

When Trouble Travels: Peer Effects of Criminal Exposure in Schools

Hikaru Kawarazaki*

[Click here for the most recent version](#)

Institute for Fiscal Studies and University College London

October 31, 2026

Abstract

First contact with the criminal justice system peaks in early adolescence, when children spend much of their time with peers at school. While a large literature documents that peers shape children's development and behavior, much less is known about the consequences of exposure to schoolmates with a criminal record. Using administrative data linking criminal justice and education records for nearly all students in England, I study whether and to what extent criminal behavior spreads within schools through exposure to juvenile offenders. Exploiting the plausibly exogenous timing of school transfers by students with a criminal record, I find that exposure increases a student's probability of starting a criminal career by 12.9% within four months, with effects persisting for at least three years. Exposure also lowers educational achievement; most of this decline is accounted for by exposed students' subsequent criminal behavior. These peer effects account for 5% of the total social cost of youth crime in England. Because the effects vary substantially across school characteristics, a simulation shows that reassigning transfer offenders to other schools within three miles could reduce peer-induced crime by 36%.

Keywords: human capital, youth crime, peer effects

JEL Classification: I21, K14, K42

*hikaru.kawarazaki@ifs.org.uk, the Institute for Fiscal Studies, 2 Marylebone Road, London NW1 4DF, United Kingdom, and the Department of Economics, University College London, 30 Gordon Street, London WC1H 0AN, United Kingdom. I appreciate feedback from Sonia Bhalotra, Manudeep Bhuller, Nicholas Bloom, Nils Braakmann, Jack Britton, Pedro Carneiro, Sarah Cattán, Arnaud Chevalier, Daniel Cunha Byström, Monica Costa-Dias, Emilia Del Bono, Magdalena Domínguez, Christian Dustmann, Dita Eckardt, Seth Gershenson, Hans Grönqvist, Hidehiko Ichimura, Ingo E. Isphording, Takuma Kamada, Hiroyuki Kasahara, Tom Kirchmaier, Lucy Kraftman, Sonya Krutikova, Stephen Machin, Masaki Takahashi, Costas Meghir, Matthew Nibloe, Emily Owens, Matthias Parey, Aureo de Paula, Damian Phelan, Luigi Pistaferri, Miika Päällysaho, Carmen Villa-Llera, Imran Rasul, Nick Ridpath, Luca Rizzotti, Yasuyuki Sawada, Hitoshi Shigeoka, Andrea Smurra, Imran Tahir, Yasuka Tateishi, Michael Vlassopoulos, Diego Zambiasi, Ben Zranko, and seminar participants at UCL, IFS, IZA/ECONtribute, CREPE at the University of Tokyo, Royal Economic Society, Summer Workshop on Economic Theory, Econometric Society World Congress, Workshop on Education Economics and Policy, Sophia University, National Graduate Institute for Policy Studies, European Association of Young Economists Annual Meeting, Centre for Economic Performance at LSE, Essen Health Conference, Annual Conference of the European Society for Population Economics, Workshop on Labour and Family Economics at University of York, Kyoto Summer Workshop on Applied Economics, Workshop on the Economics of Risky Behavior, European Association of Labour Economists, Newcastle University, and Network Science & Economics for their comments. I acknowledge financial support from the ESRC-UBEL studentship, the Japan Center for Economic Research, the Nuffield Foundation, the Uehiro Foundation on Ethics and Education, the Shihou Kyokai, and the Suntory Foundation. The data source is the Office for National Statistics (ONS) (2026). The results have been given "SRS Output Clearance" by Secure Research Service (SRS), together with the Data Owners' approvals from the Ministry of Justice (MoJ) and the Department for Education (DfE). This does not imply ONS-SRS, MoJ, or DfE's acceptance of the validity of any methods used or the output itself, and they do not accept responsibility for any onward use of the output. This paper was awarded the Best Paper Award at the Essen Health Conference 2026. The previous title was "A Rotten Apple: Peer Effects of Criminal Exposure in Schools." All errors are my own.

1 Introduction

Early adolescence marks the peak period of first contact with the criminal justice system; the most common age at first contact is 15 in England (Cattan et al., 2023a,b, 2025). Youth crime imposes high costs on individuals. A large body of theoretical and empirical work shows that criminal justice involvement can hinder offenders’ human capital accumulation and ultimately worsen their labor market outcomes.¹ Youth crime also generates substantial social and economic costs through policing, prevention, victim compensation, and other channels. In England and Wales, the annual cost of serious youth violence is estimated to be as high as £1.3 billion (\$1.8 billion) (Irwin-Rogers, Muthoo, and Billingham, 2020). Reducing youth crime is therefore a key policy priority.

Preventing youth crime, however, is challenging if a policy to reduce crime targets only individuals. A growing literature shows that criminal behavior can propagate through social networks.² For adolescents, schools are important places for social interaction, and recent evidence documents their essential role in peer effects of crime (Billings, Deming, and Ross, 2019; Díaz et al., 2021; Lee et al., 2021; Billings and Hoekstra, 2023).

Despite growing evidence of peer effects on crime in school, little is known about how exposure to offending peers affects the initiation of criminal behavior among early teenagers, when most juvenile offenders start committing crime. Existing studies tend to focus on older youth or adults and intensive margins, even though understanding the onset of offending is crucial for assessing its long-run consequences for human capital. Moreover, much of the existing literature captures the combined influence of schools and neighborhoods, leaving the relative importance of these two settings unclear. Distinguishing between them is important for policy because interventions targeted at schools differ substantially from those targeted at neighborhoods.³

This paper advances the literature by studying the peer effects of exposure to known offenders in secondary schools in England, aged 11 to 16. I examine whether exposure to an offender transferring into a school-grade causes previously nonoffending peers to begin committing crimes, by comparing their behavior before and after the offender’s move. I also study whether the transfer affects overall criminal activity, including students who have already begun offending. I discuss potential mechanisms by showing heterogeneity across individuals’ and schools’ characteristics. Beyond criminal behavior, I examine effects on academic achievement at age 14 and conduct a statistical mediation analysis to decompose them into two channels: a direct effect of exposure to a disruptive offender and an indirect effect operating through the exposed student’s own entry into

¹An extensive literature investigates the effects of crime on human capital. See Becker (1962); Pager (2003); Holzer, Raphael, and Stoll (2006); Holzer (2007); Hjalmarsson (2008); Aizer and Doyle Jr. (2015); Mueller-Smith (2015); Agan and Starr (2017); Doleac and Hansen (2020); Eren and Mocan (2021); Rose (2021); Cullen, Dobbie, and Hoffman (2023); Burton and Wasser (2025); Garin et al. (2025); Lavy and Yancu (2025).

²See Gavrilova and Puca (2022) for a review.

³For example, Dobbie and Fryer Jr. (2011) find that attending the Harlem Children’s Zone charter schools raises achievement enough to close the Black–white gap in mathematics, and that the Zone’s community programs contribute little to these gains.

criminal activity. Finally, I conduct a counterfactual exercise in which a social planner changes the destination schools of transfer offenders to assess how alternative school assignments could mitigate crime spillovers, exploiting the substantial heterogeneity in peer effects across schools.

Identification relies on variation in the timing of offenders' school transfers. Conditional on a rich set of observable characteristics, the remaining variation in when students are first exposed to an offender transferring into their school and grade is plausibly as good as random. Because the timing of offenders' transfers differs across individuals, I use a staggered event-study design to estimate average treatment effects on the treated, comparing students who have already been exposed to an incoming offender with students who have not yet been exposed, following [Callaway and Sant'Anna \(2021\)](#).

A natural concern for identification is that schools that accept offenders may differ from those that do not. To avoid any contamination, I focus on those exposed to an offender's transfer during secondary school and compare those who are exposed with those who are not yet. To ensure external validity, I show this sample selection has little effect by comparing those exposed with those never exposed. Another important concern is that those exposed earlier might be more disadvantaged. To address this concern, I compare the individuals' characteristics of early-treated and late-treated to show that there are no systematic differences in observable characteristics.

Using rich administrative data sets in England, the National Pupil Database (NPD) on demographics and educational attainments and the Police National Computer (PNC) on crime records for about a million individuals, I estimate the peer effects of offense initiation in secondary school. When an offender transfers into a student's school and cohort, the probability that the student commits a first offense within the following four months increases by 12.9%. This magnitude corresponds to approximately one additional new offender for every three offender transfers. This suggests that the peer effects are intense enough to drag non-offenders into a criminal life at a very early stage of life. The cumulative effect grows over the following two years and remains persistent, suggesting potentially long-lasting consequences for human capital. An offender's transfer also increases the number of crimes committed by students in the recipient school, including those who have initiated offending, by 24.3% over the following four months. Together, these findings indicate that incoming offenders increase both the initiation of criminal activity and overall offending.

Next, I show that the intensity of exposure to offenders matters by presenting heterogeneous results. First, I show that when the treatment intensity is larger, defined by the fraction of the number of transfer offenders to that of students in the school cohort, the treatment effect is more substantial. A large treatment intensity suggests more opportunities for peers to interact with transfer offenders, which may lead to a sizable effect. Second, since children have more intense social interactions in school than in their neighborhoods, I compare offenders' transfers into the same school but different neighborhoods with transfers into a different school within the same neighborhood. I show that school-based effects are substantially larger than neighborhood effects. Consistent with this interpretation, I also show that most estimated peer effects are concentrated

in crimes committed during school terms rather than during holidays or weekends. These results show that intensity of exposure matters for the peer effects.

To investigate the underlying mechanisms, I examine two potential channels: skill transmission and role-modeling. First, I find little evidence of skill transmission. I show descriptively that transfer offenders rarely offend after transfer. Consistent with this, I show that peers do not commit crime together with transfer offenders. Further, peers do not commit the same type of crime as transfer offenders have committed before the transfer. This evidence excludes skill transmission from the potential mechanisms.

Instead, the evidence points to role-modeling: peers might view the transfer offender as a new role model and imitate the role model’s behavior. First, I show that the existing concentration of offenders at the destination school matters. In the least “offender-dense” schools, where the fraction of incumbent offenders in a school-grade is small, the effect on offense initiation is large, while the effect is insignificant in the most offender-dense schools. One interpretation is that a transfer offender in the least dense schools might become a new role model because offenders are rare there, whereas in offender-dense schools, offenders are prevalent and an additional offender is less likely to be a new role model. In contrast, offender-dense schools amplify the effect on the number of crimes committed. Consistent with evidence from juvenile correctional facilities (Bayer, Hjalmarsson, and Pozen, 2009), this pattern suggests that concentrating offenders may strengthen peer interactions among students who are already engaged in criminal activity, inspiring existing offenders. Moreover, I show that transfer offenders’ seniority matters. The effect is larger when the transfer offender is older than their new peers than when they are younger, consistent with older students being more influential as role models. Further, I show school discipline may play an important role. I focus on English academy schools, which have more autonomy and stricter discipline (Machin and Sandi, 2020). In such schools, I show no peer effects of crime. This result is consistent with the finding of Duchini, Lavy, and Machin (2020): converting a non-academy school into an academy school reduces crime. Together with the larger effects of more intense exposure, these results suggest that role-modeling is an important mechanism.

Further, I examine whether the effects are larger for disadvantaged students. The effects are substantially larger for students with lower prior achievement, those eligible for free school meals or with special educational needs, and those living in deprived areas. Because these characteristics are associated with lower human capital, and hence a lower opportunity cost of crime, this pattern is consistent with the idea that a low opportunity cost makes students more likely to commit crime (Becker, 1968; Lochner, 2004).

I conduct a series of robustness checks. Most importantly, a placebo test shows that transfers of students who have not yet offended, but who go on to offend later, have no effect on their new peers. Because these students are likely similarly disruptive but have no criminal record at the time of the transfer, this result suggests that the estimated effects reflect the transmission of criminal behavior rather than general disruptiveness. The main results are also robust to different

specifications and sample selection.

To quantify the economic magnitude of these effects, I calculate the total cost of youth crime attributable to school-based criminal peer effects using estimates of the social cost of different crime types from [Heeks et al. \(2018\)](#). I estimate that approximately £20.0 million (\$26.4 million), equivalent to 5% of the total cost of youth crime during my sample period, can be attributed to peer effects in schools.

Next, I examine academic achievement as an additional negative externality of exposure to offending peers. I focus on the age-14 test scores and compare students exposed to an incoming offender up to one year before the exam with students exposed only after the exam. Exposure before the exam reduces age-14 test scores by 0.016 standard deviations. A statistical mediation analysis suggests that about one-third of this effect operates directly through exposure to an offending peer, while the remaining two-thirds operates indirectly through induced offending: exposure causes some students to become offenders before the exam, which is tied to lower educational achievement.

Finally, to assess how school assignment can mitigate crime spillovers, I conduct a simulation exercise in which a social planner reassigns transfer offenders across destination schools, exploiting the substantial heterogeneity in peer effects by the existing density of offenders. Holding the estimated effects fixed and restricting reassignments to schools within three miles of the original destination, the statutory walking distance for pupils aged eight and over in England ([Department for Education, 2026](#)), I find that an assignment rule designed to minimize total crime could reduce the number of crimes generated by peer effects by 36%. The choice of objective, however, is crucial. Because offender-dense schools dampen the initiation of offending but amplify offending among students already involved in crime, an assignment rule that minimizes newly induced offenders concentrates transfers in such schools and may increase the total number of crimes by 31%.

This paper builds on a growing literature on peer effects in crime. The closest study is [Billings, Deming, and Ross \(2019\)](#), which documents social spillovers in crime when disadvantaged students attending the same school are also concentrated within small neighborhoods. Their setting captures the combined influence of schools and neighborhoods, whereas I separately identify each contribution. This distinction matters for policy because interventions aimed at schools and neighborhoods differ in both design and implementation. Their analysis also focuses on joint offending, while peer effects do not necessarily operate through committing crime together. Finally, their offense records begin at age 16, whereas first contact with the criminal justice system in England peaks at age 15 ([Cattan et al., 2023a,b, 2025](#)). Studying younger adolescents therefore provides new evidence on how criminal behavior begins.

Another closely related study is [Lavy and Yancu \(2025\)](#), which finds the peer effects of violent peers on peers' involvement in fighting, focusing on transition from primary to middle school for causal identification. Their focus is on self-reported violence instead of the universe of reported crimes with police involvement. Investigating peer effects of a wider range of known crimes is an

important complement to prevent students' misbehavior.

Lee et al. (2021) is also close to this study. They identify key players in offenders' networks and show through simulation that removing a small fraction of these individuals can substantially reduce crime. While that exercise identifies which offenders are most influential, this paper focuses on who begins offending because of exposure to such individuals, providing a complementary measure of criminal spillovers. Billings and Hoekstra (2023) studies the effects of schoolmates whose parents have been arrested on students' subsequent criminal behavior. Their treatment captures exposure to potentially disruptive peers more broadly, while I focus directly on the transmission of criminal behavior following exposure to an offender.

More generally, this paper contributes to the literature on how schools and education affect crime. A large body of evidence shows that additional schooling, education, and youth clubs reduce criminal activity (Lochner and Moretti, 2004; Machin, Marie, and Vujčić, 2011; Berthelon and Kruger, 2011; Dobbie and Fryer Jr., 2015; Cook and Kang, 2016; Landersø, Nielsen, and Simonsen, 2017; Bacher-Hicks, Billings, and Deming, 2024; Bell, Costa, and Machin, 2022; Villa, 2024), as does access to early childhood education (Depew and Eren, 2016; Ando, Mori, and Yamaguchi, 2025). Several studies examine these relationships through the lens of human capital accumulation (Aizer and Doyle Jr., 2015; Mueller-Smith, 2015; Eren and Mocan, 2021), while others model youth behavior (Becker, 1968; Lochner, 2004).⁴ I contribute to this literature by identifying a factor within schools, i.e., exposure to a criminal peer, which can increase criminal activity, and estimate such endogenous peer effects with a plausible identification strategy.⁵

The paper also adds to evidence on the transmission of criminal behavior across peers. Existing studies have documented peer effects in prisons and juvenile correctional facilities (Bayer, Hjalmarsson, and Pozen, 2009; Stevenson, 2017; Ouss and Philippe, 2025; Damm and Gorinas, 2020; Johnsen and Khoury, 2024), neighborhoods (Kling, Ludwig, and Katz, 2005; Dustmann and Landersø, 2021; Billings and Schnepel, 2022; Calderon, 2026), gangs (Sviatschi, 2022; Lindquist et al., 2024; Domínguez, 2026; Disney et al., 2026), and families (Bhuller et al., 2018a,b; Norris, Pecenco, and Weaver, 2021; Byström, 2025; Ahrsjö et al., forthcoming).⁶ I provide evidence on the transmission of criminal behavior within schools. The school environment is important because it

⁴Related work in criminology and psychology also studies youth crime (Sabates, 2008; McAra and McVie, 2010; Welch and Payne, 2012; McGloin and Thomas, 2019).

⁵The paper also relates to the broader literature on peer effects, much of which focuses on academic achievement in schools. See, for example, Hoxby (2000); Sacerdote (2001); Zimmerman (2003); Kang (2007); Figlio (2007); Ding and Lehrer (2007); Carrell, Fullerton, and West (2009); Ammermueller and Pischke (2009); Carrell and Hoekstra (2010); Duflo, Dupas, and Kremer (2011); Arcidiacono et al. (2012); Lavy, Silva, and Weinhardt (2012); Burke and Sass (2013); Angrist (2014); Lu and Anderson (2015); Feld and Zölitz (2017); Murphy and Weinhardt (2020); Carneiro et al. (forthcoming); Tincani (2026); Amit (2026). Related literature studies school segregation through the lens of peer effects (Billings, Deming, and Rockoff, 2014).

⁶Related studies in school settings examine spillovers in problematic behaviors such as drinking and smoking and generally find negative peer effects (Powell, Tauras, and Ross, 2005; Nakajima, 2007; Carrell, Malmstrom, and West, 2008; Kremer and Levy, 2008; Balsa et al., 2011; Fletcher, 2012; Eisenberg, Golberstein, and Whitlock, 2014).

allows me to observe the transition from a non-offender to an offender at a young age, a margin that cannot be studied in prison settings. In addition, compulsory schooling requires adolescents to interact with peers much more often in school than in the neighborhood, making schools a potentially important setting for the transmission of criminal behavior.

Finally, this paper contributes to the literature on school assignment. A large literature develops mechanisms for allocating students to schools based on their preferences ([Abdulkadiroğlu and Sönmez, 2003](#); [Abdulkadiroğlu, Pathak, and Roth, 2005](#); [Abdulkadiroğlu et al., 2005](#); [Abdulkadiroğlu, Pathak, and Roth, 2009](#); [Pathak and Sönmez, 2013](#); [Kamada and Kojima, 2015, 2024](#)). A separate literature examines the consequences of excluding disadvantaged or disruptive students from school and how their allocation across schools can mitigate negative externalities ([McAra and McVie, 2010](#); [Balestra, Eugster, and Liebert, 2022](#)). This paper connects these two strands by applying ideas from school assignment to a counterfactual exercise aimed at reducing crime spillovers across students.

The remainder of the paper is organized as follows. Section 2 describes the institutional background, and Section 3 presents the data. Section 4 describes the empirical strategy and identification assumptions, and Section 5 presents the main results. Section 6 reports the counterfactual assignment exercise. Section 7 concludes. The appendix provides additional details on the data, robustness checks, and heterogeneity analyses.

2 Institutional Background

2.1 Education System in England

In the UK, about 94% of children are enrolled in state schools ([Kynaston and Green, 2019](#)), which are equivalent to “public” or non-private schools in other countries. Compulsory education in England begins at age 5 in primary school and continues through age 16 in secondary school.⁷⁸ Pupils take four national assessments known as Key Stages (KS), typically at ages 7 for KS1, 11 for KS2, 14 for KS3, and 16 for KS4. KS4, also known as the General Certificate of Secondary Education (GCSE), is particularly important for progression to further education. Contrary to the US, grade retention is extremely rare in the UK. After the minimum school-leaving age of 16, some children continue their studies in upper secondary education until age 18, while others pursue alternative pathways, such as vocational training. At the end of upper secondary education, students take A-Levels (KS5), which are an important qualification for entry into higher education, including colleges and universities.

⁷Prior to elementary school, some go to nursery schools or use an area-based holistic support called Sure Start. See [Carneiro et al. \(2025\)](#) for the effects of Sure Start on juvenile criminal behaviors.

⁸In 2013, the minimum leaving age was lifted to 17 in England, and further to 18 in 2015, through the Education and Skills Act 2008. Currently, after age 16, children must be in some form of education or training, although they do not have to attend a state school.

In this paper, I focus on secondary schools for students ages 11 to 16 for two reasons. First, this period falls entirely within compulsory schooling, meaning that children are required to attend school regularly, which limits selection into education. Second, under English law, individuals are criminally responsible from age 10. Since most offenders begin engaging in criminal activity considerably later (Cattan et al., 2023a,b, 2025), focusing on ages 11 to 16 allows me to capture nearly the full range of juvenile crime trajectories.⁹

Under the compulsory education system, children can transfer to another state school when necessary. To do so, their parents must apply either directly to the school in which they want their child to enroll or to the local authority responsible for the preferred school. In principle, if a place is available, a child should be admitted regardless of their background, including their prior offense history.¹⁰

The application process takes time. Parents must apply two months before the transfer, and decisions are typically announced within two to three weeks. Although school principals should be informed about the backgrounds of incoming students, teachers responsible for their classrooms are not informed in advance that they will receive additional students. As a result, they cannot adjust their teaching practices in anticipation of the characteristics of incoming students, including whether those students have a history of offending. Similarly, parents of children in the receiving class are not informed about the transfer in advance. Therefore, students at the destination school have little opportunity to respond to an incoming student’s transfer before it occurs. Most importantly, because the transfer application process takes time, students at the receiving school cannot move to another school in anticipation of the arrival of an offending student. Transfers typically occur at the beginning of a school term: September for the autumn term, January for the spring term, and April for the summer term.¹¹¹²

⁹The timing of progression to a secondary school could be used as another type of exogenous exposure to offenders (Lavy and Yancu, 2025). In my data, we see few offenders at age 11 (Figure 1b), so I do not adopt this identification strategy in this paper.

¹⁰Under Section 86(5)18 of the School Standards and Framework Act 1998, a school can refuse a child’s school transfer if the child is eligible for special education needs (SEN) or has a physical disability that is not compatible with the school’s existing resources (Department for Children, Schools and Families, 2009). However, in principle, schools cannot refuse them on the basis of their behavioral issues. If more transfer requests exceed the school’s available slots, the school must follow strict admission criteria. The criteria typically include whether a child is on a look-after status, whether they have a sibling at the same school, and their commuting distance. However, the offense history should be irrelevant to the criteria.

¹¹In some schools, students are assigned to classrooms primarily according to academic ability, particularly in math, while behavioral issues play little role in classroom assignment. Because offenders tend to exhibit lower educational achievement (Cattan et al., 2023a,b, 2025), classrooms containing students with lower academic achievement may have a higher concentration of offenders than classrooms containing higher-achieving students.

¹²Another way to relocate students to another school is through exclusion. If a student misbehaves severely at school, school principals have the right to exclude the student. However, on only a very few occasions is the student permanently excluded from the school (Cattan et al., 2023a,b). During compulsory education, students need to move to another school. I do not distinguish the effects induced by regular school transfers from those by permanent exclusions, as the number of latter cases is too small to affect the results.

2.2 Criminal Justice System in England

In England, individuals are criminally responsible for their offenses from age 10. If a child aged 10 or older commits an offense, they are subject to police intervention. For relatively minor offenses, the police issue a caution and release the child without further court proceedings. Even in such cases, the offense is formally recorded. More serious offenses may result in the child being accused, charged, and referred to court for further proceedings. If there is sufficient evidence to prosecute, the child may be convicted and receive a sanction. Punishments range from fines to community orders, probation, or incarceration. In most cases, children receive only a caution. The likelihood of imprisonment or a youth rehabilitation order following a juvenile offense is less than 1% (Cattan et al., 2023a,b, 2025). Juvenile offenders may also be excluded from school, but such exclusions are generally temporary. As a result, most juvenile offenders return to regular classroom instruction shortly after committing an offense.

Once a person commits an offense, is caught by the police, and is found guilty, the record is stored in the Police National Computer (PNC) and retained until the person’s 100th birthday. Employers may ask job applicants to provide a criminal record certificate through the government service known as the Disclosure and Barring Service (DBS). During my sample period, these certificates included all youth convictions for offenses committed within the previous 5.5 years and all youth cautions for offenses committed within the previous 2 years. Convictions for particularly serious offenses, including murder, rape, kidnapping, assault on a child, and many sexual offenses and offenses against children, were disclosed regardless of when they were committed. Therefore, a criminal record can have lasting consequences for children’s human capital.¹³

3 Data

This study uses linked administrative records on educational outcomes and criminal justice involvement for nearly the entire population of children in England. The linked dataset is provided through the Data First initiative, a data-linking program led by the Ministry of Justice (MoJ).

3.1 Offense Records

The criminal justice outcomes are drawn from the Police National Computer (PNC). The PNC contains information on individuals who were cautioned or convicted for a recordable offense in England and Wales. The data includes the offenses they were cautioned or convicted for, their

¹³Since November 2020, youth cautions are not automatically disclosed by DBS, although they are still stored in PNC, and some jobs and activities require the thorough checks directly with PNC, such as national security, the police, judicial appointments, and firearms certificates. See the UK government’s website for the recent updates (<https://www.gov.uk/government/publications/filtering-rules-for-criminal-record-check-certificates/filtering-rules-for-dbs-certificates-criminal-record-checks>, last accessed: October 31st, 2026).

criminal history and the sentences they received. My raw PNC extract covers recordable offenses committed between 2000 and 2017, with 1.4 million individuals and 13.7 million panel records. Because the cohorts I focus on in this paper were born between 1995/96 and 1998/99 and stayed in England until the age 17, all individuals have complete records of recordable offenses between ages 10 and 17, covering their entire secondary-school period. Throughout the paper, I define offenses as those recorded in the PNC and offenders as individuals with at least one offense recorded in the PNC, committed by the end of secondary school at the age of 16.

The PNC contains highly detailed information on offense types, with as many as 2,820 distinct categories. Following [Cattan et al. \(2023a,b, 2025\)](#), I aggregate these into ten categories: theft/robbery, battery, drug-related, violent, fraud, weapons, arson/damage, sexual, summary motoring, and a residual category comprising other miscellaneous or unknown offenses, loosely following the MoJ’s classification.¹⁴¹⁵

3.2 Education Records

The education records come from the National Pupil Database (NPD), provided by the Department for Education (DfE). These data cover all students attending state schools in England, who account for approximately 94% of all children ([Kynaston and Green, 2019](#)).¹⁶ Part of the NPD is drawn from the school census, which contains information on students’ demographic characteristics and school enrollment. My extract contains three school censuses per academic year from 2007/08 onward, while it contains only one or two censuses per year in earlier periods.¹⁷ Therefore, in the main analysis, I focus on cohorts for which three census observations are available in every year of secondary school, namely the 1995/96 to 1998/99 cohorts.

The NPD contains information on children’s demographic characteristics, family backgrounds, and educational outcomes. Demographic characteristics include gender, self-reported ethnicity, year and month of birth, whether English is their first language, and whether they are registered as having special educational needs (SEN). In most analyses, I aggregate ethnicity into 12 broader categories to ensure sufficient observations within each group.¹⁸ Measures of family background include children’s area of residence and socioeconomic status. Residential location is measured using lower layer super output areas (LSOAs), which are comparable to US census tracts and

¹⁴See Table A5 for the most common offenses within each crime type.

¹⁵In analyses by offense type, I further aggregate these into four broader categories to ensure sufficient statistical power: property crimes (theft/robbery, fraud, and arson/damage), violent crimes (battery, violent, and sexual), drug-related crimes, and others (weapons, summary motoring, and a residual group).

¹⁶Strictly speaking, students in private schools are assigned a Pupil Matching Reference in NPD if they take a GCSE, but basically they do not have information on enrollment.

¹⁷The autumn Census in my extract starts in 2007, the spring one in 2002, and the summer one in 2006.

¹⁸The 12 categories are a summary of the original 99 detailed categories and loosely based on the DfE’s classification. I use the same categories in my previous papers ([Cattan et al., 2023a,b, 2025](#)), which are “White British,” “Other White,” “African Black,” “Caribbean Black,” “Other Black,” “Indian,” “Pakistani,” “Bangladeshi,” “Other Asian,” “Mixed,” “Others,” and “Unknown.”

contain approximately 1,000 to 3,000 residents each. Socioeconomic status is proxied by eligibility for free school meals (FSMs), which roughly identifies the bottom quartile of children in the household income distribution. SEN and FSM status vary across school censuses, while the other demographic characteristics are effectively time-invariant. For SEN and FSM status, I construct indicators for whether a child was ever eligible by age 11, prior to entering secondary school. For the time-invariant characteristics, I use the latest available record for each individual.

Educational outcomes include school enrollment histories and measures of academic achievement at each Key Stage. School enrollment is observed at the school-census level, with censuses conducted on the first Thursday in October for the autumn census, the third Thursday in January for the spring census, and the third Thursday in May for the summer census. Because most school transfers occur at the beginning of each term, these census records are likely to capture nearly all such transfers.

Coverage of Key Stage assessments differs across cohorts, as shown in Table A1. In the main analysis, I focus on 2.4 million children born between 1995/96 and 1998/99, most of whom have test-score records at ages 7 and 11. However, none of these cohorts have test-score records at age 14. This creates an important limitation for studying the effects on educational attainment. To compare students who have already been exposed to an offender’s transfer with those who have not yet been exposed during secondary school, I need an educational outcome that is observed for both already-treated and not-yet-treated students during secondary school. The age-14 assessment is the only educational measure that satisfies this requirement. Therefore, I use the earlier 1990/91–1993/94 cohorts, for which age-14 test scores are available, when analyzing educational outcomes in Subsection 5.6.¹⁹ One limitation of these earlier cohorts is that they have only one school record per year, so school transfers can be identified only at the annual level.

3.3 Data Linkage

The PNC and NPD are linked using anonymized individual identifiers provided by the MoJ. These identifiers are constructed through deterministic matching based on an individual’s name, sex, date of birth, and postal code of residence. Table A2 reports the linkage rate and offense rate in my sample. As shown in Column (1), each cohort contains approximately 600,000 individuals with complete information on the full set of controls, yielding a total sample of 2.4 million individuals across my target cohorts, 1995/96–1998/99.²⁰

Column (2) reports the fraction of individuals with an offense ID, indicating that they had contact with the criminal justice system at some point, including outside my data period, which varies from 4.5% to 12% for the target cohorts. Columns (3) and (4) report, among individuals

¹⁹I show the descriptives and the replication of the main results using these cohorts in Appendix G. Despite the aforementioned differences, the main results still hold.

²⁰See Appendix A for the data details. Table A1 shows the original sample size, and about 95% of the observations have full information on the controls.

with an offense ID, the fraction with at least one offense recorded during my data period, that is, those successfully matched to an offense record between 2000 and 2017. Column (3) includes offenses committed at any age, while Column (4) restricts attention to offenses committed by the end of secondary school. Roughly half of individuals with an offense ID have at least one offense recorded by the end of secondary school.

Column (5) reports the overall offense rate for each cohort, calculated as Column (2) $\times$ Column (4). The offense rate ranges from 2.8% to 5.9% and exhibits a clear downward trend, consistent with the patterns documented in [Youth Justice Board and Ministry of Justice \(2023\)](#); [Cattan et al. \(2025\)](#). The final column reports the precision of the linkage. Approximately 80% of linked records are based on an exact match on an individual’s name, sex, date of birth, and postal code of residence across the NPD and PNC.²¹

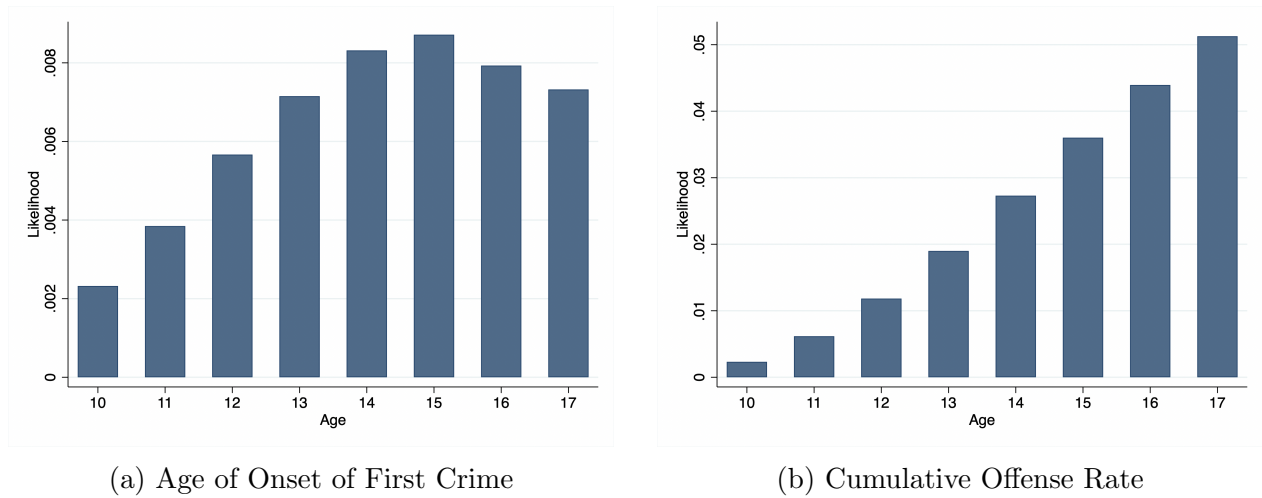


Figure 1. Age of the First Entry into the Criminal Justice System

Notes: Offenders in this figure are defined as individuals in the 1995/96–1998/99 cohorts in the NPD with at least one recordable offense in the PNC committed between ages 10 and 17 that resulted in a caution or conviction. I exclude individuals with an offense record before age 10. Data sources: Ministry of Justice and Department for Education.

3.4 Data Summary

In this subsection, I describe the summary statistics for the resulting dataset. I begin by describing the characteristics of offenders. Figure 1a shows the distribution of age at first offense. The incidence of first offending peaks at age 15, when 0.83% of all children in England commit their first offense.²² Figure 1b shows the cumulative offense rate. By age 16, when secondary school ends, 4.4% of children have committed at least one offense. Thus, the secondary school period

²¹The remaining 20% is based on a fuzzier match, which does not match at least one of these four identifiers. However, even the fuzziest matching is based on the exact forename, surname, and date of birth, ensuring high linkage quality.

²²These figures and some of the subsequent tables follow the style of those in [Cattan et al. \(2025\)](#). Numbers differ from them because I focus on a subset of their cohorts and include both boys and girls. Notably, the offense rate in my sample is significantly lower because younger cohorts and girls have lower offense rates.

covers both the early stages of criminal onset and the age at which initiation peaks, underscoring its importance for studying juvenile crime.

In the main analysis, I restrict the sample to individuals with complete school enrollment records throughout secondary school in England, from the autumn census at age 11 through the summer census at age 16. I exclude students who ever attended a very small secondary school, defined as one in which the number of pupils in a school-grade is fewer than 20. Table A3 describes this sample-selection process in detail. Together, these two restrictions remove 11.8% of observations with complete control variables, and the students remaining in the sample are less disadvantaged.²³²⁴

I further restrict the sample to students who are exposed to an offender’s school transfer at some point by age 16, excluding the never-treated group. This restriction makes the overlap assumption for identification more plausible, as it requires the probability that a school receives an offender to be bounded in $(0, 1)$. Excluding never-treated students ensures that the estimation sample consists only of individuals with a positive probability of exposure. The first half of Table 1 describes this final step of the sample-selection process.²⁵ Columns (1) to (4) compare students who are ever exposed to an offender’s school transfer with those who are never exposed. Students who are ever exposed are slightly more disadvantaged. However, these differences are small after conditioning on individual demographic characteristics, suggesting that this restriction is unlikely to substantially reduce external validity relative to the broader population. Within the resulting estimation sample, the treatment group consists of students who have already been exposed to an offender’s transfer, while the control group consists of students who have not yet been exposed but will be later during secondary school. The final sample contains 1,050,567 individuals. and 15,758,505 panel records.²⁶

The second half of the table compares students who ever transfer schools (“movers”) with those who do not (“non-movers”). Movers are less likely to be male or belong to an ethnic minority, more likely to be native English speakers, and slightly more disadvantaged along several dimensions.²⁷

²³Columns (1) to (4) of Table A3 highlight the differences in observable characteristics, among those in the educational dataset (NPD) with a full set of control variables, between those with a complete record of school censuses during secondary school and those without it. Those who do not have a complete record are more disadvantaged in many ways than those with a complete record. This could be for many reasons: one major possibility is that those without a complete record might be immigrants joining the English school system mid-secondary school, reflected in high minority status and low native-English speaker status. I omit those without a complete record, who account for about 10% of the original sample. Columns (5) to (8) describe the differences between those who had ever been in a very small school where the number of pupils in a school-grade is less than 20 and those who had not, among those with a complete record of school censuses during secondary school. The former group is more disadvantaged, although its sample size is small. Omitting them reduces the sample size further by 1.6%.

²⁴In a robustness check, I further limit the sample to individuals who have never transferred to another school to fully control for school fixed effects in the staggered event study design, at the risk of ignoring many pupils who move schools.

²⁵Columns (9) to (12) of Table A3 show the same differences but in more detail.

²⁶I conduct a robustness check including the “never-treated” group, and the results do not change.

²⁷Since movers are more likely to be offenders and more disruptive, a valid concern is that estimates of peer

Table 1. Summary Statistics and Transfer Experience

	Ever Exposed to An Offender's School Transfer				Among Those in Column (1) Transfer by Age 16			
	Exposed	Never Exposed	Uncond. Diff. [(1) = (2)]	Cond. Diff. [(1) = (2)]	Non-movers	Movers	Uncond. Diff. [(5) = (6)]	Cond. Diff. [(5) = (6)]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Demographics								
Male	.518	.494	.024*** (.001)	.004*** (.001)	.522	.507	.015*** (.001)	.040*** (.002)
Minority	.185	.231	-.046*** (.001)	.000 (.000)	.193	.160	.033*** (.001)	-.013*** (.001)
English as the first language	.902	.882	.019*** (.000)	.001 (.000)	.897	.918	-.021*** (.001)	-.012*** (.001)
Ever with SEN by age 11	.365	.316	.049*** (.001)	.003*** (.001)	.352	.409	-.057*** (.001)	-.061*** (.002)
Ever eligible for FSM by age 11	.296	.243	.054*** (.001)	.004*** (.001)	.273	.368	-.095*** (.001)	-.131*** (.002)
Offender by age 16	.038	.021	.016*** (.000)	.003*** (.000)	.031	.060	-.029*** (.001)	-.048*** (.001)
Any school transfer by age 16	.243	.176	.067*** (.001)	-.006*** (.001)	0	1	-	-
N of school transfers by age 16 any transfer	1.18 [.445]	1.14 [.410]	.036*** (.001)	-.012*** (.002)	-	1.18 [.445]	-	-
B. Education								
Age 5 overall score	-.011 [.967]	.059 [.985]	-.070*** (.007)	.016 (.008)	.019 [.959]	-.107 [.985]	.126*** (.012)	.032 (.019)
Age 7 test score	-.030 [.961]	.109 [.979]	-.139*** (.001)	-.002 (.001)	.001 [.953]	-.126 [.976]	.128*** (.002)	.058*** (.003)
Age 11 test score	-.066 [.983]	.140 [.981]	-.206*** (.002)	-.007*** (.002)	-.016 [.972]	-.225 [1.00]	.209*** (.003)	.092*** (.004)
Age 16 test: taken	1.00	1.00	.000 (.000)	.000 (.000)	1.00	1.00	.000*** (.000)	.001*** (.000)
Age 16 test: 5 A* – G taken	.857	.912	-.056*** (.000)	-.006*** (.001)	.878	.790	.088*** (.001)	.082*** (.002)
Age 16 test: 5 A* – C taken	.557	.676	-.119*** (.001)	-.010*** (.001)	.591	.450	.140*** (.001)	.108*** (.002)
Age 16 test score: point score taken	.018 [.912]	.178 [.913]	-.160*** (.001)	.001 (.001)	.083 [.887]	-.182 [.959]	.265*** (.002)	.320*** (.003)
Sample size	1,050,567	1,038,164	2,088,731	2,088,731	795,011	255,556	1,050,567	1,050,567

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. Conditional differences control for gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. All test scores are standardized within each cohort, except for the indicator variables for the age 16 tests. Early-years overall scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month (Department for Education, Schools Analysis and Research Division, 2010). The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 in constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

On average, movers transfer schools only once during secondary school.

