School peers' ability and youth custody

Richard Dorsett
University of Westminster, 35 Marylebone Road, London NW1 5LS

Emma Gorman
University of Westminster, 35 Marylebone Road, London NW1 5LS and IZA, Bonn, Germany

Greta Morando
UCL Social Research Institute, University College London, 55-59 Gordon Square, London WC1H 0NU and IZA, Bonn, Germany

Veruska Oppedisano*
University of Westminster, 35 Marylebone Road, London NW1 5LS, IFS and IZA, Bonn, Germany

Min Zhang
University of Westminster, 35 Marylebone Road, London NW1 5LS

1 July 2022

Abstract: In this paper we study whether the probability of young people experiencing prison is affected by the academic ability of their schoolmates. We use linked administrative data on the full population of English secondary school leavers for the summers of 2011 to 2014. To address the endogeneity of peer ability – the reflection problem – we employ an instrumental variables strategy. Our pupil-level instrument is defined as the prior attainment of those children who were at school with the pupil's peers but who themselves have never been the pupil's schoolmates; the so-called peers-of-peers. The estimating regression includes a rich set of pupil-level characteristics and also school-level characteristics to control for non-random selection into secondary schools. We find that having higher ability peers reduces the probability of youth custody. The effects are more pronounced among those living in more disadvantaged areas, and among those with lower ability at primary school. Positive effects on attainment and attendance are also seen, providing an insight into the mechanisms behind the impact on prison.

Keywords: prison, economics of crime; risky behaviours; schooling; peer effects.
JEL codes: C31, I2, I20, K42

Acknowledgements: This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. The authors thank the Department for Education for providing access to the data and seminar participants at the Melbourne Institute, University College Dublin and the Royal Economic Society conference. The project has been funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org.

* *Corresponding author:* Veruska Oppedisano: University of Westminster, School of Organisations, Economy and Society, 35 Marylebone Road, London NW1 5LS (e-mail: v.oppedisano@westminster.ac.uk).

Introduction

Schools are social places, where children's behaviour and outcomes may be influenced by their peers. These interactions are important, with the recursive nature of peer effects raising the potential for education interventions to have multiplier effects (Glaser et al., 2003). This in turn implies that simple rearrangements of students across schools or classes could improve pupil outcomes, potentially in a more cost-effective and straightforward way than alternative interventions on class size, teacher quality or other inputs to the education production function.

Recognising this, previous research has attempted to understand the nature of peer influence. Typically, the focus has been on the impact on attainment, with the general finding that this is increased when pupils have high-attaining peers (Epple and Romano, 2011 and Sacerdote, 2011, 2014). However, peers may also affect other outcomes such as risky behaviour, including criminal activity. The probability of children engaging in risky behaviour is increased where such behaviour is more common in their neighbourhood (Case and Katz, 1991) or among their peers (Gaviria and Raphael, 2001).

In this paper, we focus on a different question: how the probability of experiencing incarceration at age 17 or 18 – 'youth custody' – is affected by peer ability. This highlights the possibility that the social multiplier arising from ability of peers may not be limited to capturing the impact on academic attainment, but may extend also to broader outcomes. Hence, our results can provide a fuller understanding of the impacts of peer ability.

We measure peer ability with academic achievement in mathematics and English at secondary school, a standard measure of attainment used in the literature (Gibbons and Telhaj, 2016; Mendolia et al., 2018). Academic ability encompasses several student's characteristics, such as innate ability, cognitive and noncognitive skills (e.g. Heckman et al 2006, Papageorge et al. 2019), which are often unobserved traits in administrative/secondary data. How well students perform at school also correlates with several socio-economic and demographic

characteristics, such as gender and ethnicity (Codiroli McMaster, 2017; Strand, 2021). In this paper, we net out some of the correlated characteristics of peer academic ability, such as peer ethnicity, gender, socio-economic background, prior academic achievement, learning difficulties, and whether a child has been in state care. We hence interpret our measure of peer academic ability as a composite indicator of peer innate ability, cognitive skills, and some other unobservable characteristics such as preferences towards learning, socio-emotional skills like externalizing and risky behaviour, which are reflected in academic attainment.

We use linked administrative data for entire cohorts of English secondary school leavers in the summers of 2011-2014; more than two million children. Estimating causal effects arising from social interactions poses formidable challenges. Key among these is the simultaneity arising from the fact that while peers may exert an influence on a pupil, that same 'focal' pupil in turn exerts an effect on the peers; the so-called 'reflection' problem (Manski, 1993).[1] In addition, self-selection into peer groups is non-random. Individuals have some degrees of choice over where to live and which school to attend.

We address the reflection problem using an instrumental variable strategy similar to that of Mendolia et al. (2018). We rely on the idea that a pupil's ability at secondary school is partly influenced by their former classmates at primary school. For each individual pupil, we wish to capture the effect of the mean ability of other pupils within the same school year group. This is a leave-one-out mean and therefore varies for all pupils within the school. Because secondary school pupils come from multiple primary schools, it will be influenced by the ability of former peers of the other children. We use as our instrument the mean prior attainment of these 'peers of peers': the primary school peers of the focal pupil's secondary school peers, with the added restriction that we exclude those peers-of-peers who ever attended the same

[1] This reflection problem is less apparent than when studying same-outcome peer effects (such as the effect on attainment of peer attainment). Nevertheless, the simultaneity exists both directly as a result of peer ability being influenced by the ability of the focal pupil and indirectly if ability/attainment mediates the impact on youth custody.

school as the focal pupil. We also address the issue of self-selection into schools. To do this, we follow Nicoletti et al. (2018) and include school peers' means of set of covariates among the control variables. These variables control for other peers' characteristics that could be correlated to academic achievement.

