paap Wald F stat	12.78	12.78	12.78	12.78	11.95	11.95	11.95
Excl Y8, physical assault	-0.4106*	0.0763	-0.1060+	0.0396	0.0956	-0.0837	0.0350
	[0.171]	[0.174]	[0.059]	[0.069]	[0.066]	[0.064]	[0.058]
Kleibergen-Paap Wald F stat	23.29	23.29	23.29	23.29	23.79	23.79	23.79
Excl Y9, physical assault	-1.1623*	-2.1013*	-0.3326*	-0.3195*	-0.1933	0.3276*	-0.2201+
	[0.482]	[0.828]	[0.147]	[0.154]	[0.122]	[0.155]	[0.118]
Kleibergen-Paap Wald F stat	7.796	7.796	7.796	7.796	7.434	7.434	7.434
Excl Y10, physical assault	-0.2856	-1.1296*	-0.2094+	-0.0558	-0.3079	0.2115	0.2021
	[0.272]	[0.569]	[0.125]	[0.110]	[0.193]	[0.161]	[0.179]
Kleibergen-Paap Wald F stat	6.637	6.637	6.637	6.637	4.550	4.550	4.550
Excl Y8, verbal abuse	-1.0698*	0.1987	-0.2761	0.1032	0.2340	-0.2048	0.0856
	[0.525]	[0.457]	[0.170]	[0.180]	[0.172]	[0.164]	[0.144]
Kleibergen-Paap Wald F stat	8.855	8.855	8.855	8.855	9.893	9.893	9.893
Excl Y9, verbal abuse	-2.9703	-5.3701	-0.8501	-0.8165	-0.4246	0.7195	-0.4835
	[2.110]	[3.739]	[0.619]	[0.608]	[0.335]	[0.481]	[0.348]
Kleibergen-Paap Wald F stat	2.166	2.166	2.166	2.166	2.755	2.755	2.755
Excl Y10, verbal abuse	-0.2733	-1.0809*	-0.2003+	-0.0534	-0.2425+	0.1666	0.1592
	[0.252]	[0.481]	[0.108]	[0.103]	[0.125]	[0.111]	[0.128]
Kleibergen-Paap Wald F stat	11.23	11.23	11.23	11.23	11.51	11.51	11.51
Observations	2,035,656	2,035,656	2,035,656	2,035,656	1,630,561	1,630,561	1,630,561

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A5 – 2SLS estimates of exclusions, using change in AP filled spaces within 15 km

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N	Mean outcome
KS4 Maths						
Coeff	-0.0103+	-0.0461*	-0.0031	Yes	2,035,656	0.103
[se]	[0.005]	[0.019]	[0.006]			
Kleibergen-Paap Wald F stat	41.28	8.252	17.63			
KS4 English						
Coeff	0.0030	-0.0864**	-0.0241**	Yes	2,035,656	0.130
[se]	[0.006]	[0.032]	[0.009]			
Kleibergen-Paap Wald F stat	41.28	8.252	17.63			
Level 2						
Coeff	-0.0005	-0.0159*	-0.0039+	Yes	2,035,656	0.831
[se]	[0.002]	[0.006]	[0.002]			
Kleibergen-Paap Wald F stat	41.28	8.252	17.63			
Level 3						
Coeff	0.0039	-0.0151*	-0.0018	Yes	2,035,656	0.621
[se]	[0.002]	[0.007]	[0.003]			
Kleibergen-Paap Wald F stat	41.28	8.252	17.63			
Employed at 21						
Coeff	0.0004	-0.0085+	-0.0095**	Yes	1,630,561	0.797
[se]	[0.002]	[0.005]	[0.003]			
Kleibergen-Paap Wald F stat	41.75	7.222	16.50			
NEET at 21 - coeff						
Coeff	0.0005	0.0124*	0.0065*	Yes	1,630,561	0.171
[se]	[0.002]	[0.006]	[0.003]			
Kleibergen-Paap Wald F stat	41.75	7.222	16.50			

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A6 – 2SLS estimates of exclusions, using 50% smaller schools