Table A4 compares non-offenders and offenders. The patterns are consistent with those documented in Cattan et al. (2023a,b, 2025): offenders are substantially more likely to be male and White British, to have English as their first language, to come from disadvantaged backgrounds, and to have lower educational achievement. Offenders are also more likely to transfer to another school and, conditional on ever transferring, tend to make slightly more transfers.

Table 2 compares two types of offenders: (1) those who have transferred to another school *after* committing an offense and (2) those who have not, that is, offenders who never transferred schools or who transferred *only before* their first offense. I refer to the former as treaters, because their transfers generate treatment for other students, and to the latter as non-treaters.²⁸ Panel A shows that treaters are more likely to be male, to belong to an ethnic minority, to have English as an additional language, and to be less disadvantaged than non-treaters. They commit fewer offenses overall (Panel B) and commit their first offense at an older age (Panel C). There are only small differences in the types of offenses they commit, both for their first offense and overall, but treaters are more likely to receive a caution, the least severe sanction, rather than a more serious penalty (Panel D).²⁹ These patterns suggest that the peer effects identified from offenders' school transfers may provide a lower bound on the peer effects generated by offenders more generally.

Column (5) of Panel D reports descriptive statistics for treaters based only on offenses committed before their first transfer, providing a measure of the offending history that could plausibly be transmitted to peers. By construction, these offenses occur at younger ages and are more likely to result in a caution. Otherwise, the patterns are broadly similar to those in Columns (1) and (2). This comparison further suggests that the estimated peer effects might be a lower bound.

Table 3 presents the number of school transfers at each age separately for offenders and non-offenders. Offenders here are defined as students who had committed an offense before a given school transfer, while non-offenders are those who had not committed an offense before that transfer, including students who never offend. Columns (1) to (3) of Panel A report school transfers by offenders, while Columns (4) to (6) report transfers by non-offenders. The first term at age 11 marks the start of secondary school and is therefore excluded from my definition of a school transfer. The modal age of transfer is 13 for non-offenders and 14 for offenders. This difference partly reflects the fact that many children have not yet begun offending by age 13 (Figure 1b).

Panel B reports the age at which students are first exposed to a peer's school transfer. Columns (1) to (3) show the timing of first exposure to an offender's transfer, that is, the timing

effects of offenders' transfers might be contaminated by the effects of disruptive students' transfers, not necessarily those of offenders (Billings and Hoekstra, 2023). Therefore, I conduct a robustness check by examining the effects of transfers of non-offenders with disruptive/disadvantaged characteristics, and it confirms that most peer effects are from offenders, not from disruptive non-offenders.

²⁸This sample includes all offenders, not trimmed by the sample selection above, to minimize potential attrition and measurement error in transfer students' characteristics.

²⁹Table A5 lists common youth crimes across different crime types.

Table 2. Offenders

	Offending Status by Age 16				
	Treaters	Non-treaters	Uncond. Diff. [(1) = (2)]	Cond. Diff. [(1) = (2)]	Treaters, Before First Transfers
	(1)	(2)	(3)	(4)	(5)
A. Demographics					
Male	.723	.706	.017*** (.004)	.013*** (.004)	-
Minority	.226	.185	.041*** (.003)	-.003 (.003)	-
English as the first language	.907	.936	-.030*** (.002)	-.007*** (.002)	-
Ever with SEN by age 11	.629	.682	-.054*** (.004)	-.033*** (.004)	-
Ever eligible for FSM by age 11	.607	.688	-.081*** (.004)	-.056*** (.004)	-
B. Crime Onset and Offense Numbers					
Number of offenses	2.65 [3.76]	3.66 [4.99]	-1.02*** (.037)	-1.03*** (.039)	-
Number of offenses per offending day	1.22 [.658]	1.20 [.564]	.019*** (.005)	.012** (.005)	-
C. First Offense					
Age at first offense	13.8 [1.64]	12.5 [1.46]	1.28*** (.012)	1.30*** (.012)	-
Caution	.820	.882	-.062*** (.003)	-.051*** (.003)	-
Theft/Robbery	.327	.332	-.006 (.004)	-.015*** (.004)	-
Battery	.181	.199	-.017*** (.003)	-.010*** (.003)	-
Drug Related	.061	.027	.035*** (.001)	.031*** (.002)	-
Violent	.137	.133	.005 (.003)	.008*** (.003)	-
Fraud	.020	.015	.005*** (.001)	.004*** (.001)	-
Weapons	.037	.039	-.002 (.002)	-.006*** (.002)	-
Arson/Damage	.158	.195	-.038*** (.003)	-.033*** (.003)	-
Sexual	.026	.024	.002 (.001)	.001 (.001)	-
Summary Motoring	.019	.010	.008*** (.001)	.008*** (.001)	-
Others/Unknown	.035	.026	.009*** (.001)	.012*** (.001)	-
Sample size	72,859	20,953	93,812	93,812	-
D. All Offenses					
Offense age	14.1 [1.51]	13.6 [1.62]	.530*** (.008)	.551*** (.008)	12.8 [1.46]
Caution	.526	.466	.061*** (.002)	.064*** (.002)	.690
Theft/Robbery	.317	.317	-.000 (.002)	-.010*** (.002)	.336
Battery	.169	.175	-.006*** (.002)	-.003 (.002)	.190
Drug Related	.064	.045	.019*** (.001)	.015*** (.001)	.029
Violent	.131	.132	-.001 (.002)	.003 (.002)	.129
Fraud	.022	.020	.002*** (.001)	.002*** (.001)	.017
Weapons	.030	.027	.002*** (.001)	-.000 (.001)	.031
Arson/Damage	.151	.166	-.015*** (.002)	-.009*** (.002)	.191
Sexual	.019	.016	.004*** (.001)	.004*** (.001)	.021
Summary Motoring	.024	.022	.002*** (.001)	.002*** (.001)	.015
Others/Unknown	.073	.080	-.007*** (.001)	-.004*** (.001)	.042
Sample size	155,327	62,143	217,470	217,470	35,109

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. Conditional differences control for gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. Offenders in this table are defined as individuals in the NPD with at least one recordable offense in the PNC committed between the September ages of 11 and 16 that resulted in a caution or conviction. Treaters are offenders who transferred to another school after committing their first offense. Non-treaters are offenders who either never transferred during secondary school or committed their first offense only after their final transfer to another secondary school, so that no subsequent transfer could induce treatment. Column (5) reports summary statistics for offenses committed by treaters before their first school transfer. Offense types are based loosely on the offense groups defined by the MoJ. The order of offenses is determined as follows. First, offenses are sorted chronologically for each offender based on the offense date. An offense committed on the earliest date is classified as the first offense, and subsequent offenses are ordered analogously. If multiple offenses occur on the same date, they are ordered according to the severity of the penalty, the length of any custodial sentence, and the amount of any fine, in that order. If these are all identical, one offense is selected at random. Thus, only one offense is retained for each offense date in this table. Panel D includes only offenses committed during the secondary-school period. Data sources: Ministry of Justice and Department for Education.

Table 3. Age of School Transfers

Age in Sept.	Offenders' Transfers			Non-Offenders' Transfers		
	Autumn (1)	Spring (2)	Summer (3)	Autumn (4)	Spring (5)	Summer (6)
A. Transfer						
11	-	423	722	-	18,856	30,998
12	2,059	1,131	1,388	87,990	22,859	31,001
13	4,308	2,074	1,733	146,986	24,126	19,517
14	4,428	2,510	1,838	88,152	20,340	14,235
15	3,104	2,944	936	40,334	16,364	10,217
B. First Contact to Others' Transfers						
11	-	31,518	59,783	-	1,269,477	434,327
12	98,326	75,242	102,565	270,010	44,740	19,738
13	122,645	106,842	112,271	20,164	8,026	2,775
14	115,099	70,539	77,019	3,798	1,704	682
15	55,483	14,881	8,354	2,255	10* ¹	

Notes: Offenders are those who have committed a crime by the time of their school transfer, while non-offenders are those who have never committed a crime or have not yet committed a crime by the time of their school transfer. Panel A counts school transfers rather than individuals; if an individual transfers schools multiple times, each transfer is counted separately, while Panel B counts individuals only once.

*1: There are 10 transfers for Spring and Summer at age 15 combined. Showing them separately is disclosive.

of treatment, while Columns (4) to (6) show first exposure to a non-offender's transfer. Most students are first exposed to a non-offender's transfer at the beginning of secondary school, whereas first exposure to an offender's transfer is more evenly distributed across ages.

Finally, I describe the characteristics of the origin and destination schools of transfer students. Columns (1) to (4) of Table 4 show the origin and the destination schools of offenders' transfers. Offenders tend to transfer to schools with a higher concentration of boys, white British students, children whose first language is English, students eligible for FSM and SEN by age 11, students with lower educational achievement, students with a prior offense history, and fewer students. If I control for other variables, all magnitudes become much smaller, except for the number of students: offenders tend to move to larger schools. Although the disparity is not large, offenders tend to move to a slightly more disadvantaged school.

Columns (5) to (8) report the corresponding statistics for non-offenders' transfers. Differences between origin and destination schools are generally small, except for school size: non-offenders also tend to move to larger schools. Columns (9) and (10) compare the transfer patterns of offenders and non-offenders. Most differences are small, with the notable exception of school size: offenders tend to move to smaller schools than non-offenders. Overall, transfer students tend to move to slightly more disadvantaged schools. This pattern is reasonable because schools accepting new students must have available places. Importantly, as discussed in the next section, my main identification strategy exploits variation in the timing of school transfers by comparing treated

Table 4. Summary Statistics on Schools and School Transfers

	Transfer (school-term) level										
	Offenders' transfers					Non-offenders' transfers				Difference-in-differences	
	All school (1)	Origin school (2)	Destination school (3)	Uncond. Diff. [(2)=(3)] (4)	Cond. Diff. [(2)=(3)] (5)	Origin school (6)	Destination school (7)	Uncond. Diff. [(6)=(7)] (8)	Cond. Diff. [(6)=(7)] (9)	Uncond. Diff. [(4)=(8)] (10)	Cond. Diff. [(4)=(8)] (11)
Male	.531 [.174]	.536 [.133]	.544 [.122]	-.008*** (.001)	-.002** (.001)	.515 [.126]	.516 [.125]	-.000* (.000)	-.001*** (.000)	-.008*** (.001)	-.002 (.001)
Minority	.216 [.256]	.193 [.235]	.185 [.222]	.008*** (.002)	-.002** (.001)	.192 [.242]	.184 [.231]	.008*** (.000)	.000** (.000)	.000 (.002)	-.003*** (.001)
English as the first language	.885 [.174]	.890 [.167]	.899 [.153]	-.009*** (.002)	-.004*** (.001)	.891 [.174]	.896 [.165]	-.005*** (.000)	-.001*** (.000)	-.004** (.002)	-.003*** (.001)
Ever with SEN by age 11	.416 [.199]	.422 [.143]	.444 [.158]	-.022*** (.001)	-.014*** (.001)	.387 [.122]	.383 [.123]	.003*** (.000)	-.009*** (.000)	-.026*** (.001)	-.004*** (.001)
Ever eligible for FSM by age 11	.331 [.197]	.378 [.188]	.397 [.200]	-.019*** (.002)	-.006*** (.001)	.325 [.178]	.319 [.173]	.006*** (.000)	-.005*** (.000)	-.025*** (.002)	.000 (.001)
Age 5 overall score	-.152 [.686]	-.013 [.060]	-.015 [.049]	.003 (.003)	-.004 (.003)	-.006 [.046]	-.005 [.045]	-.000 (.000)	-.000 (.000)	.003 (.003)	-.002 (.003)
Age 7 test score	-.210 [.744]	-.132 [.307]	-.168 [.338]	.036*** (.003)	.002 (.002)	-.048 [.306]	-.042 [.316]	-.005*** (.001)	.004*** (.000)	.041*** (.003)	-.007*** (.002)
Age 11 test score	-.180 [.590]	-.201 [.283]	-.227 [.283]	.026*** (.003)	.008*** (.002)	-.138 [.280]	-.127 [.281]	-.010*** (.001)	.001*** (.000)	.036*** (.003)	.003 (.002)
Offender	.042 [.101]	.054 [.103]	.091 [.150]	-.036*** (.001)	-.028*** (.001)	.022 [.035]	.027 [.047]	-.005*** (.000)	-.006*** (.000)	-.032*** (.001)	-.027*** (.001)
N of students	154 [75.6]	169 [73.1]	179 [92.8]	-9.28*** (.800)	-16.2*** (.669)	164 [65.0]	210 [94.3]	-46.3*** (.153)	-45.4*** (.143)	37.0*** (.815)	31.1*** (.695)
Sample size	4,559	21,803	21,803	43,606	43,606	559,419	559,419	1,118,838	1,118,838	1,162,444	1,162,444

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. Offenders in this table are defined as individuals in the NPD who, by the time of their school transfer, had committed at least one recordable offense in the PNC between the September ages of 11 and 16 that resulted in a caution or conviction. Non-offenders are individuals who had not committed such an offense by the time of their transfer, including those who never committed an offense. Column (1) reports averages of school-level characteristics, calculated by first taking the average of each outcome within each school. Columns (2), (3), (6), and (7) report leave-one-out averages of school-level outcomes. The unit of observation is a school transfer. Conditional differences control for school-level measures of gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. All test scores are standardized within each cohort. Early-years test scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month (Department for Education, Schools Analysis and Research Division, 2010). The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 in constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

students with not-yet-treated students, allowing me to control for school characteristics.

4 Identification

This section outlines the empirical strategy used to identify the causal peer effects of exposure to others' criminal behavior at school. Exposure to offenders is generally endogenous. For example, a student may be bullied and subsequently commit an offense against the students who bullied them in retaliation. In such a case, exposure to the retaliatory offense is clearly not exogenous. Identifying causal peer effects therefore requires plausibly exogenous variation in exposure to offenders.

This paper exploits the timing of offenders' school transfers as a plausibly exogenous shock. As discussed in Subsection 2.1, these transfers cannot be anticipated by other students, their parents, or even the teachers responsible for the receiving classrooms. Therefore, conditional on an offender transferring into a school-cohort, the timing of peers' exposure to that offender is plausibly exogenous to their behavior prior to the transfer at the school-term level.³⁰ In other words, after conditioning on a rich set of observable characteristics, the remaining variation in exposure generated primarily by the timing of offenders' school transfers is plausibly random.³¹

I formally define the treatment and control groups. The treatment group includes children who are already exposed to an offender's school transfer into their school-grade in secondary school. The control group contains those not yet exposed to an offender's school transfer into their school-grade but will be later during secondary school. If treatment timing is random, this comparison gives us a causal estimate of the peer effects of exposure to an offender.

Because the treatment and control groups differ in composition, and those in the control group eventually move to the treatment group, a simple balancing test between the two groups is difficult. Therefore, to see balancing, I compare those who are treated earlier and later. The former group is more likely to stay in the treatment group as they are treated earlier, while the latter are more likely to be used as the control group. I split the sample into approximately equal groups: those who received an offender by the Age 13 Autumn Census, and those exposed later, by the end of secondary school.

Table 5 compares the pre-secondary school characteristics between these groups. Although there are some row differences, there is no systematic difference remaining after conditioning on demographic variables, which I use as control variables in the main specification.³²

³⁰As explained in Subsection 2.1, the standard timing of school transfers is at the beginning of school terms. The randomness I exploit here is across the timing of school transfers at the start of terms over secondary school.

³¹Balsa et al. (2023) uses the same strategy in their analysis of the peer effects of students' mental disorders. It is close to the identification in Bayer, Hjalmarsson, and Pozen (2009), which views the timing of juvenile correctional facility inmates' entry into facilities as a random shock to incumbent inmates. It is also close to Fadlon and Nielsen (2019, 2021), which view the timing of a health shock as exogenous variation.

³²There are statistically significant differences in FSM and Age 11 test scores, but the magnitude is very small and not economically meaningful.

Table 5. Balancing between Early and Late Treated

	Treated Earlier (1)	Treated Later (2)	Uncond. Diff. (3)	Cond. Diff. (4)
Male	.520	.517	.003*** (.001)	-.001 (.001)
Minority	.174	.194	-.020*** (.001)	.001 (.001)
English as the first language	.907	.897	.010** (.001)	.000 (.000)
Ever with SEN by age 11	.376	.356	.021*** (.001)	-.001 (.001)
Ever eligible for FSM by age 11	.313	.282	.031*** (.001)	-.002** (.001)
Age 5 overall score	-.039 [.979]	.009 [.957]	-.048*** (.010)	-.023 (.016)
Age 7 test score	-.059 [.963]	-.002 [.956]	-.057*** (.002)	-.003 (.002)
Age 11 test score	-.102 [.987]	-.033 [.978]	-.069*** (.002)	-.007*** (.003)
Sample size	490,079	560,488	1,050,567	1,050,567

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and standard errors clustered at the treatment level are reported in parentheses. I use the earliest available secondary-school record. Conditional differences control for gender, ethnicity, an indicator for English as the first language, and SEN and FSM status at age 11, year and month of birth, and secondary school ID, unless the variable is itself the outcome. All test scores are standardized within each cohort. Early-years overall scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month (Department for Education, Schools Analysis and Research Division, 2010). The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 in constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

The conceptual framework can be written as follows.³³ Because the timing of offenders' school transfers varies across individuals and the treatment effects may be heterogeneous, I adopt a staggered event-study design, following Callaway and Sant'Anna (2021).

Let $Y_{i,t}(g)$ denote the potential outcome for individual i at time $t = 1, \dots, T$ if i were treated at time $g > 1$, and $Y_{i,t}(0)$ denote the potential outcome without treatment. Let G_i be the treatment time for individual i . The average treatment effect on the treated (ATT) at time t for those treated at g can be written as:

$$ATT(g, t) = E[Y_{i,t}(g) - Y_{i,t}(0) | G_i = g].$$

Under the assumption I describe below, $ATT(g, t)$ can be identified as

$$ATT(g, t) = E[Y_{i,t} - Y_{i,g-1} | G_i = g] - E[Y_{i,t} - Y_{i,g-1} | G_i > t],$$

³³See Appendix B for the formal setup of the estimation.

for $g \leq t$.

Then, to obtain the outcome of interest, the ATT with length of exposure to the treatment $e = t - g \geq 0$, I aggregate $ATT(g, t)$ as

$$\theta(e) = \sum_{g \in \mathcal{G}_e} \left(\frac{N_g}{\sum_{g' \in \mathcal{G}_e} N_{g'}} \right) ATT(g, g + e),$$

where $N_g = \sum_i \mathbf{1}\{G_i = g\}$, the number of children exposed at time g , and $\mathcal{G}_e = \{g : g + e \leq T\}$, the collection of treatment group g observed at time e .

In the main specification, I condition on individual characteristics X_i including gender and FSM and SEN status by age 11; an indicator for whether English is the individual's first language; a minority indicator; a linear age term; and a linear month-of-birth term. Standard errors are clustered at the level of exposure, defined by the school-cohort-timing cell. I drop records at the time of transfer to avoid conflating the effects of exposure to offenders with those of entering a new school environment. I estimate the aggregated ATT by a doubly robust estimator based on stabilized inverse probability weighting and ordinary least squares (Sant'Anna and Zhao, 2020).³⁴

To identify the average treatment effect on the treated, I require several assumptions. First, I assume that post-treatment outcomes would follow parallel trends in the absence of treatment conditional on observed covariates: for $t \geq g$,

$$E[Y_{i,t}(0) - Y_{i,t-1}(0) | X_i, G_i = g] = E[Y_{i,t}(0) - Y_{i,t-1}(0) | X_i, G_i > t].$$

In my context, without exposure to an offender's transfer, behavioral changes of children treated at g between time $t - 1$ and t would be the same as those of children not yet treated at t . Following standard practice, I examine pre-treatment trends and use their similarity as supporting evidence for the plausibility of parallel counterfactual trends after treatment (Roth, 2026).

Second, I assume no treatment anticipation: for $t < g$,

$$E[Y_{i,t}(g) | X_i, G_i = g] = E[Y_{i,t}(0) | X_i, G_i = g].$$

That is, children and their parents cannot predict whether an offender will transfer into their school and grade and therefore cannot adjust their behavior in advance. As discussed in Subsection 2.1, children, their parents, and even the teachers responsible for the receiving classroom do not know whether they will receive a transfer student before the transfer occurs. This institutional setting

³⁴Although the data form a panel, I estimate the model using the repeated cross-section version of the Callaway and Sant'Anna (2021) estimator. I omit each individual's observation in the period of their own transfer, so the panel is unbalanced by construction: every transferred individual is missing one period. A balanced-panel estimator would therefore discard every transferred individual. This is a substantial loss of information given that offending is relatively rare and offenders are more likely to move (Table A4). The repeated cross-section estimator instead uses all available observations. Under the assumptions described below, the treatment effects on the treated can be identified.

makes the no-anticipation assumption plausible. Moreover, any anticipatory behavioral response would be reflected in the pre-treatment trend.³⁵

Third, I assume overlap: the probability of each school receiving an offender is bounded in $(0, 1)$. I exclude the never-treated group from the sample so that the estimation sample consists only of students who are exposed at some point during secondary school, making the overlap assumption more plausible.

Finally, I assume treatment irreversibility, meaning that once students are exposed to an offender, their treatment status cannot revert to untreated. In this setting, this assumption is natural because exposure cannot be undone once contact with an offender has occurred.³⁶

5 Results

5.1 Peer effects of crime

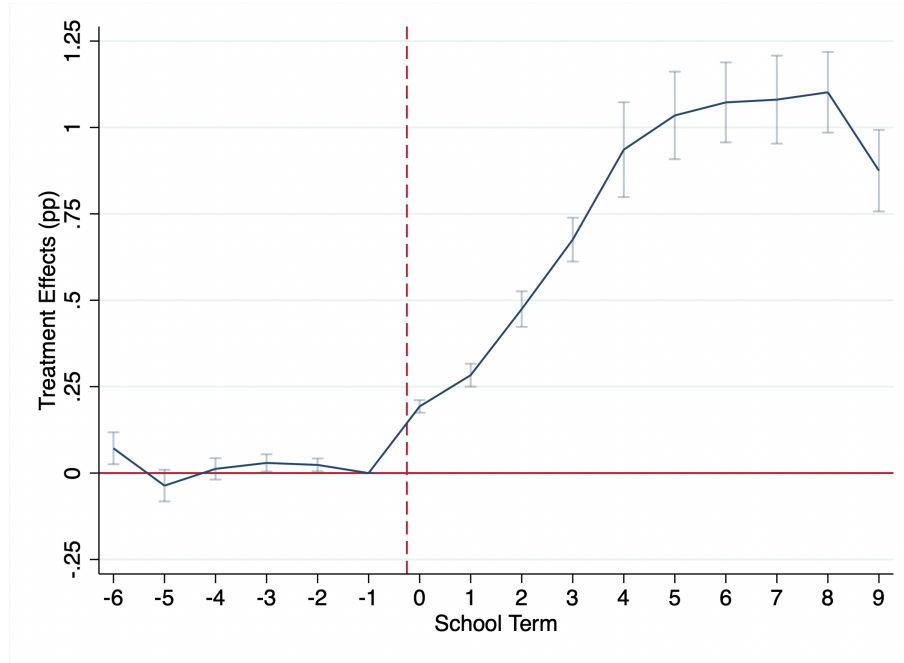


Figure 2. Effects of Receiving an Offender's Transfer on Offense Initiation

Note: This figure reports the treatment effects of exposure to offenders transferring into students' school-grades on their peers' criminal behavior. The treatment effect is normalized at $t = -1$ and multiplied by 100 for visibility. The treatment group consists of students who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#). The bars show 95% confidence intervals based on standard errors clustered at the treatment level. Data sources: Ministry of Justice and Department for Education.

³⁵I check their behaviors in offending and school transfers to avoid treatment anticipation, which would show up as a pre-trend, and there is no evidence of anticipation, supporting plausibility of this assumption.

³⁶As a robustness check, I repeat the analysis censoring the panel records if pupils are treated a second time, and the results do not change. This robustness suggests the validity of the irreversibility assumption.

Having discussed the richness of the data and the plausibility of the identification strategy, I now turn to the main findings on the peer effects of crime in schools. I begin with the visual evidence. Figure 2 plots the treatment effects on offense initiation over six pre-treatment and nine post-treatment school terms, spanning approximately two years before and three years after an offender transfers into a peer’s school and grade. I use a two-year pre-treatment window and a three-year post-treatment window because treatment is most common around the Autumn Census at age 13 (Table 3), and these periods approximately cover the full duration of secondary school for this population. I normalize the treatment effect at -1 and multiply the estimates by 100, so the y-axis in Figure 2 can be interpreted as the percentage-point change in the probability of offense initiation. First, there is no evidence of a systematic pre-trend difference between the treatment and control groups. This finding supports the plausibility of both the parallel-trends assumption in the absence of treatment and the no-anticipation assumption. Turning to the main result, exposure to an offender’s school transfer increases the probability that peers begin offending. At the time of the transfer, the probability of becoming an offender increases by 0.193 percentage points. This magnitude implies that approximately one new offender is induced for every three offender transfers.³⁷ The effect continues to grow for approximately two years, reaching 1.1 percentage points. Thereafter, it levels off but does not disappear, indicating a persistent gap in cumulative offense initiation between treated and not-yet-treated students.³⁸

Figure 3 shows the effects on overall criminal activity: whether pupils commit an offense in each term and the number of offenses they commit in each term, including among students who had already begun offending. Although the pre-trends are slightly noisier, there is a clear increase at time 0.³⁹ Once pupils are exposed to an offender’s school transfer at time 0, the treatment effects on both the probability of committing an offense and the number of offenses increase sharply. These effects continue to rise for approximately one year, remain positive for around two years, and fade toward zero within three years.

There are two possible explanations for why the treatment effect on overall criminal activity fades over time. One is that the increase in offending among treated pupils is temporary. The other is that pupils in the control group begin offending before they themselves are treated, so that treatment primarily shifts the timing of criminal activity forward. Taken together with the offense-initiation results in Figure 2, the evidence is more consistent with the former interpretation.

³⁷This number comes from a back-of-the-envelope calculation: (the effect size) $\times$ (the number of students in the destination school in Table 4) = $0.00193 \times 178 = 0.345$, meaning that there is a new offender at the destination school upon the arrival of one offender with 34.5% probability.

³⁸If we examine the effects separately on caution and conviction experiences in Figure F1, we see the same pattern.

³⁹According to Table 4, the destination school is slightly more disadvantaged. This small pre-trend might reflect this fact. To assess how much this pre-trend matters, I conduct a robustness check using the method proposed by Rambachan and Roth (2023), which bounds the treatment effects based on the pre-trend patterns. Figures C7 and C8 show the result. Even under a very conservative assumption and regardless of the method, the results are robust and we see clear treatment effects.

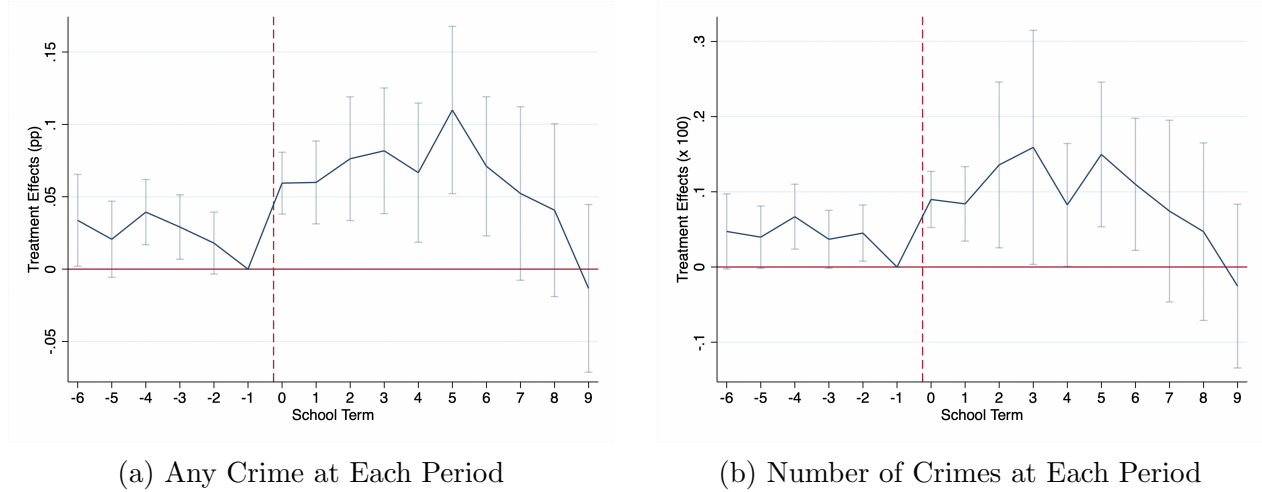


Figure 3. Effects of Receiving an Offender's Transfer on Overall Criminal Activities

Note: This figure reports the treatment effects of exposure to offenders transferring into students' school-grades on their peers' criminal behavior. The treatment effect is normalized at $t = -1$ and multiplied by 100 for visibility. The treatment group consists of students who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. I omit observations when they move to another school. The estimation follows [Callaway and Sant'Anna \(2021\)](#). The bars show 95% confidence intervals based on standard errors clustered at the treatment level. Data sources: Ministry of Justice and Department for Education.

Table 6. Effects of Offenders' Transfers

Effects on:	Crime Onset (1)	Any Crime (2)	Number of Crimes (3)
Treatment Effects at $t = 0$	.193*** (.009)	.059*** (.011)	.090*** (.019)
Romano-Wolf Step-down p-value	[.000]	[.000]	[.000]
Pre-Treatment Mean	1.50	.273	.369
Relative Effect Size	12.9%	21.8%	24.3%
Sample Size (Panel Count)	15,458,144	15,458,144	15,458,144
Number of Individuals	1,050,567	1,050,567	1,050,567

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. p-values based on the Romano-Wolf step-down procedure are reported in brackets ([Romano and Wolf, 2005](#); [Clarke, Romano, and Wolf, 2020](#)). Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. Relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

During the first two years after treatment, the effect on offense initiation continues to grow while the effect on overall offending remains positive. Once the latter approaches zero, the initiation effect also stops increasing. This pattern suggests that exposure to an offender produces a relatively short-lived increase in overall criminal activity but a persistent difference in the probability of acquiring an offense record. Such an effect can have longer-run consequences for human capital and potentially for labor market outcomes.⁴⁰

Table 6 summarizes the treatment effects at $t = 0$, together with the pre-treatment mean, calculated as the mean of each outcome one school term before treatment.⁴¹ All coefficients are multiplied by 100 for visibility. Column (1) reports the effect on whether pupils have begun offending by each school term. The contemporaneous treatment effect within one school term, approximately four months, is 0.193 percentage points, corresponding to a 12.9% increase relative to the pre-treatment mean of 1.50%. This number seems massive, but it is not an outlier in the literature: Billings, Deming, and Ross (2019) show that a one-standard-deviation increase in disadvantaged students (8.3 students) among the same school peers increases the probability of ever being arrested by 3.9 percentage points: 0.47 percentage points per disadvantaged student, equivalent to a 23% increase. Columns (2) and (3) report the effects on whether pupils commit an offense in a given term and on the number of offenses committed in that term, respectively. The treatment effect on the former is 0.00059, representing a 21.8% increase relative to the pre-treatment mean of 0.00273, while the treatment effect on the latter is 0.00090, corresponding to a 24.3% increase.

Because the three outcome variables are likely to be highly correlated, testing them separately raises the probability of making a Type I error. Therefore, I apply the Romano-Wolf stepdown procedure to control the family-wise error rate (Romano and Wolf, 2005; Clarke, Romano, and Wolf, 2020). I report the adjusted p-values, and the main results remain statistically significant at the 1% level. Overall, the findings show that exposure to an offender’s school transfer produces a sizable immediate increase in criminal activity and a persistent increase in the probability of entering the criminal justice system. In most of the analyses that follow, I focus on the effect on crime onset because it most directly captures whether exposure induces a student to acquire an offense record.

5.1.1 Intensity of exposure

I have shown that exposure to an offender generates crime spillovers. A natural next question is whether intensity of exposure to offenders amplifies crime transmission.

⁴⁰As discussed in Subsection 2.2, during my data period, crime records could be shared with employers at the job applications upon their request.

⁴¹Since the offending activities are monotonically increasing almost always during secondary school, according to Figure 1, these means could be slightly underestimated. However, this is still the best possible reference point for understanding how large the treatment effects are, so I show them despite this caveat.

Table 7. Treatment Intensity

Tercile in treatment intensity:	Lowest (1)	Middle (2)	Highest (3)
Treatment Intensity (%)	.357	.510	.859
Treatment Effects at $t = 0$	.099*** (.014)	.123*** (.015)	.325*** (.022)
Pre-Treatment Mean	1.28	1.41	1.79
Relative Effect Size	7.68%	8.73%	18.1%
Sample Size (Panel Count)	5,165,416	5,027,003	5,265,725
Number of Individuals	350,482	333,455	348,441

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Treatment intensity is measured as the ratio of the number of offenders transferring into a school-grade to the number of students in that school-grade. Data sources: Ministry of Justice and Department for Education.

First, I examine whether a larger fraction of transfer offenders amplifies the spillovers. I define treatment intensity as the fraction of offenders transferring into a school-grade relative to the number of pupils in that grade and divide the sample into terciles. Higher treatment intensity means a larger share of students are incoming offenders, increasing the likelihood of peers interacting with the transfer offender.⁴²

Table 7 reports the results. The top row shows large heterogeneity in mean treatment intensity across quartiles. In the lowest-intensity group, transfer offenders account for 0.357% of students in the school-grade, compared with 0.859% in the highest-intensity group. The second row reports the contemporaneous treatment effects on offense initiation. The effects increase monotonically with treatment intensity. This result means that if peers are exposed to offenders' transfers with a greater intensity, they are more likely to initiate offending.⁴³

Since children spend more time on social interactions in school than in the neighborhood, exposure to offenders' transfers might be more intense if offenders move into school than in cases where offenders move into the neighborhood. Therefore, I compare the effects in schools and in the neighborhoods.

I use two geographic units for this analysis. The first is the Middle-layer Super Output Area (MSOA), which contains roughly 2,000 to 6,000 households and 5,000 to 15,000 residents. Given that the average secondary school size is about 150 (Table 4), MSOAs are broadly comparable in size to school catchment areas. The second is the Lower-layer Super Output Area (LSOA), a finer geographic unit containing 400 to 1,200 households and 1,000 to 3,000 residents. LSOAs are similar in size to US census tracts.

Column (1) replicates Table 6, while Columns (2) and (5) report the effects of offenders' transfers into the same locality.⁴⁴ Comparing Columns (1), (2), and (5), there are some neighborhood peer effects, consistent with prior studies (Dustmann and Landersø, 2021; Billings and Schnepel, 2022), but the magnitudes are smaller.

To make the comparison clearer, I next distinguish the two channels more directly by comparing offenders who move into the same school but a different neighborhood with those who move into the same neighborhood but a different school.⁴⁵ The estimated school effects in Columns (3) and (6) are substantially larger than the neighborhood effects in Columns (4) and (7) in both the MSOA and LSOA comparisons. These results suggest that the peer effects of crime operate

⁴²In most cases, the number of transfer offenders is one, so the variation mainly comes from the denominator, the number of students.

⁴³On the other hand, the main results may be driven by unusually large numbers of transfer students entering a school-grade at the same time. To address this concern, I exclude observations in which the number of transfer students is above the 95th percentile of the distribution and repeat the main analysis. Figure C14 confirms that the results are unchanged.

⁴⁴The move into the neighborhood is measured by the change in the LSOA registered in the school census. Therefore, I can measure only up to three times a year. If they move more than once during the school census period, this will introduce measurement error, although moving to another place within 4 months is very unlikely.

⁴⁵The LSOA analysis has relatively few observations because children living very close to one another are also likely to attend the same school, making the estimates slightly noisier.

Table 8. Locality

Locality Measure:	School	MSOA			LSOA		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment Effects at $t = 0$	.193*** (.009)	.138*** (.009)	.141*** (.027)	.070** (.030)	.168*** (.018)	.130*** (.023)	.050 (.073)
Pre-Treatment Mean	1.50	1.51	1.35	1.25	2.16	1.59	1.46
Relative Effect Size	12.9%	9.12%	10.4%	5.60%	7.78%	8.17%	3.44%
School	Same	Any	Same	Different	Any	Same	Different
Locality	Any	Same	Different	Same	Same	Different	Same
Sample Size (Panel Count)	15,458,144	16,617,101	3,555,911	2,194,080	5,969,939	5,636,557	502,607
Number of Individuals	1,050,567	1,127,226	221,504	141,222	406,544	347,076	31,942

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. MSOA stands for Middle Layer Super Output Area, a geographic unit containing approximately 5,000 to 15,000 residents. LSOA stands for Lower Layer Super Output Area, a geographic unit containing approximately 1,000 to 3,000 residents. In Columns (3) and (6), I exclude students who are treated at the neighborhood level before or at the same time as they are treated at the school level. Similarly, in Columns (4) and (7), I exclude students who are treated at the school level before or at the same time as they are treated at the neighborhood level. Data sources: Ministry of Justice and Department for Education.

through an intense exposure to offenders in school.

If crime transmission occurs mainly through interactions in schools, spillovers should be stronger during school terms and on weekdays. Table D1 confirms this pattern: treatment effects are substantially larger during school terms than holidays and larger on weekdays than weekends. This provides further evidence that transmission occurs primarily through school interactions.⁴⁶ This set of results shows that intensity of exposure to offenders increases crime spillovers, suggesting that interactions with transfer offenders matter.