Our key findings show non-negligible effects of peer ability on the probability of entering youth custody at ages 17 or 18 years. Increasing attainment among lower-attaining schools, such that the fraction of pupils with at least grade C in KS4 is brought to the level of the median in England as a whole, is estimated to reduce the probability of youth custody by 14%. These impacts suggest that efforts to improve attainment in lower-attaining schools will have the knock-on benefit of reduced youth custody. In terms of heterogeneity, we do not detect differences in effects by low-income status. However, impacts are considerably more negative among those living in areas of increasing deprivation, and among those in the bottom tertile of the prior ability distribution.

Our paper offers several new contributions. First, while several studies have assessed the effects of disruptive peers on subsequent criminal activity of a focal student, there is little evidence directly linking peer ability with the propensity to engage in criminal activities to the extent that imprisonment ensues. Second, our novel data allow us to study a reliable indicator of custody in young people - a rare but important outcome which is not typically feasible to study in surveys due to smaller samples and self-report biases. Finally, our empirical approach builds on and improves the identification approach used in previous similar studies in the UK due to the features of our large administrative dataset. Our results show that peer ability is a mechanism for reducing young people's probability of admissions to custody. As a result, policies that decrease residential and school segregation will – all else equal – decrease crime through the formation of more virtuous peer networks.

Reducing the criminal activity that leads to custody would bring wider benefits. First, crime generates enormous costs to society.[2] Second, incarceration during adolescence may interrupt human and social capital accumulation at a critical time, leading to greater criminal activity and reduced future wages (Aizer and Doyle, 2015).[3] Finally, those in custody are the most disadvantaged young people; in England they are disproportionately likely to come from a minority ethnic background,[4] a large proportion have previously been in care, and more than a third have a diagnosed mental health disorder (Taylor, 2016). Hence, there is a strong equity argument in support of efforts to reduce youth custody.

Related empirical literature on peer effects

There are two possible channels through which peers' characteristics may affect youth custody. First, peers' ability can affect individual's educational attainment. Higher educational attainment can lead to reductions in adult criminal behaviour by increasing the opportunity cost of engaging in crime (Lochner, 2004), by altering time discounting (Becker and Mulligan 1997), and by building non-cognitive skills that contribute to better self-control and decision-making as adults (Fudenberg and Levine 2006). Second, peers' characteristics can affect the propensity for disruptive and risky behaviours (Case and Katz, 1991; Gaviria and Raphael, 2001). Disruptive behaviour in youth is a predictor of criminal behaviour in adulthood (Tremblay et al., 1992).

[2] The total costs of crime in England and Wales in the 2015/16 are estimated to be approximately £50bn for crimes against individuals and £9bn for crimes against businesses. The economic and social costs of crime second edition, 2018. Gov.uk.

[3] Sixty-four per cent of children given a Youth Rehabilitation Order by the court, and sixty-nine per cent of those sentenced to custody, go on to reoffend within a year. Ministry of Justice (2016) Proven reoffending statistics quarterly: July 2013 to June 2014, Table C1b. https://www.gov.uk/government/statistics/proven-reoffending-statistics-quarterly-january-to-december2014

[4] The proportion of youth prisoners from Black, Asian and Minority Ethnic backgrounds has risen from 25% to 41% in the decade 2006-2016 (Lammy, 2017).

Most of the existing literature on peer effects in education focuses on the influence of peers on individuals' achievements or schooling outcomes such as the choice of college major (Lyle, 2007), university enrolment (Mendolia et al., 2018) or drop out (Gaviria and Raphael, 2001). Epple and Romano (2011) and Sacerdote (2011, 2014) provide an extensive literature on peer effects on education. Results from the multitude of research studies encompass a fairly large range of estimates, reflecting the variety of educational settings and non-linearities in peer effects. Non-linearities suggest that the impact of peer effects varies significantly at the bottom, middle and top of the distribution, with students at the bottom of the score distribution benefitting more from higher ability students and students at the top of the score distribution benefitting more from the addition of students who are also at the top of the distribution (Hoxby and Weingarth, 2005; Imberman et al., 2009; Gibbons and Telhaj, 2016). Lavy et al. (2012) find that low ability peers at the bottom of the ability distribution are more influential than those of average or high ability. Carrell and Hoekstra (2008) show significant negative effects emanating from the presence of children exposed to domestic violence. These papers are consistent with Lazear's (2001) suggestion that the most significant effects are the negative ones coming from disruptive peers.

In the UK, Mendolia et al. (2018) use survey data from the Longitudinal Study of Young People in England to study the impact of peers on attainment at ages 16 and 18 and on university attendance. They find an impact on age 18 scores but not on age 16 scores nor on the probability of going on to university. Effects are greater for pupils with low grades. Using the same data, Dickerson et al. (2018) find evidence of peers affecting educational aspirations in high school. Battiston et al. (2020), using administrative data from one cohort of the National Pupil Database (NPD), analyse educational trajectory choice and find that the probability of choosing a vocational choice is negatively affected by the presence of high ability peers. Gibbons and Telhaj (2016) exploit across cohorts variation in the composition

of students, using four cohorts of the NPD. They look at the effect of peer quality on achievement at age 14, finding relatively small effects.

An emerging literature documents the negative influence of criminally active and risk-inclined peers in a variety of settings. The effects of such peers on criminal outcomes or risky behaviour for youth seem to be larger than on education (Sacerdote, 2011). Because of its illicit nature, the criminal sector exploits informal social networks and interactions to proliferate. Evidence indicates that family background, peer influences and social interactions influence criminal behaviour, especially for less-serious crimes (Glaeser et al., 1996).