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N	Mean outcome
KS4 Maths						
Coeff	-0.0022	-0.0188**	-0.0021	Yes	1,011,688	0.0884
[se]	[0.004]	[0.006]	[0.004]			
Kleibergen-Paap Wald F stat	51.28	24.04	33.05			
KS4 English						
Coeff	0.0024	-0.0325**	-0.0090+	Yes	1,011,688	0.1154
[se]	[0.004]	[0.011]	[0.005]			
Kleibergen-Paap Wald F stat	51.28	24.04	33.05			
Level 2						
Coeff	0.0014	-0.0041*	-0.0018	Yes	1,011,688	0.819
[se]	[0.001]	[0.002]	[0.002]			
Kleibergen-Paap Wald F stat	51.28	24.04	33.05			
Level 3						
Coeff	0.0003	-0.0017	0.0004	Yes	1,011,688	0.6093
[se]	[0.002]	[0.002]	[0.002]			
Kleibergen-Paap Wald F stat	51.28	24.04	33.05			
Employed at 21 - coeff						
Coeff	0.0036*	-0.0034	-0.0015	Yes	811,729	0.7857
[se]	[0.002]	[0.002]	[0.002]			
Kleibergen-Paap Wald F stat	55.44	22.38	30.66			
NEET at 21						
Coeff	-0.0029+	0.0049*	0.0026	Yes	811,729	0.1810
[se]	[0.002]	[0.002]	[0.002]			
Kleibergen-Paap Wald F stat	55.44	22.38	30.66			

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A7 - 2SLS estimates of exclusions, with three treatments in the same regression

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N	Mean outcome	Kleibergen-Paap Wald F stat
KS4 Maths							
Coeff	-0.0091+	-0.0245***	0.0021	Yes	2,035,656	0.103	7.029
[se]	[0.005]	[0.007]	[0.005]				
KS4 English							
Coeff	0.0030	-0.0427***	-0.0043	Yes	2,035,656	0.130	7.029
[se]	[0.007]	[0.011]	[0.008]				
Level 2							
Coeff	-0.0022	-0.0061**	-0.0013	Yes	2,035,656	0.831	7.029
[se]	[0.002]	[0.002]	[0.002]				
Level 3							
Coeff	0.0013	-0.0066*	0.0009	Yes	2,035,656	0.621	7.029
[se]	[0.002]	[0.003]	[0.002]				
Employed at 21							
Coeff	0.0026+	-0.0027	-0.0029	Yes	1,630,561	0.797	6.109
[se]	[0.002]	[0.002]	[0.002]				
NEET at 21 -							
Coeff	-0.0025	0.0061*	0.0008	Yes	1,630,561	0.171	6.109
[se]	[0.002]	[0.003]	[0.002]				

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row set is a regression output. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A8 - Falsification Test: Effects of placebo treatments

Outcome	Y10 exclusion using Y9 instrument	Y9 exclusion using Y8 instrument	Y9 exclusion using Y10 instrument	Y8 exclusion using Y9 instrument	Controls, School and cohort FE	N	Mean outcome
KS4 Maths							
Coeff	-0.1099	-0.29	-0.0152	-0.5640	Yes	2,035,656	0.103
[se]	[0.105]	[1.148]	[0.016]	[2.054]			
Kleibergen-Paap Wald F stat	1.131	0.0646	3.873	0.0755			
KS4 English							
Coeff	-0.1958	0.0462	-0.0602	-1.0047	Yes	2,035,656	0.130
[se]	[0.186]	[0.216]	[0.037]	[3.659]			
Kleibergen-Paap Wald F stat	1.131	0.0646	3.873	0.0755			
Level 2							
Coeff	-0.0291	-0.0683	-0.0112	-0.1494	Yes	2,035,656	0.831
[se]	[0.028]	[0.271]	[0.008]	[0.544]			
Kleibergen-Paap Wald F stat	1.131	0.0646	3.873	0.0755			
Level 3							
Coeff	-0.0283	0.0306	-0.0030	-0.1451	Yes	2,035,656	0.621
[se]	[0.028]	[0.129]	[0.006]	[0.529]			
Kleibergen-Paap Wald F stat	1.131	0.0646	3.873	0.0755			
Employed at 21 - coeff							
Coeff	-0.0138	0.0392	-0.0135	-0.1131	Yes	1,630,561	0.797
[se]	[0.015]	[0.085]	[0.009]	[0.626]			
Kleibergen-Paap Wald F stat	1.293	0.229	3.961	0.0333			
NEET at 21 -							
Coeff	0.0257	-0.0341	0.0093	0.2104	Yes	1,630,561	0.171
[se]	[0.024]	[0.075]	[0.007]	[1.157]			
Kleibergen-Paap Wald F stat	1.293	0.229	3.961	0.0333			