5.2 Potential Mechanisms

So far, I have shown that an offender's school transfer increases peers' criminal activity, and that the intensity of exposure to offenders matters for crime transmission. This subsection investigates the potential mechanisms behind these effects. I consider two main channels: skill transmission and role modeling. First, peers exposed to an offender may acquire crime-specific skills, as documented in juvenile correctional facilities ([Bayer, Hjalmarsson, and Pozen, 2009](#)). Second, peers may view an incoming offender as a role model and imitate them. An extensive literature documents role-

⁴⁶Children might commit a different type of crime in and out of school. See [Boshoff, Machin, and Sandi \(2026\)](#) for a discussion of this.

model effects in contexts where role models are scarce.⁴⁷ The evidence suggests that role modeling is an important channel, whereas skill transmission is not.

5.2.1 Skill Transmission not as a major channel

First, I show that skill transmission is not a major channel. I start by describing the transfer offenders' behaviors after arrival. I examine whether offenders keep committing crime after moving. I use the same event-study specification for transfer offenders, with the event as their transfer. One important note is that the decision to move is endogenous. Therefore, I cannot interpret these estimates as causal. Another is that I focus on children who have committed a crime at the time of transfer, and the comparison is between those who have moved and those who will move later. Therefore, the comparison group for the long-run effects contains many non-offenders, who will commit a crime later. Therefore, by construction, the long-run estimate might be biased downwards. Despite these caveats, documenting offenders' behavior is informative.

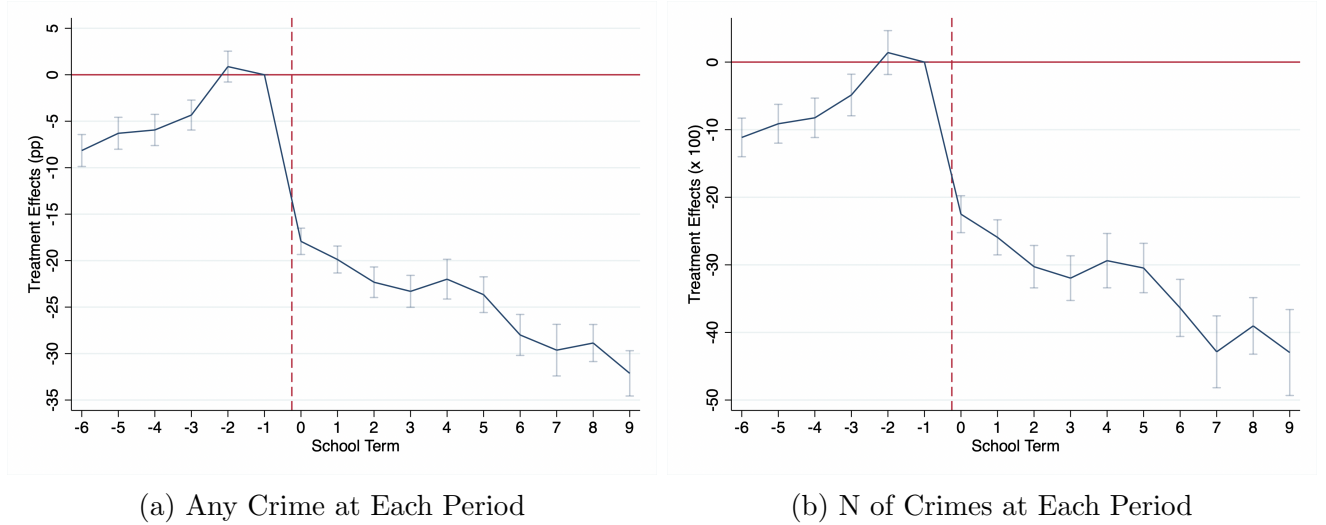


Figure 4. Descriptive Evidence: Behavior of Transfer Offenders After Their Move

Note: This figure reports the treatment effects of moves on transfer offenders themselves. The treatment effect is standardized at $t = -1$. The treatment group is offenders who moved at $t = 0$, while the control group is those who will move later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).

Figure 4 shows the probability and number of crimes at each period before and after the school transfer. Clearly, offenders stop committing crimes after the school transfer. This suggests that peers learn how to commit crime by witnessing transfer offenders' criminal activities is unlikely to be a major channel.

I also examine whether joint crime increases. Peers might commit crime together with transfer

⁴⁷Most of these studies focus on female role models in politics, educational choices, and mathematics, where gender gaps are large ([Neumark and Gardecki, 1998](#); [Beaman et al., 2009, 2012](#); [Bhalotra, Clots-Figueras, and Iyer, 2018](#); [Porter and Serra, 2020](#); [Riley, 2024](#)). [Serra \(2025\)](#) reviews role-model effects more broadly in developing countries.

offenders, such as organized crimes.⁴⁸ Joint crime is defined as peers’ crime committed on the same day as transfer offenders in the same school-cohort, after offenders’ arrival.⁴⁹ I do not require the offenses to be of the same type, since offenders in organized crime may take on different roles and therefore commit different types of offenses. I examine whether peers initiate offending with a joint crime. Figure D1 shows that this measure accounts for only a tiny share of the overall treatment effect, providing little evidence that coordinated offending with transfer offenders is a major mechanism.

If exposure to offenders helps peers accumulate criminal capital by teaching how to commit crimes, exposed students should be more likely to commit the same types of crimes that the transfer offenders committed before their transfers. I examine whether this mechanism explains the treatment effects. To ensure sufficient statistical power, I aggregate the 10 crime categories into four broader groups: violent crimes (battery, violence, and sexual offenses), property crimes (theft, arson/damage, and fraud), drug-related crimes, and others (weapons, motoring, and other offenses).

Table 9. Crime Types

Crime Types:	All Crimes		Violent Crimes		Property Crimes		Drug Crimes		Other Crimes	
	Same (1)	Different (2)	Same (3)	Different (4)	Same (5)	Different (6)	Same (7)	Different (8)	Same (9)	Different (10)
Treatment Effects at $t = 0$	.089*** (.007)	.082*** (.008)	.068*** (.011)	.104*** (.014)	.106*** (.010)	.082*** (.009)	.019 (.034)	.179** (.072)	-.027 (.018)	.054 (.046)
Pre-Treatment Mean	.714	.806	.661	1.30	1.08	.714	.121	2.21	.155	1.96
Relative Effect Size	12.4%	10.2%	10.2%	7.98%	9.82%	11.5%	15.8%	8.08%	-17.7%	2.77%
Sample Size (Panel Count)	12,184,487	12,184,487	6,182,281	6,182,281	8,351,676	8,351,676	699,820	699,820	1,628,495	1,628,495
Number of Individuals	827,507	827,507	420,119	420,119	567,629	567,629	47,586	47,586	110,664	110,664

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender’s school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. I exclude observations that were simultaneously exposed to an offender with a different type of offense history at the time of treatment. The estimation follows Callaway and Sant’Anna (2021), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table 9 shows the results. The first two columns summarize the overall effects: whether peers start committing the same/different type of crimes. We do not see a large difference. Even if I look into the heterogeneity across crime types, we do not see such a tendency. These pieces of evidence show that it is unlikely that transfer offenders commit a crime with peers or teach them how to commit a crime.

5.2.2 Role-modeling as a potentially important mechanism

The remaining mechanism is role modeling. Even if transfer offenders neither engage in criminal activity nor teach criminal skills, their presence may affect peers indirectly: peers may imitate

⁴⁸Previous work shows that criminal networks can be sustained by a small number of “key players” (Lee et al., 2021), and that gang recruitment in school is an important channel for offense participation (Byström, 2025).

⁴⁹Therefore, I cannot define this for the pre-treatment period.

transfer offenders and begin offending. I show several heterogeneity analyses whose results are consistent with this channel, together with the results of exposure intensity.

First, I examine heterogeneity by the presence of existing offenders before treatment. Schools differ substantially in their existing offender density, which I define as the number of existing offenders divided by the number of children in the school-grade in the term before treatment. If there are few offenders, a transfer offender may become a new role model because such students are rare in the school, which may increase the treatment effect. By contrast, if offenders make up a large share of the grade, a transfer offender may be less likely to become a role model because many such students are already present, making the treatment effects small.

Table 10. Offender-Dense School

Quartile in offender density:	Lowest (1)	Middle (2)	Highest (3)
<i>Panel A: Effects on Crime Onset</i>			
Treatment Effects at $t = 0$	.023*** (.009)	.243*** (.055)	.034 (.128)
Pre-Treatment Mean	.324	1.06	3.11
Relative Effect Size	7.15%	22.9%	1.11%
<i>Panel B: Effects on Number of Crimes</i>			
Treatment Effects at $t = 0$	-.012 (.015)	.041 (.029)	.284*** (.097)
Pre-Treatment Mean	.100	.274	.738
Relative Effect Size	-12.0%	14.8%	38.5%
Sample Size (Panel Count)	5,101,909	5,084,637	5,049,704
Number of Individuals	344,858	344,858	344,858

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Offender density is calculated as the ratio of the number of offenders in a school-grade to the total number of students in that school-grade, measured one school term before treatment. Data sources: Ministry of Justice and Department for Education.

Table 10 reports the results. I divide the sample into terciles based on the density of offenders in school cohorts at the treatment timing. The pre-treatment means in Panel A illustrate the

differences across terciles: in the least offender-dense schools, only 0.324% of pupils have an offense record, compared with 3.11% in the most offender-dense schools. Panel A shows the effects on offense initiation. The effects on offense initiation are large and statistically significant except in the most offender-dense schools. This result is consistent with the above interpretation: a transfer offender becomes a new role model in the low offender-dense school and induces a large treatment effect, while the treatment effect is zero in the high offender-dense school, as there are many such role models.⁵⁰

The arrival of a new offender affects not only non-offender peers, but also offender peers. Panel B shows the impact on overall criminal activity, including those who have initiated offending at the time of treatment. We see a large treatment effect in high offender-density schools. This suggests that transfer offenders could inspire existing offenders. This pattern is consistent with evidence from prison and juvenile correction facility settings that offenders can increase other offenders' criminal activity (Bayer, Hjalmarsson, and Pozen, 2009; Damm and Gorinas, 2020). We do not see such a pattern in low offender-dense schools, as there are not many existing offenders to be affected.⁵¹

To further support the mechanism of role-modeling, I examine heterogeneity by seniority of transfer offenders. If an offender transfers to the same school but a different grade, peers have fewer opportunities to interact with the offender.⁵² At the same time, such older offenders may be more salient role models than younger offenders, who would be more influential as role models.

Table 11 reports treatment effects when offenders transfer into peers' own grade (Column (3)), one or two cohorts below (Columns (1) and (2)), and one or two cohorts above (Columns (4) and (5)).⁵³ The effects are stronger for older offenders than for younger offenders. As the grade gap widens, the effects generally weaken, likely because opportunities for interaction become more limited. Notably, however, the point estimate is larger when the incoming offender is one cohort older than when the offender is in the same grade. This pattern provides further evidence consistent with the role-model channel.⁵⁴⁵⁵

⁵⁰The non-monotonicity may reflect weaker role-modeling effects in the least offender-dense schools, where an incoming offender may differ too much from peers to be a good role model.

⁵¹Some may be concerned that differences in demographics in these schools make the difference in the treatment effects. To address this concern, I partial out the outcome variables first, then use the residualized outcome variables for the analysis. As noted in Footnote 73, the interpretation of the coefficient is not straightforward, but the result pattern is the same, which means that the difference in demographics does not drive the results.

⁵²Remember that grade retention is extremely rare in the UK, so a difference in cohorts translates to a difference in grades.

⁵³To ensure the SUTVA assumptions, I omit those who have received an offender transfer in their cohorts prior to the targeted cohorts in each column.

⁵⁴More experienced offenders may also be more influential. Cattani et al. (2023a,b, 2025) document substantial heterogeneity in the number of crimes committed during the juvenile years. Table D4 divides transfer offenders into three groups according to the number of crimes committed before transfer: one crime, two to five crimes, or more than five crimes. The treatment effects increase with prior offending experience, suggesting that more experienced offenders generate stronger peer effects.

⁵⁵Relatedly, I examine whether the age at which peers are exposed to an offender matters. Offenders who begin offending at very young ages may be more disruptive (Cattani et al., 2025) and therefore more influential, making

Table 11. Treatment Seniority Heterogeneity

Treatment Cohort:	Two Cohorts Younger (1)	One Cohort Younger (2)	Same Cohort (3)	One Cohort Older (4)	Two Cohorts Older (5)
Treatment Effects at $t = 0$	.015 (.011)	.051*** (.011)	.183*** (.034)	.204*** (.015)	.073*** (.026)
Pre-Treatment Mean	.400	.697	1.39	1.93	2.78
Relative Effect Size	3.65%	7.38%	13.1%	10.6%	2.61%
Sample Size (Panel Count)	6,704,483	9,790,276	7,321,414	4,737,409	1,398,616
Number of Individuals	453,765	663,904	497,302	322,733	95,296

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender’s school transfer. The treatment group consists of students who are exposed to an offender transferring into the school and grade specified in each column, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. In Columns (1), (2), (4), and (5), I exclude students who are exposed to an offender’s transfer into their cohort before or at the same time as the treatment. In Column (3), I exclude students who are exposed to an offender’s transfer into a different cohort before or at the same time as the treatment. The estimation follows [Callaway and Sant’Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. Data sources: Ministry of Justice and Department for Education.

Some schools impose disciplinary sanctions when students commit offenses, with exclusion being a common punishment.⁵⁶ In such schools, transfer offenders might not become role models. In England, some state schools, known as “academies,” have greater autonomy and tend to enforce stricter discipline ([Machin and Sandi, 2020](#)); students in these schools commit fewer crimes ([Duchini, Lavy, and Machin, 2020](#)). In Table D2, I examine heterogeneity by school type. It shows no significant treatment effects in academies, suggesting that strict discipline might prevent a transfer offender from becoming a role model and help stop crime spillover.⁵⁷⁵⁸

Role-modeling may be accelerated if homophily strengthens peer interactions ([Billings, Deming, and Ross, 2019](#); [Lee et al., 2021](#)). Some studies show that sharing characteristics is an important channel for role-modeling ([Beaman et al., 2009, 2012](#); [Bhalotra, Clots-Figueras, and Iyer, 2018](#)). I therefore examine heterogeneity by gender, first-language status, and ethnicity.⁵⁹ Table D3 shows

them better role models. At the same time, as children grow older, offending becomes more common, which may lower the psychological barrier to initiating crime. Table D5 reports the results. Because the control mean increases monotonically with age, as shown in Figure 1b, the relative effect size is particularly informative. The treatment effects follow a U-shaped pattern: relative effects are larger at ages 11 and 14 than at ages 12 and 13. This pattern is consistent with two offsetting mechanisms. The influence of particularly disruptive early offenders may decline with age, while the psychological barrier to initiating crime may also fall as offending becomes more prevalent. Together, these forces can generate the observed U-shaped relationship, consistent with the role-model explanation.

⁵⁶Descriptive evidence shows that the timing of offending and exclusion is highly correlated ([Cattan et al., 2023a,b](#)).

⁵⁷Foundation schools also show no treatment effects, but the estimate is much less precise, so we cannot conclude there is no effect.

⁵⁸Not only strict school discipline, but also police intervention in schools may work, which is common in the US. See [Owens \(2017\)](#) for the analysis on this aspect.

⁵⁹Although I do not have information on friendship in the data, the richness of demographic information allows me to investigate how important homophily in terms of observable characteristics is in peer effects. Based on a

that the treatment effects is generally larger if the transfer offender share characteristics of gender and the first language.⁶⁰⁶¹ The set of results, together with the results of exposure intensity, suggests that role-modeling is an important channel.

5.2.3 Low opportunity cost of crime as an accelerator of the treatment effects

So far, I have documented peer effects of crime arising from offenders who transfer into their peers' school and grade, and an important potential mechanism may be through role-modeling. Some individuals' characteristics might change the treatment effects, especially the opportunity cost of committing a crime.⁶²

I first examine heterogeneity by educational achievement before secondary school, which serves as a proxy for the opportunity cost. Table 12 reports the effects on offense initiation. Panel A shows heterogeneity by age-7 test scores, while Panel B uses age-11 test scores. I divide students into quartiles based on standardized educational attainment at each age. Both panels show a clear pattern: treatment effects decline monotonically with educational achievement. This is consistent with higher-achieving students facing a greater opportunity cost of acquiring a criminal record.⁶³

Another explanation for the monotonicity by educational achievement is that students with higher test scores may also have better self-control and therefore be less likely to offend. To investigate this channel, I use measures of non-cognitive skills at age 5.⁶⁴ Because the sample is relatively small, I divide students into only two groups based on whether their scores are above or below the median.⁶⁵ Table D6 shows that treatment effects are generally smaller among students with stronger non-cognitive skills. Across the detailed domains, the clearest difference appears in "Per-

nationwide UK household longitudinal study, Understanding Society, 83.3% of those between 11 and 16 years old have their best friends in the same ethnic group, and 50% have ones in the same ethnic-gender-age groups. This suggests that homophily is crucial for friendship. So, it is reasonable to think that if a transfer student already has the same observable characteristics as children in the destination school, they are more likely to interact with the transfer student.

⁶⁰A notable asymmetry emerges for opposite-gender exposure: male offenders increase offense initiation among girls, whereas female offenders *decrease* offense initiation among boys, although the estimate is a little imprecise. One possible explanation is that female offenders are relatively rare and may therefore be less likely to generate the same type of peer interaction or role-model effect among boys.

⁶¹In contrast to previous findings on ethnic homophily in Germany and the US (Leszczensky and Pink, 2015; Patacchini and Zenou, 2016), I find little evidence that ethnic matching plays an important role in the peer effects of crime. This might be because racial segregation is less pronounced in the UK (Jenkins, Micklewright, and Schnepf, 2008; Iceland, Mateos, and Sharp, 2011).

⁶²Becker (1968); Lochner (2004) discuss the framework of committing a crime with a model-based approach.

⁶³As mentioned in Footnote 11, in some schools, classroom allocation is based on pupils' educational level. Since offenders tend to have lower achievement (Cattan et al., 2023a,b, 2025), those in the lower quartile might have more intense exposure to offenders' transfers. However, if I compare the treatment effects among those with similar treatment intensity, we still observe this monotonicity by educational level.

⁶⁴Although non-cognitive abilities are measured at age 5, they may remain relatively stable over time (Cornelissen and Dustmann, 2019), making them a useful proxy for non-cognitive abilities during secondary school.

⁶⁵I have a record of non-cognitive measures at the age of 5 only from the 1997/8 cohort onwards. See Table A1. Also, it is only for a 10% quasi-random sample of children, as the outcomes are recorded for those born on the 5th, 15th, and 25th of each month (Department for Education, Schools Analysis and Research Division, 2010).

Table 12. Heterogeneity in Educational Achievement

Quartile in Educational Achievement:	Highest (1)	2nd Highest (2)	3rd Highest (3)	Lowest (4)
<i>Panel A: Measured by Key Stage 1 at Age 7</i>				
Treatment Effects at $t = 0$	.074*** (.014)	.141*** (.017)	.192*** (.020)	.354*** (.026)
Pre-Treatment Mean	.769	1.23	1.51	2.57
Relative Effect Size	9.65%	11.5%	12.7%	13.8%
Sample Size (Panel Count)	3,687,155	3,679,891	3,674,242	3,666,407
Number of Individuals	249,863	249,861	249,863	249,859
<i>Panel B: Measured by Key Stage 2 at Age 11</i>				
Treatment Effects at $t = 0$	.028*** (.011)	.083*** (.015)	.139*** (.020)	.198*** (.024)
Pre-Treatment Mean	.582	1.12	1.78	2.82
Relative Effect Size	4.85%	7.39%	7.84%	7.03%
Sample Size (Panel Count)	2,982,887	2,973,715	2,966,242	2,955,983
Number of Individuals	201,761	201,760	201,761	201,758

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. All test scores are standardized within each cohort. The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 when constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Data sources: Ministry of Justice and Department for Education.

sonal social and emotional development.” This component measures children’s skills to manage emotions and self-control, and to persist and wait for what they want and direct attention as necessary, all of which are consistent with placing greater weight on future consequences (lower time discount factor).⁶⁶⁶⁷ Overall, these patterns suggest that self-control may reduce susceptibility to crime spillovers by increasing the perceived future cost of offending.

I further examine heterogeneity by peers’ individual and family backgrounds. I first focus on FSM and SEN eligibility by age 11. As shown in Table D7, the treatment effects are concentrated among disadvantaged students, particularly those eligible for FSM. This pattern suggests that students from more disadvantaged backgrounds may be more susceptible to exposure to offenders, potentially because they face a lower opportunity cost of offending. This interpretation is also consistent with evidence on intergenerational mobility showing that children growing up in lower-income households have lower predicted adult income (Krutikova et al., 2023).

Next, related to FSM, I examine heterogeneity by local deprivation. Although the peer effects operate primarily through schools rather than neighborhoods, the opportunity cost of offending may vary with local economic conditions. I measure local deprivation using the 2007 Income Deprivation Affecting Children Index (IDACI). IDACI measures the proportion of children aged 0 to 15 living in income-deprived households receiving certain means-tested benefits (Noble et al., 2008). I divide the sample into quartiles based on the IDACI value of peers’ LSOA of residence at the time of treatment.⁶⁸⁶⁹

The first rows of Table 13 report treatment effects on all crime types. Both the absolute and relative effects increase monotonically with local deprivation, consistent with the previous results for FSM. The second and third rows focus on theft and fraud. Because receiving stolen goods is the most common fraud offense in my data (Table A5), these two crime types have more direct monetary returns. The result shows the same broad pattern, but strikingly, the treatment effects on theft and fraud are zero in the least deprived areas, and the effect on fraud appears only in the most deprived areas. These results are consistent with monetary returns making offending relatively more attractive by lowering the *net* opportunity cost.⁷⁰

⁶⁶See the UK government’s website for the details (<https://help-for-early-years-providers.education.gov.uk/areas-of-learning/personal-social-and-emotional-development>, last accessed: October 31st, 2026).

⁶⁷Another important point is that those with high “Knowledge and understanding of the world” are *more* likely to start committing crime. This measure evaluates how children make sense of their physical world and community, reflecting their curiosity. Therefore, we can interpret this result as those with higher curiosity about others being more likely to be affected.

⁶⁸See Figure A1 for the distribution of deprived areas across England and London.

⁶⁹Relatedly, some might be interested in the geographical difference in the treatment effects. First, Table F4 shows the heterogeneity with respect to urban and rural areas. The treatment effects tend to be higher in urban areas. This is consistent with the fact that the crime rate is higher in urban areas in England (Hall et al., 2026).

⁷⁰I show the results on the crime types that directly address the potential returns to crime with the measure in 2007. However, the overall results are the same with other crimes and with the measure in 2004. See Tables F1 and F2.

Table 13. Local Deprivation

Quartile in Deprivation Index:	Least Deprived (1)	2nd Least Deprived (2)	3rd Least Deprived (3)	Most Deprived (4)
All Crimes	.073*** (.016)	.119*** (.017)	.209*** (.023)	.309*** (.021)
Pre-Treatment Mean	.784	1.10	1.59	2.22
Relative Effect Size	9.35%	10.8%	13.1%	13.9%
Theft	.014 (.010)	.048*** (.010)	.082*** (.015)	.131*** (.014)
Pre-Treatment Mean	.326	.462	.655	.934
Relative Effect Size	4.31%	10.4%	12.6%	14.0%
Fraud	.002 (.002)	.001 (.003)	.005 (.006)	.008* (.004)
Pre-Treatment Mean	.011	.026	.028	.049
Relative Effect Size	21.1%	1.96%	18.2%	15.7%
Sample Size (Panel Count)	3,299,643	3,503,179	3,970,134	4,685,188
Number of Individuals	223,594	237,727	270,064	319,182

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Deprivation is measured using the 2007 Income Deprivation Affecting Children Index (IDACI), defined as the proportion of children aged 0 to 15 living in income-deprived families at the LSOA level. LSOA stands for Lower Layer Super Output Area, a geographic unit containing approximately 1,000 to 3,000 residents. Data sources: Ministry of Justice and Department for Education.

5.3 Robustness Checks

We have seen that an offender's school transfer accelerates peers' criminal activity. Here, I present robustness checks showing that the effects are attributable to offenders' school transfers and the result is robust to specifications.

If peers start offending due to the influence of offenders, they do not start committing crime because of the effects of disruptive non-offenders. To examine this channel, I focus on not-yet-offenders, who have not committed a crime at the time of school transfer but will have done so by the end of secondary school. Such children are likely to be as disruptive as offenders, but do not have any offending experience. This focus allows me to compare the effects from those who have similar characteristics but have different crime histories.⁷¹ Figure C1 shows no significant effects of not-yet offenders school transfer. This evidence suggests that exposure to offending history is important, not exposure to disruptive children.

Not-yet offenders may nevertheless be less disruptive than offenders. Thus, I examine transfers of non-offenders with disadvantaged or disruptive characteristics. Disadvantage is proxied by eligibility for free school meals and special education needs before transfer, while disruptiveness is proxied by prior school suspensions. Table C2 shows small positive and significant effects,

⁷¹Also, according to Tables 3 and C1, the distribution of the timing of exposure to offenders' school transfers is closer to that of not-yet offenders' school transfers than that of non-offenders. Even with this similarity, I use the weight to adjust for differences in exposure timing across grades and seasons of school censuses.

consistent with the literature on disruptive peers (Billings, Deming, and Ross, 2019; Billings and Hoekstra, 2023).⁷² However, the magnitudes are much smaller than the main treatment effects. Moreover, conditioning on offenders' disadvantaged or disruptive characteristics changes the estimates little. Together, these results suggest that prior offending is the key characteristic driving crime transmission.

In the main analysis, I include students who change schools in order to capture the overall effects of exposure. However, if a student in the control group transfers into a school-grade that has already been treated, that student may be surrounded by treated peers despite not yet having been directly exposed to an offender's transfer. Although this does not necessarily violate SUTVA, it may create contamination between treated and control groups. To address this concern, I censor each student's panel at their first school transfer, at the risk of losing the panel structure of the data. Figure C2 shows that the results are unchanged.

As Table 3 shows, some pupils are exposed to multiple offender transfers. Because the irreversibility assumption treats exposure as absorbing, subsequent treatments should not alter treatment status. Therefore, I censor each student's panel at their second treatment. Figure C3 shows that the results are unchanged, supporting the plausibility of the irreversibility assumption.

Although my identification relies on the conditional parallel-trends assumption and the plausibility of the exogenous shock, it is also informative to report unconditional estimates. Importantly, conditional parallel trends do not imply unconditional parallel trends. Figure C4 presents the results without controls. The effect on crime onset in Figure C4a exhibits a pre-treatment trend difference. However, this difference is approximately linear, with a clear jump at time zero, providing additional suggestive evidence of peer effects. To assess the sensitivity of the estimates to this pre-trend, I apply the method of Rambachan and Roth (2023), which allows for deviations from parallel trends based on the observed pre-treatment pattern. Figures C5a and C6a show that the estimated effects remain robust after accounting for the pre-trend.⁷³ The effects on overall criminal activities with the unconditional specification in Figures C4b and C4c do not show any pre-trend difference, and the main result still holds.⁷⁴

Next, I conduct additional checks of the identifying assumptions and sample restrictions. Because the main analysis excludes never-treated students, I re-estimate the model using the never-treated group as the control to assess external validity.⁷⁵ Figure C10 shows similar immediate treatment effects. Although the effect on overall activity remains positive even after three years, the main result is unchanged. This suggests that excluding the never-treated group is unlikely

⁷²Column (1) of Panel A in Table C2 shows the effects of not-yet offenders' transfer, which is a replication of Figure C1.

⁷³As a further robustness check, I use outcome variables residualized with respect to demographic characteristics. The pattern of results remains unchanged in Figure C9.

⁷⁴Nevertheless, I run the honest difference-in-differences. Figures C5b, C6b, C5c, and C6c confirm the results are robust to the specification.

⁷⁵Since the sample size is too large for the memory that I work with, I draw a random 50% of individuals for this analysis.

to drive the findings and supports the external validity of the main finding to the rest of the population.

If peers transfer to another school in response to an offender’s arrival, the treatment effects of exposure might be underestimated. To assess this possibility, I examine whether offenders’ school transfers increase the probability that incumbent students transfer out. Figure C11 shows no evidence of such a response.

I also examine what happens when offenders themselves leave a school. Because transfer-outs may be endogenous, these estimates should be interpreted as descriptive rather than causal. Table C3 shows that a peer’s departure is associated with an increase in offense initiation *regardless of* whether the departing student is an offender, not-yet offender, or non-offender, while there is no association with overall criminal activity when offenders or non-offenders leave. Thus, there is no evidence that removing offenders reduces crime transmission.⁷⁶ Consistent with this result, Figure C12 shows that the main estimates are unchanged when I restrict the sample to schools from which no offender leaves.

A further concern is that offender transfers may be endogenous because offenders or their parents choose to move.⁷⁷ This may threaten identification if peers can anticipate which offender will enter their school-grade before the transfer occurs. To examine this possibility, I study heterogeneity by whether, and how far, the offender changes residence around the time of the school transfer. Transfers accompanied by longer-distance residential moves are more likely to be unexpected by peers at the destination school. Such moves may also be more likely to reflect parental circumstances, such as job changes, rather than factors directly related to the offender. Table C4 shows slightly larger treatment effects when offenders move farther away, consistent with this interpretation. However, sizable effects remain even when there is no residential move. This provides little evidence that anticipatory information spillovers within neighborhoods drive the main results.

In relatively rare cases, an offender who transfers out later returns to the original school, in which case exposure is less plausibly a new shock. Excluding these cases leaves the results unchanged (Figure C13).

Finally, because the main outcomes are binary and crime is relatively rare, I examine whether the linear specification drives the findings. I estimate Poisson and logit models following Wooldridge (2023, 2025). Table C5 shows qualitatively the same pattern, confirming that the main findings are robust to alternative functional forms.

⁷⁶There are two possible explanations. First, even after an offender moves out, the remaining peers have already been exposed to them. Because peers may begin offending with a lag after exposure (Figure 2), the offender’s departure may appear to trigger the effects even if it does not. Second, when a child leaves the peer group, the remaining peers may begin to misbehave. Lavy and Sand (2019) find that having more reciprocal friends in the classroom reduces students’ own involvement in violence, which suggests that losing a friend, regardless of whether that friend has a criminal record, may increase misbehavior. Either mechanism could explain the increase in criminal activity following a student’s departure.

⁷⁷Another case of an offender’s school transfer is due to permanent exclusion. However, permanent exclusion is a very rare punishment and does not affect my main analysis (Cattan et al., 2023a,b).

5.4 Social Cost of Peer Effects of Crime

To assess the economic importance of these peer effects, I estimate their social cost. First, I calculate the social cost of youth crime in my data using monetary cost estimates from [Heeks et al. \(2018\)](#). I classify all 2,820 crime types into 13 categories with non-negative costs, following [Cattan et al. \(2025\)](#).⁷⁸ Table 14 summarizes the calculation. For each crime category, Column (3) reports the total social cost, calculated as the number of youth offenses in Column (1) multiplied by the social cost per offense from [Heeks et al. \(2018\)](#) in Column (2).⁷⁹

Then, I re-estimate the peer effects separately for each crime category (Table F3) and use these estimates to calculate the number of additional crimes attributable to peer effects, reported in Column (4).⁸⁰ Column (5) converts these additional offenses into social costs by multiplying Column (2) by Column (4). Finally, I divide Column (5) by Column (3) to calculate the share of the total social cost attributable to peer effects, shown in Column (6). Overall, approximately 5% of the social cost of youth crime in my data period, equivalent to about 20 million pounds (26.4 million dollars), is attributable to these peer effects.⁸¹

5.5 Effects on Other Behavioral Responses

So far, I have focused on criminal behavior. Because the NPD also contains rich information on behavioral problems, I next examine whether exposure to offenders affects milder forms of misconduct. Even if children do not commit a crime, they may still respond to exposure through other problematic behaviors. I focus on suspension (fixed-term exclusion) and unauthorized absence. Offenders are more likely to experience longer suspensions and more unauthorized absences ([Cattan et al., 2023a,b, 2025](#)), making these outcomes useful proxies for less severe behavioral problems.

Figure F2 presents the results. Figures F2a and F2b show that offenders' transfers increase both the probability and duration of exclusion in the short run.⁸² Figure F2c also shows an

⁷⁸I omit the categories of homicides and cybercrimes because of potentially disclosive low counts. Homicides have the highest estimated cost among the crimes considered in ([Heeks et al., 2018](#)), at £3,217,740 per offense. Excluding them may lead me to underestimate the total social cost of youth crime and may partly explain the difference between my estimate and that in [Irwin-Rogers, Muthoo, and Billingham \(2020\)](#).

⁷⁹The social cost per crime for all crimes is not reported directly in [Heeks et al. \(2018\)](#). I calculate it by dividing the total cost, obtained as the sum across crime types in Column (3), by the total number of crimes in Column (1).

⁸⁰I replace the point estimate with zero if it is not statistically significantly different from zero at a 10% significance level.

⁸¹An alternative approach to calculating the social cost draws on [Cohen, Piquero, and Jennings \(2010\)](#), who show the present value of the lifetime costs of juvenile crime in the US. Although crime patterns may differ substantially between the US and England, their estimates are useful for gauging the long-run social cost. They estimate that a criminal career costs at least \$2.1 million, which implies a total cost of \$419 million ($\$2.1 \text{ million} \times 0.00019 \times 1,050,567$ individuals in my data). Alternatively, their most conservative estimate, for an offender who commits only one crime and using a 2% discount rate, is \$53,000, which implies a total cost of \$10.6 million ($\$53,000 \times 0.00019 \times 1,050,567$). Although the two estimates differ in magnitude, both point to substantial long-run social costs of youth crime generated through peer effects.

⁸²The effects become *negative* in the long-run. This may be because of substitution from disruptive behavior

Table 14. Social Costs

	Number of Youth Crimes (1)	Social Cost per Crime (2)	Social Cost of the Crime Type (3)	Increase in the Number Crimes due to Peer Effects (4)	Increase in Social Cost due to Peer Effects (5)	Fraction of Increase (6)
All Crimes	61,538	6,456	397,267,948	2,922	20,000,725	5.03%
Cost of Crime for Each Crime Type						
Violence with Injury	16,233	14,050	228,073,648	355	4,990,295	2.19%
Violence without Injury	4,125	5,930	24,461,250	940	5,571,984	22.8%
Rape	426	39,360	16,767,360	78.9	3,103,717	18.5%
Other Sexual Offence	2,595	6,520	16,919,400	123	799,296	4.72%
Robbery	2,964	11,320	33,552,480	-42.0	-475,684	-1.42%
Domestic Burglary	4,025	5,930	23,868,250	513	3,041,085	12.7%
Theft of Vehicle	711	10,290	7,316,190	247	2,542,533	34.8%
Theft from Vehicle	653	870	568,110	-119	-103,151	-18.2%
Theft from Person	17,080	1,380	23,570,400	449	618,974	2.63%
Criminal Damage – Arson	720	8,420	6,062,400	-83.7	-704,632	-11.6%
Criminal Damage – Other	10,612	1,350	14,326,200	444	598,877	4.18%
Fraud	1,344	1,290	1,733,760	0	0	0%
Commercial Theft	50	970	48,500	18.0	17,430	35.9%

Notes: Column (1) reports the number of crimes committed by individuals in the final sample described in Sections 3 and A. The crimes listed in the second part of the table constitute a subset of all crimes in the data. To calculate the social cost of crime, I map the 2800 crime types recorded in the PNC to the 13 categories used by Heeks et al. (2018). Some crimes recorded in the PNC, such as weapon offenses, do not have a corresponding category. I assign a cost of zero to these crimes and exclude them from the list. This approach therefore provides a lower-bound estimate of the social cost, since these offenses are likely to impose some social cost. The cost for each crime is calculated by multiplying the cost per crime by the number of crimes induced by peer effects, as shown in Table F3. If the treatment effect is not statistically significant at the 10% level, I set the effect equal to zero. Data sources: Ministry of Justice and Department for Education.

increase in the share of unauthorized absence sessions. These patterns are broadly consistent with those observed for offenders. Overall, the results suggest that exposure to offenders increases problematic behavior even among students whose behavior does not escalate to criminal offending.

5.6 Effects on Educational Achievement

Table 15. Effects on Educational Outcomes at Age 14

Transfer Student:	Offender		Non-offender	
	(1)	(2)	(3)	(4)
Any Offender's Transfer	-.041*** (.010)	-.016*** (.005)	-.129*** (.010)	.004 (.009)
Controls		x		x
Sample Size	862,553	862,553	69,297	69,297

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors clustered at the treatment level are reported in parentheses. Control variables include FSM and SEN status before the age 14 exam, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, and month-of-birth fixed effects. The age 11 and 14 overall scores are calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Data sources: Ministry of Justice and Department for Education.

at school toward absence from school. Indeed, the unauthorized absences and fixed exclusions do not decrease (Figures F2d and F2e), which suggests that a decrease in exclusion does *not* mean that exposure to offenders improves peers' behavior in the long run.

In the school context, educational achievement is a key outcome. However, the cohorts used in the main analysis, born between 1995/96 and 98/99, do not have measures of educational achievement during secondary school in my data extract. Therefore, I cannot compare the outcomes of students who have already been exposed to an offender with those of students who have not yet been exposed.⁸³ I instead focus on earlier cohorts born between 1990/91 and 1993/94, for whom educational achievement at age 14 is observed.⁸⁴ A major limitation is that these cohorts have only one school census per year in the spring census, so the exact timing of school transfers cannot be observed. This introduces potential measurement error in treatment status. Despite this limitation, these cohorts allow me to examine an important educational outcome.⁸⁵

Because I use the age-14 standardized test score (KS3), I define the treatment group as students exposed to an offender transferring into their school during the academic year beginning in September at age 13; that is, up to one year before the age-14 exam. The control group consists of students exposed after the exam, up to one year later. I estimate:

$$y_{ics} = \alpha + \beta d_{cs} + \mathbf{x}_i^T \boldsymbol{\delta} + \eta_c + \nu_s + \varepsilon_{ics}, \quad (1)$$

where y_{ics} is the age-14 test score for individual i in cohort c and school s , d_{cs} is the treatment indicator, $\mathbf{x}_i$ is a vector of individual characteristics, and η_c and ν_s are cohort and secondary-school fixed effects, respectively. Standard errors are clustered at the level of exposure, defined by the school-cohort-timing cell.⁸⁶

Table 15 presents the results. Column (1) reports the unconditional difference, while Column (2) reports the conditional estimate. Exposure to an offender reduces educational attainment by 0.016 standard deviations. As a robustness check, I instead focus on non-offenders' transfers on educational achievement. Column (4) shows that there is no adverse impact from a non-offender's school transfer.

Understanding the source of this negative effect is important: I decompose the decline in educational achievement into an indirect effect operating through offense initiation and a direct effect of exposure itself. To do so, I compare children who differ in both the timing of exposure to an offender's transfer and the timing of their first offense.

I define four groups, Groups A, B, C, and D. Groups A and B are exposed to an offender at age 13 and form the treatment group, while Groups C and D are exposed at age 14 and form the

⁸³I do not use the GCSE exam at age 16 because people may leave state school at 16, and the choice is likely correlated with educational achievement and offense history, which may cause selection bias if I adopt a difference-in-differences approach.

