Does School Choice Increase Crime?

Andrew Bibler, University of Nevada, Las Vegas

Stephen B. Billings, University of Colorado - Boulder

Stephen Ross, University of Connecticut & NBER

January 23, 2024

Abstract

School choice lotteries are an important tool for allocating access to high-quality and oversubscribed public schools. While prior evidence suggests that winning a school lottery decreases adult criminality, there is little evidence for how school choice lotteries impact non-lottery students who are left behind at their neighborhood school. We leverage variation in actual lottery winners conditional on expected lottery winners to link the displacement of middle school peers to adult criminal outcomes. We find that non-applicant boys are more likely to be arrested as adults when applicants from their neighborhood win the school choice lottery. These effects are concentrated among boys who are at low risk of being arrested based on observables. Finally, we confirm existing evidence that students who win the lottery have lower adult criminality. Putting these results together, we show that, on net, school choice lotteries increase overall arrests and days incarcerated for young men.

We thank seminar participants at University of Nevada – Las Vegas, London School of Economics, University of Bocconi, University of Wyoming, NBER Winter Economics of Crime Meeting 2024, APPAM 2023, the 2022 Urban Economics Association Meetings, the 2022 Society for Institutional and Organizational Economics Meetings, and the 2019 NYU Wagner School Conference on Race, Crime, and Policing. We are grateful to the North Carolina Education Research Data Center, Mecklenburg County Sheriff's Office and Charlotte-Mecklenburg School District for providing data.

Introduction

School choice is an increasingly popular tool for public school districts to better compete with private and charter schools and stem the loss of students to neighboring school districts (Brunner et al., 2012; Tuttle, Gleason & Clark, 2012). School choice also has the potential to delink residence from schools which may address recent social trends of increasing residential segregation (Schwartz, Voicu & Mertens Horn, 2014; Hess, 2021). One of the byproducts of school choice is the oversubscription of high quality and specialized schools for which school choice lotteries are used to assign limited classroom space to students.

Using the random assignment mechanism from school choice lotteries, several scholars (Abdulkadiroglu et al., 2011; Abdulkadiroglu, Pathak, & Walters, 2018; Cullen, Jacob, & Levitt, 2006; Dobbie and Fryer, 2015; Hastings, Kane, & Staiger, 2006; Hastings & Weinstein, 2008; Imberman, 2011; Deming et al., 2014; Muralidharan & Sundararaman, 2015; Mills & Wolf, 2017) estimate effects of winning a school choice lottery on later end-of-grade exams and other academic outcomes, finding mixed evidence. In contrast, scholars consistently find benefits in terms of non-academic outcomes with lower self-reported disciplinary issues, arrest, and incarceration for lottery winners (Cullen, Jacob, & Levitt, 2006; Deming, 2011; Imberman, 2011; Dobbie and Fryer, 2015). To our knowledge no previous research examines the effect of winning a lottery on the subsequent school discipline and criminal justice outcomes of the students left behind. One related paper (Lavy, 2010) highlights the aggregate impacts of introducing a school choice system, finding a net reduction in violence and classroom disruption in Tel-Aviv.[1]

[1] Dills and Hernandez-Julian (2010) find that metropolitan areas with more school choice have lower teenage arrest rates, using rivers and streams as an instrument. However, their analysis differs from ours and Lavy's in that school choice lotteries disrupt existing social networks, while jurisdictional fragmentation and rivers/streams separate populations and prevent associations from ever being formed.

In the study presented here, we examine the impact of school choice lotteries on the applicant's neighborhood peers, as well as the direct effects on lottery winners. Thus, we provide novel evidence on the adult criminality of students who experience the loss of a lottery winner in their neighborhood middle school. We then compare the indirect effects (on peers left behind) with the direct effects (on lottery winners) to assess of the impact of school choice lotteries on the arrest and incarceration rates of teens and young adult males.

While prior literature suggests that winning a lottery decreases adult criminality, the expected effects on non-applicants is less clear. If lottery winners are positive school peers and/or their families offer positive parental inputs to neighborhood schools, then non-applicants may be worse off in terms of criminality. Furthermore, the relationship between lottery winners and non-participants may be even more complicated by the behavioral response of parents and the sorting of teachers to high-achieving students (Pop-Eleches & Urquiola, 2013).[2] Alternatively, a number of papers (e.g. Glaeser, Sacerdote & Scheinkman, 1996; Bayer, Hjalmarsson & Pozen, 2009; Billings, Deming, & Ross, 2019; Damm & Gorinas, 2020; Kim & Fletcher, 2018; Billings & Hoekstra, 2022) highlight the negative influence of peers on adult crime. Therefore, if lottery winners are negative or neutral peers, then students who are left behind may be less likely to engage in future criminal activity.