Case and Katz (1991) find that the prevalence of crime, drug use and gang membership within a neighbourhood has a large positive impact on the probability that young people will themselves engage in such risky activities. Damn and Dustnamm (2014) find strong evidence that a higher proportion of offenders convicted for crimes committed in the neighbourhood leads to higher crime conviction rates later. Bayer et al. (2009) document that exposure to peers with a history of committing a particular crime increases the probability that an individual who has already committed the same type of crime recidivates with that crime. Rotger and Galster (2019) show that the criminal record of youngsters sharing the same social housing positively impact the probability of committing offences.

A smaller number of studies have investigated the relationship between dimensions of school quality and crime. Gaviria and Raphael (2001) find strong evidence of peer-group effects at the school level for engagement in risky activities such as drug use, alcohol drinking, cigarette smoking, church-going and dropping out of high school. Using high school lotteries, Deming's (2011) results suggest that school quality predicts subsequent criminal behaviour. Billing et al. (2019) show that peer effects on criminal behaviour only arise when school peers (of the same race and gender) live less than a half-mile from each other, suggesting that school and residential segregation facilitate the formation of denser criminal networks. While there is

strong evidence of externalities for criminal behaviour among similar peers, also in schools, no study has yet investigated the possibility that virtuous high ability peers may decrease the propensity to engage in criminal activities.

Fewer studies still examine the effect of peer ability on criminality or related outcomes. Elsner and Isphording (2018) provide evidence that that a student's ordinal rank in the ability distribution of her high school cohort is an important determinant of risky behaviour. While related, this is distinct from the effect of mean peer ability. We are not aware of any studies that explore that relationship and it is here that this paper aims to make its contribution.

Institutional background

In England, schooling is organised in blocks of years called Key Stages (KS). KS1 spans from age 5 to 7; KS2 from 7 to age 10-11 (Year 6) when children leave primary education and enter secondary education with a change of school. At the end of KS2 (age 11) and KS3 (age 14) children are assessed in standard national tests in three core subjects (Maths, Science and English). KS3 tests were abolished after 2009. Compulsory secondary school is completed at the age of 16 (Year 11) with the GCSE qualification. Students do not repeat grades or retake these KS tests.

Local Authorities (LA) are responsible for organizing the admission policies for primary and secondary schools. Parents have choice over which school to enrol their children, and they rank schools in order of preference. If the schools are under-subscribed, pupils can attend that school regardless of where they live. If the schools are over-subscribed, places are allocated according to published criteria, which typically include whether the pupil has been in state care, has special educational needs, older siblings at the school, attended feeder schools, living in catchment areas, proximity to the school, and in some cases religion. We study so-called comprehensive, or mixed-ability, schools in England - in which admission based on

academic ability is prohibited. These schools represent 90 percent of state schools. School numbers and group sizes are reported in Table 1. In the period under study, the sample of schools consists of 16,506 primary schools and 3,272 secondary schools. The admission process can reflect differences in the cohort size across the years, in the number of children in need and with siblings in the school and in parental preferences, which guarantees some components of exogenous within-school variation in peer groups characteristics across the cohorts.

Table 1 Descriptive statistics – school counts

	Mean	SD
Number of primary schools	16,506	
Number of secondary schools	3,272	
Pupil counts by primary schools & cohorts	42.52	25.10
Pupil counts by primary schools	165.74	97.59
Pupil counts by secondary schools	698.65	262.75
Pupil counts by secondary schools & cohorts	178.91	64.20
Pupil counts by primary-to-secondary transition	19.36	40.03
N	380,630	

Notes: Data source is the NPD.

In secondary schools, students are grouped with different peers for different subjects, so they interact with many students in the same secondary school year group. Sometimes students are allocated in groups according to their ability level to attend some specific subject in preparation for their GCSE. This feature of the secondary educational system makes it likely that students in secondary schools will encounter most of the other students in their cohort at some point, justifying our measure of peer group quality on the basis of the group of students in the same cohort who attend the same secondary school. On average, in each cohort, there are around 179 students in a secondary school age cohort and the average primary–secondary transition group size is 19 students. For the average student, around 90% of the secondary school peer group is composed of new peers from other primary schools.

Data

This paper makes use of data for the entire cohorts of secondary school leavers in the summers of 2011-2014. The National Pupil Database (NPD) provides information about students and their academic achievements in English schools. It also records whether the individual is in youth custody at the age of 17/18. This custody information is sourced from the National Client Caseload Information System (NCCIS), one of the component tables of the NPD.

Table 2 reports on the characteristics of the analytical sample, split by whether the pupil was observed to be in custody at age 17 or 18 years. Young people who are male, of a Black ethnic origin, eligible for Free School Meals, have Special Educational Needs, or have been in care, are overrepresented among those in custody. Similarly, prior attainment (KS2) at age 11 years is lower on average, and individuals in this group are more likely to live in deprived areas as measured by the IDACI index of multiple deprivation. Our 'treatment' variable is peer ability, as measured by the KS4 scores of the focal pupil's secondary school peers. Peer tests scores are on average lower among those who enter custody.

Table 2 Pupil-level descriptive statistics by custody status

	In custody		Not in custody	
	N=2,025		N=2,118,366	
	Mean	SD	Mean	SD
Female	0.06	0.24	0.50	0.50
Ethnicity:				
White British	0.63	0.48	0.79	0.41
Black	0.15	0.36	0.04	0.20
Asian	0.07	0.26	0.08	0.27
Mixed	0.09	0.29	0.03	0.18
Other	0.06	0.24	0.05	0.21
FSM (up to 11)	0.69	0.46	0.27	0.44
Ever looked after (up to 11)	0.05	0.22	0.01	0.08
English additional language (up to 11)	0.18	0.39	0.12	0.33
SEN (up to 11)	0.72	0.45	0.34	0.47

IDACI score 11 deciles	7.68	2.35	5.49	2.88
KS2 English point score	3.89	0.77	4.52	0.71
KS2 Maths point score	4.04	0.78	4.56	0.80
KS2 Sciences score	4.31	0.73	4.79	0.66
KS4 peers' Maths	36.69	3.76	39.16	4.54
KS4 peers' English	37.48	3.80	39.70	4.19

Notes: Data sources are the NPD and the NCCIS.