⁸⁴The data linkage information is in Table A2. The linkage rate is similar to the cohorts born between 1995/96–98/99, although the offense rate is higher, which is consistent with the fact that the offense rate is decreasing (Youth Justice Board and Ministry of Justice, 2023; Cattani et al., 2025).

⁸⁵Appendix G replicates the summary statistics and main results using these cohorts. The results are robust to the choice of the cohorts.

⁸⁶Because there is only one school census per year for these cohorts, clustering at the school-cohort level is equivalent to clustering at the school-cohort level.

control group. Within each exposure group, I further distinguish children by their own offending behavior. Children in Groups A and C commit their first offense in the same year as exposure, at ages 13 and 14, respectively, while those in Groups B and D do not commit an offense by age 14.⁸⁷

These comparisons allow me to separate the association between educational attainment and becoming an offender following exposure from the direct effect of earlier exposure itself. Panel A of Table 16 summarizes the four groups and reports their educational achievement at age 14. Students who begin offending following exposure have lower educational attainment than those who do not, and earlier exposure is also associated with lower test scores, consistent with my prior paper (Cattan et al., 2025).

Table 16. Mediation Analysis

Panel A: Age 14 Educational Achievements						
Group	Criteria			Mean	SD	Sample Size
	Age at which they are exposed to an offender	First crime at the age 13 or 14	Age of the first crime (13 or 14)			
A:	13	Yes	13	-.570	[.905]	8,623
B:	13	No	-	-.011	[.914]	421,568
C:	14	Yes	14	-.469	[.896]	9,663
D:	14	No	-	.028	[.912]	387,368

Panel B: Mediation					
Parameter:	$\gamma + \delta$ (1)	$\beta + \delta$ (2)	γ (3)	β (4)	δ (5)
Conditional Difference	-.225*** (.007)	-.066*** (.018)	-.190*** (.007)	-.016*** (.002)	-.034*** (.010)
Controls	x	x	x	x	x
Sample Size (Panel Count)	430,191	18,286	397,031	808,936	827,222

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors clustered at the treatment level are reported in parentheses. Control variables include FSM and SEN status before the age 14 exam, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, and month-of-birth fixed effects. The age 14 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. Data sources: Ministry of Justice and Department for Education.

I now formalize these comparisons using the following regression model. For simplicity, I omit

⁸⁷One caveat of this specification is that I compare students who are exposed to an offender at age 13 but do not commit a crime for the following year (Group B) with those who are exposed at age 14 but do not commit a crime only during that same year (Group D). This difference arises from the limited frequency of the school information and cannot be fully avoided in the mediation analysis. To address this concern partially, as a robustness check, I exclude from Group D students who commit a crime at age 15, so that Groups B and D are required to remain nonoffending for the same length of time following exposure. This alternative definition conditions group membership on a future outcome and should therefore be interpreted only as a robustness exercise. The results are unchanged under either definition of the groups.

controls and the school and cohort subscripts, s and c , as well as the corresponding fixed effects:

$$Y_i = \alpha + \beta T_i + \gamma C_i + \delta T_i * C_i + \varepsilon_i, \quad (2)$$

where Y_i is individual i 's educational achievement at age 14, T_i is the treatment indicator, C_i is a “complier” indicator identifying students who are vulnerable to the shock of exposure to an offender, and ε_i is an error term. I include T_i to capture the direct effect of exposure, which may arise, for example, through classroom disruption or psychological distress; previous studies document adverse effects of exposure to violent crime on educational achievement (Brown and Velásquez, 2017; Koppensteiner and Menezes, 2021; Cabral et al., 2026). I include C_i because compliers may differ systematically from never-takers, given that their offending behavior responds to the relatively small shock of an offender’s school transfer. Finally, the interaction $T_i * C_i$ identifies students who become offenders following treatment. Thus, δ captures the indirect effect of exposure operating through offense initiation, while β captures the direct effect of exposure on educational achievement.

For each group, the following applies.

Group A	Group B	Group C	Group D
$\alpha + \beta + \gamma + \delta$	$\alpha + \beta$	$\alpha + \gamma$	α

To estimate β , I compare Groups B and D, while I estimate δ using the difference-in-differences $(A - B) - (C - D)$. In all specifications, I include school and cohort fixed effects, as well as observed individual characteristics that may be correlated with treatment and complier status.

Panel B of Table 16 presents the results. Exposure to an offender directly reduces educational achievement by 0.016 standard deviations. This suggests that exposure itself adversely affects learning, even without inducing the student to offend. One possible mechanism is that disruptive peers interfere with the learning environment or require teachers to redirect resources toward them, consistent with the literature on disruptive students (Carrell, Hoekstra, and Kuka, 2018; Carneiro et al., forthcoming; Tincani, 2026). In addition to this direct effect, becoming an offender following exposure reduces test scores by a further 0.034 standard deviations, approximately twice the magnitude of the direct exposure effect.⁸⁸ Thus, the indirect effect operating through offense initiation appears larger than the direct effect of exposure. This pattern is consistent with my previous findings in Cattani et al. (2025) – students who offend before the exam have substantially lower educational achievement. One possible explanation is that acquiring an offense record reduces the expected return to further human-capital investment, thereby weakening incentives to exert effort in school.

⁸⁸In this regression, the sample is slightly different from Table 15, because I exclude from the mediation analysis those who are exposed at age 13 but start committing a crime at age 14, and those who are exposed at age 14 but start committing a crime at age 13. I also add fixed effects and controls in all regressions. Therefore, the point estimates in Table 16 do not sum up to those in Table 15.

6 Simulation

So far, I have shown that peer effects of crime are substantial and contribute meaningfully to youth offending and its social cost. Now I consider a potential policy aimed at reducing crime through the peer-composition channel. Given the substantial heterogeneity in treatment effects documented above, reallocating transfer students across schools may reduce criminal spillovers. Because my identification strategy relies on offenders' school transfers, a natural counterfactual is to consider how alternative school assignments could mitigate these negative peer effects. Any such policy would raise important ethical and practical concerns, but quantifying the potential reduction in crime provides a useful benchmark for policy discussion.

Suppose there were a centralized school-transfer market in which all pupils wishing to move schools registered in advance, and a social planner assigned them to schools based on their characteristics, preferences, and available capacity.⁸⁹ Using a matching algorithm from mechanism design, the planner can construct a stable allocation that reduces predicted criminal spillovers.⁹⁰

For simplicity, I focus only on contemporaneous treatment effects and abstract from the dynamic effects documented earlier.⁹¹ Incorporating the full dynamics would require a substantially more complex model in which students jointly make decisions about school transfers, friendship formation, and criminal activity in every single school term. Such a complex model would be difficult to implement in practice, and a simpler framework remains informative for quantifying the potential gains from reassignment.

I also focus on a single source of treatment-effect heterogeneity. Although the previous analysis documents heterogeneity along many dimensions, there is no natural way to summarize them in a single index without imposing additional assumptions. Moreover, a simpler rule is more transparent and potentially more relevant for policy implementation. Accordingly, I treat offenders and non-offenders as homogeneous and consider a policy that swaps the destination school of an offender with that of a non-offender.

6.1 Simulation Algorithm

I now describe the simulation algorithm step by step and discuss the desirable properties of the resulting allocation.

Step 1: Interpolate the treatment effect. The heterogeneity by offender density reported in Table 10 is estimated by tercile. So, I interpolate the treatment effects over the

⁸⁹There are many other possible alternative policies as a thought experiment, ethically right or not, such as denying offenders' school transfer, for example. However, in this analysis, I assume all pupils who want to move can move to another school. Another alternative would be assigning offenders and victims (if any) to different schools if there were linked data on crimes and victims, as theoretically discussed in [Kasuya \(2016\)](#).

⁹⁰See Appendix E for the detail.

⁹¹I also abstract any feedback effects, so essentially I focus on a partial equilibrium, not a general equilibrium.

distribution of offender density. Specifically, I linearly interpolate both the point estimates and the 95% confidence intervals, and set the treatment effect to zero whenever the confidence interval includes zero. Figure E1 provides a visual illustration. I restrict the analysis to the range of offender density between the top and bottom terciles to avoid extrapolating treatment effects.

Step 2: Calculate the potential treatment effect of those who move. I identify students who transfer to another school at the beginning of each school term. I assume that each departure creates one available slot in the school-grade the student leaves, so the number of available slots equals the number of departing students. For every student in the transfer market, I calculate the predicted peer effect associated with assigning that student to each school-grade with an available slot, based on the student’s characteristics and the characteristics of the potential destination school. Since I assume those who wish to move can move, I exclude any effect at the origin school.⁹² Because non-offenders’ transfers do not increase crime at any tercile (Figure E2), I assign them a predicted treatment effect of zero.

Step 3: Stack the peer effects at each school into one vector for each student. For each transfer student, I collect the predicted peer effects associated with all feasible destination school-grades into a column vector. If a school has multiple available slots, I duplicate the corresponding row. I exclude a school when the school is too far from the student’s original destination school. Distance is measured between the student’s original destination school and each school with an available slot.⁹³ I consider distance thresholds of 1, 3, 5, 10, and 25 miles.

Table 17 provides an example. Suppose a transfer market contains one offender (A) and a non-offender (B), and there are two schools. Originally, A is assigned to School 1 and B to School 2, as shown in the gray-shaded cells. If Offender A is assigned to Schools 1 and 2, the predicted increases in crime are 3 and 1, respectively. Non-offender B does not increase crime. Assume that all schools lie within the allowed distance from A and B’s original destinations.

Table 17. Example of Vectors of Preferences for Simulation

	Offender A	Non-offender B
School 1	3	0
School 2	1	0

⁹²Offenders’ leaving is endogenous, and I cannot causally estimate the effects of offenders’ leaving on stayers. However, a correlational analysis might be informative. Table C3 shows that, although we see association of similar magnitudes *regardless* of the leaver’s type, there is no clear association between offenders leaving a school and peers’ offending.

⁹³The location is identified as the mode of the LSOA of students’ residences.

Step 4: Interpret the vector as an individual’s preference. I interpret each individual’s vector as their *preference* over school choices, with lower predicted peer effects corresponding to higher preferences. For example, for Offender A, School 2 is the most preferred, and School 1 is the least preferred. For Non-offender B is indifferent among all schools.⁹⁴ The outside option is defined as transferring to the individual’s original destination school. I assume that schools have no preferences over students.⁹⁵

Step 5: Apply an algorithm from mechanism design theory. Given the set of preferences, I apply a random priority mechanism: I assign a random number to each individual and assign the most preferred school from the pool of available slots in ascending order of that random number.

The resulting assignment has several attractive properties (Bogomolnaia and Moulin, 2001). First, the mechanism is strategy-proof, so no student has an incentive to misreport their preferences. It also satisfies equal treatment of equals. Finally, it is ex post efficient: no reassignment can make one student better off without making another worse off.

In the example above, the status quo generates three additional crimes. Suppose the random-priority order is A and then B. The social planner then assigns A to School 2, and B to School 1. This alternative allocation reduces the predicted number of additional crimes by two. This example illustrates how reassignment can reduce crime by reallocating transfer students across schools.

6.2 Simulation Results

Table 18 presents the simulation results. Panel A reports outcomes when the planner minimizes the number of crimes, while Panel B reports outcomes when the planner minimizes the number of offenders. Each panel contains three sets of two rows. The first reports the number of offenders reassigned away from their original destination school and the corresponding share of all offender transfers. The second and third report the number of crimes and offenders generated by peer effects under the alternative assignment, together with the percentage change relative to the status quo.

Column (1) reports the status quo: the number of crimes and offenders generated by contemporaneous peer effects and the number of offender transfers. Columns (2) to (6) report simulations in which alternative schools must lie within an m -mile radius of the original destination, where $m \in \{1, 3, 5, 10, 25\}$. Column (7) removes the distance restriction and allows reassignment across

⁹⁴In the empirical exercise, I impose strict preferences by randomly ordering schools that are otherwise equally preferred.

⁹⁵I make the technical assumption that, even if a student enters the reassignment market and thereby creates an available slot at their original school, they may still choose the outside option if the resulting assignment is worse. In that case, they are allowed to remain at their original destination school even if the slot they created has already been assigned to another incoming student; they are not forced to move to another school.

Table 18. Simulation Results

		Maximum Distance between the Originally-assigned and the Re-assigned Schools (miles)					
	Original (1)	1 (2)	3 (3)	5 (4)	10 (5)	25 (6)	Any (7)
<i>Panel A: Minimizing N of Crimes</i>							
N of offenders swapped	3661	310	1649	2283	3071	3543	3657
Fraction of offenders swapped		8%	45%	62%	84%	97%	100%
N of crimes	379	350	244	193	137	113	91
Δ crimes		-8%	-36%	-49%	-64%	-70%	-76%
N of offenders	887	893	917	932	924	918	826
Δ offenders		1%	3%	5%	4%	3%	-7%
<i>Panel B: Minimizing N of Offenders</i>							
N of offenders swapped	3661	309	1579	2224	2961	3470	3653
Fraction of offenders swapped		8%	43%	61%	81%	95%	100%
N of crimes	379	392	498	609	805	1089	1244
Δ crimes		3%	31%	61%	113%	187%	228%
N of offenders	887	848	686	617	574	603	648
Δ offenders		-4%	-23%	-30%	-35%	-32%	-27%

Notes: This table shows the simulation results based on the random priority mechanism ([Bogomolnaia and Moulin, 2001](#)). I drop the school outside the interpolation window in Figure E1. I repeat the simulation 10 times and take the average.

England. The statutory walking distance for secondary-school students under the Education Act 1996 is 3 miles, making this a useful benchmark ([Department for Education, 2026](#)).

Under the status quo, peer effects generate 379 additional crimes and 887 additional offenders. When the planner minimizes the number of crimes, allowing reassignments within 3 miles reduces peer-effect-induced crimes by 36%. Achieving this requires swapping the destination schools of 1,649 offenders with those of 1,649 non-offenders. If reassignment is allowed across England, the reduction reaches 76%. The number of additional offenders, however, changes relatively little.

Panel B highlights an important trade-off. When the planner instead minimizes the number of offenders, a 3-mile reassignment reduces the number of additional offenders by 23%, but increases the number of crimes by 31%. When reassignment is allowed across England, the number of crimes more than doubles relative to the status quo. This occurs because assigning offenders to offender-dense schools reduces offense initiation but increases criminal activity among existing offenders (Table 10). Thus, reassignment can substantially reduce crime spillovers, but the choice of objective function is crucial: a policy that improves one margin may worsen another.

7 Conclusion

This paper investigates peer effects in offending among secondary-school students in England by examining whether exposure to a peer with a prior offense affects students' subsequent criminal behavior. I examine both offense initiation and contemporaneous overall criminal activity, as well as heterogeneity in these effects and the mechanisms underlying them. To identify causal effects, I exploit the plausibly exogenous timing of school transfers by students who had committed an offense before transferring. Using linked administrative data from the National Pupil Database (NPD) and the Police National Computer (PNC), I estimate the effects in a staggered event-study design comparing already-treated students with those not yet exposed.

Exposure to an offender's transfer increases the probability of offense initiation by 12.9% within one school term, with the effect persisting for several years; approximately every third offender's transfer induces one new offender. Heterogeneity analyses suggest that role-model effects play an important role in the mechanism. The effects extend beyond criminal behavior: exposure also increases other forms of problematic behavior and reduces educational attainment.

The economic magnitude of these effects is substantial. Using crime-specific social cost estimates from [Heeks et al. \(2018\)](#), I estimate that peer effects account for approximately 5% of the total social cost of youth crime in my data, equivalent to about 20 million pounds. A counterfactual reassignment exercise further suggests that reallocating transfer students across schools could reduce crimes generated by peer effects by 36% under a 3-mile reassignment constraint. However, the choice of the planner's objective is crucial: minimizing the number of offenders can increase the total number of crimes, highlighting an important trade-off in the design of such policies.

These findings also have implications for school exclusion policies following student offending (Cattan et al., 2023a,b). In the case of permanent exclusion, a student must transfer to another school because of compulsory education requirements. Previous studies show that exclusion can increase subsequent offending among excluded students (McAra and McVie, 2010; Bacher-Hicks, Billings, and Deming, 2024). My results add another concern: when an offender transfers to a new school, students at the destination school become more likely to offend. At the same time, I find there is no association between removing offenders and reduction of criminal behavior among peers at the origin school. Taken together, these findings suggest that exclusion policies based on criminal behavior may shift negative externalities across schools rather than eliminate them. Policies addressing youth offending should therefore account not only for the consequences for excluded students, but also for the spillovers their reassignment may generate for other students.

References

- Abdulkadiroğlu, Atila, and Tayfun Sönmez.** 2003. “School Choice: A Mechanism Design Approach.” *American Economic Review*, 93(3): 729–747.
- Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez.** 2005. “The Boston public school match.” *American Economic Review*, 95(2): 368–371.
- Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth.** 2005. “The New York City High School Match.” *American Economic Review*, 95(2): 364–367.
- Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth.** 2009. “Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match.” *American Economic Review*, 99(5): 1954–1978.
- Agan, Amanda, and Sonja Starr.** 2017. “The Effect of Criminal Records on Access to Employment.” *American Economic Review*, 107(5): 560–564.
- Ahrsjö, Ulrika, Costas Meghir, Mårten Palme, and Marieke Schnabel.** forthcoming. “The Effect of Education Policy on Crime: an Intergenerational Perspective.” *Journal of Political Economy*.
- Aizer, Anna, and Joseph J. Doyle Jr.** 2015. “Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges.” *Quarterly Journal of Economics*, 130(2): 759–803.
- Amit, Inbar.** 2026. “Preference for Peers and Educational Inequality: Evidence from School Admission Lotteries.” Working Paper.

- Ammermueller, Andreas, and Jörn-Steffen Pischke.** 2009. “Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study.” *Journal of Labor Economics*, 27(3): 315–348.
- Ando, Michihito, Hiroaki Mori, and Shintaro Yamaguchi.** 2025. “Universal Early Childhood Education and Adolescent Risky Behavior.” *Journal of Public Economics*, 245: 105353.
- Angrist, Joshua D.** 2014. “The Perils of Peer Effects.” *Labour Economics*, 30: 98–108.
- Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster, and Josh Kinsler.** 2012. “Estimating Spillovers Using Panel Data, with an Application to the Classroom.” *Quantitative Economics*, 3(3): 421–470.
- Bacher-Hicks, Andrew, Stephen B. Billings, and David J. Deming.** 2024. “The School-to-Prison Pipeline: Long-run Impacts of School Suspensions on Adult Crime.” *American Economic Journal: Economic Policy*, 16(4): 165–193.
- Balestra, Simone, Beatrix Eugster, and Helge Liebert.** 2022. “Peers with Special Needs: Effects and Policies.” *Review of Economics and Statistics*, 104(3): 602–618.
- Balsa, Ana I., Jenny F. Homer, Michael T. French, and Edward C. Norton.** 2011. “Alcohol Use and Popularity: Social Payoffs from Conforming to Peers’ Behavior.” *Journal of Research on Adolescence*, 21(3): 559–568.
- Balsa, Ana, Monica Costa-Dias, Anders Holm, Freja Thim, and Marcos Vera Hernandez.** 2023. “Mental Health Spillovers in Primary Schools.” Mimeo.
- Bayer, Patrick, Randi Hjalmarsson, and David Pozen.** 2009. “Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections.” *Quarterly Journal of Economics*, 124(1): 105–147.
- Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova.** 2012. “Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India.” *Science*, 335(6068): 582–586.
- Beaman, Lori, Raghavendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova.** 2009. “Powerful Women: Does Exposure Reduce Bias?” *Quarterly Journal of Economics*, 124(4): 1497–1540.
- Becker, Gary S.** 1962. “Investment in human capital: A theoretical analysis.” *Journal of Political Economy*, 70(5, Part 2): 9–49.
- Becker, Gary S.** 1968. “Crime and Punishment: An Economic Approach.” *Journal of Political Economy*, 76(2): 169–217.

- Bell, Brian, Rui Costa, and Stephen Machin.** 2022. “Why Does Education Reduce Crime?” *Journal of Political Economy*, 130(3): 732–765.
- Berthelon, Matias E., and Diana I. Kruger.** 2011. “Risky Behavior among Youth: Incapacitation Effects of School on Adolescent Motherhood and Crime in Chile.” *Journal of Public Economics*, 95(1-2): 41–53.
- Bhalotra, Sonia, Irma Clots-Figueras, and Lakshmi Iyer.** 2018. “Pathbreakers? Women’s Electoral Success and Future Political Participation.” *Economic Journal*, 128(613): 1844–1878.
- Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad.** 2018a. “Incarceration Spillovers in Criminal and Family Networks.” National Bureau of Economic Research Working Paper.
- Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad.** 2018b. “Intergenerational Effects of Incarceration.” *AEA Papers and Proceedings*, 108: 234–40.
- Billings, Stephen B., and Kevin T. Schnepel.** 2022. “Hanging Out with the Usual Suspects: Neighborhood Peer Effects and Recidivism.” *Journal of Human Resources*, 57(5): 1758–1788.
- Billings, Stephen B., and Mark L. Hoekstra.** 2023. “The Effect of School and Neighborhood Peers on Achievement, Misbehavior, and Adult Crime.” *Journal of Labor Economics*, 41(3): 643–685.
- Billings, Stephen B., David J. Deming, and Jonah Rockoff.** 2014. “School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg.” *Quarterly Journal of Economics*, 129(1): 435–476.
- Billings, Stephen B., David J. Deming, and Stephen L. Ross.** 2019. “Partners in Crime.” *American Economic Journal: Applied Economics*, 11(1): 126–150.
- Bogomolnaia, Anna, and Hervé Moulin.** 2001. “A New Solution to the Random Assignment Problem.” *Journal of Economic Theory*, 100(2): 295–328.
- Boshoff, Janine, Stephen Machin, and Matteo Sandi.** 2026. “Youth Crime and Delinquency in and out of School.” *Economic Journal*, ueag081.
- Brown, Ryan, and Andrea Velásquez.** 2017. “The Effect of Violent Crime on the Human Capital Accumulation of Young Adults.” *Journal of Development Economics*, 127: 1–12.
- Burke, Mary A., and Tim R. Sass.** 2013. “Classroom Peer Effects and Student Achievement.” *Journal of Labor Economics*, 31(1): 51–82.

- Burton, Anne M., and David N. Wasser.** 2025. “Revisiting the Unintended Consequences of Ban the Box.” *Journal of Public Economics*, 250: 105473.
- Byström, Daniel Cunha.** 2025. “From Friends to ‘Family’: Schools, Neighborhoods, and Gang Recruitment.” Working Paper.
- Cabral, Marika, Bokyoung Kim, Maya Rossin-Slater, Molly Schnell, and Hannes Schwandt.** 2026. “Trauma at School: The Impacts of Shootings on Students’ Human Capital and Economic Outcomes.” *Review of Economic Studies*, 93(1): 327–365.
- Calderon, Daniel Jaramillo.** 2026. “The Spillover Effects of Prisoner Releases: Evidence from an Ecuadorian Mass Pardon.” Working Paper.
- Callaway, Brantly, and Pedro H.C. Sant’Anna.** 2021. “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics*, 225(2): 200–230.
- Carneiro, Pedro, Sarah Cattán, Gabriella Conti, Claire Crawford, Christine Farquharson, and Nick Ridpath.** 2025. “The Short- and Medium-term Effects of Sure Start on Children’s Outcomes.” IFS Report.
- Carneiro, Pedro, Yyannu Cruz-Aguayo, Francesca Salvati, and Norbert Schady.** forthcoming. “Disruption in the Classroom: Experimental evidence from Ecuador.” *American Economic Journal: Applied Economics*.
- Carrell, Scott E., and Mark L. Hoekstra.** 2010. “Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone’s Kids.” *American Economic Journal: Applied Economics*, 2(1): 211–28.
- Carrell, Scott E., Frederick V. Malmstrom, and James E. West.** 2008. “Peer Effects in Academic Cheating.” *Journal of Human Resources*, 43(1): 173–207.
- Carrell, Scott E., Mark L. Hoekstra, and Elira Kuka.** 2018. “The Long-run Effects of Disruptive Peers.” *American Economic Review*, 108(11): 3377–3415.
- Carrell, Scott E., Richard L. Fullerton, and James E. West.** 2009. “Does Your Cohort Matter? Measuring Peer Effects in College Achievement.” *Journal of Labor Economics*, 27(3): 439–464.
- Cattán, Sarah, Monica Costa-Dias, Hikaru Kawarazaki, and Imran Rasul.** 2023a. “The Schooling System and Teen Crime: New Evidence from Administrative Data from England for Boys.” Mimeo.

- Cattan, Sarah, Monica Costa-Dias, Hikaru Kawarazaki, and Imran Rasul.** 2023*b*. “The Schooling System and Teen Crime: New Evidence on Gender Difference from Administrative Data from England.” Mimeo.
- Cattan, Sarah, Monica Costa-Dias, Hikaru Kawarazaki, and Imran Rasul.** 2025. “Prolific Juvenile Offenders.” Mimeo.
- Clarke, Damian, Joseph P. Romano, and Michael Wolf.** 2020. “The Romano–Wolf Multiple-Hypothesis Correction in Stata.” *Stata Journal*, 20(4): 812–843.
- Cohen, Mark A., Alex R. Piquero, and Wesley G. Jennings.** 2010. “Estimating the Costs of Bad Outcomes for At-risk Youth and the Benefits of Early Childhood Interventions to Reduce Them.” *Criminal Justice Policy Review*, 21(4): 391–434.
- Cook, Philip J., and Songman Kang.** 2016. “Birthdays, Schooling, and Crime: Regression-discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation.” *American Economic Journal: Applied Economics*, 8(1): 33–57.
- Cornelissen, Thomas, and Christian Dustmann.** 2019. “Early School Exposure, Test Scores, and Noncognitive Outcomes.” *American Economic Journal: Economic Policy*, 11(2): 35–63.
- Cullen, Zoë, Will Dobbie, and Mitchell Hoffman.** 2023. “Increasing the Demand for Workers with a Criminal Record.” *Quarterly Journal of Economics*, 138(1): 103–150.
- Damm, Anna Piil, and Cédric Gorinas.** 2020. “Prison as a Criminal School: Peer Effects and Criminal Learning Behind Bars.” *Journal of Law and Economics*, 63(1): 149–180.
- Department for Children, Schools and Families.** 2009. *School Admissions Code*. The Stationery Office.
- Department for Education.** 2026. “Travel to school for children of compulsory school age Statutory guidance for local authorities.” Department for Education.
- Department for Education, Schools Analysis and Research Division.** 2010. “Achievement of Children in the Early Years Foundation Stage Profile.” Department for Education DFE-RR034.
- Depew, Briggs, and Ozkan Eren.** 2016. “Born on the Wrong Day? School Entry Age and Juvenile Crime.” *Journal of Urban Economics*, 96: 73–90.
- Díaz, Carlos, Eleonora Patacchini, Thierry Verdier, and Yves Zenou.** 2021. “Leaders in Juvenile Crime.” *Journal of Economic Behavior & Organization*, 192: 638–667.

- Ding, Weili, and Steven F. Lehrer.** 2007. “Do Peers Affect Student Achievement in China’s Secondary Schools?” *Review of Economics and Statistics*, 89(2): 300–312.
- Disney, Richard, Tom Kirchmaier, Stephen Machin, and Carmen Villa.** 2026. “Gangs of London and Public Housing.” Centre for Economic Performance, Discussion Paper.
- Dobbie, Will, and Roland G. Fryer Jr.** 2011. “Are High-quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children’s Zone.” *American Economic Journal: Applied Economics*, 3(3): 158–187.
- Dobbie, Will, and Roland G. Fryer Jr.** 2015. “The Medium-term Impacts of High-achieving Charter Schools.” *Journal of Political Economy*, 123(5): 985–1037.
- Doleac, Jennifer L., and Benjamin Hansen.** 2020. “The Unintended Consequences of “Ban the Box”: Statistical Discrimination and Employment Outcomes When Criminal Histories are Hidden.” *Journal of Labor Economics*, 38(2): 321–374.
- Domínguez, Magdalena.** 2026. “Sweeping Up Gangs: The Effects of Tough-on-Crime Policies from a Network Approach.” *American Economic Journal: Economic Policy*.
- Duchini, Emma, Victor Lavy, and Stephen Machin.** 2020. “Youth Crime in the Era of School Takeovers. Evidence from London Secondary School Academies.”
- Duflo, Esther, Pascaline Dupas, and Michael Kremer.** 2011. “Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya.” *American Economic Review*, 101(5): 1739–1774.
- Dustmann, Christian, and Rasmus Landersø.** 2021. “Child’s Gender, Young Fathers’ Crime, and Spillover Effects in Criminal Behavior.” *Journal of Political Economy*, 129(12): 3261–3301.
- Eisenberg, Daniel, Ezra Golberstein, and Janis L. Whitlock.** 2014. “Peer Effects on Risky Behaviors: New Evidence from College Roommate Assignments.” *Journal of Health Economics*, 33: 126–138.
- Eren, Ozkan, and Naci Mocan.** 2021. “Juvenile Punishment, High School Graduation, and Adult Crime: Evidence from Idiosyncratic Judge Harshness.” *Review of Economics and Statistics*, 103(1): 34–47.
- Fadlon, Itzik, and Torben Heien Nielsen.** 2019. “Family Health Behaviors.” *American Economic Review*, 109(9): 3162–3191.
- Fadlon, Itzik, and Torben Heien Nielsen.** 2021. “Family Labor Supply Responses to Severe Health Shocks: Evidence from Danish Administrative Records.” *American Economic Journal: Applied Economics*, 13(3): 1–30.

- Feld, Jan, and Ulf Zölitz.** 2017. “Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects.” *Journal of Labor Economics*, 35(2): 387–428.
- Figlio, David N.** 2007. “Boys Named Sue: Disruptive Children and Their Peers.” *Education Finance and Policy*, 2(4): 376–394.
- Fletcher, Jason M.** 2012. “Peer Influences on Adolescent Alcohol Consumption: Evidence Using an Instrumental Variables/Fixed Effect Approach.” *Journal of Population Economics*, 25(4): 1265–1286.
- Garin, Andrew, Dmitri Koustas, Carl McPherson, Samuel Norris, Matthew Pecenco, Evan K. Rose, Yotam Shem-Tov, and Jeffrey Weaver.** 2025. “The Impact of Incarceration on Employment, Earnings, and Tax Filing.” *Econometrica*, 93(2): 503–538.
- Gavrilova, Evelina, and Marcello Puca.** 2022. “Peer Effects in Crime.” In *A Modern Guide to the Economics of Crime*. 227–242. Edward Elgar Publishing.
- Hall, Stephen, Sarah Harriss, Beth Kerwin, and Martin Fowell.** 2026. “Statistical Digest of Rural England: 4 – Communities and Households.” Department for Environment, Food & Rural Affairs.
- Heeks, Matthew, Sasha Reed, Mariam Tafsiri, and Stuart Prince.** 2018. “The Economic and Social Costs of Crime, Second Edition.” Home Office Research Report99.
- Hjalmarsson, Randi.** 2008. “Criminal Justice Involvement and High School Completion.” *Journal of Urban Economics*, 63(2): 613–630.
- Holzer, Harry J.** 2007. “Collateral Costs: The Effects of Incarceration on the Employment and Earnings of Young Workers.” IZA Discussion Paper 3118.
- Holzer, Harry J., Steven Raphael, and Michael A. Stoll.** 2006. “Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practices of Employers.” *Journal of Law and Economics*, 49(2): 451–480.
- Hoxby, Caroline.** 2000. “Peer Effects in the Classroom: Learning from Gender and Race Variation.” National Bureau of Economic Research.
- Iceland, John, Pablo Mateos, and Gregory Sharp.** 2011. “Ethnic Residential Segregation by Nativity in Great Britain and the United States.” *Journal of Urban Affairs*, 33(4): 409–429.
- Irwin-Rogers, Keir, Abhinay Muthoo, and Luke Billingham.** 2020. “Youth Violence Commission Final Report.” Youth Violence Commission.

- Jenkins, Stephen P., John Micklewright, and Sylke V. Schnepf.** 2008. “Social Segregation in Secondary Schools: How does England Compare with Other Countries?” *Oxford Review of Education*, 34(1): 21–37.
- Johnsen, Julian, and Laura Khoury.** 2024. “Peer Effects in Prison.” IZA Discussion Paper No. 17114.
- Kamada, Yuichiro, and Fuhito Kojima.** 2015. “Efficient Matching under Distributional Constraints: Theory and Applications.” *American Economic Review*, 105(1): 67–99.
- Kamada, Yuichiro, and Fuhito Kojima.** 2024. “Fair Matching Under Constraints: Theory and Applications.” *Review of Economic Studies*, 91(2): 1162–1199.
- Kang, Changhui.** 2007. “Classroom Peer Effects and Academic Achievement: Quasi-randomization Evidence from South Korea.” *Journal of Urban Economics*, 61(3): 458–495.
- Kasuya, Yusuke.** 2016. “Anti-bullying School Choice Mechanism Design.” Mimeo.
- Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz.** 2005. “Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment.” *Quarterly Journal of Economics*, 120(1): 87–130.
- Koppensteiner, Martin Foureaux, and Livia Menezes.** 2021. “Violence and Human Capital Investments.” *Journal of Labor Economics*, 39(3): 787–823.
- Kremer, Michael, and Dan Levy.** 2008. “Peer Effects and Alcohol Use among College Students.” *Journal of Economic Perspectives*, 22(3): 189–206.
- Krutikova, Sonya, Lindsey Macmillan, David Sturrock, and Laura van der Erve.** 2023. “Intergenerational mobility in the UK.” Institute for Fiscal Studies: London.
- Kynaston, David, and Francis Green.** 2019. *Engines of Privilege: Britain’s Private School Problem*. Bloomsbury Publishing Plc.
- Landersø, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen.** 2017. “School Starting Age and the Crime-age Profile.” *Economic Journal*, 127(602): 1096–1118.
- Lavy, Victor, and Assaf Yancu.** 2025. “Violent Peers at School: Impacts and Mechanisms.” National Bureau of Economic Research.
- Lavy, Victor, and Edith Sand.** 2019. “The Effect of Social Networks on Students’ Academic and Non-cognitive Behavioural Outcomes: Evidence from Conditional Random Assignment of Friends in School.” *Economic Journal*, 129(617): 439–480.

- Lavy, Victor, Olmo Silva, and Felix Weinhardt.** 2012. “The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools.” *Journal of Labor Economics*, 30(2): 367–414.
- Lee, Lung-Fei, Xiaodong Liu, Eleonora Patacchini, and Yves Zenou.** 2021. “Who is the Key Player? A Network Analysis of Juvenile Delinquency.” *Journal of Business & Economic Statistics*, 39(3): 849–857.
- Leszczensky, Lars, and Sebastian Pink.** 2015. “Ethnic Segregation of Friendship Networks in School: Testing a Rational-choice Argument of Differences in Ethnic Homophily between Classroom-and Grade-level Networks.” *Social Networks*, 42: 18–26.
- Lindquist, Matthew J., Eleonora Patacchini, Michael Vlassopoulos, and Yves Zenou.** 2024. “Spillovers in Criminal Networks: Evidence from Co-offender Deaths.” CEPR Discussion Paper No. 19159.
- Lochner, Lance.** 2004. “Education, Work, and Crime: A Human Capital Approach.” *International Economic Review*, 45(3): 811–843.
- Lochner, Lance, and Enrico Moretti.** 2004. “The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports.” *American Economic Review*, 94(1): 155–189.
- Lu, Fangwen, and Michael L. Anderson.** 2015. “Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups.” *Journal of Labor Economics*, 33(1): 91–122.
- Machin, Stephen, and Matteo Sandi.** 2020. “Autonomous Schools and Strategic Pupil Exclusion.” *Economic Journal*, 130(625): 125–159.
- Machin, Stephen, Olivier Marie, and Sunčica Vujić.** 2011. “The Crime Reducing Effect of Education.” *Economic Journal*, 121(552): 463–484.
- McAra, Lesley, and Susan McVie.** 2010. “Youth Crime and Justice: Key Messages from the Edinburgh Study of Youth Transitions and Crime.” *Criminology & Criminal Justice*, 10(2): 179–209.
- McGloin, Jean Marie, and Kyle J. Thomas.** 2019. “Peer influence and delinquency.” *Annual Review of Criminology*, 2: 241–264.
- Mueller-Smith, Michael.** 2015. “The Criminal and Labor Market Impacts of Incarceration.” Working Paper.
- Murphy, Richard, and Felix Weinhardt.** 2020. “Top of the Class: The Importance of Ordinal Rank.” *Review of Economic Studies*, 87(6): 2777–2826.

- Nakajima, Ryo.** 2007. “Measuring Peer Effects on Youth Smoking Behaviour.” *Review of Economic Studies*, 74(3): 897–935.
- Neumark, David, and Rosella Gardecki.** 1998. “Women Helping Women? Role Model and Mentoring Effects on Female Ph. D. Students in Economics.” *Journal of Human Resources*, 33(1): 220–246.
- Noble, Michael, David McLennan, Kate Wilkinson, Adam Whitworth, Helen Barnes, and Chris Dibben.** 2008. “The English Indices of Deprivation 2007.” Communities and Local Government: London.
- Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver.** 2021. “The Effects of Parental and Sibling Incarceration: Evidence from Ohio.” *American Economic Review*, 111(9): 2926–2963.
- Ouss, Aurélie, and Arnaud Philippe.** 2025. “Prison Peer Effects and the Public Safety Trade-off.” Working Paper.
- Owens, Emily G.** 2017. “Testing the School-to-Prison Pipeline.” *Journal of Policy Analysis and Management*, 36(1): 11–37.
- Pager, Devah.** 2003. “The Mark of a Criminal Record.” *American Journal of Sociology*, 108(5): 937–975.
- Patacchini, Eleonora, and Yves Zenou.** 2016. “Racial Identity and Education in Social Networks.” *Social Networks*, 44: 85–94.
- Pathak, Parag A., and Tayfun Sönmez.** 2013. “School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation.” *American Economic Review*, 103(1): 80–106.
- Porter, Catherine, and Danila Serra.** 2020. “Gender Differences in the Choice of Major: The Importance of Female Role Models.” *American Economic Journal: Applied Economics*, 12(3): 226–254.
- Powell, Lisa M., John A. Tauras, and Hana Ross.** 2005. “The Importance of Peer Effects, Cigarette Prices and Tobacco Control Policies for Youth Smoking Behavior.” *Journal of Health Economics*, 24(5): 950–968.
- Rambachan, Ashesh, and Jonathan Roth.** 2023. “A More Credible Approach to Parallel Trends.” *Review of Economic Studies*, 90(5): 2555–2591.
- Riley, Emma.** 2024. “Role Models in Movies: The Impact of Queen of Katwe on Students’ Educational Attainment.” *Review of Economics and Statistics*, 106(2): 334–351.