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

Table A9 - Falsification Test: Placebo outcomes

Outcome	Excl Y8	Excl Y9	Excl Y10	Controls, School and cohort FE	N
KS2 Math					
Coeff	0.003	-0.061	0.049	Yes	2,035,656
[se]	[0.039]	[0.053]	[0.037]		
Kleibergen-Paap Wald F stat	67.98	31.81	49.38		
KS2 Eng					
Coeff	-0.04	-0.053	0.064	Yes	2,035,656
[se]	[0.041]	[0.057]	[0.039]		
Kleibergen-Paap Wald F stat	67.98	31.81	49.38		
Absences					
Coeff	0.000	-0.004	0.003	Yes	2,008,868
[se]	[0.003]	[0.004]	[0.003]		
Kleibergen-Paap Wald F stat	67.88	32.13	49.29		

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.10. School fixed effect linear regression estimated. Each row-column combination is a separate regression. Standard errors in brackets clustered at school and cohort level. Controls: KS2 Maths, KS2 English, EAL, FSM (up to 11), Statement or EHC plan (up to 11), SEN (up to 11), Ever looked after up to age 16, Pupil had ever been in need, Female, Unauthorized absence up to 12, Ethnicity dummies, urban dummy, IDACI score decile dummies, year dummies, dummies for missing KS2 and missing unauthorized absences.

References

Ammermueller, A., and Pischke, J.S. (2009). Peer effects in European primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics, 27*(3), 315-348.

Angrist, J.D. (2014). The perils of peer effects. Labour Economics*, 30*, 98-108.

Angrist, J.D., Graddy, K., & Imbens, G.W. (2000). The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish. The Review of Economic Studies, 67 (3), 499-527

Angrist, J.D. and Pischke, J.S., 2009. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Bacher-Hicks, A., Billings, S.B. and Deming, D.J., 2024. The School-to-Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime. American Economic Journal: Economic Policy, 16(4), pp.165-193.

Buchmueller, G., Cathro, C., Dorsett, R., Oppedisano, V., Sutherland, A. and Tagliaferri, G., 2025. School Exclusion and Youth Custody. Journal of Quantitative Criminology, pp.1-28.

Burgess, S., Cantillon, E., Cavallo, M., Greaves, E. and Zhang, M., 2023. School admissions in England: The rules schools choose on which pupils to admit (No. 2013/356676). ULB--Universite Libre de Bruxelles.

Carrell, S.E. and Hoekstra, M.L., 2010. Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. American Economic Journal: Applied Economics, 2(1), pp.211-228.

Carrell, S.E., Hoekstra, M. and Kuka, E., 2018. The long-run effects of disruptive peers. American Economic Review, 108(11), pp.3377-3415.

Cuellar, A.E. and Markowitz, S., 2015. School suspension and the school-to-prison pipeline. International Review of Law and Economics, 43, pp.98-106.

Department for Education, 2023/24. Suspensions and permanent exclusions in England, retrieved from:
https://explore-education-statistics.service.gov.uk/find-statistics/suspensions-and-permanent-exclusions-in-england/2023-24

Department for Education, 2024/25 Suspensions and permanent exclusions in England, retrieved from:
https://explore-education-statistics.service.gov.uk/find-statistics/suspensions-and-permanent-exclusions-in-england/2024-25-autumn-term

Department for Education, 2024. National behaviour survey Findings from academic year 2022 to 2023, retrieved from:
https://assets.publishing.service.gov.uk/media/6836f84b8ade4d13a63236c0/National_behaviour_survey_AY_2022_2023_report.pdf