Empirical approach

Estimating causal effects arising from social interactions is challenging. As Manski (1993) pointed out, estimation of these effects needs to deal with problems of endogenous effects from group behaviour; contextual effects from group characteristics, and correlated effects from unobservables that influence members of the group in common. The relationship between peers' ability and the probability of experiencing youth incarceration is explained by similarities between the individual and his peers' behaviour (endogenous effects), by the socio-economic composition of the reference group (contextual effects) and by the shared school environment they are exposed to (correlated effects). To clean the association between peers' ability and the probability of custody of these correlated observed and unobserved factors, we make use of the peers-of-peers ability instrument. A related challenge is individual self-selection into peer groups. Individuals have some degrees of choice over where to live, which school to enrol and with whom to associate. Correlations between individual behaviours and peers' characteristics may reflect some unobserved variables associated with the choice of location, school or peer group selection.

Methodologically, our identification strategy is similar to that adopted by Bramoullé et al. (2009), De Giorgi et al. (2010), Nicoletti et al. (2018) and Nicoletti and Rabe (2019), and is based on the presence of intransitivity in the network of peers. Intransitivity occurs if a person interacts with her peers - but not with all of the peers of her peers. In our application we have

intransitivity because secondary school peers have attended a variety of primary schools and therefore primary school peers network do not interact. In the educational literature, the studies of Mendolia et al. (2018), Dickerson et al. (2018), Battiston et al. (2020) are the closest, methodologically, to our study as they employ the 'peers-of-peers' strategy to primary and secondary school students.

Consider an initial specification which relates custody outcomes to peer attributes, and other background characteristics, as in Equation (1):

$$y^*_{ist} = \vartheta_0 + \vartheta_1 a_{ist} + \vartheta_2 \bar{a}_{s-i,t} + \vartheta_3 X_{ist} + \vartheta_4 \bar{Z}_{s-i} + \tau_t + u_{ist} \quad (1)$$

Where y^*_{ist} is a latent variable determining whether individual *i* who went to secondary school *s* in cohort *t* is in custody. Ability is represented by a_{ist} and $\bar{a}_{s-i,t}$ is the average ability of the secondary school peer group excluding the individual *i* (the 'leave-one-out' mean), so that ϑ_2 is the coefficient of interest. X_{ist} is a vector of pupil-level controls, $\bar{Z}_{s-1}$ a vector of secondary school peers' means characteristics and τ_t denotes cohort effects. Estimates of ϑ_2 in this specification will likely be biased due to endogeneity between peer groups and custody outcomes.

To address this issue, we use an instrumental variable approach. In line with previous studies, we rely on the idea that mean ability of a pupil's peers is partly influenced by their former classmates at primary school. Because secondary school pupils come from a number of different primary schools, ability of a pupils' peers will be partly affected by an independent influence (that of the peers who attended a primary school different from that of the pupil). Following Mendolia et al. (2018), Hedges and Speckesser (2017) and Dickerson et al. (2018), we use as our instrument the KS2 attainment of these 'peers of peers': the primary school peers of an individual's secondary school peers who attended a different primary and secondary school.

This approach, while controlling for the reflection problem, does not control for the possibility that students and families self-select into schools and peer groups that are similar to them. Primary to secondary school fixed effects would eliminate potential sorting and selection effects and control for unobserved factors of all cohorts of pupils who moved from the same primary to the same secondary school. However, we cannot include fixed effects in a probit model and we include instead cohort-specific school peers' means of set of covariates among the control variables: Index of Area Deprivation, prior attainment (KS1, KS2) in English, Maths and Reading, ethnicity, English as an Additional Language, Free School Meal eligibility, Special Educational Needs, ever been looked after by the state, ever classified as "in need", sex and absences (up to age 11 years).[5] Including these aggregate effects would produce exclusion bias arising from the fact that the assignment of peers is done without replacement: *i* cannot be his own peer. When including these aggregate effects, the exclusion of *i* from the pool of *i*'s peers creates a small sample negative relationship between *i*'s characteristics and that of his peers. To control for exclusion bias, we follow Caeyers and Fafchamps (2016) and compute school level aggregates as deviations from the school mean using the leave-one-out school mean. These variables control for peers' characteristics that could be correlated to academic achievement, reducing the concern that the link between peers' ability and the probability of youth custody is driven by peers' observable characteristics.

The strategy assumes that peer effects exist in primary schools or that children maintain relationships with ex-primary schoolmates. While there are no studies examining this link in the UK, there is extensive literature on peer effects in primary schools in the USA and in other European countries (Hoxby, 2000; Boozer and Cacciola, 2001; Ammermueller and Pischke, 2009). The first stage equation is:

[5] The linear probability model reported included in the Appendix (Table A1) as a sensitivity analysis includes primary to secondary school fixed effects along with leave-one-out school mean.

$$\bar{a}_{s-i,t} = \delta_1 \bar{a}_{pp,t} + \delta_2 X_{ist} + \delta_3 \bar{Z}_{s-i} + \tau_t + u_{it} \quad (4)$$

Average secondary school peer ability, $\bar{a}_{s-i,t}$, depends on $\bar{a}_{pp,t}$, which is the average ability of the peers-of-peers, as defined above. The reason for excluding those who have ever been at school with pupil i from the calculation of $\bar{a}_{pp,t}$ is to make sure that the ability of the peers of peers does not directly affect the pupil's probability of youth custody at the age of 17-18, other than via the on pupil i's secondary school peers.