- Romano, Joseph P., and Michael Wolf.** 2005. “Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing.” *Journal of the American Statistical Association*, 100(469): 94–108.
- Rose, Evan K.** 2021. “Does Banning the Box Help Ex-offenders Get Jobs? Evaluating the Effects of a Prominent Example.” *Journal of Labor Economics*, 39(1): 79–113.
- Roth, Jonathan.** 2026. “Interpreting Event-studies from Recent Difference-in-differences Methods.” *Japanese Economic Review*, 77(2): 275–288.
- Sabates, Ricardo.** 2008. “Educational Attainment and Juvenile Crime: Area-level Evidence Using Three Cohorts of Young People.” *British Journal of Criminology*, 48(3): 395–409.
- Sacerdote, Bruce.** 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *Quarterly Journal of Economics*, 116(2): 681–704.
- Sant’Anna, Pedro H.C., and Jun Zhao.** 2020. “Doubly Robust Difference-in-differences Estimators.” *Journal of Econometrics*, 219(1): 101–122.
- Serra, Danila.** 2025. “Role Models in Developing Countries.” In *Handbook of Experimental Development Economics*. 123–135. Edward Elgar Publishing.
- Stevenson, Megan.** 2017. “Breaking Bad: Mechanisms of Social Influence and the Path to Criminality in Juvenile Jails.” *Review of Economics and Statistics*, 99(5): 824–838.
- Sviatschi, María Micaela.** 2022. “Spreading Gangs: Exporting US Criminal Capital to El Salvador.” *American Economic Review*, 112(6): 1985–2024.
- Tincani, Michela M.** 2026. “Peer Effects and Rank Concerns in the Classroom.” *Journal of Political Economy*, 134(10): 000–000.
- Villa, Carmen.** 2024. “The Effects of Youth Clubs on Education and Crime.” IFS Working Paper.
- Welch, Kelly, and Allison Ann Payne.** 2012. “Exclusionary School Punishment: The Effect of Racial Threat on Expulsion and Suspension.” *Youth Violence and Juvenile Justice*, 10(2): 155–171.
- Wooldridge, Jeffrey M.** 2023. “Simple Approaches to Nonlinear Difference-in-differences with Panel Data.” *Econometrics Journal*, 26(3): C31–C66.
- Wooldridge, Jeffrey M.** 2025. “Two-way Fixed Effects, the Two-way Mundlak Regression, and Difference-in-differences Estimators.” *Empirical Economics*, 69(5): 2545–2587.

Youth Justice Board, and Ministry of Justice. 2023. “Youth Justice Statistics: 2021 to 2022.”

Zimmerman, David J. 2003. “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment.” *Review of Economics and Statistics*, 85(1): 9–23.

A Data Details

The NPD extract I received contains 164.6 million panel records for 1.1 million individuals. To make the dataset manageable and to facilitate tracking the analysis, I restrict the sample to the 1995/96–1998/99 cohorts and exclude records outside England.⁹⁶ I then restrict the sample to individuals observed at ages 10 and 15 in September so that I can track them throughout secondary school. Table A1 in the Appendix summarizes the resulting data structure, drawing on the Appendix tables in Cattán et al. (2023a,b, 2025).

My raw PNC extract contains 8,844,526 recorded offenses committed by 1,431,443 individuals, together with 15,805,405 penalty records. I first exclude offenses with missing dates or dates before 2000 or after 2017, leaving 8,749,087 offenses, 15,647,338 penalty records, and 1,423,472 individuals.⁹⁷ I then exclude offenses that did not result in a finding of guilt, leaving 6,836,862 offenses, 13,683,785 penalty records, and 1,422,793 individuals.

Using these cleaned datasets, I merge the NPD and PNC through the unique offender identifiers provided by the MoJ. I then drop 558 individuals with offense records before age 10.⁹⁸ The resulting analysis sample contains 1,957,990 individuals and 17,519,984 panel records, of whom 281,005 have an offense record. Table A2 in the Appendix reports the linkage rates.

Column (1) reports the number of individuals with a complete set of controls, including year and month of birth, secondary-school ID, SEN status, FSM status, an indicator for whether English is their first language, and ethnicity. Most individuals have complete information on these controls, although completeness rates are slightly lower among minority groups (i.e., those who are not White British), replicating the pattern in Column (2) of Table A1.

Among individuals in Column (1), Column (2) reports the fraction with a unique offender identifier. Among those in Column (2), Column (3) reports the fraction with an actual offense record in my PNC extract, while Column (4) restricts this measure to offenses committed by the end of secondary school. Column (5) calculates the juvenile offense rate as Column (2) times Column (4). Finally, Column (6) reports the fraction of linked observations matched at the most precise level, using first and last names, gender, postcode, and date of birth.

⁹⁶Also, before 1995 cohorts, the school ID was collected annually, not for each census. This would cause the analysis based on school transfers to be much more imprecise.

⁹⁷Some cases have multiple offense start dates for the same case and offense. In this cleaning process, I keep cases and offenses if they have at least one record between 2000 and 2017.

⁹⁸I dropped these observations after merging the two data sets because I calculated the offense age using the year and month of birth in NPD but not PNC.

Table A1. Data Structure

Year of Birth Cohort	Available test scores/exams:						Sample size		
	Early Yrs Age 5	KS1 Age 7	KS2 Age 11	KS3 Age 14	KS4 Age 16	KS5 Age 18	(1) NPD Records	(2) NPD Records with control	(2)/(1)
1990/91			X	X	X	X	659,546	637,949	96.7%
1991/92			X	X	X	X	661,511	635,280	96.0%
1992/93			X	X	X	X	645,005	613,926	95.2%
1993/94			X	X	X	X	649,414	612,748	94.4%
1994/95		X	X		X	X	641,308	599,006	93.4%
1995/96		X	X		X	X	638,645	592,244	92.7%
1996/97		X	X		X	X	654,837	602,703	92.0%
1997/98	X	X	X		X	X	641,619	589,475	91.9%
1998/99	X	X			X	X	632,221	583,545	92.3%

Notes: We report details of the data linkage between the National Pupil Database (NPD) and the Police National Computer (PNC), where we refer to the PNC as shorthand for linked data from the Police National Computer, the Home Office Court Appearance System, and the National Offender Management Information System. Each row shows information by year of birth cohort. The controls referred to in Column (2) are whether the individual is ever classified as having special educational needs, is ever eligible for free school meals, whether English is a first language, their minority status (as captured through 12 categories: White British, Other White, Black African, Black Caribbean, Other Black, Indian, Pakistani, Bangladeshi, Other Asian, Mixed, Others and Unknown), cohort, month of birth, and the secondary school identifiers.

Table A2. Data Linkage

Birth Cohort	NPD sample with controls	Share of (1) with offense ID	Share of (2) with		Share of Individuals with offenses by end of sec. school	Share of (5) with the perfect match
			any offense	offenses by end of sec. school		
	(1)	(2)	(3)	(4)	(5)	(6)
1990/91	637,949	25.0%	81.9%	45.5%	11.4%	80.7%
1991/92	635,280	23.1%	82.6%	48.5%	11.2%	80.8%
1992/93	613,926	20.7%	82.8%	50.5%	10.4%	81.2%
1993/94	612,748	17.9%	83.1%	51.2%	9.15%	81.0%
1994/95	599,006	15.0%	83.2%	50.9%	7.65%	80.2%
1995/96	592,244	11.9%	83.1%	49.5%	5.90%	79.3%
1996/97	602,703	9.15%	83.8%	49.3%	4.51%	79.4%
1997/98	589,475	6.66%	85.9%	51.8%	3.45%	77.5%
1998/99	583,545	4.53%	91.3%	62.3%	2.82%	76.3%

Notes: I report details of the data linkage between the National Pupil Database (NPD) and the Police National Computer (PNC) where I refer to the PNC as shorthand for linked data from the Police National Computer, the Home Office Court Appearance System, and the National Offender Management Information System. Each row shows information by year of birth cohort. Column (6) indicates the quality of data linkages: we show the share of links that are perfect matches using exact information on names, date of birth, zip code of residence, and sex.

Table A3. Sample Selection

	Among NPD sample with controls Complete Records during Secondary School				Among Those in Column (1) Experienced a School-grade less than 20 Students				Among Those in Column (5) Ever Exposed to An Offender's School Transfer			
	Complete	Not Complete	Uncond. Diff. [(1) = (2)]	Cond. Diff. [(1) = (2)]	Never Experienced	Experienced	Uncond. Diff. [(5) = (6)]	Cond. Diff. [(5) = (6)]	Exposed	Never Exposed	Uncond. Diff. [(9) = (10)]	Cond. Diff. [(9) = (10)]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Demographics												
Male	.510	.540	-.031*** (.001)	-.019*** (.001)	.506	.710	-.203*** (.002)	-.124*** (.004)	.518	.494	.024*** (.001)	.004*** (.001)
Minority	.207	.500	-.293*** (.001)	-.084*** (.001)	.208	.191	.016*** (.002)	.009*** (.003)	.185	.231	-.046*** (.001)	.000 (.000)
English as the first language	.892	.650	.242*** (.001)	.058*** (.001)	.892	.908	-.016*** (.002)	-.018*** (.002)	.902	.882	.019*** (.000)	.001 (.000)
Ever with SEN by age 11	.351	.353	-.002** (.001)	.006*** (.001)	.341	.941	-.600*** (.001)	-.296*** (.004)	.365	.316	.049*** (.001)	.003*** (.001)
Ever eligible for FSM by age 11	.274	.322	-.049*** (.001)	-.010*** (.001)	.270	.508	-.238*** (.003)	-.175*** (.005)	.296	.243	.054*** (.001)	.004*** (.001)
Offender by age 16	.214	.237	-.024*** (.001)	-.025*** (.001)	.210	.451	-.241*** (.003)	-.555*** (.004)	.243	.176	.067*** (.001)	-.006*** (.001)
Any school transfer by age 16	1.17 [.439]	1.26 [.559]	-.091*** (.002)	-.070*** (.003)	1.16 [.431]	1.30 [.607]	-.144*** (.005)	-.220*** (.008)	1.18 [.445]	1.14 [.410]	.036*** (.001)	-.012*** (.002)
N of school transfers by age 16 any transfer	.031	.135	-.104*** (.001)	-.110*** (.001)	.030	.124	-.094*** (.002)	-.204*** (.004)	.038	.021	.016*** (.000)	.003*** (.000)
B. Education												
Age 5 overall score	.017 [.992]	-.268 [1.08]	.285*** (.015)	.080*** (.013)	.029 [.978]	-1.51 [1.49]	1.54*** (.057)	.425*** (.054)	-.011 [.967]	.059 [.985]	-.070*** (.007)	.016 (.008)
Age 5 score: personal, social and emotional development	.018 [.989]	-.282 [1.12]	.300*** (.015)	.132*** (.014)	.030 [.975]	-1.47 [1.53]	1.50*** (.059)	.480*** (.057)	-.005 [.970]	.055 [.977]	-.061*** (.007)	.018** (.009)
Age 5 score: communication, language, and literacy	.019 [.992]	-.290 [1.07]	.309*** (.015)	.075*** (.012)	.030 [.982]	-1.41 [1.27]	1.44*** (.049)	.383*** (.048)	.070 [.978]	.070 [.983]	-.095*** (.007)	.023*** (.008)
Age 5 score: mathematical development	.016 [.990]	-.257 [1.12]	.274*** (.015)	.065*** (.013)	.028 [.975]	-1.48 [1.58]	1.51*** (.061)	.416*** (.058)	.062 [.977]	.062 [.972]	-.079*** (.007)	.020** (.009)
Age 5 score: knowledge and understanding the world	.012 [.992]	-.191 [1.11]	.204*** (.015)	.042*** (.014)	.022 [.980]	-1.27 [1.56]	1.29*** (.060)	.393*** (.060)	-.006 [.983]	.042 [.977]	-.048*** (.007)	.031*** (.009)
Age 5 score: physical development	.011 [.992]	-.178 [1.11]	.189*** (.015)	.051*** (.014)	.022 [.975]	-1.41 [1.78]	1.43*** (.068)	.443*** (.068)	-.001 [.973]	.039 [.976]	-.040*** (.007)	.033*** (.009)
Age 5 score: creative development	.012 [.993]	-.190 [1.09]	.202*** (.015)	.040*** (.014)	.022 [.980]	-1.31 [1.61]	1.33*** (.062)	.330*** (.059)	-.009 [.982]	.045 [.977]	-.054*** (.007)	.022** (.009)
Age 7 overall score	.026 [.988]	-.375 [1.09]	.401*** (.003)	.099*** (.002)	.039 [.972]	-1.85 [1.42]	1.89*** (.012)	.504*** (.012)	-.030 [.961]	.109 [.979]	-.139*** (.001)	-.002 (.001)
Age 7 score (English, math, science)	.025 [.990]	-.357 [1.08]	.382*** (.003)	.094*** (.002)	.037 [.973]	-1.84 [1.49]	1.88*** (.013)	.503*** (.012)	-.028 [.959]	.105 [.983]	-.133*** (.001)	-.003 (.001)
Age 7 score: English reading	.025 [.988]	-.367 [1.10]	.392*** (.003)	.098*** (.002)	.037 [.976]	-1.62 [1.26]	1.66*** (.011)	.428*** (.011)	-.030 [.974]	.105 [.973]	-.135*** (.001)	-.000 (.001)
Age 7 score: English writing	.025 [.987]	-.365 [1.11]	.391*** (.003)	.103*** (.002)	.037 [.975]	-1.68 [1.23]	1.71*** (.011)	.472*** (.011)	-.028 [.970]	.103 [.975]	-.131*** (.001)	-.002 (.001)
Age 7 score: English listening and speaking	.020 [.992]	-.293 [1.06]	.314*** (.003)	.065*** (.003)	.031 [.980]	-1.60 [1.41]	1.64*** (.012)	.418*** (.012)	-.025 [.967]	.090 [.989]	-.115*** (.001)	-.003 (.002)
Age 7 score: overall English	.025 [.989]	-.364 [1.09]	.389*** (.003)	.092*** (.002)	.038 [.974]	-1.78 [1.36]	1.82*** (.012)	.476*** (.012)	-.030 [.964]	.107 [.979]	-.137*** (.001)	-.002 (.001)
Age 7 score: math	.022 [.991]	-.318 [1.07]	.340*** (.003)	.089*** (.003)	.033 [.978]	-1.65 [1.45]	1.68*** (.013)	.471*** (.013)	-.025 [.969]	.093 [.984]	-.118*** (.001)	-.002 (.001)
Age 7 score: science	.020 [.993]	-.292 [1.06]	.312*** (.003)	.076*** (.003)	.031 [.980]	-1.58 [1.49]	1.61*** (.013)	.424*** (.013)	-.022 [.966]	.085 [.991]	-.107*** (.001)	-.003 (.002)
Age 11 overall score	.026 [.989]	-.387 [1.08]	.412*** (.004)	.148*** (.003)	.029 [.987]	-1.14 [1.15]	1.17*** (.019)	.418*** (.019)	-.066 [.983]	.140 [.981]	-.206*** (.002)	-.007*** (.002)
Age 11 score: English	.025 [.990]	-.373 [1.07]	.398*** (.003)	.128*** (.003)	.028 [.987]	-1.21 [1.11]	1.23*** (.018)	.412*** (.017)	-.067 [.983]	.142 [.980]	-.209*** (.002)	-.008*** (.002)
Age 11 score: math	.024 [.992]	-.351 [1.05]	.375*** (.003)	.138*** (.003)	.027 [.990]	-1.06 [1.08]	1.08*** (.016)	.404*** (.017)	-.058 [.987]	.128 [.984]	-.186*** (.002)	-.005*** (.002)
Age 11 score: science	.026 [.986]	-.380 [1.12]	.407*** (.004)	.161*** (.003)	.031 [.982]	-1.18 [1.26]	1.21*** (.017)	.464*** (.018)	-.047 [.990]	.123 [.965]	-.170*** (.002)	-.004** (.002)
Age 16 score: taken	1.00 (.001)	.805 (.001)	.195*** (.001)	.198*** (.001)	1.00 (.000)	.997 (.000)	.002*** (.001)	.003*** (.001)	1.00 (.001)	1.00 (.000)	.000 (.000)	.000 (.000)
Age 16 score: 5 A* – G	.872 (.001)	.536 (.001)	.336*** (.001)	.330*** (.001)	.884 (.001)	.150 (.002)	.734*** (.002)	.432*** (.004)	.857 (.000)	.912 (.000)	-.056*** (.000)	-.006*** (.001)
Age 16 score: 5 A* – C	.607 (.001)	.324 (.001)	.283*** (.001)	.249*** (.001)	.616 (.001)	.051 (.003)	.565*** (.001)	.209*** (.003)	.557 (.000)	.676 (.001)	-.119*** (.001)	-.010*** (.001)
Age 16 score: point score taken	.083 [.932]	-.932 [1.24]	1.01*** (.003)	.994*** (.003)	.098 [.916]	-2.00 [.713]	2.09*** (.006)	1.17*** (.009)	.018 [.912]	.178 [.913]	-.160*** (.001)	.001 (.001)
Sample size	2,123,028	244,939	2,367,967	2,367,967	2,088,731	34,297	2,123,028	2,123,028	1,050,567	1,038,164	2,088,731	2,088,731

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. The sample in Columns (1)–(4) consists of observations in the NPD with the full set of control variables. The sample size corresponds to the sum of the 1995/96–1998/99 cohorts reported in Column (2) of Table A1 and Column (1) of Table A2. Conditional differences control for gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. All test scores are standardized within each cohort, except for the indicator variables for the age 16 tests. Early-years test scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month (Department for Education, Schools Analysis and Research Division, 2010). The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 in constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

Table A4. Summary Statistics on Non-offenders and Offenders

	Offending Status by Age 16			
	Non-offenders	Offenders	Uncond. Diff.	Cond. Diff.
	(1)	(2)	[(1) = (2)] (3)	[(1) = (2)] (4)
A. Demographics				
Male	.511	.697	-.186*** (.002)	-.157*** (.002)
Minority	.185	.171	.014*** (.002)	-.007*** (.001)
English as the first language	.901	.930	-.029*** (.001)	-.007*** (.001)
Ever with SEN by age 11	.356	.604	-.248*** (.003)	-.148*** (.002)
Ever eligible for FSM by age 11	.285	.591	-.306*** (.003)	-.235*** (.002)
Any school transfer by age 16	.238	.387	-.149*** (.002)	-.091*** (.002)
N of school transfers by age 16 any transfer	1.17 [.434]	1.30 [.587]	-.130*** (.005)	-.083*** (.005)
Sample size	1,010,977	39,590	1,050,567	1,050,567
B. Age 5 Education				
Age 5 overall score	.018 [.939]	-.328 [.964]	.346*** (.031)	.033 (.028)
Age 5 score: personal, social and emotional development	.024 [.941]	-.325 [1.02]	.349*** (.032)	.066** (.032)
Age 5 score: communication, language, and literacy	.004 [.957]	-.400 [.977]	.404*** (.031)	.030 (.028)
Age 5 score: mathematical development	.012 [.945]	-.344 [1.05]	.356*** (.033)	.059 (.031)
Age 5 score: knowledge and understanding the world	.017 [.959]	-.178 [1.02]	.194*** (.032)	-.012 (.033)
Age 5 score: physical development	.022 [.941]	-.098 [.983]	.120*** (.031)	-.088*** (.032)
Age 5 score: creative development	.016 [.955]	-.245 [1.05]	.261*** (.033)	-.001 (.034)
Sample size	35,736	1,013	36,749	36,749
C. Age 7 Education				
Age 7 overall score	.010 [.927]	-.444 [.953]	.454*** (.005)	.117*** (.004)
Age 7 score (English, math, science)	.010 [.925]	-.400 [.942]	.410*** (.005)	.104*** (.004)
Age 7 score: English reading	.007 [.948]	-.476 [.994]	.483*** (.005)	.135*** (.004)
Age 7 score: English writing	.009 [.944]	-.460 [1.01]	.469*** (.005)	.119*** (.005)
Age 7 score: English listening and speaking	.009 [.941]	-.365 [.949]	.374*** (.005)	.091*** (.004)
Age 7 score: overall English	.009 [.934]	-.458 [.962]	.467*** (.005)	.120*** (.004)
Age 7 score: math	.008 [.940]	-.327 [.966]	.336*** (.005)	.081*** (.005)
Age 7 score: science	.011 [.938]	-.310 [.944]	.321*** (.005)	.083*** (.005)
Sample size ²	961,786	37,914	999,700	999,700
D. Age 11 Education				
Age 11 overall score	-.042 [.976]	-.577 [.971]	.535*** (.006)	.226*** (.005)
Age 11 score: English	-.042 [.975]	-.616 [.972]	.574*** (.006)	.217*** (.005)
Age 11 score: math	-.038 [.982]	-.462 [.971]	.424*** (.005)	.168*** (.005)
Age 11 score: science	-.022 [.979]	-.529 [1.03]	.507*** (.006)	.240*** (.005)
Sample size ³	783,184	31,952	815,136	815,136
E. Age 16 Education				
Age 16 score: taken	1.00	.998	.001*** (.000)	.001*** (.000)
Age 16 score: 5 A* – G taken	.871	.491	.380*** (.003)	.275*** (.002)
Age 16 score: 5 A* – C taken	.573	.143	.430*** (.002)	.238*** (.002)
Age 16 score: point score taken	.078 [.868]	-.924 [.995]	1.00*** (.005)	.720*** (.005)
Sample size	1,010,795	39,526	1,050,321	1,050,321

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. Conditional differences control for gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. Offenders in this table are defined as individuals in the NPD with at least one recordable offense in the PNC committed between the September ages of 11 and 16 that resulted in a caution or conviction. All test scores are standardized within each cohort, except for the indicator variables for the age 16 tests. Early-years test scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month (Department for Education, Schools Analysis and Research Division, 2010). The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 in constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

Table A5. Detailed Common Offense Descriptions

Rank	Count	Detailed offense description
Theft Offenses		
1	23640	Theft from shops
2	6554	Other stealing and unauthorised taking: Offences under the Theft Act 1968 S.1 not classified elsewhere
3	6169	Robbery
4	4735	Burglary in a building other than a dwelling with intent to steal/infect grievous bodily harm/commit damage – triable either way
5	4370	Other burglary in a dwelling
6	2572	Theft of pedal cycle
7	2392	Stealing from the person of another
8	1589	Theft in a dwelling other than from automatic machine or meter
9	1442	Theft from a motor vehicle
10	2143	Others
Battery Offenses		
1	30631	Common assault and battery
2	101	Assaulting a designated or accredited person, or person assisting him or her, in the execution of his or her duty
3	16	Resisting or wilfully obstructing a designated or accredited person, or person assisting, in the execution of his or her duty
4	15	Others
Drug-related Offenses		
1	6231	Having possession of a controlled drug – class B (cannabis, including cannabis resin, cannabinalol and cannabinalol derivatives)
2	264	Having possession of a controlled drug – class B (other)
3	192	Possession of a controlled drug with intent to supply – class B (cannabis, including cannabis resin, cannabinalol and cannabinalol derivatives)
4	116	Supplying or offering to supply a controlled drug – class B (cannabis, including cannabis resin, cannabinalol and cannabinalol derivatives)
5	105	Having possession of a controlled drug – class C (cannabis, including cannabis resin, cannabinalol and cannabinalol derivatives) – historic
6	66	Having possession of a controlled drug – class A (cocaine)
7	57	Production of or being concerned in production of a controlled drug – class B (cannabis, including cannabis resin, cannabinalol and cannabinalol derivatives)
8	51	Possession of a controlled drug with intent to supply – class A (heroin)
9	48	Possession of a controlled drug with intent to supply – class A (crack cocaine)
10	525	Others
Violent Offenses		
1	7401	Assaults occasioning actual bodily harm
2	5242	Causing harassment, alarm or distress
3	1952	Causing fear or provocation of violence
4	1386	Assault on a constable
5	1252	Breach of Anti-Social Behaviour Order and Interim ASBO (order made to protect from alarm, distress or harassment)
6	1148	Affray
7	1121	Causing intentional harassment, alarm or distress
8	596	Wound / inflict grievous bodily harm without intent
9	309	Racially or religiously aggravated common assault or beating
10	1822	Others
Fraud Offenses		
1	1294	Receiving stolen goods
2	1079	Undertaking or assisting in the retention, removal, disposal or realisation of stolen goods, or arranging to do so
3	567	Dishonestly making a false representation to make a gain for oneself or another or to cause loss to another or to expose another to a risk
4	178	Travelling by railway without paying correct fare, failing to show ticket, failing to give name and address, etc.
5	81	Attempting to pervert the course of public justice (fabrication of false evidence, causing person to be wrongly convicted, interference with witness) – indictable only
6	35	Communicating false information alleging the presence of bombs
7	31	Possession etc. of articles for use in frauds
8	23	Passing etc. counterfeit coin or note as genuine
9	23	Obtaining pecuniary advantage by deception (except railway frauds)
10	74	Others
Weapon Offenses		
1	1774	Having an article with a blade or point in a public place
2	1323	Possession of offensive weapons without lawful authority or reasonable excuse
3	588	Having an article with a blade or point on school premises
4	232	Possession of offensive weapons without lawful authority or reasonable excuse on school premises
5	158	Possess imitation firearm or loaded/unloaded air weapon in public place
6	152	Carrying a loaded or unloaded or imitation firearm or air weapon in public place
7	98	Possession of a firearm or imitation firearm, with intent to cause fear of violence (Group I)
8	67	Possessing anything with intent to destroy or damage property
9	52	Possessing or distributing prohibited weapons designed for discharge of noxious liquid etc. (Group I)
10	148	Others
Damage/Arson Offenses		
1	21708	Causing damage to an allotment through negligence or any unlawful act
2	3410	Other criminal damage
3	1503	Arson not endangering life
4	114	Threat to commit criminal damage
5	105	Arson endangering life
6	70	Unlawfully and maliciously causing or conspiring to cause explosion likely to endanger life or property
7	60	Racially or religiously aggravated criminal damage
8	43	Racially aggravated other criminal damage
9	17	Others
Sexual Offenses		
1	635	Sexual assault on a female aged 13 or over – no penetration
2	365	Sexual assault on a female child aged under 13 – touching
3	245	Taking, permitting to be taken or making, distributing or publishing indecent photographs or pseudo photographs of children
4	185	Rape of a female child aged under 13 by a male
5	177	Exposure
6	142	Sexual assault on a male child aged under 13 – touching
7	107	Rape of a male child aged under 13 by a male
8	103	Sexual assault on a female child aged under 13 – penetration
9	100	Sexual activity with a female child aged 13 to 15 – offender aged under 18 – penetration
10	1070	Others
Motoring Offenses		
1	1161	Unauthorised taking or driving or attempt taking or driving of a motor vehicle
2	457	Being carried knowing vehicle to have been taken or driven away
3	457	Interference with motor vehicle/trailer
4	344	Using motor vehicle uninsured against third party risks
5	247	Taking or riding a pedal cycle without consent etc.
6	195	Driving causing or permitting a person to drive other than in accordance with a licence (full or provisional) (except HGV)
7	124	Unauthorised taking or driving of conveyance other than motor vehicle or pedal cycle
8	53	Aiding, abetting, causing or permitting reckless driving
9	50	Driving whilst disqualified from holding or obtaining a licence
10	196	Others
Other Offenses		
1	3692	[description missing in source]
2	1385	Absconding by person released on bail
3	1222	Being guilty, while drunk, of disorderly behaviour
4	708	Resisting or obstructing a constable in execution of duty – offences under the Police Act 1996
5	495	Racially or religiously aggravated intentional harassment or alarm or distress – words or writing
6	485	Offence of harassment
7	311	Improper use of public electronic communications network
8	300	Racially or religiously aggravated harassment or alarm or distress – words or writing
9	277	Being on enclosed premises for an unlawful purpose
10	1671	Others

Table A6. Summary Statistics on Schools and School Transfers, Not-yet-offenders

	All school (1)	Transfer (school-term) level									
		Offenders' transfers				Not-yet-offenders' transfers				Difference-in-differences	
		Origin school (2)	Destination school (3)	Uncond. Diff. [(2)=(3)] (4)	Cond. Diff. [(2)=(3)] (5)	Origin school (6)	Destination school (7)	Uncond. Diff. [(6)=(7)] (8)	Cond. Diff. [(6)=(7)] (9)	Uncond. Diff. [(4)=(8)] (10)	Cond. Diff. [(4)=(8)] (11)
Male	.531 [.174]	.536 [.133]	.544 [.122]	-.008*** (.001)	-.002** (.001)	.531 [.136]	.530 [.123]	.000 (.001)	-.001 (.001)	-.009*** (.002)	-.002 (.002)
Minority	.216 [.256]	.193 [.235]	.185 [.222]	.008*** (.002)	-.002** (.001)	.213 [.255]	.201 [.243]	.011*** (.002)	.002*** (.001)	-.003 (.003)	-.004*** (.001)
English as the first language	.885 [.174]	.890 [.167]	.899 [.153]	-.009*** (.002)	-.004*** (.001)	.880 [.177]	.887 [.168]	-.007*** (.002)	.000 (.001)	-.002 (.002)	-.003*** (.001)
Ever with SEN by age 11	.416 [.199]	.422 [.143]	.444 [.158]	-.022*** (.001)	-.014*** (.001)	.412 [.128]	.416 [.129]	-.003*** (.001)	-.011*** (.001)	-.019*** (.002)	-.002** (.001)
Ever eligible for FSM by age 11	.331 [.197]	.378 [.188]	.397 [.200]	-.019*** (.002)	-.006*** (.001)	.366 [.183]	.365 [.182]	.001 (.002)	-.002** (.001)	-.020*** (.003)	-.002 (.002)
Early years overall score	-.152 [.686]	-.013 [.060]	-.015 [.049]	.003 (.003)	-.004 (.003)	-.008 [.069]	-.008 [.042]	-.000 (.003)	-.001 (.002)	.003 (.004)	-.002 (.004)
Age 7 overall score	-.210 [.744]	-.132 [.307]	-.168 [.338]	.036*** (.003)	.002 (.002)	-.107 [.295]	-.119 [.313]	.011*** (.003)	.011*** (.002)	.024*** (.004)	-.009*** (.003)
Age 11 overall score	-.180 [.590]	-.201 [.283]	-.227 [.283]	.026*** (.003)	.008*** (.002)	-.182 [.285]	-.188 [.270]	.006** (.003)	.011*** (.002)	.020*** (.004)	-.002 (.003)
Offender	.042 [.101]	.054 [.103]	.091 [.150]	-.036*** (.001)	-.028*** (.001)	.027 [.054]	.036 [.071]	-.009*** (.001)	-.011*** (.001)	-.027*** (.001)	-.017*** (.001)
N of students	154 [75.6]	169 [73.1]	179 [92.8]	-9.28*** (.800)	-16.2*** (.669)	166 [66.5]	190 [83.3]	-24.4*** (.694)	-25.0*** (.633)	15.1*** (1.06)	9.51*** (.921)
Sample size	4,559	21,803	21,803	43,606	43,606	23,590	23,590	47,180	47,180	90,786	90,786

Notes: *** denotes significance at the 1

Table A7. Detailed Common Offense Descriptions that Lead to Caution Marginally

Rank	Count	Detailed offense description
Theft Offenses		
1	101,395	Burglary in a building other than a dwelling with intent to steal/inflict grievous bodily harm/commit damage – triable either way
2	46,714	Theft from a motor vehicle
3	31,128	Theft of pedal cycle
4	23,280	Theft in a dwelling other than from automatic machine or meter
5	12,962	Theft of a motor vehicle (excluding aggravated vehicle taking) – triable either way
6	594	Theft from a vehicle – other than a motor vehicle
Battery Offenses		
1	476,403	Common assault and battery
Drug-related Offenses		
1	4,638	Having possession of a controlled drug – class A (MDMA)
2	777	Production or being concerned in production of a controlled drug – class C (cannabis, including cannabis resin, cannabinal and cannabinal derivatives) (historic)
3	458	Supplying or offering to supply a controlled drug – class C (other)
Violent Offenses		
1	115,566	Causing harassment, alarm or distress
2	1,404	Making threats to kill
3	625	Sending letters etc. with intent to cause distress or anxiety
Fraud Offenses		
1	10,746	Failure to comply with requirements concerning approval and signing of abbreviated accounts
2	8,341	Travelling by railway without paying correct fare, failing to show ticket, failing to give name and address, etc.
3	1,240	Passing etc. counterfeit coin or note as genuine
4	1,008	Possession etc. of articles for use in frauds
5	476	Forgery or copying false instrument
Weapon Offenses		
1	56,984	Possession of offensive weapons without lawful authority or reasonable excuse
2	3,457	Possession of a firearm or imitation firearm, with intent to cause fear of violence (Group I)
3	1,716	Carrying a loaded or unloaded or imitation firearm or air weapon in public place
Sexual Offenses		
1	684	Sexual activity with a female child aged 13 to 15 – offender aged 18 or over – penetration
2	287	Exposing the person in any street, etc. or in view thereof, or in any place or public resort with intent to insult any female
Motoring Offenses		
1	23,994	Being carried knowing vehicle to have been taken or driven away
Other Offenses		
1	48,702	Being guilty, while drunk, of disorderly behaviour
2	12,358	Offence of harassment
3	4,330	Improper use of public electronic communications network
4	2,773	Contravention of a direction given by a constable under Section 30(4) [dispersal of groups and removal of persons under 16 to their place of residence]
5	1,583	Offences against private Acts relating to railways
6	1,056	Being found drunk in a highway or other public place whether a building or not, or a licensed premises
7	613	Railway byelaws as to the conduct of passengers
8	468	False alarms of fire (Fire Services Act 1947)
9	205	Offences against the State and Public Order: Other offences under which proceedings were taken
10	202	Possessing fireworks etc. in, or when entering, a designated sports ground

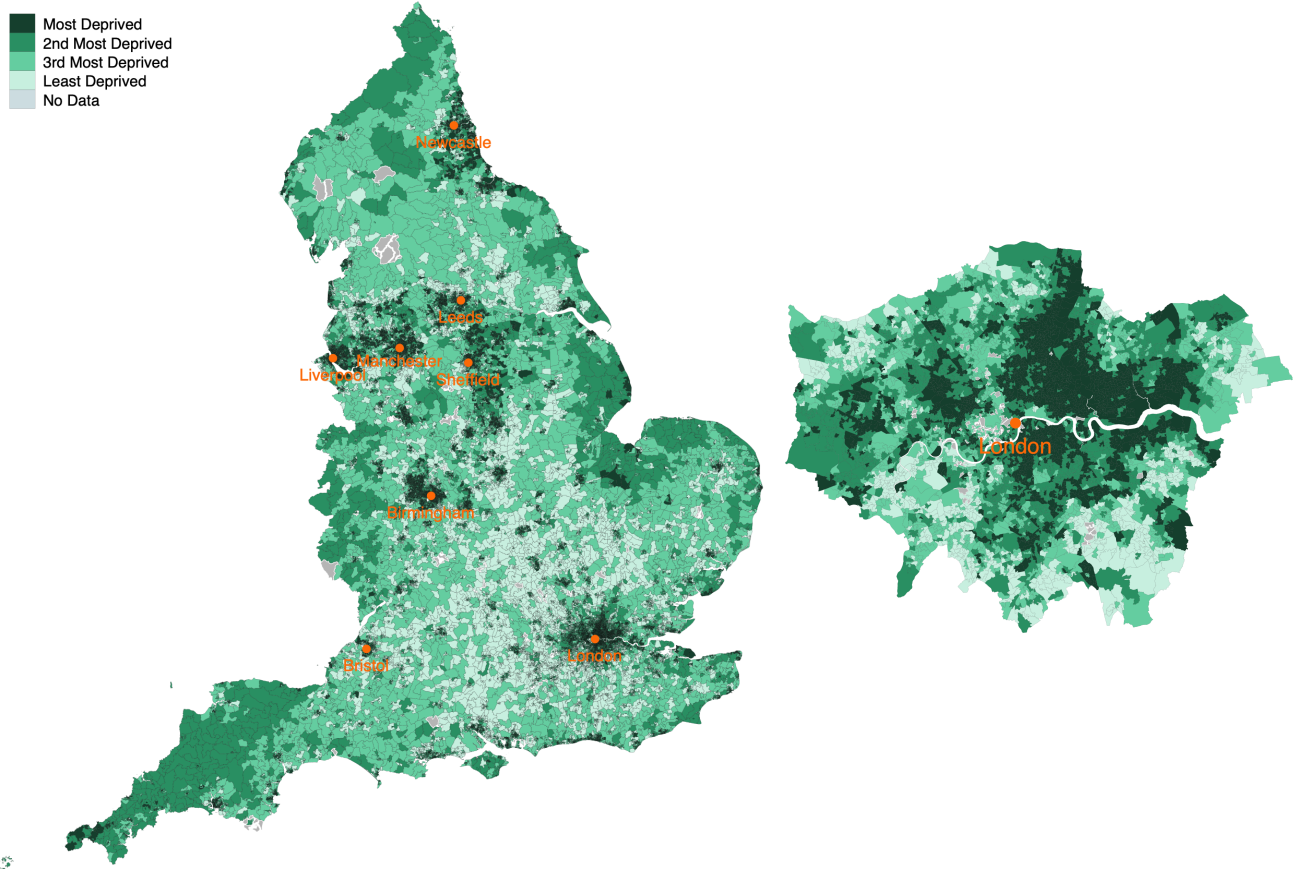


Figure A1. Deprivation Index across England and London

Note: The location of London in the right panel is that of the Prime Minister's Office, 10 Downing Street.

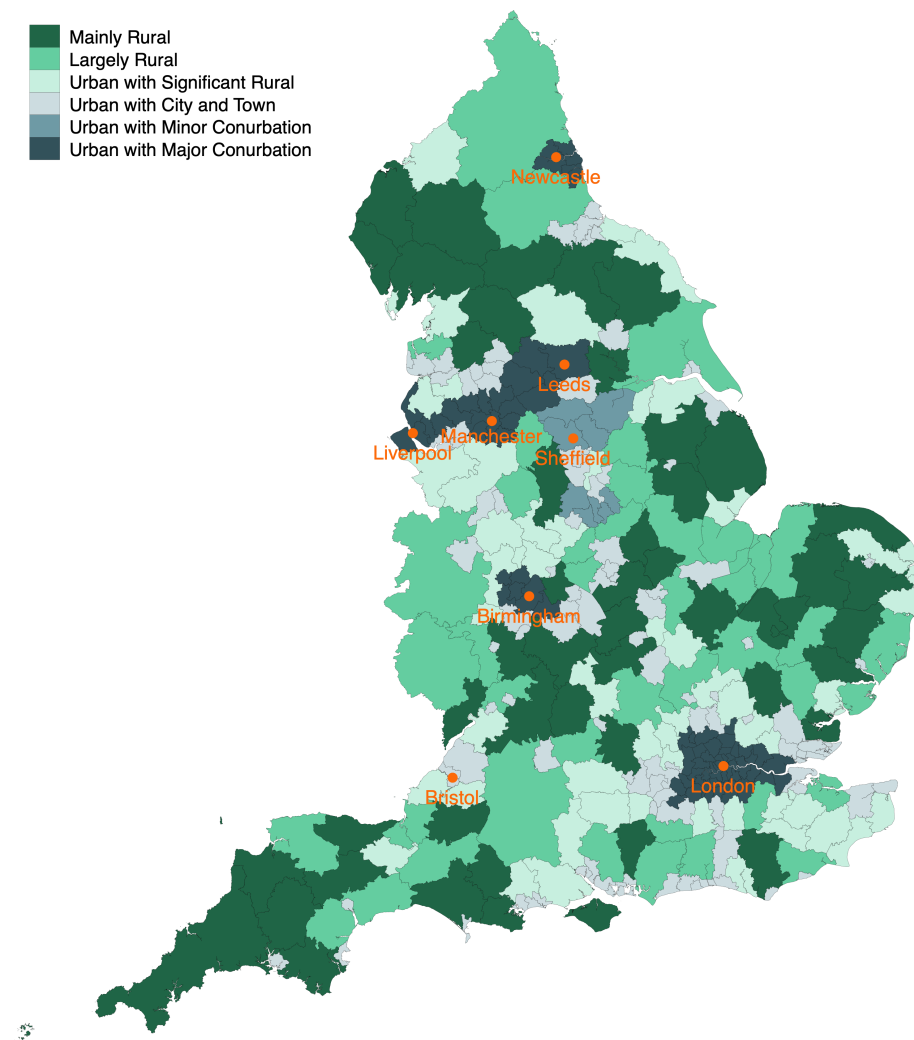


Figure A2. Rural and Urban Classification

B Details in Estimation

The estimation framework is as follows. There are $\mathcal{T}$ periods, indexed by $t = 1, 2, \dots, \mathcal{T}$. Each period corresponds to a school census season, so the interval between periods is approximately four months. Let D_{it} denote the treatment indicator, where $D_{it} = 1$ if individual i has been exposed to an offender's transfer by time t . Because I exclude transfers occurring in the first term at age 11, $D_{i1} = 0$. I assume that treatment is irreversible: once a child is exposed to an offender's transfer, the treatment status cannot be reversed. Thus, $D_{i,t-1} = 1$ implies $D_{i,t} = 1$ a.s. for $t = 2, \dots, \mathcal{T}$.