Department for Education, Suspensions and permanent exclusions in England, 2023-24, DfE, https://explore-education-statistics.service.gov.uk/find-statistics/suspensions-and-permanent-exclusions-in-england/2023-24

Department for education Arranging Alternative Provision A Guide for Local Authorities and Schools, 2025, DfE, retrieved from: https://assets.publishing.service.gov.uk/media/67a1ee367da1f1ac64e5fe2c/Arranging_Alternative_Provision_-_A_Guide_for_Local_Authorities_and_Schools.pdf

Dieterle, S. and Snell, A., 2014. It's hip to be square: Using quadratic first stages to investigate instrument validity and heterogeneous effects. Unpublished manuscript.

Goulas, S., Griselda, S., Megalokonomou, R. and Zenou, Y., 2024. Disruptive Peers and Academic Performance: Short-and Long-Term Outcomes.

Imbens, G.W. and Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. Journal of economic literature, 47(1), pp.5-86.

Imberman, S.A., Kugler, A.D. and Sacerdote, B.I., 2012. Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. American Economic Review, *102*(5), pp.2048-2082.

Institute of Education and National Foundation for Educational Research. (2014, July). School exclusion trial evaluation. Department for Education. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/331795/RR364_-_School_Exclusion_Trial_Final_Report.pdf

Kristoffersen, J.H.G., Krægpøth, M.V., Nielsen, H.S. and Simonsen, M., 2015. Disruptive school peers and student outcomes. Economics of Education Review, *45*, pp.1-13.

Lacoe, J. and Steinberg, M.P., 2018. Rolling back zero tolerance: The effect of discipline policy reform on suspension usage and student outcomes. Peabody Journal of Education, 93(2), pp.207-227.

Lavy, V., Silva, O. and Weinhardt, F., 2012. The good, the bad, and the average: Evidence on ability peer effects in schools. Journal of Labor Economics, 30(2), pp.367-414.

Lavy, V. and Yancu, A., 2025. Violent Peers at School: Impacts and Mechanisms (No. w34482). National Bureau of Economic Research.

Madia, J.E., Obsuth, I., Thompson, I., Daniels, H. and Murray, A.L., 2022. Long-term labour market and economic consequences of school exclusions in England: Evidence from two counterfactual approaches. British journal of educational psychology, *92*(3), pp.801-816.

McAra, L. and McVie, S., 2010. Youth crime and justice: Key messages from the Edinburgh Study of Youth Transitions and Crime. Criminology & Criminal Justice, 10(2), pp.179-209.

Pope, N.G. and Zuo, G.W., 2023. Suspending suspensions: The education production consequences of school suspension policies. The Economic Journal, 133(653), pp.2025-2054.

Sacerdote, B., 2011. Peer effects in education: How might they work, how big are they and how much do we know thus far?. In Handbook of the Economics of Education (Vol. 3, pp. 249-277). Elsevier.

Sanders, J., Liebenberg, L. and 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), pp.386-403.

Sorensen, L.C., Bushway, S.D. and Gifford, E.J., 2022. Getting tough? The effects of discretionary principal discipline on student outcomes. Education Finance and Policy, 17(2), pp.255-284.

Sutherland, A. and Eisner, M., 2014. The treatment effect of school exclusion on unemployment. Available at SSRN 2380956.

Thomson, D., 2019. Timpson Review reflections, part one: Not all pupils who end up in alternative provision have been permanently excluded. FFTLab, available here: https://ffteducationdatalab.org.uk/2019/05/timpson-review-reflections-part-one-not-all-pupils-who-end-up-in-alternative-provision-have-been-permanently-excluded/

Valdebenito, S., Gaffney, H., Arosemena-Burbano, M.J., Hitchcock, S., Jolliffe, D. and Sutherland, A., 2025. School-Based Interventions for Reducing Disciplinary School Exclusion. An Updated Systematic Review. Campbell Systematic Reviews, 21(4), p.e70063.

Zhao, L. and Zhao, Z., 2021. Disruptive peers in the classroom and students' academic outcomes: Evidence and mechanisms. Labour Economics, 68, p.101954.