We conduct pupil-level instrumental variable regression of custody on peer group ability. The outcome variable – whether or not the pupil experienced custody at age 17 or 18 – is binary, so we adopt a probit specification in order to constrain estimated probabilities to lie between 0 and 1. Our observed outcome, the indicator of whether pupil i in school s and cohort t is in custody at age 17/18 is defined $y_{ist} = y^*_{ist} > 0$.

We define peer ability in two ways. Our preferred measure is the proportion who achieve at least one grade C (now grade 4) at GCSE Mathematics (and English in robustness checks). The reason for this preference is that, although grade 5 is the new standard for school performance tables, a C (grade 4) remains the effective passport to post-16 education and training, such as A-Levels and equivalent level vocational studies, including apprenticeships. This reflects the long history of regarding at least a C grade as a so-called 'good' pass. This measure captures the extent to which a contingent of peers performs poorly at GCSE, failing to secure a good qualification. This measure may provide an indicator of the extent to which groups of peers are excluded from education and training. It may be that such peers exert a negative influence more than the group of peers as a whole – as measured by mean attainment – tend to do. This tallies with Lavy et al. (2012) and Mendolia et al. (2018) who examine peer effects on attainment in England and both no effect of mean ability but a significant effect of the proportion of 'bad peers'. Nonetheless, mean attainment is a natural measure to consider

so we also present results showing how peers' mean GCSE Mathematics attainment affects custody.

Results

Table 3 reports the main results. The first column shows the impact when measuring peer ability by the proportion achieving at least one C grade at GCSE while the second column measures peer ability by mean attainment at GCSE Mathematics. The 'first stage' coefficients from the regression of each measure of peer ability on the instrument (peers of peers' ability) are also reported and their high level of statistical significance suggests it is a strong instrument.

Turning to the key findings, column (1) shows a negative impact of peer ability that is statistically significant at the 95% level. The bottom panel of Table 2 translates the probit coefficient into a more substantively relevant figure, showing how the predicted probability of youth custody would change from its baseline of 0.1% were it possible to adjust peer ability in lower attaining schools such that the proportion achieving at least one C grade at GCSE was raised to the level of the median school. Doing this is estimated to reduce the probability of youth custody by 14%.

Column 2 shows that, when measuring peer ability by mean attainment at GCSE Mathematics, having higher-attaining peers significantly reduces the probability of experiencing youth custody. The bottom panel reports that a one-standard deviation increase in peer attainment would be predicted to reduce the probability of youth custody by 31%.[6]

[6] We report in the appendix alternative estimates using IV linear regression (a linear probability model, similar to Mendolia et at, 2018) and aggregated data (similar to Gibbons and Tehaj, 2016). Both approaches likewise find that higher peer ability reduces the probability of youth custody.

Table 3 Impact of peers' ability on the probability of custody

	(1)	(2)
Measure of peer ability:	*Proportion of peers with grade C+ at KS4*	*Mean KS4 of peers*
Peer ability	-3.72**	-0.456**
(s.e.)	(1.786)	(0.228)
First stage:		
Peers of peers' ability	0.00621***	0.0503***
(s.e.)	(0.00120)	(0.00677)
N	2120391	2120391
Pupil controls	x	x
Cohort effects	x	x
Aggregate controls	x	x
Predicted custody:		
- baseline	0.096 %	0.096 %
- raising % with C to median	0.082 %	
- increasing mean by 1 SD		0.066 %
- difference from baseline	-0.013 %-pts	-0.030 %-pts
- difference as % of baseline	-14 %	-31 %

Notes: Data sources are the NPD and the NCCIS. Specifications includes controls for pupil characteristics at age 11 (gender; ethnicity; English as an additional language (EAL); free school meals (FSM); Education Health and Care (EHC) plan; special educational needs (SEN); ever a child looked after (CLA); ever a child in need (CiN); unauthorised absences; attainment at KS1 and KS2), cohort and aggregated age-11 characteristics (proportion female; proportion EAL; proportion FSM; proportion EHC; proportion SEN; proportion ever CLA; % ever CiN, mean unauthorised absences; mean attainment at KS1 and KS2). Standard errors (in parentheses) are clustered by secondary school. Asterisks denote statistical significance: *=90%, **=95%, ***=99%

The remaining results focus on our preferred measure of peer ability; the proportion of peers with grade C at KS4 Mathematics. Table 4 shows how impacts vary across subgroups of pupils. We consider three dimensions: free school meal status (an indicator of low-income); whether the areas in which a school's pupils live have become more deprived between 2011 and 2014; and attainment level at Key Stage 2. For FSM status, we see no evidence that the impact on those receiving FSM differs from that on those not receiving FSM. However, for area deprivation we see that impacts are considerably more negative among those living in areas of increasing deprivation. The IDACI index measures the overall relative measure of

deprivation across local authorities on different domains, including crime. Hence, while individual impacts do not seem to vary by own-disadvantage, they do seem to vary by contextual disadvantage. This can be explained by more deprived areas offering larger opportunities for committing crime. There is also a relationship with KS2 attainment. Impacts are concentrated among those in the bottom third of the prior attainment distribution.