Let G_i denote the period in which individual i first becomes treated. If an individual is never treated, let $G_i = \infty$.⁹⁹ Define $G_{i,g}$ as an indicator taking the value one if individual i is first treated in period g , such that $G_{i,g} = \mathbb{1}\{G_i = g\}$. Define C_i as an indicator taking the value one if individual i is never treated, i.e., $C_i = \mathbb{1}\{G_i = \infty\} = 1 - D_{i,\mathcal{T}}$. Let $\bar{g} = \max_{i=1, \dots, n} G_i$ denote the latest treatment period observed in the data.

Let $p_{i,g,s}(X_i) = P(G_{i,g} = 1 | X_i, G_{i,g} + (1 - D_{i,s})(1 - G_{i,g}) = 1)$ denote the generalized propensity score, that is, the probability of first being treated in period g , conditional on pre-treatment covariates X_i and on either being treated in period g ($G_{i,g} = 1$) or belonging to the “not-yet-treated” group ($G_{i,g} = 0$ and $D_{i,s} = 0$). Also define $p_{i,g}(X_i) = p_{i,g,\mathcal{T}}(X_i) = P(G_{i,g} = 1 | X_i, G_{i,g} + C_i = 1)$, which is the probability of first being treated in period g , conditional on X_i and on either being treated in period g ($G_{i,g} = 1$ and $C_i = 0$) or belonging to the “never-treated” group ($G_{i,g} = 0$ and $C_i = 1$). Let $\mathcal{G} = \text{supp}(G_i) \setminus \{\bar{g}\} \subseteq \{2, 3, \dots, \mathcal{T}\}$, let $\mathcal{X} = \text{supp}(X_i) \subseteq \mathbb{R}^k$, and, for a generic $\delta \geq 0$, let $\mathcal{G}_\delta = \mathcal{G} \cap \{2 + \delta, 3 + \delta, \dots, \mathcal{T}\}$.

Finally, for $g = 2, \dots, \mathcal{T}$, let $Y_{i,t}(g)$ denote the potential outcome that individual i would experience at time t if first treated in period g . The observed outcome $Y_{i,t}$ can then be expressed as

$$Y_{i,t} = Y_{i,t}(0) + \sum_{g=2}^{\mathcal{T}} (Y_{i,t}(g) - Y_{i,t}(0)) \times G_{i,g}. \quad (3)$$

Based on this expression, the average treatment effect on the treated (ATT) for the pair (g, t) is

$$ATT(g, t) = \mathbb{E}[Y_{i,t}(g) - Y_{i,t}(0) | G_{i,g} = 1], \quad (4)$$

which is the main building block of the framework.

In this paper, Y_{it} primarily takes three forms.

The first is an indicator equal to one if individual i has committed an offense *by* time t . This outcome allows me to examine whether previously nonoffending students begin offending following an offender's school transfer. As emphasized above, this transition is particularly important when

⁹⁹In this paper, I omit the never-treated individuals, but I keep this notation for completeness.

considering potential consequences for human capital, relative to the literature studying peer effects among individuals who are already involved in crime, such as those in incarceration settings.

The second is an indicator equal to one if individual i commits an offense *at* time t . The third is the number of offenses committed by individual i *at* time t . I use these latter two outcomes to examine the effect of exposure on contemporaneous criminal activity.

Following the assumptions in [Callaway and Sant’Anna \(2021\)](#), I impose the following conditions. First, I assume random sampling, such that $\{Y_{i,1}, Y_{i,2}, \dots, Y_{i,T}, X_i, D_{i,1}, D_{i,2}, \dots, D_{i,T}\}_{i=1}^n$ is independently and identically distributed.

Second, under limited treatment anticipation, I assume that there is a known $\delta \geq 0$ such that

$$\mathbb{E}[Y_{i,t}(g) \mid X_i, G_{i,g} = 1] = \mathbb{E}[Y_{i,t}(0) \mid X_i, G_{i,g} = 1] \quad a.s.$$

for all $g \in \mathcal{G}, t \in \{1, \dots, T\}$ s.t. $t < g - \delta$. In the context of this paper, this assumption requires that children do not change their criminal behavior in anticipation of an offender’s transfer before the relevant anticipation window. This assumption is plausible if children cannot predict when an offender will enter their school and grade, as discussed above.

Third, I impose conditional parallel trends using either (i) a “never-treated” group or (ii) a “not-yet-treated” group. Let δ be defined as above. Under (i), for each $g \in \mathcal{G}$ and $t \in \{2, \dots, T\}$ such that $t \geq g - \delta$,

$$\mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0) \mid X_i, G_{i,g} = 1] = \mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0) \mid X_i, C_i = 1] \quad a.s.,$$

and under (ii), for each $g \in \mathcal{G}$ and $(s, t) \in \{2, \dots, T\} \times \{2, \dots, T\}$ such that $t \geq g - \delta$ and $t + \delta \leq s < \bar{g}$,

$$\mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0) \mid X_i, G_{i,g} = 1] = \mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0) \mid X_i, D_{i,s} = 0, G_{i,g} = 0] \quad a.s.$$

These assumptions require that, conditional on X_i , untreated potential outcomes would evolve similarly across the treatment and relevant control groups in the absence of treatment. I consider both control-group definitions and show that the results are robust to the choice of control group.

Finally, I impose an overlap assumption. For each $t \in \{2, \dots, T\}, g \in \mathcal{G}$, there exists some $\varepsilon > 0$ such that $P(G_{i,g} = 1) > 0$ and $p_{i,g,t}(X_i) < 1 - \varepsilon$ *a.s.* This assumption requires a positive probability of first receiving treatment in each relevant treatment period while ruling out treatment with probability one conditional on the observed covariates. In the empirical setting, a school cannot reject a transfer solely because the student is an offender, nor can it ensure that it will receive an offender with certainty, making this assumption plausible.

Under these assumptions, the doubly robust estimand can be written as

$$ATT(g, t; \delta) = \mathbb{E} \left[\left(\frac{G_{i,g}}{\mathbb{E}[G_{i,g}]} - \frac{\frac{p_{i,g}(X_i)C_i}{1-p_{i,g}(X_i)}}{\mathbb{E} \left[\frac{p_{i,g}(X_i)C_i}{1-p_{i,g}(X_i)} \right]} \right) (Y_{i,t} - Y_{i,g-\delta-1} - m_{i,g,t,\delta}^{nev}(X_i)) \right], \quad (5)$$

under the specification with the “never-treated” group, where $m_{i,g,t,\delta}^{nev}(X_i) = \mathbb{E}[Y_{i,t} - Y_{i,g-\delta-1} | X_i, C_i = 1]$, and

$$ATT(g, t; \delta) = \mathbb{E} \left[\left(\frac{G_{i,g}}{\mathbb{E}[G_{i,g}]} - \frac{\frac{p_{i,g,t+\delta}(X_i)(1-D_{i,t+\delta})(1-G_{i,g})}{1-p_{i,g,t+\delta}(X_i)}}{\mathbb{E} \left[\frac{p_{i,g,t+\delta}(X_i)(1-D_{i,t+\delta})(1-G_{i,g})}{1-p_{i,g,t+\delta}(X_i)} \right]} \right) (Y_{i,t} - Y_{i,g-\delta-1} - m_{i,g,t,\delta}^{ny}(X_i)) \right], \quad (6)$$

under the specification with the “not-yet-treated” group, where $m_{i,g,t,\delta}^{ny}(X_i) = \mathbb{E}[Y_{i,t} - Y_{i,g-\delta-1} | X_i, D_{i,t+\delta} = 0, G_{i,g} = 0]$.

I then aggregate the group-time treatment effects by event time. The treatment effect e periods after exposure is

$$\theta_{es}(e) = \sum_{g \in \mathcal{G}} \mathbb{1}\{g + e \leq \mathcal{T}\} P(G = g | G + e \leq \mathcal{T}) ATT(g, g + e), \quad (7)$$

where the primary parameter of interest in this paper is $\theta_{es}(0)$, which captures the treatment effect immediately following an offender’s transfer.

C Robustness Checks

Table C1. Age of Not-Yet-Offenders' School Transfers

Age in Sept.	Not-yet-offenders' Transfers		
	Autumn (1)	Spring (2)	Summer (3)
A. Transfer			
11	-	1,430	1,997
12	4,687	1,861	1,948
13	5,301	1,696	1,254
14	2,717	1,065	617
15	751	383	60
B. First Contact to Others' Transfers (Treatment)			
11	-	173,126	180,795
12	210,195	121,780	105,445
13	119,196	67,767	62,753
14	58,013	34,375	22,418
15	16,761	3,380	517

Notes: Offenders are those who have committed a crime by the time of their school transfer, while non-offenders are those who have never committed a crime at all or those not yet by the time of their school transfer.

Table C2. Effects of Disadvantaged and Disruptive Students' Transfers

Characteristics of Transfer Students:	Any	FSM	SEN	FSM and SEN	Suspension	FSM, SEN, and Suspension
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Offenders' Transfers						
Treatment Effects at $t = 0$	.125 (.115)	.075*** (.008)	.043*** (.009)	.114*** (.010)	.106*** (.008)	.122*** (.012)
Pre-Treatment Mean	1.56	.632	.640	.744	1.22	1.39
Relative Effect Size	8.02%	11.9%	6.79%	15.3%	8.72%	8.74%
Sample Size (Panel Count)	9,455,532	14,409,248	14,286,249	13,819,441	12,551,511	10,406,529
Number of Individuals	642,916	979,095	970,886	939,161	852,880	707,334
Panel B: Offenders' Transfers						
Treatment Effects at $t = 0$	.193*** (.009)	.208*** (.012)	.205*** (.012)	.218*** (.014)	.192*** (.013)	.224*** (.017)
Pre-Treatment Mean	1.50	1.52	1.52	1.53	1.50	1.52
Relative Effect Size	12.9%	13.7%	13.4%	14.2%	12.8%	14.8%
Sample Size (Panel Count)	15,458,144	12,327,449	11,676,655	9,720,359	10,750,223	7,022,390
Number of Individuals	1,050,567	837,843	793,664	660,754	730,468	477,164

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to the school transfer of an offender or a disadvantaged or disruptive student. The treatment group consists of students who are exposed to an individual with the characteristics indicated in each column transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. In Panel A, I exclude students who are exposed to an offender's transfer at the same time as the transfer of the specific type of individual considered in each column. Column (1) in Panel A adjusts the estimation using sampling weights based on treatment timing. In Columns (2) to (6) in Panel B, I exclude students who were exposed to the transfer of an offender without the specified characteristic before the transfer of the specific type of individual considered in each column. Data sources: Ministry of Justice and Department for Education.

Table C3. Effects of Students' Transferring Out on Stayers

Students Transferring Out:	Offender (1)	Not-yet-offender (2)	Non-offender (3)
Panel A: Effects on Crime Onset			
Treatment Effects at $t = 0$	.137*** (.010)	.118*** (.012)	.105*** (.027)
Pre-Treatment Mean	1.21	.763	.404
Relative Effect Size	11.3%	15.5%	25.9%
Panel B: Effects on Number of Crimes			
Treatment Effects at $t = 0$	.014 (.021)	.095*** (.027)	.014 (.040)
Pre-Treatment Mean	.328	.222	.155
Relative Effect Size	4.32%	42.9%	9.04%
Sample Size (Panel Count)	12,072,353	15,788,824	14,858,403
Number of Individuals	819,109	1,070,911	1,006,868

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table C4. Transfer with a Residence Move

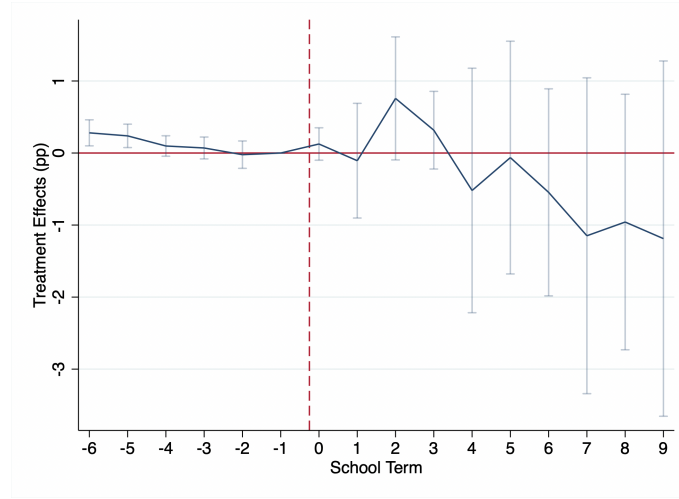
	No Residence Move	Residence Move				
	(1)	Any (2)	0-25 miles (3)	25-50 miles (4)	50-100 miles (5)	More than 100 miles (6)
Treatment Effects at $t = 0$	.182*** (.012)	.261*** (.021)	.277*** (.028)	.346*** (.089)	.086 (.055)	.276*** (.064)
Pre-Treatment Mean	1.50	1.53	1.53	1.53	1.63	1.58
Relative Effect Size	12.1%	17.0%	18.1%	22.6%	5.26%	17.4%
Sample Size (Panel Count)	10,837,428	4,939,848	3,130,345	731,958	568,499	597,536
Number of Individuals	736,373	335,947	212,850	49,722	38,635	40,797

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Residential moves are identified based on changes in LSOA recorded in the NPD. Distance is measured between the population-weighted centroids of LSOAs, as provided by the Office for National Statistics. LSOA stands for Lower Layer Super Output Area, a geographic unit containing approximately 1,000 to 3,000 residents. Data sources: Ministry of Justice and Department for Education.

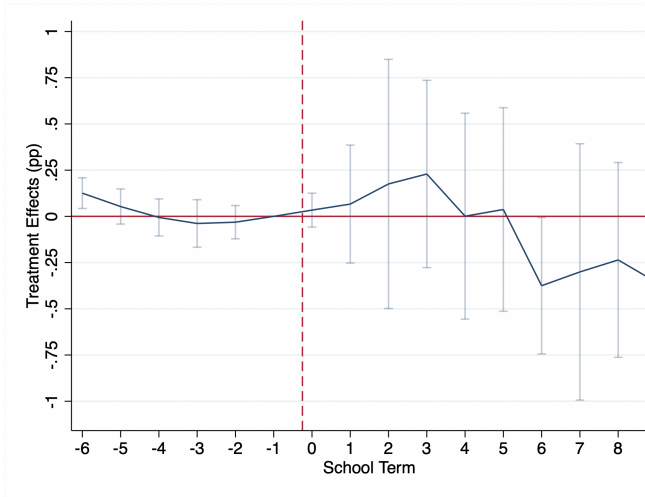
Table C5. Different Specifications

Effects on:	Crime Onset (1)	Any Crime (2)	Number of Crimes (3)
Poisson Regression	.130*** (.010)	.021*** (.007)	.030** (.013)
Logit Regression	14.55*** (1.38)	13.28*** (2.41)	-
Sample Size	15,262,170	15,262,170	15,262,170

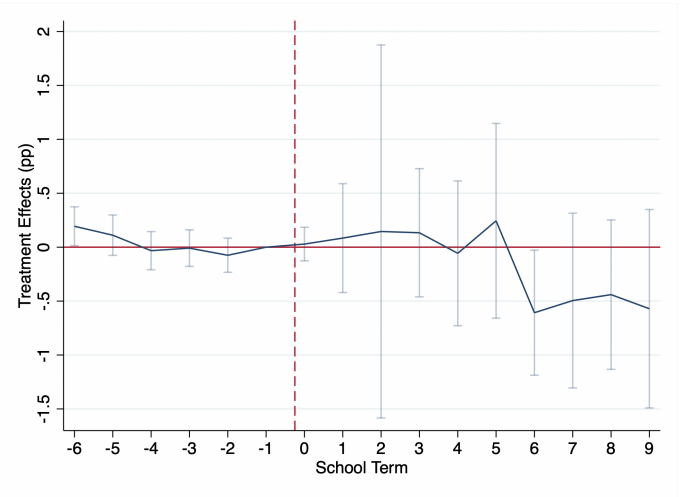
Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Wooldridge \(2023, 2025\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.



(a) Effects on Offending Initiation



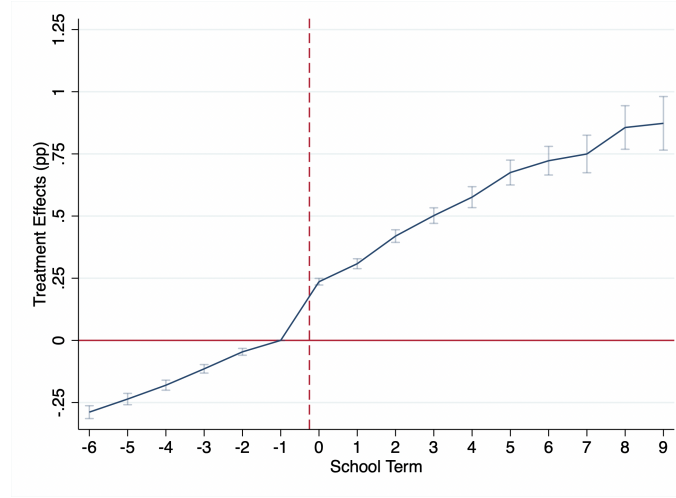
(b) Effects on Any Offending



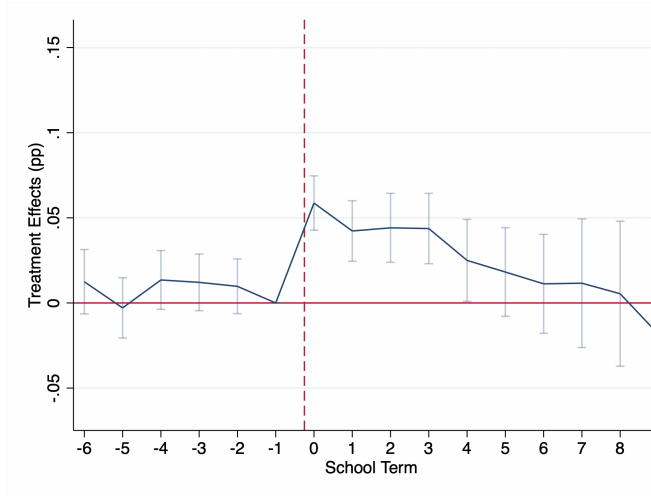
(c) Effects on Number of Crimes

Figure C1. Placebo Test: Receiving a Not-Yet Offender's Transfer

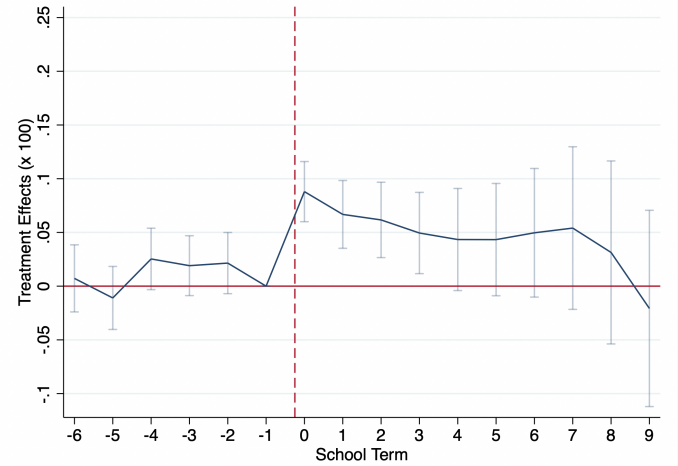
Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#). I use the weight to adjust for differences in exposure timing across grades and school census seasons between offenders' and not-yet-offenders' transfers.



(a) Effects on Offending Initiation



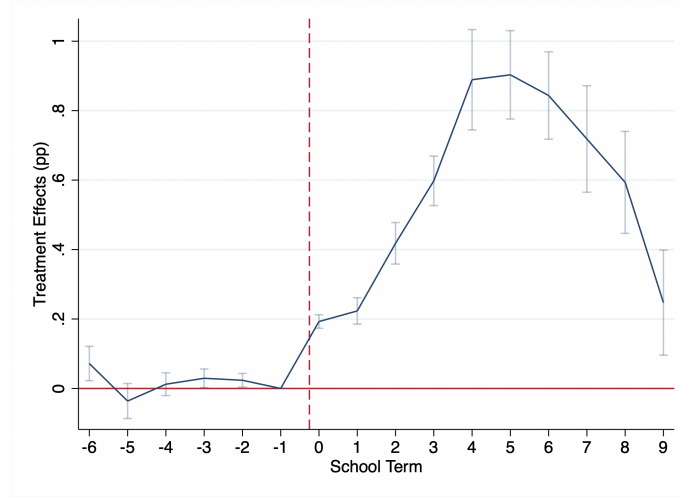
(b) Effects on Any Offending



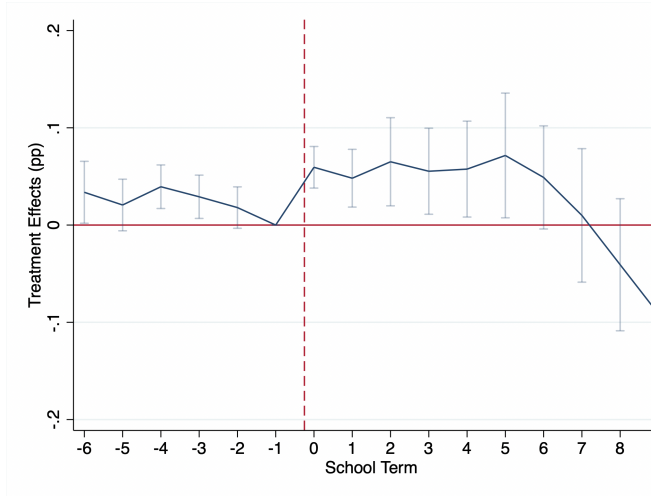
(c) Effects on Number of Crimes

Figure C2. Treatment Effects: Censored at the First Transfer

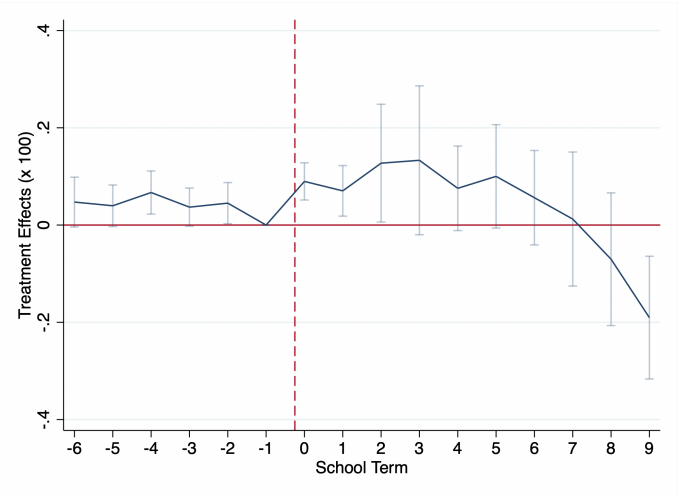
Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#); [Rambachan and Roth \(2023\)](#).



(a) Effects on Offending Initiation



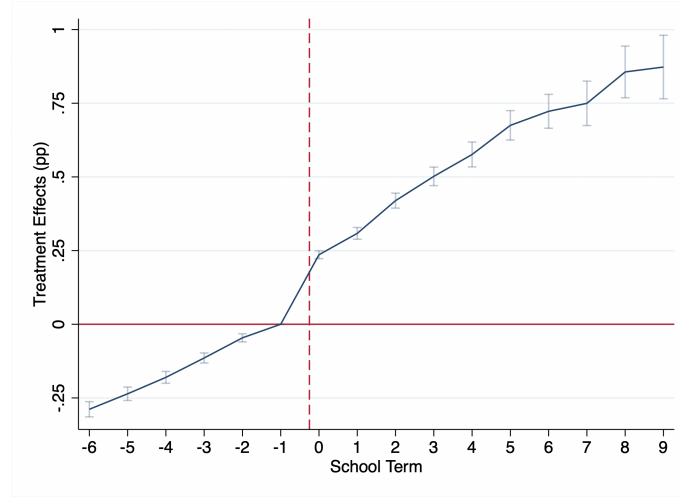
(b) Effects on Any Offending



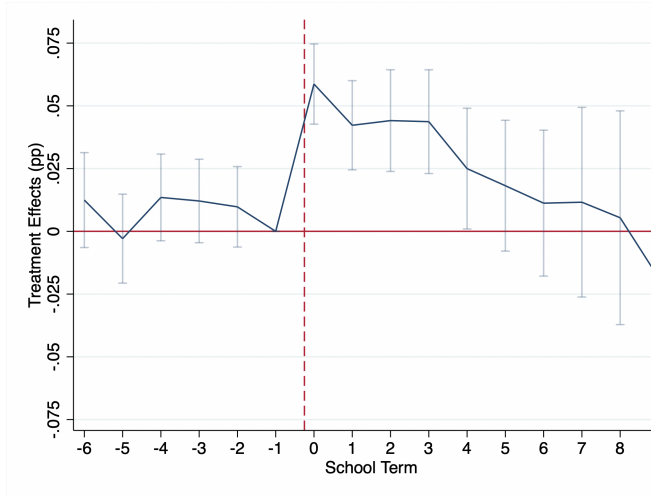
(c) Effects on Number of Crimes

Figure C3. Censored at the Second Treatment

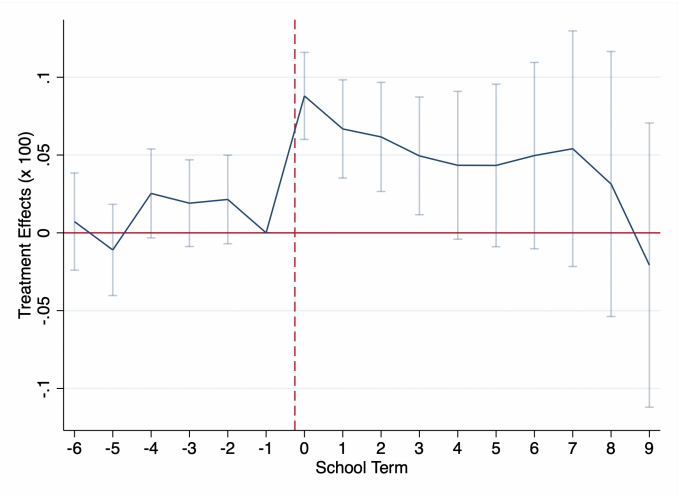
Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#); [Rambachan and Roth \(2023\)](#).



(a) Effects on Offending Initiation



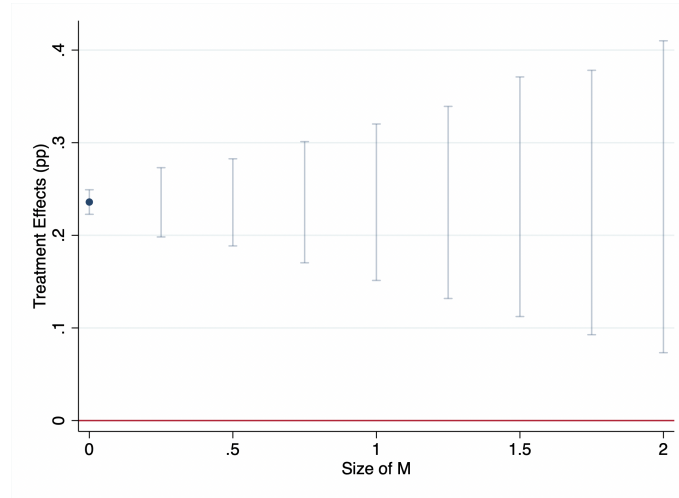
(b) Effects on Any Offending



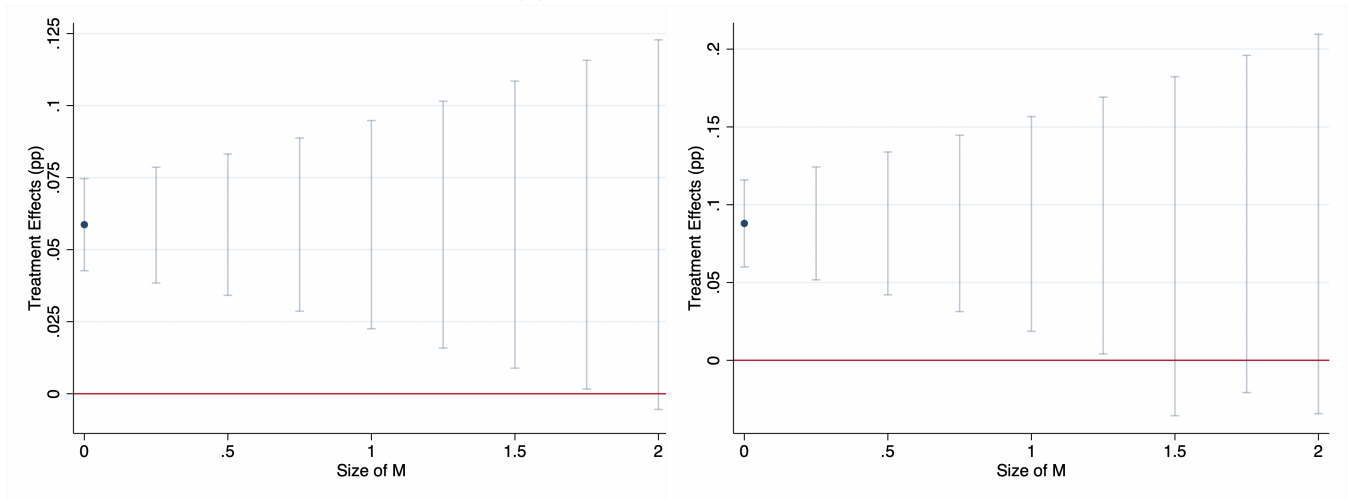
(c) Effects on Number of Crimes

Figure C4. Main Results without Control Variables

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation

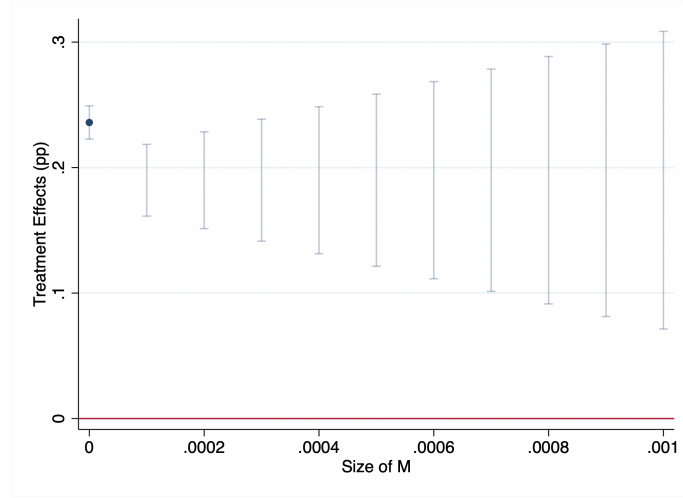


(b) Any Crime at Each Period

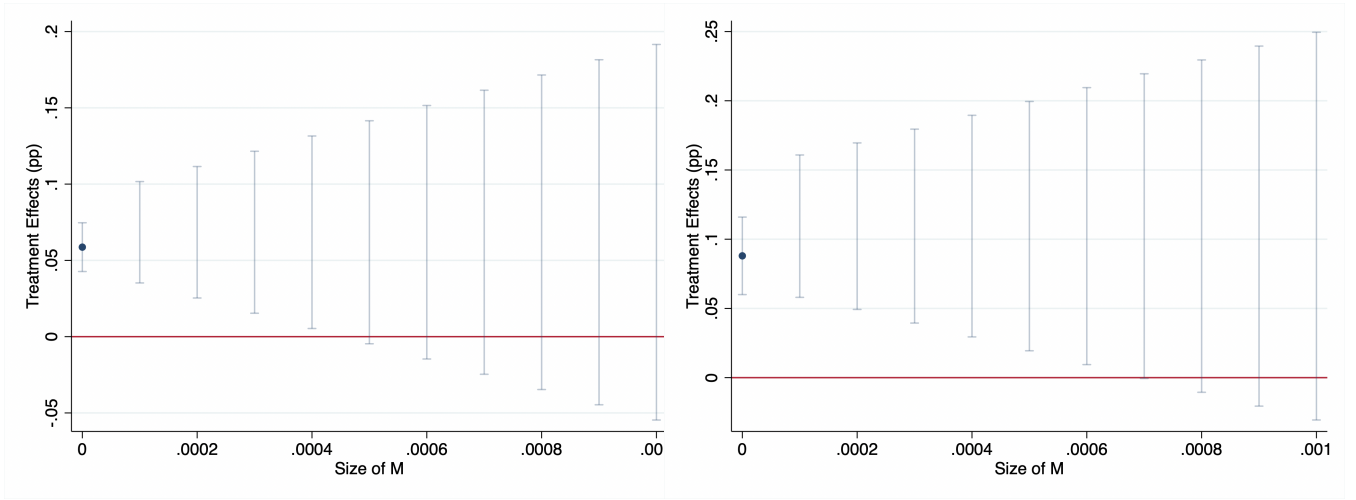
(c) N of Crimes at Each Period

Figure C5. Honest Difference-in-Differences without Control Variables, Relative Magnitude Restrictions

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#); [Rambachan and Roth \(2023\)](#).



(a) Offending Initiation

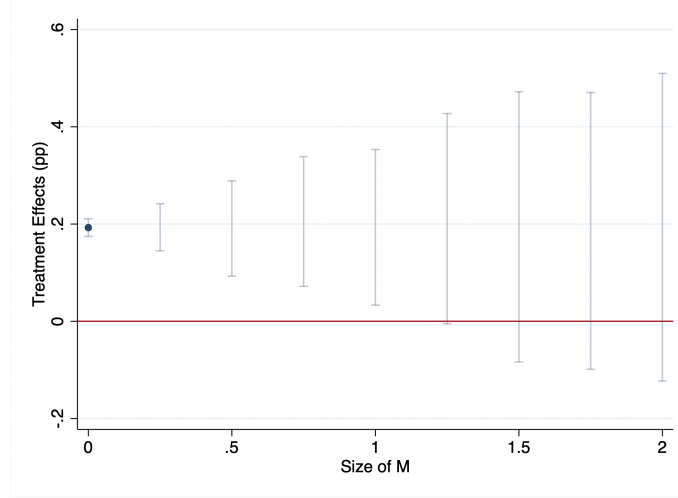


(b) Any Crime at Each Period

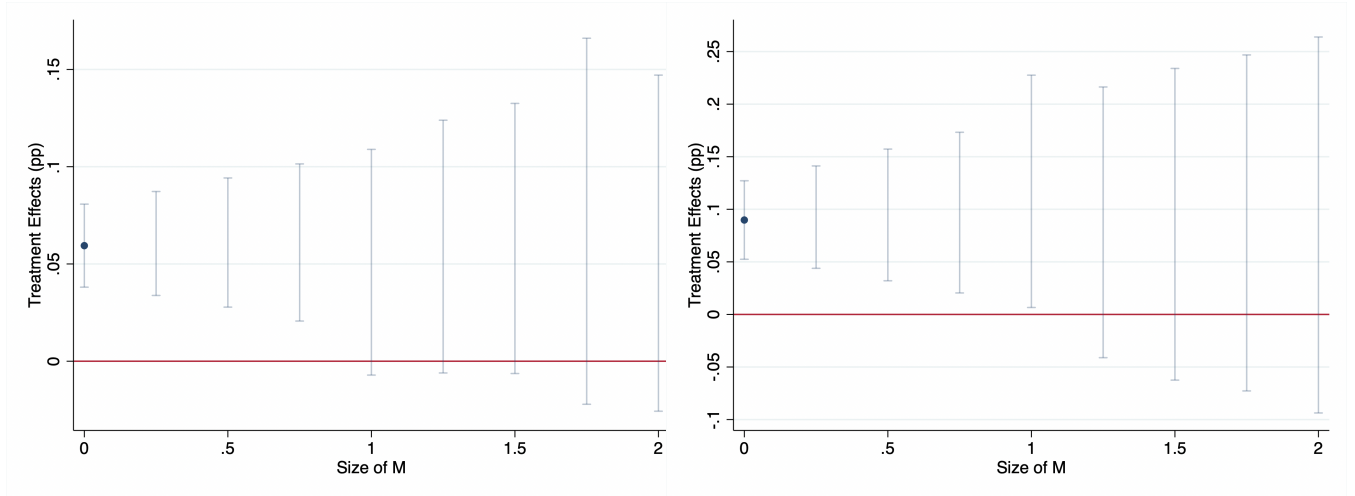
(c) N of Crimes at Each Period

Figure C6. Honest Difference-in-Differences without Control Variables, Second Difference Restrictions

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#); [Rambachan and Roth \(2023\)](#).