Table 4 Impacts by subgroup - FSM status, area gentrification and KS2 tertiles

	FSM		Change in mean area deprivation, 2011-14		KS2 tertile		
	Not FSM	FSM	No more deprived	More deprived	KS2 low	KS2 mid	KS2 high
Peer ability	-3.705	-4.297	-2.336	-7.683***	-5.588***	-1.023	-2.931
	(2.513)	(2.648)	(2.239)	(2.932)	(2.004)	(2.912)	(5.680)
N	1,554,886	565,505	1,743,680	376,711	737,472	683,263	699,330

Notes: Data sources are the NPD and the NCCIS. Specifications includes controls for pupil characteristics at age 11 (gender; ethnicity; English as an additional language (EAL); free school meals (FSM); Education Health and Care (EHC) plan; special educational needs (SEN); ever a child looked after (CLA); ever a child in need (CiN); unauthorised absences; attainment at KS1 and KS2), cohort and aggregated age-11 characteristics (proportion female; proportion EAL; proportion FSM; proportion EHC; proportion SEN; proportion ever CLA; % ever CiN, mean unauthorised absences; mean attainment at KS1 and KS2). Standard errors (in parentheses) are clustered by secondary school. Asterisks denote statistical significance: *=90%, **=95%, ***=99%

Table 5 examines the effects of peer ability on other outcomes which may act as mediators for the impact on custody. We consider two channels through which these custody impacts might arise. First, if being among higher ability peers increases an individual's own educational attainment, this may act to increase the opportunity cost of criminal behaviour. Second, higher-attaining peers are themselves likely to be better behaved and therefore less likely to lead individuals to be disruptive and/or engage in risky behaviour. To understand which mechanism is likely to be at play, we estimate our preferred regression, replacing the dependent variable with educational attainment (average point scores in Maths and English at

KS4) and then with two measures of disruptive behaviours: unauthorized absences and permanent exclusion.

We find statistically significant effects of peer ability on KS4 maths and English average point scores. These effects are sizeable and statistically significant, implying a one-standard deviation increase in peer ability increases maths and English attainment by about one-quarter of a standard deviation.[7] This is in contrast to Mendolia et al. (2018) who find no significant effect on KS4 attainment (although this may reflect their much smaller sample size), although they do find impacts on attainment at KS3 and KS5. Positive impacts on attainment are consistent with a human capital channel; exposure to high-ability peers may increase expectations, aspirations or have direct effects via peer learning on human capital development, which, – in turn, would increase the opportunity cost of engaging in criminal activity.

There is also evidence of a behavioural channel. Having higher ability peers reduces the proportion of teaching sessions missed due to unauthorised absence. A one standard deviation increase in ability reduces unauthorised absence by 1.6 percentage points. Given the mean proportion is 1.8 percent, this represents a substantial impact. By contrast, the results indicate a positive impact on the probability of permanent exclusion. While perhaps surprising, this effect is actually very small. A one-standard deviation increase in ability increases the probability of exclusion by 0.002 percentage points. Relative to the mean exclusion rate of 0.2 percent, this represents an increase of less than 1 percent. Hence, of these two measures of disruption, the effect on unauthorised absence dominates.

[7] The standard deviations for maths and English are 11.9 and 10.7, respectively. Dividing the estimated coefficients by these standard deviations yields effect sizes of 26.8% and 24.9%, respectively.

Table 5 IV regressions on educational and behavioural outcomes

	KS4 Maths	KS4 English	Proportion of sessions missed due to unauthorised absence	Permanent exclusion
	IV regression	IV regression	IV regression	IV probit
Peer ability	3.195***	2.673***	-0.115***	2.795**
	(0.370)	(0.575)	(0.0300)	(1.369)
N	2120391	2120391	2120391	2120391

Notes: Data source is the National Pupil Database and the NCCIS. Specifications includes controls for pupil characteristics at age 11 (gender; ethnicity; English as an additional language (EAL); free school meals (FSM); Education Health and Care (EHC) plan; special educational needs (SEN); ever a child looked after (CLA); ever a child in need (CiN); unauthorised absences; attainment at KS1 and KS2), cohort and aggregated age-11 characteristics (proportion female; proportion EAL; proportion FSM; proportion EHC; proportion SEN; proportion ever CLA; % ever CiN, mean unauthorised absences; mean attainment at KS1 and KS2). Standard errors (in parentheses) are clustered by secondary school. Asterisks denote statistical significance: *=90%, **=95%, ***=99%

Concluding remarks

An emerging literature documents the negative influence of criminal activity and risky behaviour among peers in a variety of settings. The effects of such peers on criminal outcomes or risky behaviour for youth seem to be larger than peer effects on education. Despite the strong arguments linking education and crime, there is very little evidence on the relationship between peer quality at school and crime. In this paper, we used a novel administrative data comprising education records and custody outcomes to study the impacts of peer ability on youth custody.

We use an instrumental variables strategy that relies on the idea that mean ability of a pupil's peers is partly influenced by the mean ability of their former classmates. For each pupil, mean peer ability is uniquely defined as a leave-one-out mean and the instrument is defined in such a way that excludes those peers-of-peers who attended the same school as the focal pupil. Hence, peers-of-peers ability can influence peer ability but is exogenous to outcomes of the focal pupil. We further strengthen our empirical approach by controlling for school peers' means of pupil characteristics and so addressing non-random selection into schools. This

removes the concern that the link between peers' ability and the probability of youth custody is driven by peers' observable characteristics.

Our main results show non-negligible effects of peer ability on youth custody. Increasing attainment among lower-attaining schools such that the fraction of pupils with at least grade C in KS4 is brought to the level of the median is estimated to reduce the probability of youth custody by 14%. While we do not detect differences in the effects by low-income status, we do see that negative impacts are considerably stronger among those living in areas of increasing deprivation, and among those in the bottom tertile of prior ability. Both individual and contextual characteristics mediate the causal relationship between peers' ability and the probability of being in youth custody. Students with low academic ability are those with a lower opportunity-cost of engaging in criminal activities; students living in increasing deprived areas, are those more likely to be exposed to a criminal environment and, hence, have more opportunities to commit crime themselves.