(a) Offending Initiation

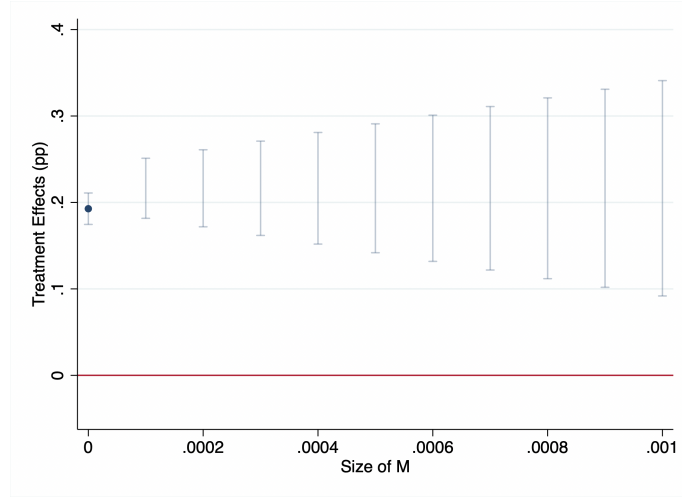


(b) Any Crime at Each Period

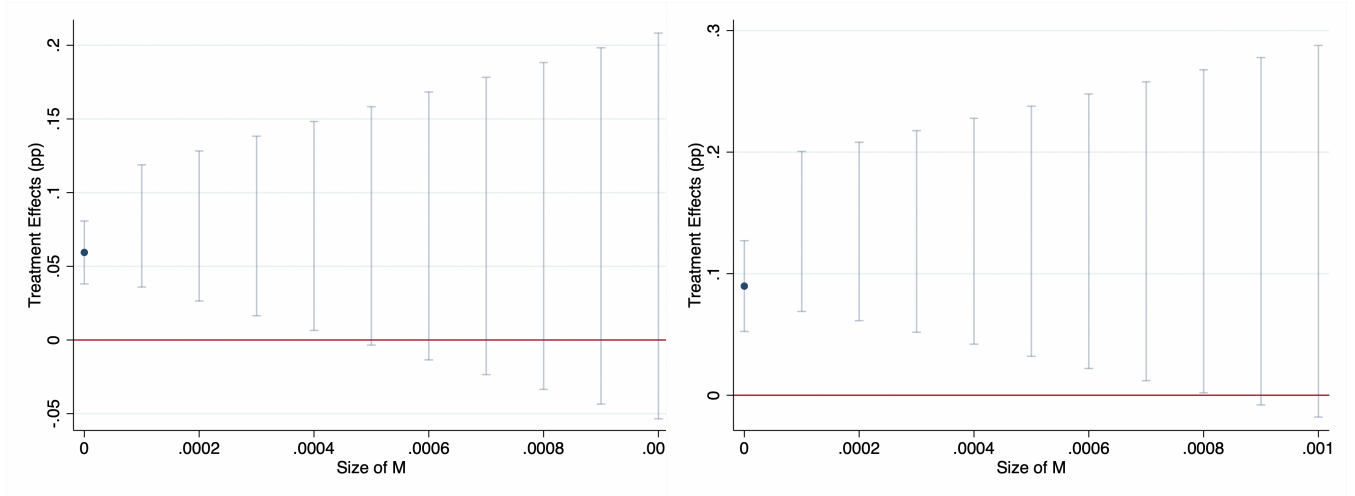
(c) N of Crimes at Each Period

Figure C7. Honest Difference-in-Differences, Relative Magnitude Restrictions

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation

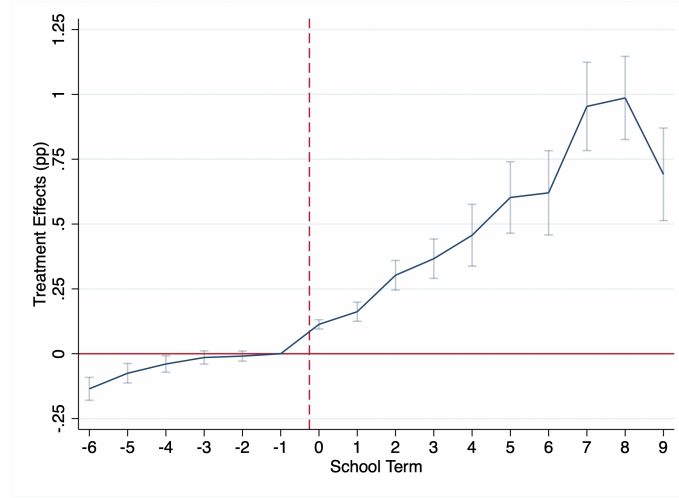


(b) Any Crime at Each Period

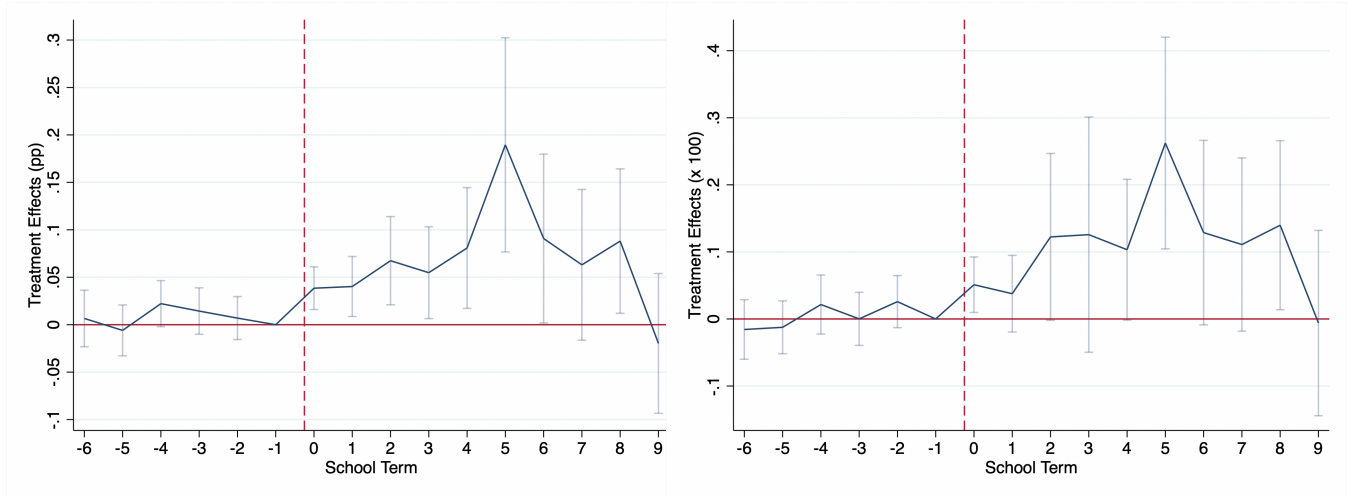
(c) N of Crimes at Each Period

Figure C8. Honest Difference-in-Differences, Second Difference Restrictions

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation

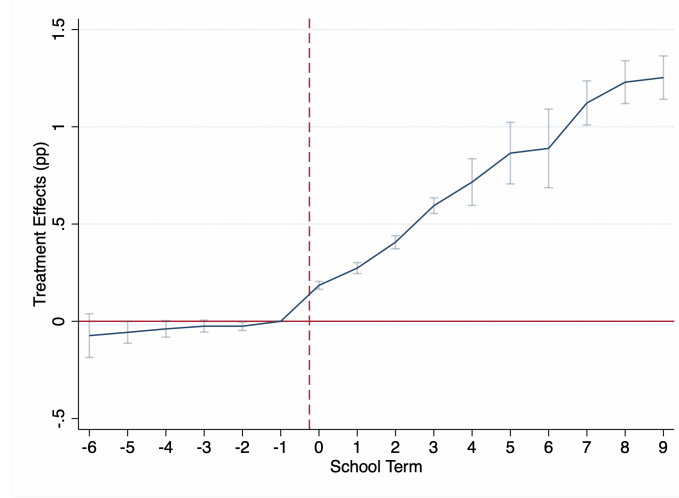


(b) Any Crime at Each Period

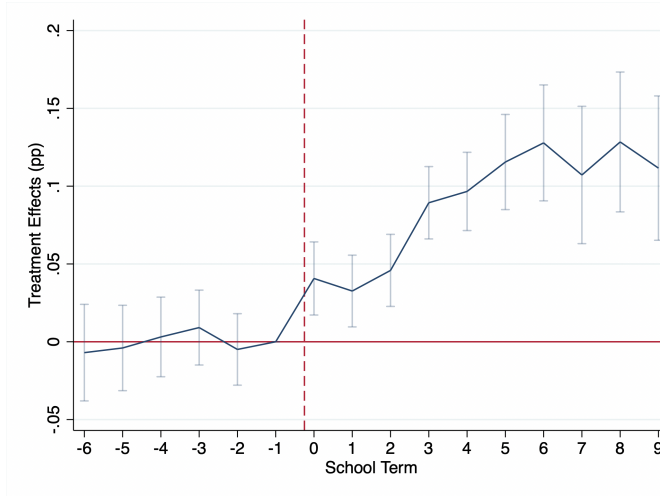
(c) N of Crimes at Each Period

Figure C9. Treatment Effects, Partial out with Demographics

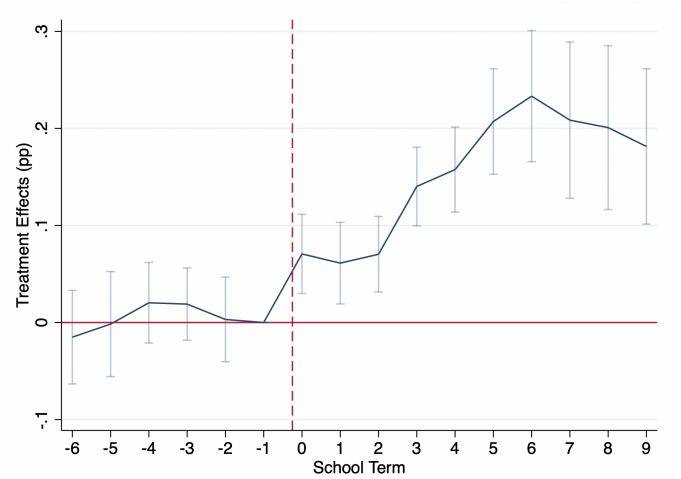
Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation



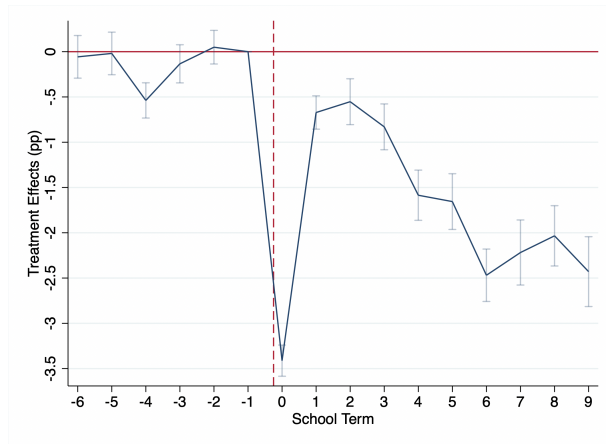
(b) Any Crime at Each Period



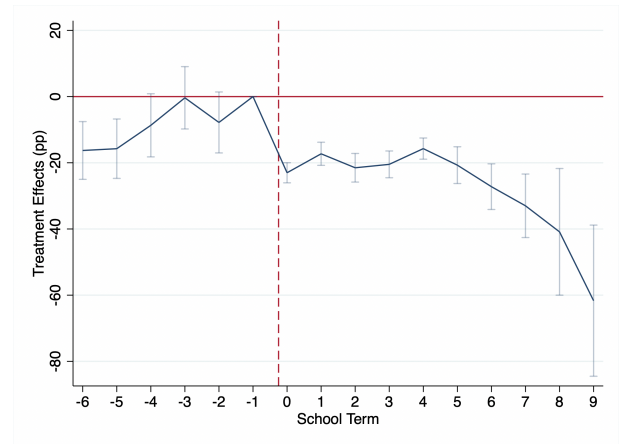
(c) N of Crimes at Each Period

Figure C10. Treatment Effects Including Never Treated Group

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



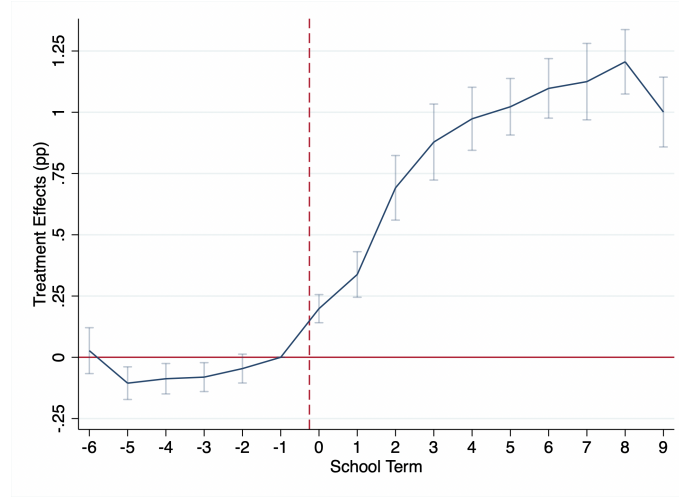
(a) Effects of Receiving an Offender Transfer



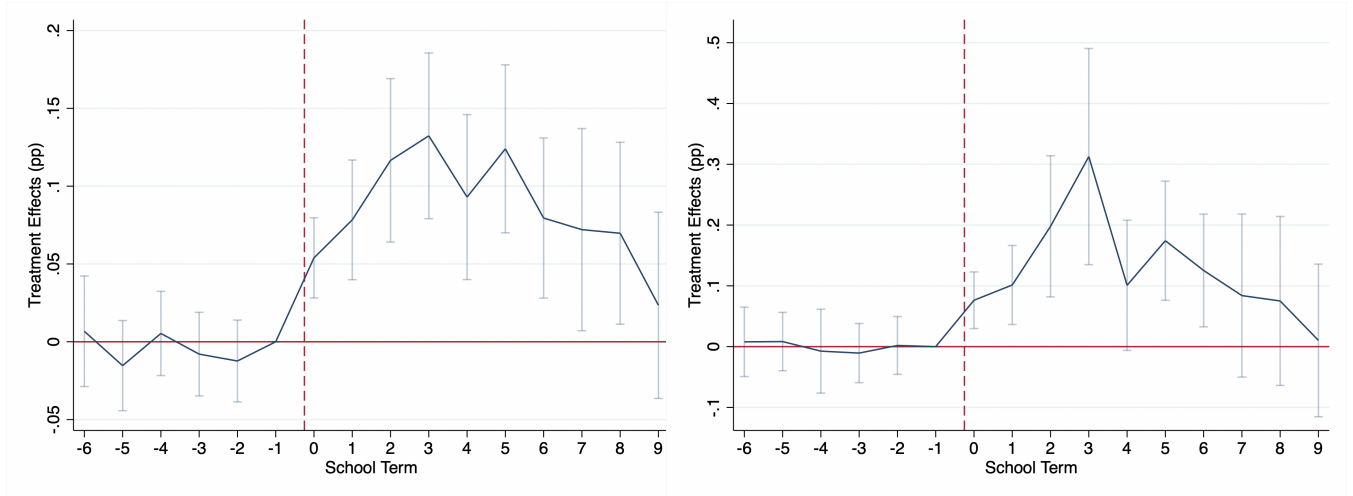
(b) Effects of Receiving a Non-offender Transfer

Figure C11. Effects on Transfer

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation

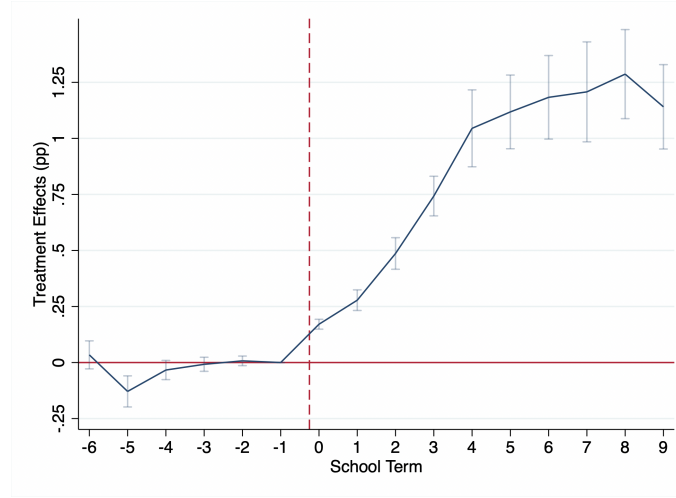


(b) Any Crime at Each Period

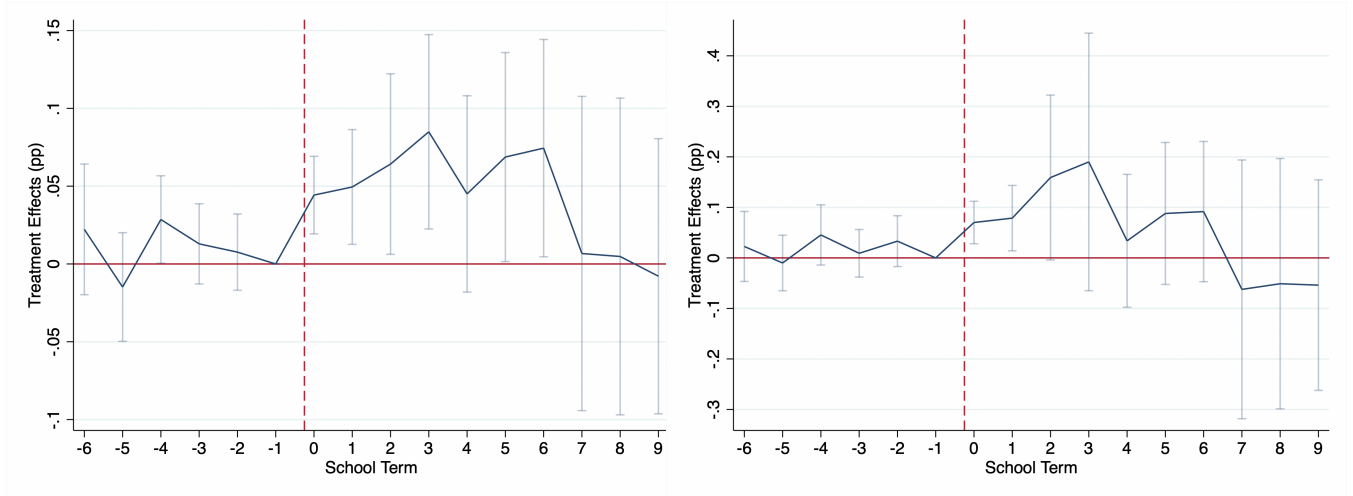
(c) N of Crimes at Each Period

Figure C12. Treatment Effects Focusing on Schools from which No Offender Left

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation

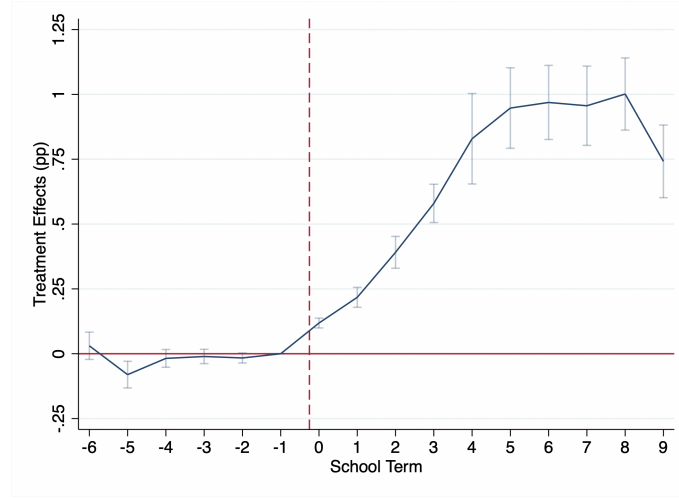


(b) Any Crime at Each Period

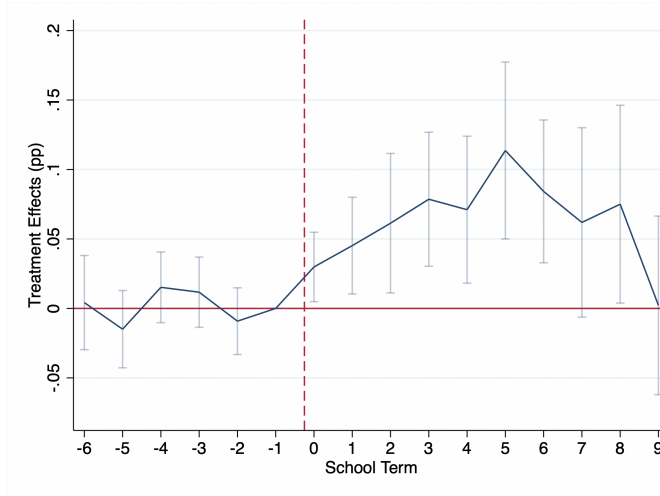
(c) N of Crimes at Each Period

Figure C13. Treatment Effects Excluding Those Who Move Back to the Previous School

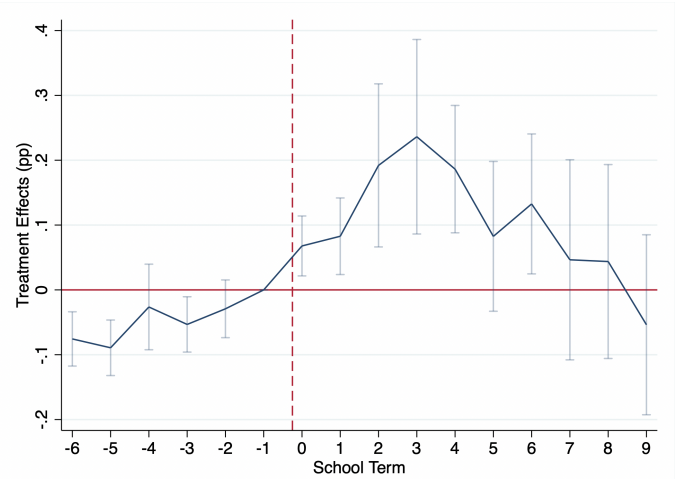
Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



(a) Offending Initiation



(b) Any Crime at Each Period



(c) N of Crimes at Each Period

Figure C14. Treatment Effects Excluding Those Who Have Been Exposed to a Mass Transfer

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).

D Heterogeneity in Detail

Table D1. Crimes during School Terms

	During School Term* ¹ (1)	During Holidays* ¹ (2)	Weekdays (3)	Weekends (4)
<i>Panel A: Effects on becoming an offender by each term</i>				
Treatment Effects at $t = 0$	.136*** (.008)	.090*** (.007)	.150*** (.009)	.073*** (.006)
Pre-Treatment Mean	.997	.665	1.14	.501
Relative Effect Size	13.6%	13.5%	13.1%	14.5%
<i>Panel B: Effects on the number of crimes committed in each term</i>				
Treatment Effects at $t = 0$	.050*** (.010)	.016** (.008)	.046*** (.011)	.020*** (.006)
Pre-Treatment Mean	.191	.121	.226	.086
Relative Effect Size	26.3%	13.1%	20.2%	23.6%
Sample Size (Panel Count)	15,458,144	15,458,144	15,458,144	15,458,144
Number of Individuals	1,050,567	1,050,567	1,050,567	1,050,567

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. School term dates vary slightly across local authorities in England. This table uses the school term dates for Brighton & Hove. Data sources: Ministry of Justice and Department for Education.

Table D2. School Types

School Types:	Academy Converter (1)	Academy Sponsor Led (2)	Community School (3)	Foundation School (4)	Voluntary Aided School (5)	Voluntary Controlled School (6)
<i>Panel A: Effects on Crime Onset</i>						
Treatment Effects at $t = 0$	-.027 (.035)	.104 (.066)	.139*** (.014)	-.023 (.301)	.649* (.386)	.467** (.229)
Pre-Treatment Mean	1.12	2.14	1.50	1.61	1.15	1.37
Relative Effect Size	-2.39%	4.86%	9.21%	-1.40%	56.5%	34.1%
<i>Panel B: Effects on Number of Crimes</i>						
Treatment Effects at $t = 0$	-.025 (.088)	-.122 (.095)	.049* (.029)	-.136 (.303)	.782*** (.230)	.247* (.147)
Pre-Treatment Mean	.334	.521	.353	.373	.341	.314
Relative Effect Size	-7.40%	-23.4%	13.9%	-36.3%	229%	78.4%
Sample Size (Panel Count)	2,000,104	989,257	5,710,743	3,747,099	1,272,572	285,724
Number of Individuals	134,819	68,313	386,644	252,803	85,623	19,495

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. School type is measured at the time of treatment. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table D3. Heterogeneity in Demographic Characteristics

<i>Heterogeneity in Gender</i>			
Population	All (1)	Boys (2)	Girls (3)
<i>Panel A: Effects of Any Offenders' Transfers</i>			
Treatment Effects at $t = 0$	.193*** (.009)	.249*** (.016)	.126*** (.011)
Pre-Treatment Mean	1.50	1.98	.979
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	12.9%	12.6%	12.9%
Sample Size (Panel Count)	15,458,144	8,013,566	7,444,578
Number of Individuals	1,050,567	544,283	506,284
<i>Panel B: Effects of Transfers of Offenders of the Same Gender</i>			
Treatment Effects at $t = 0$	.217*** (.014)	.261*** (.019)	.136*** (.022)
Pre-Treatment Mean	1.63	1.93	1.04
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	13.3%	13.5%	13.1%
Sample Size (Panel Count)	8,448,209	5,534,952	2,913,257
Number of Individuals	573,876	375,792	198,084
<i>Panel C: Effects of Transfers of Offenders of the Opposite Gender</i>			
Treatment Effects at $t = 0$	.138** (.058)	-.136 (.108)	.119*** (.014)
Pre-Treatment Mean	1.35	2.10	.941
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	10.2%	-6.50%	12.7%
Sample Size (Panel Count)	7,009,935	2,478,614	4,531,321
Number of Individuals	476,691	168,491	308,200
<i>Heterogeneity in First Language</i>			
Population	Native English Speakers (1)	Non-native English Speakers (2)	
<i>Panel A: Effects of Transfers of Offenders with the Same First Language Status</i>			
Treatment Effects at $t = 0$	.503*** (.013)	.478*** (.108)	
Pre-Treatment Mean	1.34	.540	
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	37.5%	88.6%	
Sample Size (Panel Count)	14,601,647	416,685	
Number of Individuals	834,068	21,763	
<i>Panel B: Effects of Transfers of Offenders with a Different First Language Status</i>			
Treatment Effects at $t = 0$	-.304 (.226)	.007 (.028)	
Pre-Treatment Mean	2.01	.883	
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	-15.1%	.804%	
Sample Size (Panel Count)	948,034	1,315,063	
Number of Individuals	44,262	76,774	
<i>Heterogeneity in Ethnicity</i>			
	White British (1)	Non White British (2)	
<i>Panel A: Effects of Transfers of Offenders of the Same Ethnicity</i>			
Treatment Effects at $t = 0$	-.290** (.129)	.048 (.137)	
Pre-Treatment Mean	1.54	1.11	
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	-18.8%	4.35%	
Sample Size (Panel Count)	1,856,895	1,346,979	
Number of Individuals	126,035	91,035	
<i>Panel B: Effects of Transfers of Offenders of a different Ethnicity</i>			
Treatment Effects at $t = 0$	.255*** (.037)	.176*** (.013)	
Pre-Treatment Mean	1.27	1.59	
Ratio of the Spontaneous Effects to the Pre-Treatment Mean	20.1%	11.1%	
Sample Size (Panel Count)	1,574,950	10,895,631	
Number of Individuals	107,049	741,223	

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. I exclude observations that were exposed to another type of offender prior to treatment. The estimation follows Callaway and Sant'Anna (2021), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table D4. Offender's Experience

Quartile in offender density:	One Crime (1)	Two to Five Crimes (2)	More than Five Crimes (3)
Panel A: Effects on Crime Onset			
Treatment Effects at $t = 0$	.198*** (.012)	.275*** (.024)	.539*** (.165)
Pre-Treatment Mean	1.46	1.65	2.10
Relative Effect Size	13.5%	16.7%	25.7%
Panel B: Effects on Number of Crimes			
Treatment Effects at $t = 0$	.085*** (.023)	.154*** (.041)	.474* (.259)
Pre-Treatment Mean	.365	.385	.552
Relative Effect Size	23.4%	39.9%	86.0%
Sample Size (Panel Count)	11,380,413	4,191,421	341,441
Number of Individuals	773,254	285,159	23,257

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table D5. Treatment Age Heterogeneity

Treatment Age:	Age 11 (1)	Age 12 (2)	Age 13 (3)	Age 14 (4)
Treatment Effects at $t = 0$	.084*** (.018)	.061*** (.022)	.065*** (.025)	.172*** (.064)
Pre-Treatment Mean	.834	1.31	1.69	1.99
Relative Effect Size	10.1%	4.67%	3.83%	8.63%
Sample Size (Panel Count)	3,127,728	3,081,488	3,051,872	3,085,214
Number of Individuals	91,301	276,133	341,758	262,657

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade at the specified age, while the control group consists of students who will be treated later at the same age. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. Data sources: Ministry of Justice and Department for Education.

Table D6. Heterogeneity in Early Years Educational Achievement

Educational Achievement: Treatment Effects at $t = 0$	High (1)	Low (2)
Overall Test Score	.115** (.049)	.129* (.072)
Pre-Treatment Mean	.675	1.40
Relative Effect Size	17.1%	9.23%
Personal, social and emotional development	.054 (.068)	.191*** (.069)
Pre-Treatment Mean	.703	1.38
Relative Effect Size	7.66%	13.9%
Communication, language and literacy	.117** (.048)	.138* (.074)
Pre-Treatment Mean	.692	1.39
Relative Effect Size	16.9%	9.93%
Mathematical development	.088 (.055)	.155** (.072)
Pre-Treatment Mean	.764	1.32
Relative Effect Size	11.5%	11.7%
Knowledge and understanding of the world	.168*** (.058)	.079 (.063)
Pre-Treatment Mean	.889	1.18
Relative Effect Size	18.9%	6.72%
Physical development	.112* (.066)	.151** (.061)
Pre-Treatment Mean	.914	1.16
Relative Effect Size	12.2%	13.0%
Creative development	.102* (.059)	.158** (.069)
Pre-Treatment Mean	.863	1.21
Relative Effect Size	11.8%	13.1%
Sample Size (Panel Count)	273,240	272,122
Number of Individuals	18,525	18,523

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. All test scores are standardized within each cohort. Early-years test scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month ([Department for Education, Schools Analysis and Research Division, 2010](#)). Data sources: Ministry of Justice and Department for Education.

Table D7. Heterogeneity in Individual Characteristics

	With FSM and SEN (1)	With FSM Only (2)	With SEN Only (3)	Neither FSM nor SEN (4)
Treatment Effects at $t = 0$	.570*** (.041)	.233*** (.029)	.116*** (.019)	.072*** (.008)
Pre-Treatment Mean	4.17	2.04	1.37	.578
Relative Effect Size	13.7%	11.4%	8.47%	12.4%
Sample Size (Panel Count)	2,357,011	2,199,327	3,277,133	7,624,673
Number of Individuals	161,381	150,051	222,544	516,591

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. The variable for SEN is a dummy variable that takes the value one if there is any explicit record on SEN provision before the secondary school, and zero if otherwise. The variable for FSM is also a dummy variable that takes the value one if there is any record as eligible for FSM before the secondary school and zero if otherwise. Data sources: Ministry of Justice and Department for Education.

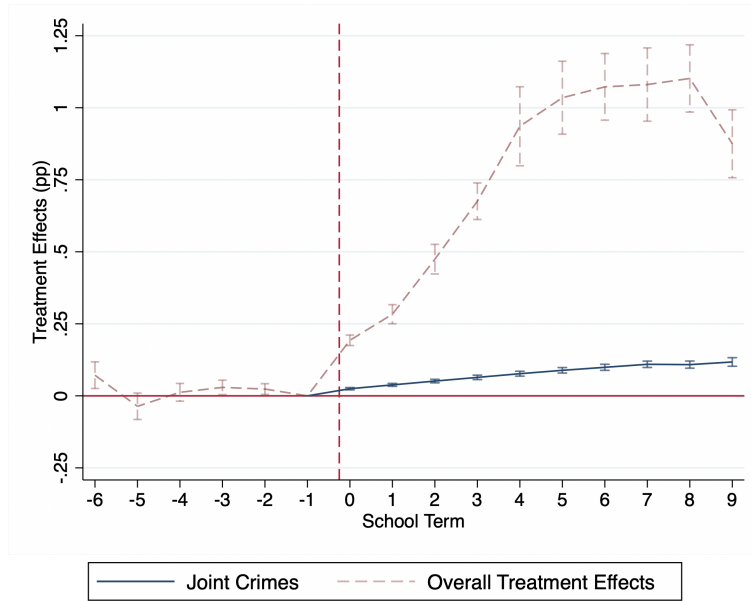


Figure D1. Any Offending with Transfer Offenders

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).

E Simulation Details

In this section, I describe the matching problem mathematically. The setting is a special case of the generalized theory discussed in [Bogomolnaia and Moulin \(2001\)](#).

Suppose that there is a finite set of transfer students $s \in S$ and a finite set of school-grades $g \in G$.¹⁰⁰ Denote $N = |S|$. I assume a null school-grade $\emptyset \in G$. In this exercise, I consider swapping students transfer *after* they are initially assigned to a school, so the outside option by assigned to $\emptyset$ represents being unmatched with an alternative destination school and ending up with the original destination school.

Each student s has a strict preference $\succ_s$ over G .¹⁰¹ For school-grade $g, g' \in G$, $g \succ_s g'$ means that student s strictly prefers g to s' . Assume $\succ_s$ satisfies completeness and transitivity. For any $s \in S$ and $g, g' \in G$, denote $g \succeq_s g'$ if $g \succ_s g'$ or $g = g'$. I assume that school-grades do not have any preference.¹⁰² Denote $\succ = (\succ_1, \succ_2, \dots, \succ_N)$ as a preference profile.

Define the probability that student $s \in S$ is assigned to $g \in G$ as P_{sg} , where $P_{sg} \in [0, 1]$ and $\sum_{g \in G} P_{sg} = 1$ for all $s \in S$, and a row vector of these probabilities for s as $P_s = (P_{sg})_{g \in G}$. P is a matrix that has P_s in each row, as $P = (P_s)_{s \in S}$.

I consider a random priority mechanism (RP mechanism, also known as a random serial dictatorship) as an assignment mechanism. The procedure is as follows:

Step 1. Assign a random number to each student $s \in S$.

Step 2. From the top student on the priority list based on their random numbers, assign the most preferred school-grade remaining in the pool to the student.

Step 3. Repeat Step 2 until all students are assigned to a school-grade (including $\emptyset$).

This RP mechanism has several desirable properties. First, it is strategy-proof. Each student chooses the best possible choice at each round, and the random number assignment is random, so there is no incentive to manipulate. It also satisfies a notion of fairness, equal treatment of equals, which is characterized as assigning equal probability to those who have the same preference, i.e., for $s, s' \in S$ and for any $g \in G$, $P_{sg} = P_{s'g}$. Further, it is ex-post efficient, meaning that once the allocation has been made, there is no way to benefit any student without sacrificing others by swapping schools.

¹⁰⁰I can discuss the theory at the school level by introducing a grade-level capacity constraint as a type-specific quota as in [Abdulkadiroğlu and Sönmez \(2003\)](#). But for simplicity of notation, I consider this at a school-grade level.

¹⁰¹In the empirical exercise, I ensure preferences are strict by ordering their preferred school by (1) distance from the original destination, (2) the size of the peer effects, and (3) random number.

¹⁰²This is equivalent to assuming that all school-grades $g \in G$ have a weak preference $\succeq_g$, but they are indifferent across all students. I.e., for any $s, s' \in S$, $s \sim_g s'$.

This allocation does not satisfy ordinary efficiency, meaning that if agents have Von Neumann-Morgenstern preferences, there is a Pareto-efficient allocation. However, there is no allocation mechanism that have all the above desirable properties in addition to the Pareto efficiency ([Bogomolnaia and Moulin, 2001](#)). There is another allocation mechanism that satisfies ex-ante efficiency and envy-free, which means that the probability assignment to each school is always the most preferred compared to others. However, it does not satisfy strategy-proofness and hence there is a room for manipulation. Therefore I use this RP mechanism.

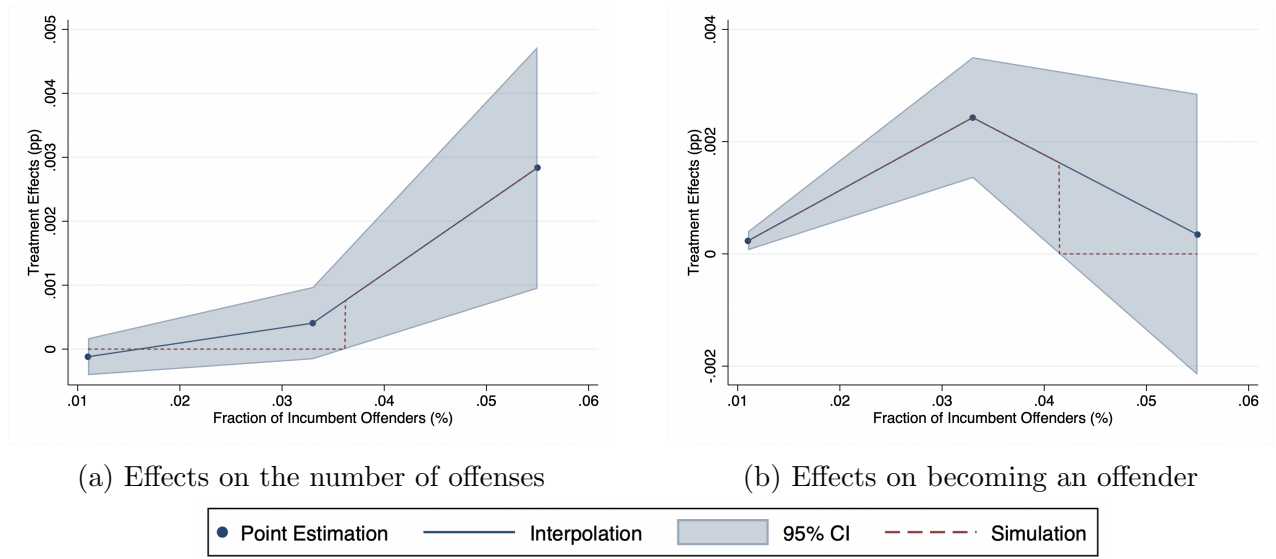


Figure E1. Interpolation of the Treatment Effects

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades, corresponding to Table 10. The estimation is based on Callaway and Sant'Anna (2021). I interpolate the treatment effects linearly.

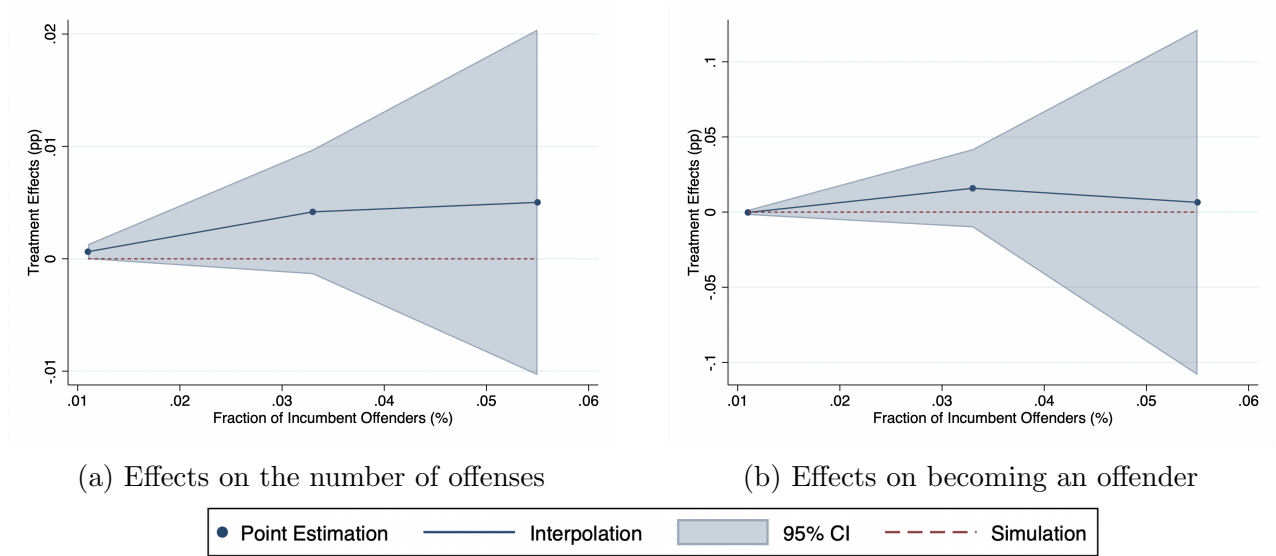


Figure E2. Placebo: Interpolation of the Treatment Effects

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades, corresponding to Table 10. The estimation is based on Callaway and Sant'Anna (2021). I interpolate the treatment effects linearly.

F Other Results

Table F1. Local Deprivation: Other Crime Types

Quartile in Deprivation Index:	Least Deprived (1)	2nd Least Deprived (2)	3rd Least Deprived (3)	Most Deprived (4)
Battery	.024*** (.008)	.040*** (.008)	.068*** (.013)	.068*** (.010)
Pre-Treatment Mean	.172	.233	.357	.460
Relative Effect Size	14.2%	17.3%	19.1%	14.9%
Drug	.013*** (.005)	.014*** (.005)	.023*** (.006)	.019*** (.005)
Pre-Treatment Mean	.024	.032	.043	.040
Relative Effect Size	53.4%	42.0%	53.1%	48.8%
Violence	.006 (.006)	.018*** (.006)	.033*** (.010)	.062*** (.009)
Pre-Treatment Mean	.083	.128	.215	.316
Relative Effect Size	7.56%	14.0%	15.2%	19.7%
Weapon	.006* (.003)	.004 (.002)	.011** (.005)	.009* (.005)
Pre-Treatment Mean	.025	.029	.048	.077
Relative Effect Size	22.6%	13.5%	22.7%	12.1%
Arson/Damage	.030*** (.008)	.025*** (.009)	.062*** (.013)	.085*** (.010)
Pre-Treatment Mean	.185	.295	.412	.604
Relative Effect Size	16.2%	8.42%	15.1%	14.0%
Sexual	-.000 (.002)	.006** (.003)	.011*** (.004)	.007** (.004)
Pre-Treatment Mean	.032	.023	.040	.053
Relative Effect Size	-7.83%	27.6%	28.3%	13.5%
Motoring	.000 (.003)	.003 (.003)	.006* (.004)	.010*** (.004)
Pre-Treatment Mean	.007	.010	.016	.035
Relative Effect Size	4.45%	30.1%	40.9%	29.9%
Others	.004 (.003)	.008** (.004)	.017*** (.004)	.012** (.006)
Pre-Treatment Mean	.023	.034	.042	.073
Relative Effect Size	19.0%	23.3%	41.1%	16.6%
Sample Size (Panel Count)	3,299,643	3,503,179	3,970,134	4,685,188
Number of Individuals	223,594	237,727	270,064	319,182

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Deprivation is measured using the 2007 Income Deprivation Affecting Children Index (IDACI), defined as the proportion of children aged 0 to 15 living in income-deprived families at the LSOA level. LSOA stands for Lower Layer Super Output Area, a geographic unit containing approximately 1,000 to 3,000 residents. Data sources: Ministry of Justice and Department for Education.

Table F2. Local Deprivation: Measure in 2004

Quartile in Deprivation Index:	Least Deprived (1)	2nd Least Deprived (2)	3rd Least Deprived (3)	Most Deprived (4)
All Crimes	.078*** (.017)	.112*** (.016)	.202*** (.023)	.317*** (.021)
Pre-Treatment Mean	.801	1.09	1.60	2.21
Relative Effect Size	9.80%	10.3%	12.6%	14.3%
Theft	.023** (.010)	.037*** (.010)	.086*** (.016)	.135*** (.014)
Pre-Treatment Mean	.342	.453	.655	.931
Relative Effect Size	6.60%	8.18%	13.1%	14.5%
Battery	.019** (.008)	.038*** (.008)	.070*** (.013)	.073*** (.010)
Pre-Treatment Mean	.176	.232	.355	.461
Relative Effect Size	10.9%	16.4%	19.8%	15.7%
Drug	.013*** (.004)	.010** (.005)	.024*** (.007)	.020*** (.006)
Pre-Treatment Mean	.023	.031	.044	.040
Relative Effect Size	57.6%	32.8%	55.4%	49.6%
Violence	.012** (.006)	.016*** (.006)	.033*** (.010)	.059*** (.009)
Pre-Treatment Mean	.083	.126	.213	.320
Relative Effect Size	14.6%	12.3%	15.7%	18.6%
Fraud	-.001 (.002)	.002 (.003)	.004 (.006)	.009** (.004)
Pre-Treatment Mean	.012	.024	.033	.046
Relative Effect Size	-4.87%	8.01%	11.3%	20.1%
Weapon	.007** (.003)	.003 (.003)	.009* (.006)	.011** (.005)
Pre-Treatment Mean	.025	.028	.050	.077
Relative Effect Size	28.7%	12.3%	18.3%	14.1%
Arson/Damage	.025*** (.009)	.030*** (.008)	.062*** (.013)	.083*** (.010)
Pre-Treatment Mean	.184	.299	.408	.605
Relative Effect Size	13.4%	10.2%	15.2%	13.7%
Sexual	-.001 (.002)	.007** (.003)	.011*** (.004)	.007** (.003)
Pre-Treatment Mean	.032	.023	.041	.052
Relative Effect Size	-2.72%	29.9%	26.6%	13.4%
Motoring	.001 (.003)	.003 (.002)	.007* (.004)	.010** (.004)
Pre-Treatment Mean	.007	.009	.016	.034
Relative Effect Size	8.58%	34.9%	40.5%	27.8%
Others	.005 (.003)	.009** (.004)	.016*** (.005)	.011* (.006)
Pre-Treatment Mean	.023	.032	.044	.072
Relative Effect Size	20.9%	27.0%	35.9%	15.8%
Sample Size (Panel Count)	3,269,879	3,558,504	3,971,545	4,658,216
Number of Individuals	221,566	241,552	270,143	317,306

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Deprivation is measured using the 2004 Income Deprivation Affecting Children Index (IDACI), defined as the proportion of children aged 0 to 15 living in income-deprived families at the LSOA level. LSOA stands for Lower Layer Super Output Area, a geographic unit containing approximately 1,000 to 3,000 residents. Data sources: Ministry of Justice and Department for Education.

Table F3. Treatment Effects on Each Crime Type

Effects on:	Violence with Injury (1)	Violence without Injury (2)	Rape (3)	Other Sexual Offence (4)	Robbery (5)	Domestic Burglary (6)	Theft of Vehicle (7)	Theft from Vehicle (8)	Theft from Person (9)	Criminal Damage (Arson) (10)	Criminal Damage (Other) (11)	Fraud (12)	Commercial Theft (13)
Time -6	.009 (.009)	.005 (.005)	-.000 (-.000)	-.006** (-.006)	.004 (.004)	.007** (.007)	.003** (.003)	.003* (.003)	-.004 (-.004)	.002 (.002)	.009 (.009)	.001 (.001)	.001* (.001)
Time -5	.005 (.005)	.005 (.005)	.000 (.000)	-.003* (-.003)	.001 (.001)	.005* (.005)	.003*** (.003)	.001 (.001)	.007 (.007)	.000 (.000)	-.003 (-.003)	.002 (.002)	.001* (.001)
Time -4	.002 (.002)	.011*** (.011)	-.000 (-.000)	-.002 (-.002)	-.000 (-.000)	.003 (.003)	.002* (.002)	.000 (.000)	.012* (.012)	-.001 (-.001)	.000 (.000)	.003 (.003)	.001* (.001)
Time -3	.008 (.008)	.009*** (.009)	.000 (.000)	-.004*** (-.004)	-.003 (-.003)	.001 (.001)	-.000 (-.000)	-.001 (-.001)	.006 (.006)	-.002** (-.002)	.005 (.005)	.001 (.001)	.001 (.001)
Time -2	-.010* (-.010)	.009*** (.009)	.000 (.000)	.001 (.001)	-.002 (-.002)	.002 (.002)	.001 (.001)	-.001 (-.001)	.010 (.010)	.000 (.000)	.006 (.006)	.003 (.003)	.000 (.000)
Time -1	0	0	0	0	0	0	0	0	0	0	0	0	0
Time 0	.014** (.014)	.010*** (.010)	.000 (.000)	.000 (.000)	-.004** (-.004)	.003 (.003)	-.001 (-.001)	-.003 (-.003)	.016** (.016)	-.000 (-.000)	.010** (.010)	-.002 (-.002)	.001* (.001)
Time 1	.019** (.019)	.011** (.011)	-.000 (-.000)	-.003 (-.003)	-.000 (-.000)	.004 (.004)	.000 (.000)	-.004 (-.004)	.012 (.012)	-.002 (-.002)	.003 (.003)	.001 (.001)	.001** (.001)
Time 2	.011 (.011)	.020* (.020)	.000 (.000)	.001 (.001)	.002 (.002)	.014* (.014)	.005** (.005)	-.003 (-.003)	.021* (.021)	-.002* (-.002)	-.005 (-.005)	-.000 (-.000)	-.001 (-.001)
Time 3	.027* (.027)	.013*** (.013)	.000 (.000)	-.002 (-.002)	-.008 (-.008)	.012*** (.012)	.004 (.004)	-.006*** (-.006)	.025 (.025)	-.001 (-.001)	.032 (.032)	-.003 (-.003)	-.000 (-.000)
Time 4	.010 (.010)	.019* (.019)	.002* (.002)	.003 (.003)	-.007 (-.007)	.005 (.005)	.001 (.001)	-.006* (-.006)	.008 (.008)	-.003 (-.003)	.007 (.007)	.001 (.001)	-.000 (-.000)
Time 5	-.009 (-.009)	.009 (.009)	.002*** (.002)	.015 (.015)	.011 (.011)	.001 (.001)	.000 (.000)	-.004 (-.004)	.037** (.037)	-.001 (-.001)	.008 (.008)	.011 (.011)	-.000** (-.000)
Time 6	.001 (.001)	.009 (.009)	.001 (.001)	.000 (.000)	-.003 (-.003)	.009* (.009)	.003 (.003)	-.001 (-.001)	.006 (.006)	-.002 (-.002)	.032** (.032)	.001 (.001)	.001 (.001)
Time 7	-.016 (-.016)	.013 (.013)	.001* (.001)	.012* (.012)	.002 (.002)	.014*** (.014)	.009*** (.009)	.000 (.000)	-.004 (-.004)	.001 (.001)	.010 (.010)	-.006 (-.006)	.002 (.002)
Time 8	-.025 (-.025)	.010 (.010)	.003* (.003)	.007 (.007)	-.001 (-.001)	.007 (.007)	.010* (.010)	-.002 (-.002)	-.008 (-.008)	.000 (.000)	.017 (.017)	.001 (.001)	.001 (.001)
Time 9	-.027* (-.027)	.016** (.016)	.001 (.001)	-.003 (-.003)	.005 (.005)	.009 (.009)	-.001 (-.001)	-.000 (-.000)	-.031* (-.031)	-.006*** (-.006)	-.008 (-.008)	.000 (.000)	.001 (.001)
Sample Size (Panel Count)	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144	15,458,144
Number of Individuals	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567	1,050,567

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. Control variables include FSM and SEN status before secondary school, an indicator for whether English is the first language, ethnicity fixed effects, cohort fixed effects, a linear month-of-birth term, and a linear age term. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table F4. Rural and Urban Areas

	Rural Areas			Urban Areas		
	Mainly Rural (1)	Largely Rural (2)	Urban with Significant Rural (3)	Urban with City and Town (4)	Urban with Minor Conurbation (5)	Urban with Major Conurbation (6)
Treatment Effects at $t = 0$	.147** (.064)	.176*** (.032)	.227** (.105)	.215*** (.019)	.267** (.118)	.200*** (.020)
Pre-Treatment Mean	1.39	1.51	1.54	1.60	1.78	1.38
Relative Effects Size	10.6%	11.7%	14.7%	13.5%	15.0%	14.5%
Sample Size (Panel Count)	1,450,874	1,991,579	2,000,476	4,549,275	732,882	4,733,058
Number of Individuals	99,039	136,555	136,240	309,107	49,785	319,841

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors reported in parentheses are clustered at the treatment level. This table reports the treatment effects of exposure to an offender's school transfer. The treatment group consists of students who are exposed to an offender transferring into their school-grade, while the control group consists of students who will be treated later. Never-treated students are excluded from the estimation. The definitions of urban and rural areas are based on the 2011 Local Authority Classification. The estimation follows [Callaway and Sant'Anna \(2021\)](#), and the table reports the contemporaneous effects at $t = 0$. The relative effect size is the ratio of the contemporaneous effect to the pre-treatment mean. All coefficients are multiplied by 100. Data sources: Ministry of Justice and Department for Education.

Table F5. School Characteristics: Offender-Dense Schools

	Lowest Dense	Middle Dense	Highest Dense	Differences between Lowest and Middle Terciles			Differences between Middle and Highest Terciles			Differences between Lowest and Highest Terciles		
				Uncond. Diff.	Cond. Diff.	Cond. Diff.	Uncond. Diff.	Cond. Diff.	Cond. Diff.	Uncond. Diff.	Cond. Diff.	Cond. Diff.
				[(1) = (2)]	[(1) = (2)]	[(1) = (2)]	[(2) = (3)]	[(2) = (3)]	[(2) = (3)]	[(1) = (3)]	[(1) = (3)]	[(1) = (3)]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Male	.506	.523	.525	.017*** (.001)	-.000*** (.000)	-.000 (.000)	.002 (.001)	-.029 (.000)	-.071 (.000)	-.026*** (.001)	-.005*** (.000)	-.000*** (.000)
Minority	.197	.201	.158	.005*** (.001)	-.004*** (.001)	.002** (.001)	-.043*** (.001)	-.029*** (.001)	.002 (.001)	-.039*** (.001)	-.033*** (.001)	.001 (.001)
English as the first language	.897	.892	.915	-.005*** (.001)	-.000*** (.000)	.000 (.000)	.022*** (.001)	-.000*** (.000)	.000 (.000)	.017*** (.001)	-.000*** (.000)	-.000 (.000)
Ever with SEN by age 11	.330	.362	.402	.032*** (.001)	-.000*** (.000)	.000*** (.000)	.040*** (.001)	-.000*** (.000)	-.000*** (.000)	.072*** (.001)	-.000*** (.000)	-.000*** (.000)
Ever eligible for FSM by age 11	.244	.292	.351	.048*** (.001)	.000*** (.000)	.000*** (.000)	.059*** (.001)	-.000*** (.000)	-.000*** (.000)	.107*** (.001)	.000*** (.000)	-.000*** (.000)
Age 5 overall score	.043 [.956]	-.029 [.966]	-.118 [.985]	-.071*** (.011)	-.026*** (.010)	-.005 (.023)	-.090*** (.015)	-.026** (.013)	-.018 (.032)	-.161*** (.015)	-.043*** (.013)	.009 (.043)
Age 7 overall score	.061 [.935]	-.018 [.957]	-.121 [.976]	-.078*** (.002)	-.040*** (.002)	-.004 (.003)	-.104*** (.002)	-.069*** (.002)	-.012*** (.003)	-.182*** (.002)	-.108*** (.002)	-.014*** (.004)
Age 11 overall score	.049 [.957]	-.041 [.980]	-.169 [.993]	-.090*** (.003)	-.061*** (.002)	.001 (.004)	-.128*** (.003)	-.097*** (.002)	-.012*** (.004)	-.218*** (.003)	-.152*** (.003)	-.007 (.006)
Cummulative N of offenses any	2.01 [2.16]	2.00 [2.31]	2.34 [2.84]	-.013 (.063)	-.038 (.063)	.011*** (.001)	.342*** (.057)	.241*** (.058)	.022*** (.001)	.329*** (.059)	.217*** (.066)	.026*** (.002)
Sample size	344,858	344,858	344,858	689,716	689,716	689,716	689,716	689,716	689,716	689,716	689,716	689,716

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. The sample in Columns (1)–(4) consists of observations in the NPD with the full set of control variables. The sample size corresponds to the sum of the 1995/96–1998/99 cohorts reported in Column (2) of Table A1 and Column (1) of Table A2. Conditional differences control for gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. All test scores are standardized within each cohort, except for the indicator variables for the age 16 tests. Early-years test scores are available only for a quasi-random 10% of the sample comprising children born on the 5th, 15th, or 25th of each month (Department for Education, Schools Analysis and Research Division, 2010). The age 7 overall score is calculated as a weighted average of teachers' assessments of Reading, Writing, Speaking and Listening, Math, and Overall Science. Within English, Reading, Writing, and Speaking and Listening are weighted 1:1:2 so that the four underlying abilities contribute equally. English, Math, and Science are then weighted 1:1:1 in constructing the overall score. The age 11 overall score is calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

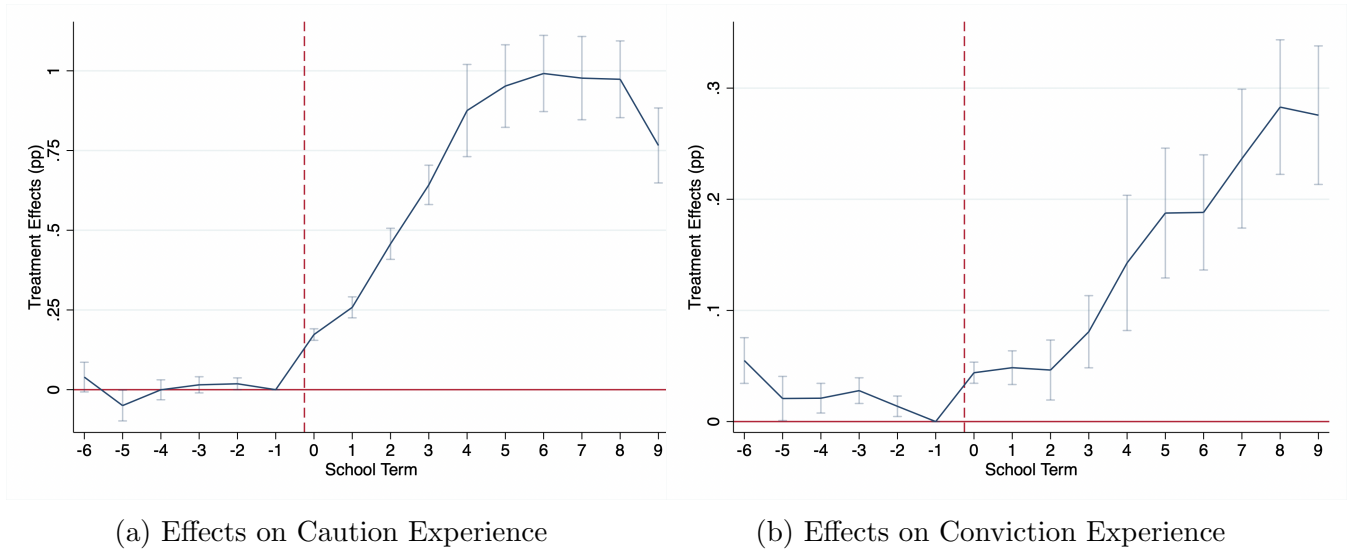
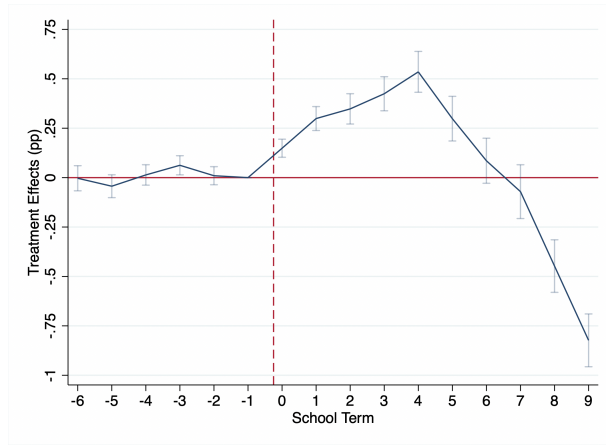
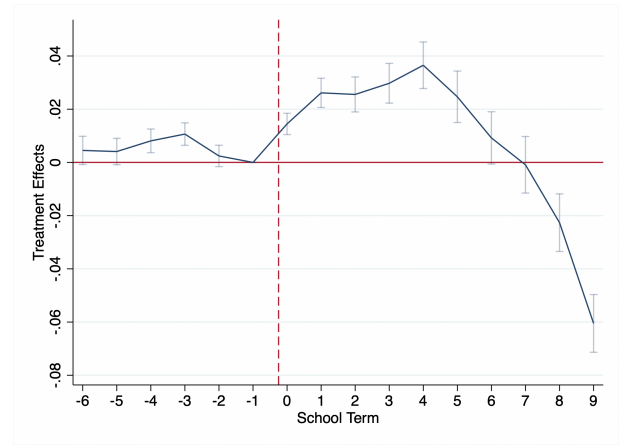


Figure F1. Treatment Effects on Cautions and Convictions Experience

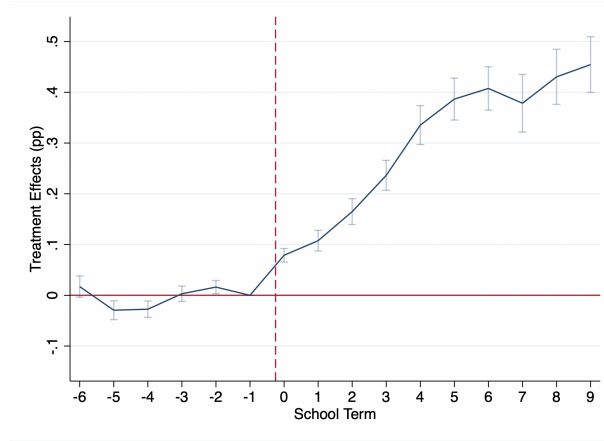
Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).



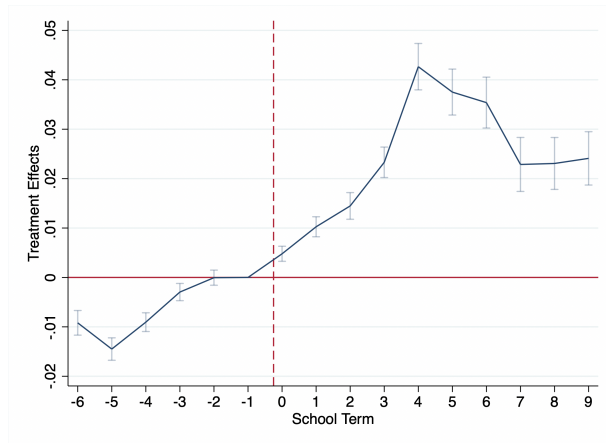
(a) Prob. of Any Fixed Exclusion



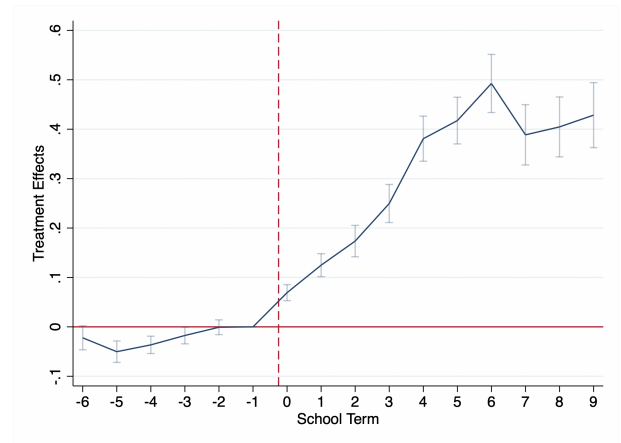
(b) N of Fixed Exclusion sessions



(c) Prob. of Unauthorized Absence



(d) Prob. of Any Fixed Exclusion or Unauthorized Absence



(e) N of Fixed Exclusion sessions or Unauthorized Absence

Figure F2. Effects on Problematic Behaviors

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).

G Results for 1991–94 cohorts

Table G1. Sample Selection

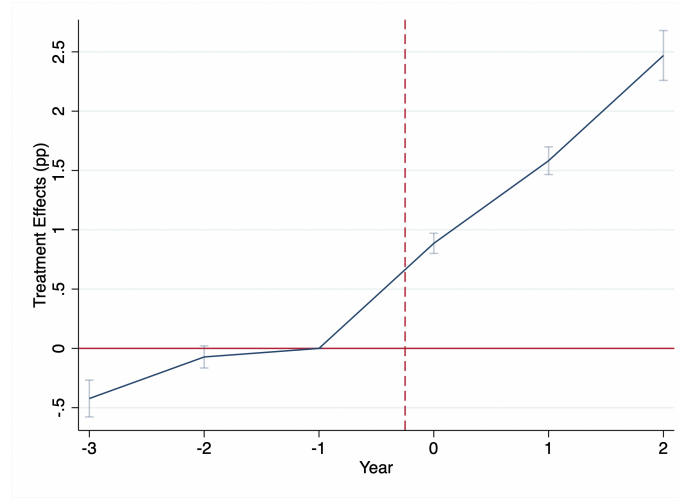
	Among NPD sample with controls Complete Records during Secondary School				Among Those in Column (1) Experienced a School-grade less than 20 Students				Among Those in Column (5) Those Treated at Age 13/14 or not			
	Complete (1)	Not Complete (2)	Uncond. Diff. [(1) = (2)] (3)	Cond. Diff. [(1) = (2)] (4)	Never Experienced (5)	Experienced (6)	Uncond. Diff. [(5) = (6)] (7)	Cond. Diff. [(5) = (6)] (8)	Not Treated at Age 13/14 (9)	Treated at Age 13/14 (10)	Uncond. Diff. [(9) = (10)] (11)	Cond. Diff. [(9) = (10)] (12)
A. Demographics												
Male	.498	.564	-.066*** (.001)	-.023*** (.001)	.498	.802	-.304*** (.009)	-.175*** (.010)	.491	.506	-.015*** (.001)	-.004*** (.001)
Minority	.161	.312	-.151*** (.001)	-.043*** (.000)	.161	.143	.018** (.008)	.014** (.007)	.173	.146	.026*** (.001)	.000 (.000)
English as the first language	.923	.799	.125*** (.001)	.040*** (.000)	.923	.952	-.028*** (.005)	-.010** (.004)	.918	.930	-.012*** (.000)	-.001*** (.000)
Ever with SEN by age 11	.236	.457	-.221*** (.001)	-.169*** (.001)	.235	.784	-.549*** (.009)	-.408*** (.010)	.225	.247	-.022*** (.001)	-.000 (.001)
Ever eligible for FSM by age 11	.179	.305	-.126*** (.001)	-.057*** (.001)	.179	.437	-.258*** (.011)	-.171*** (.012)	.171	.189	-.018*** (.001)	.000 (.001)
Offender by age 16	.080	.198	-.118*** (.001)	-.104*** (.001)	.079	.444	-.365*** (.011)	-.343*** (.012)	.074	.085	-.011*** (.000)	.001** (.000)
Any school transfer by age 16	.175	.229	-.054*** (.001)	-.026*** (.001)	.174	.830	-.655*** (.008)	-.654*** (.009)	.170	.179	-.009*** (.001)	.018*** (.001)
N of school transfers by age 16 any transfer	1.10 [.325]	1.16 [.406]	-.055*** (.001)	-.034*** (.001)	1.10 [.323]	1.24 [.500]	-.137*** (.012)	-.099*** (.013)	1.11 [.348]	1.09 [.291]	.026*** (.001)	.046*** (.001)
B. Education												
Age 11 score (English, math, science)	.071 [.960]	-.528 [1.13]	.598*** (.002)	.288*** (.002)	.072 [.959]	-.912 [1.11]	.984*** (.025)	.263*** (.023)	.130 [.962]	-.002 [.951]	.131*** (.001)	.008*** (.001)
Age 11 score: English	.082 [.953]	-.549 [1.13]	.631*** (.002)	.295*** (.002)	.083 [.952]	-.904 [1.07]	.988*** (.024)	.228*** (.022)	.141 [.955]	.010 [.943]	.131*** (.001)	.007*** (.001)
Age 11 score: math	.086 [.958]	-.537 [1.09]	.622*** (.002)	.300*** (.002)	.087 [.957]	-.755 [1.08]	.842*** (.024)	.223*** (.024)	.140 [.959]	.019 [.950]	.121*** (.001)	.008*** (.001)
Age 11 score: science	.106 [.927]	-.605 [1.17]	.711*** (.002)	.369*** (.002)	.106 [.926]	-.719 [1.15]	.825*** (.026)	.255*** (.025)	.151 [.923]	.050 [.927]	.101*** (.001)	.007*** (.001)
Age 14 score (English, math, science)	.063 [.951]	-.609 [1.23]	.672*** (.003)	.465*** (.002)	.064 [.950]	-.760 [1.15]	.824*** (.026)	.334*** (.026)	.137 [.966]	-.027 [.922]	.164*** (.001)	.009*** (.001)
Age 14 score: English	.077 [.961]	-.604 [1.10]	.681*** (.002)	.443*** (.002)	.078 [.960]	-1.06 [.978]	1.14*** (.022)	.468*** (.020)	.150 [.972]	-.013 [.937]	.163*** (.001)	.007*** (.001)
Age 14 score: math	.082 [.949]	-.485 [1.15]	.567*** (.002)	.357*** (.002)	.082 [.949]	-.150 [1.19]	.232*** (.027)	.016 (.029)	.142 [.962]	.007 [.926]	.135*** (.001)	.013*** (.001)
Age 14 score: science	.089 [.940]	-.531 [1.17]	.620*** (.002)	.453*** (.002)	.089 [.939]	-.304 [1.13]	.393*** (.025)	.244*** (.027)	.117 [.931]	.055 [.949]	.062*** (.001)	.000 (.002)
Age 16 score: taken	1.00 (.000)	.857 (.000)	.142*** (.001)	.150*** (.001)	1.00 (.000)	.998 (.001)	.002 (.001)	.002 (.001)	1.00 (.000)	1.00 (.000)	.000 (.000)	-.000 (.000)
Age 16 score: 5 A* – G	.935 (.001)	.585 (.001)	.349*** (.001)	.262*** (.001)	.935 (.001)	.335 (.011)	.600*** (.011)	.552*** (.011)	.940 (.000)	.929 (.000)	.011*** (.000)	-.002*** (.000)
Age 16 score: 5 A* – C	.620 (.001)	.261 (.001)	.359*** (.001)	.216*** (.001)	.621 (.001)	.112 (.007)	.509*** (.008)	.249*** (.008)	.649 (.001)	.586 (.001)	.063*** (.001)	.003*** (.001)
Age 16 score: point score taken	.171 [.880]	-.779 [1.14]	.950*** (.002)	.734*** (.002)	.173 [.878]	-1.37 [.956]	1.55*** (.022)	1.13*** (.023)	.214 [.874]	.120 [.881]	.095*** (.001)	.002 (.001)
Sample size	1,949,897	549,972	2,499,869	2,499,869	1,947,914	1,983	1,949,897	1,949,897	1,085,361	862,553	1,947,914	1,947,914

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. The sample in Columns (1)–(4) consists of observations in the NPD with the full set of control variables. The sample size corresponds to the sum of the 1990/91–1993/94 cohorts reported in Column (2) of Table A1 and Column (1) of Table A2. Conditional differences control for gender, minority status, an indicator for English as the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. All test scores are standardized within each cohort, except for the indicator variables for the age 16 tests. The age 11 and 14 overall scores are calculated as a weighted average of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.

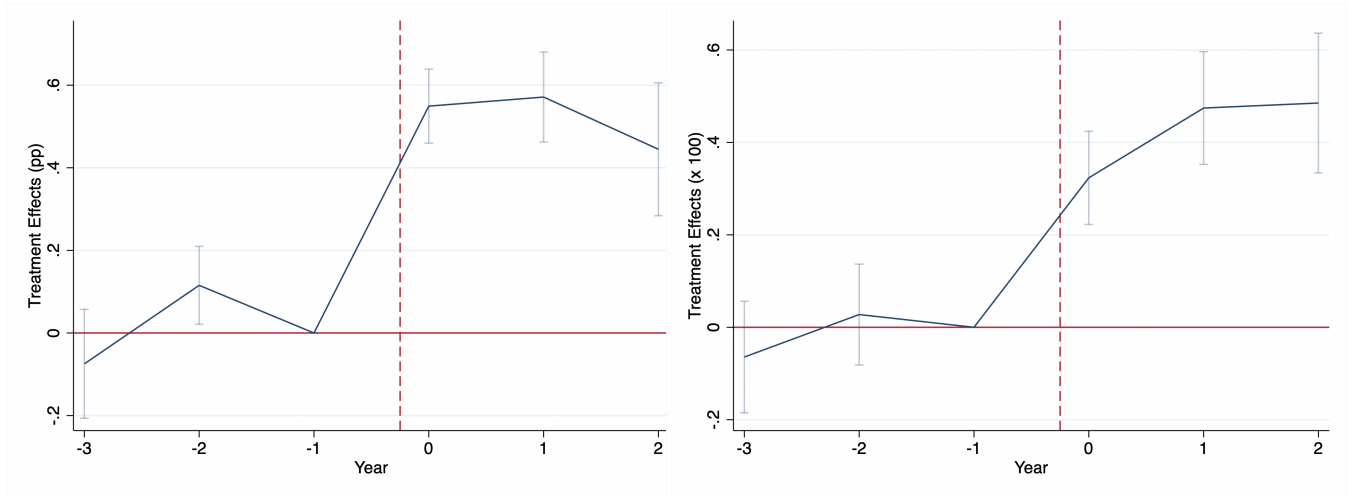
Table G2. Summary Statistics and Transfer Experience

	Treatment Status				Transfer by Age 16			
	Control	Treated	Uncond. Diff.	Cond. Diff.	Non-movers	Movers	Uncond. Diff.	Cond. Diff.
	(1)	(2)	[(1) = (2)]	[(1) = (2)]	(5)	(6)	[(5) = (6)]	[(5) = (6)]
A. Demographics								
Male	.503	.509	-.006*** (.001)	-.004*** (.001)	.508	.496	.012*** (.001)	.037*** (.002)
Minority	.154	.139	.015*** (.001)	.001 (.001)	.146	.147	-.001 (.001)	.000 (.001)
English as the first language	.926	.934	-.008*** (.001)	.001** (.000)	.929	.933	-.003*** (.001)	-.006*** (.001)
Ever with SEN by age 11	.243	.252	-.009*** (.001)	.004*** (.001)	.241	.276	-.034*** (.001)	-.048*** (.002)
Ever eligible for FSM by age 11	.186	.193	-.007*** (.001)	.004*** (.001)	.180	.231	-.051*** (.001)	-.091*** (.002)
Offender by age 16	.084	.087	-.002*** (.001)	.002** (.001)	.078	.121	-.043*** (.001)	-.078*** (.002)
Any school transfer by age 16	.224	.138	.087*** (.001)	.038*** (.001)	0	1	-	-
N of school transfers by age 16 any transfer	1.10 [.315]	1.06 [.248]	.042*** (.001)	.057*** (.002)	-	1.09 [.291]	-	-
B. Education								
Age 11 score (English, math, science)	.008 [.950]	-.011 [.951]	.019*** (.002)	-.004*** (.002)	.029 [.945]	-.140 [.965]	.169*** (.003)	.097*** (.004)
Age 11 score: English	.023 [.942]	-.001 [.944]	.025*** (.002)	-.003 (.002)	.037 [.940]	-.112 [.950]	.149*** (.003)	.081*** (.004)
Age 11 score: math	.025 [.950]	.014 [.950]	.011*** (.002)	-.005*** (.002)	.049 [.943]	-.118 [.968]	.167*** (.003)	.093*** (.004)
Age 11 score: science	.059 [.926]	.041 [.929]	.018*** (.002)	-.003 (.002)	.074 [.921]	-.059 [.951]	.133*** (.003)	.085*** (.004)
Age 14 score (English, math, science)	-.006 [.921]	-.047 [.923]	.041*** (.002)	-.005** (.002)	-.004 [.917]	-.131 [.940]	.127*** (.003)	.102*** (.004)
Age 14 score: English	.001 [.935]	-.026 [.938]	.027*** (.002)	-.014*** (.002)	.015 [.934]	-.142 [.938]	.157*** (.003)	.136*** (.004)
Age 14 score: math	.027 [.930]	-.011 [.921]	.038*** (.002)	-.000 (.003)	.020 [.921]	-.050 [.947]	.069*** (.003)	.063*** (.004)
Age 14 score: science	.070 [.946]	.041 [.951]	.029*** (.002)	.007** (.003)	.063 [.943]	.015 [.971]	.048*** (.003)	.017*** (.005)
Age 16 score: taken	1.00	1.00	.000 (.000)	.000 (.000)	1.00	.999	.001*** (.000)	.002*** (.000)
Age 16 score: 5 A* – G	.933	.926	.007*** (.001)	-.003*** (.001)	.938	.887	.051*** (.001)	.084*** (.002)
Age 16 score: 5 A* – C	.595	.577	.018*** (.001)	-.003** (.001)	.604	.501	.103*** (.001)	.121*** (.002)
Age 16 score: point score taken	.130 [.877]	.110 [.884]	.020*** (.002)	-.002 (.002)	.156 [.864]	-.045 [.934]	.201*** (.003)	.301*** (.004)
Sample size	412,447	450,106	862,553	862,553	707,893	154,660	862,553	862,553

Notes: *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard deviations are reported in brackets, and heteroskedasticity-robust standard errors clustered at the secondary-school level are reported in parentheses. I use the earliest available secondary-school record. Conditional differences control for gender, minority status, an indicator for whether English is the first language, and SEN and FSM status at age 11, unless the variable is itself the outcome. All test scores are standardized within each cohort, except for the indicator variables for the age 16 tests. The age 11 and age 14 overall scores are calculated as weighted averages of English, Math, and Science, with equal weight assigned to each subject. Some observations are missing information for certain variables. The reported sample size is the number of observations with demographic information. Data sources: Ministry of Justice and Department for Education.



(a) Offending Initiation



(b) Any Crime at Each Period

(c) N of Crimes at Each Period

Figure G1. Treatment Effects for 1991–94 cohorts

Note: This figure reports the treatment effects of exposure to offenders transferring into children's schools and grades. The treatment effect is standardized at $t = -1$. The treatment group is those who are exposed to an offender transferring into their school-grade at $t = 0$, while the control group is those who will be treated later. I exclude the never-treated group from the estimation. The estimation is based on [Callaway and Sant'Anna \(2021\)](#).