Our measure of ability is KS4 attainment, that may also reflect other unobservable peer characteristics correlated with criminal behaviour, such as peers' orientation towards learning, risky behaviour, and disruptiveness. Our strategy and data do not allow us to disentangle the role of attainment per se from the role of these other unobservable characteristics of peers, which is a topic of future research.

Inspecting the channels through which this impact may arise, we find positive impacts of peer ability on own-test scores. These results are consistent with a human capital channel; exposure to high ability peers may increase expectations, aspirations or have direct effects via peer learning on human capital development. This in turn would increase the opportunity cost of engaging in criminal activity. We also find support for a behavioural channel, with higher ability peers reducing unauthorised absence.

Our results imply that efforts to improve attainment in lower-attaining schools in particular will have the knock-on benefit of reduced youth custody. The results also underscore the broader implications of the development of virtuous peer networks when grouping pupils through school admission, setting or streaming policies.

References

Aizer, A., & Doyle Jr, J. J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. The Quarterly Journal of Economics, 130(2), 759-803.

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

Battiston, A., Hedges, S., Lazarowicz, T. and Speckesser, S., 2020. Peer effects and social influence in post-16 educational choice. CVER Research Discussion Paper 025.

Bayer, P., Hjalmarsson, R., & Pozen, D. (2009). Building criminal capital behind bars: Peer effects in juvenile corrections. The Quarterly Journal of Economics, 124(1), 105-147.

Becker, Gary S. and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference." Quarterly Journal of Economics 112(3): 729-758.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. Journal of econometrics, 150(1), 41-55.

Boozer, M., & Cacciola, S. E. (2001). Inside the'Black Box'of Project STAR: Estimation of peer effects using experimental data. *Available at SSRN 277009*.

Caeyers, B., & Fafchamps, M. (2016). Exclusion bias in the estimation of peer effects (No. w22565). National Bureau of Economic Research.

Case, A. C., & Katz, L. F. (1991). The company you keep: The effects of family and neighborhood on disadvantaged youths (No. w3705). National Bureau of Economic Research.

Carrell, S. E., & Hoekstra, M. L. (2008). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids (No. w14246). National Bureau of Economic Research.

Codiroli Mcmaster, N. (2017). Who studies STEM subjects at A level and degree in England? An investigation into the intersections between students' family background,

gender and ethnicity in determining choice. British Educational Research Journal, 43(3), 528-553.

Damn, A. P., & Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior?. American Economic Review, 104(6), 1806-32.

De Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. American Economic Journal: Applied Economics, 2(2), 241-75.

Deming, D. J. (2011). Better schools, less crime?. The Quarterly Journal of Economics, 126(4), 2063-2115.

Dickerson, A., Maragkou, K., & McIntosh, S. (2018). The causal effect of secondary school peers on educational aspirations. CVER Discussion Papers.

Elsner, B., & Isphording, I. E. (2018). Rank, sex, drugs, and crime. Journal of Human Resources, 53(2), 356-381.

Epple, D., & Romano, R. E. (2011). Peer effects in education: A survey of the theory and evidence. In Handbook of social economics (Vol. 1, pp. 1053-1163). North-Holland.

Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behavior. Review of Economics and Statistics, 83(2), 257-268.

Gibbons, S. and Telhaj, S., 2016. Peer effects: Evidence from secondary school transition in England. Oxford Bulletin of Economics and Statistics, 78(4), pp.548-575.

Glaeser, E., Sacerdote, B., and Scheinkman, J. (2003) The social multiplier, Journal of the European Economic Association, 1, 345–53.

Hedges, S., & Speckesser, S. (2017). Peer effects and social influence in post-16 educational choice. CVER Research Paper CVERDP008.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor economics, 24(3), 411-482.

Heeks, M., Reed, S., Tafsiri, M., & Prince, S. (2018). The economic and social costs of crime second edition. Home Office Research report 99.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation.

Hoxby, C. M., & Weingarth, G. (2005). Taking race out of the equation: School reassignment and the structure of peer effects (No. 7867). Working paper.

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. American Economic Review, 102(5), 2048-82.

Lammy, David. (2017). The Lammy Review: An Independent Review into the Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System. Gov.uk.

Lavy, V., Silva, O., & Weinhardt, F. (2012). The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, *30*(2), 367-414.

Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point. *The Review of Economics and Statistics*, *89*(2), 289-299.

Lochner, L. (2004). Education, work, and crime: A human capital approach. International Economic Review, 45(3), 811-843.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. The review of economic studies, 60(3), pp.531-542.

Mendolia, S., Paloyo, A. R., & Walker, I. (2018). Heterogeneous effects of high school peers on educational outcomes. Oxford Economic Papers, 70(3), 613-634.

Nicoletti, C., & Rabe, B. (2019). Sibling spillover effects in school achievement. Journal of Applied Econometrics, 34(4), 482-501.

Nicoletti, C., Salvanes, K. G., & Tominey, E. (2018). The family peer effect on mothers' labor supply. American Economic Journal: Applied Economics, 10(3), 206-34.

Papageorge, N. W., Ronda, V., & Zheng, Y. (2019). *The economic value of breaking bad: Misbehavior, schooling and the labor market* (No. w25602). National Bureau of Economic Research.

Proven Reoffending Statistics Quarterly Bulletin January to December 2014, (2016) England and Wales, Ministry of Justice Statistics Bulletin

Rotger, G. P., & Galster, G. C. (2019). Neighborhood peer effects on youth crime: natural experimental evidence. Journal of economic geography, 19(3), 655-676.

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

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward?. Annu. Rev. Econ., 6(1), 253-272.

Strand, S. (2021). Ethnic, socio-economic and sex inequalities in educational achievement at age 16: An analysis of the Second Longitudinal Study of Young People in England (LSYPE2). *Report for the Commission on Ethnic and Racial Disparities (CRED)*.

Taylor, C. (2016) Review of the Youth Justice System in England and Wales: London: MoJ.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. The Economic Journal, 113(485), F3-F33.

Tremblay, R. E., Masse, B., Perron, D., LeBlanc, M., Schwartzman, A. E., & Ledingham, J. E. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: longitudinal analyses. *Journal of consulting and clinical psychology*, *60*(1), 64.

Appendix - Alternative identification and estimation approaches

To explore the robustness of our main estimates, we show results for linear probability models, rather than probit. Table A1 presents the key results from individual-level regressions. Column 1 estimates adjust for pupil characteristics, cohort fixed effects, secondary school fixed effects and primary-secondary school combined fixed effects. Specification 2 uses IV regression (with the peers-of-peers instrument) to address endogeneity of peer ability and controls for pupil characteristics and cohort. Specification 3 again uses the peers-of-peers instrument, with controls for pupil characteristics and cohort plus pupil characteristics aggregated to the school-level. Across all specifications, we find that the probability of experiencing youth custody is reduced by the having higher ability peers.

Table A1 Linear probability models

	(1) OLS	(2) IV	(3) IV
Peer effects captured as:			
Mean KS4 of peers	-0.000325*** (0.0000834)	-0.000296*** (0.0000534)	-0.00155** (0.000752)
% of peers with grade C at KS4	-0.001733*** (0.000421)	-0.00245*** (0.00044)	-0.01164** (0.005875)
Controls:			
Student's characteristics	x	x	x
Fixed effects			
- Cohort	x	x	x
- KS4 school	x		
- KS2 school	x		
- KS2 school X KS4 school	x		
Aggregate leave-one-out means			x
Unweighted N	1999030	2120391	2120391

Notes: Pupil-level covariates included are: prior KS1 and KS2 attainment, gender, ethnicity, FSM, EAL, whether the child is looked after, whether the child has special educational needs, mainstream school, IDACI decile dummies. Standard errors (in parentheses) are clustered by secondary school. In column (1), the number of observations, N, is smaller than for the main estimates due to singleton observations being dropped when there are multiple fixed effects. Asterisks denote statistical significance: *=90%, **=95%, ***=99%.

Table A2 shows school-level estimation results, following the strategy implemented by Gibbons and Telhaj (2016). They eliminate high-dimensional fixed and trending effects without losing information on peer group changes, by estimating a transformed model based on within-school, cohort-to-cohort differences. The aggregation has also the advantage of mitigating some of the problems inherent in individual level value-added models, when test-scores are noisy measures of prior achievement (Todd and Wolpin, 2003). Using multiple cohorts of administrative pupil level data, the analysis is in practice performed on student data aggregated to primary school-by-secondary school-by-cohort cells.

To control for fixed unobservable characteristics common to students making the same primary-school-to-secondary-school transition over the years, we take the first difference of variables across adjacent cohorts, within primary-school-to-secondary-school transition. To control for unobservable fixed characteristics common to all students leaving primary school *p* in cohort *t*, we take the within group transformation differencing the variables from primary-school-by-cohort means. The estimated specification reads as follow:

$$\tilde{\Delta} y_{pst} = \vartheta_1 \tilde{\Delta} a_{pst} + \vartheta_2 \tilde{\Delta} X_{st} + \tilde{\Delta} v_{st} + \tilde{\Delta} u_{pst} \quad (A1)$$

where y_{pst} is the proportion of individuals who went to the same primary *p* and secondary school *s* in cohort *t* and are in custody. a_{pst} is the average ability of the group of peers who went to the same primary and secondary school in cohort *t*. ϑ_1 is the coefficient of interest. X_i is a vector of the same controls listed in the previous specifications (prior KS1 and KS2 achievements, gender, ethnicity, FSM, EAL, IDACI decile dummies).

The notation $\tilde{\Delta}$ is defined to the mean. For example, for the peer group variable X: $\tilde{\Delta} X_{st} = (X_{st} - \tilde{X}^p_{st}) - (X_{st-1} - \tilde{X}^p_{st-1})$ where $\tilde{X}^p_{st}$ is the average of the secondary school peer group characteristics for students who were in the same primary school *p* in a given cohort *t*.

The first component in brackets represents the difference of characteristics between a cohort of secondary school students from a given primary school and their peers in the same cohort from the same primary school who went to a different secondary school. The second term in brackets compute the same difference for the previous cohort. Identification of ϑ_1 arises from the change between cohorts in the differences in secondary peer group quality. This variation reflects between-cohort differences in observed achievements. The identifying assumption is that shocks to peer group quality are unpredictable, uncorrelated with secondary school quality and uncorrelated with unobservable determinants of custody.

Columns 1 and 2 show unweighted and weighted results, respectively. Here, the weights reflect the size of the school. The results are similar, again showing a negative impact of peer ability on the probability of experiencing custody.

Table A2 School-level regressions	(1) Unweighted	(2) Weighted
Peer effects captured as:		
Mean KS4 of peers	-0.000817*** (0.000234)	-0.000868*** (0.0000141)
% of peers with grade C at KS4	-0.00483*** (0.00119)	-0.00484*** (0.0000707)
Controls:		
Mean pupil characteristics, by KS4 x KS2 x cohort	x	x
Fixed effects		
- Cohort	x	x
- KS4 school	x	x
- KS2 school	x	x
- KS2 school X KS4 school	x	x
Unweighted N	257466	257466

Notes: Standard errors (in parentheses) are clustered by secondary school. Asterisks denote statistical significance: *=90%, **=95%, ***=99%.