To isolate the effects on non-applicants, we follow existing work by focusing our analysis of spillovers from lottery winners onto kids who live in the same neighborhood, are the same gender, are similar in age and are assigned to attend the same middle school (Billings, Deming, and Ross, 2019). Our main empirical analysis includes a sample of three cohorts of male 5th grade

[2] Existing literature on the role of positive peer effects on youth criminality is very limited and primarily based on lottery winners moving to schools with positive peer attributes (e.g. Deming, 2011; Cullen, Jacob, & Levitt, 2006).

students in Charlotte, NC in the 2005-2006 to 2007-2008 school years. The geography and cohorts included in our sample are dictated by the availability of student level lottery and non-applicant administrative data (North Carolina Education Data Research Center) matched to arrest and incarceration data (Mecklenburg County Sheriff's Office and NC Dept of Public Safety). Our primary estimates are for the effects of losing peers through the lottery on a student's adult (age 16-22) arrest and incarceration. Our primary estimation sample is restricted to students who did not apply to the school lottery and were initially assigned to their neighborhood school in 6th grade.

We estimate the effect of peer school choice lottery wins using the number of 5th grade applicants in the same cohort and same within school attendance zone neighborhood (i.e., same Census Block Group (CBG) based on 5th grade residential location and same neighborhood middle school zone) who won their first choice in the lottery, divided by the total number of other 5th grade students in this cohort and neighborhood (i.e., win share).[3] Following Borusyak and Hull (In Press), we isolate the random portion of exposure to lottery winners by conditioning on a control function for the expected win shares based on the individual applicant probabilities of winning the lottery, which is similar to Abdulkadiroglu et al.'s (2017) approach for identifying the effect on a student of winning the lottery.[4] We also condition on middle school attendance zone, block group and cohort fixed effects to address any spatial and temporal variation in expectations around expected win shares. As a pseudo first stage, we confirm that lottery winners are substantially more

[3] Deming (2011) who estimates effects of winning a school lottery on future arrests in Charlotte-Mecklenburg County also examines the effect of first choice lottery wins.

[4] For example, the expected win share captures nonrandom factors such as application choices, neighborhood, year of application as well as characteristics related to lottery priorities (e.g., Title I choice status, economic disadvantage) that determine win probabilities. Borusyak and Hull (2021) is a generalization of Abdulkadiroglu et al.'s (2017). Abdulkadiroglu et al.'s (2017) develops an approach for isolating the random variation arising from individual lottery wins, while Borusyak and Hull (2021) develop a general approach for isolating the random variation in a broad class of functions that depend upon both random shocks and the endogenous exposure of individuals to those shocks.

likely to attend their first choice school than lottery losers and less likely to attend their neighborhood school.[5]

Looking within assigned middle school, we find that male students whose immediate, same-grade neighbors win the lottery are more likely to be arrested at ages 16-22. A change from no winning peers to the average proportion of winners in a neighborhood (4.0% of same grade male students in neighborhood) implies a 0.06 increase in the number of arrests from ages 16-22, which is an 13% increase over the sample average. The effects are more precisely estimated and considerably larger for students with below median arrest risk. Among those with low predicted risk of future arrest, a change from no winners to the average win share leads to a 1.3 percentage point increase in the probability of arrest, 0.024 more violent arrests and 1.2 more days incarcerated.[6]

Lottery applicants have substantially higher test scores than non-applicants at the time of application and lottery losers who remain at the neighborhood school continue to outperform non-applicants through middle school. Thus, the loss of these high-achieving, neighborhood peers from the school environment appears to negatively affect student's later life outcomes. This effect is more pronounced among students at low a priori risk of arrest who may have been more likely to associate with high-achieving peers. Further, we show that the negative effects of lost peers emerge in the short run by documenting effects on disciplinary outcomes starting in middle school. These results are consistent with Bacher-Hicks, Billings & Deming (2019) who find that school discipline problems contribute to future arrest and incarceration as an adult.[7] We also show that the effects

[5] While we find that lottery losers are more likely to change residences to access higher-quality schools in different neighborhoods, consistent with Bibler & Billings (2020), this effect is quite small relative to the compliance of lottery winners with their first choice.

[6] While these changes may appear modest, the lower rates of arrest and incarceration for this subsample imply large percentage increases in incidence of 24.7%, 60.1% and 121% of the respective means.

[7] Sorensen, Bushway & Gifford (2022); and Fabelo et al. (2011) show that school discipline has a negative impact on several academic and behavioral outcomes.

on arrests persist and increase in absolute magnitude after age 19 when most youth are no longer in high school.[8]

Consistent with a peer effects mechanism, increases in adult criminality are stronger for non-applicants who attended elementary school with lottery winners. Further, we rule out broader effects of lotteries on school quality by showing that the effects are not related to underlying changes in neighborhood middle school characteristics and by demonstrating that our estimates are robust to the inclusion of attendance zone by cohort fixed effects, which non-parametrically controls for any school level effects unique to each cohort. By ruling out cohort wide effects as an explanation, we can point to residential proximity as the key factor driving our results, and residential proximity is an important determinant of social or peer relationships (Billings, Deming & Ross, 2019). We further show no effects on academic outcomes which suggests that our criminal justice outcomes are not simply a function of worse labor or education options as young adults.

Notably, we find no effects on the criminal outcomes for girls. These gender differences are consistent with Lavy and Sand (2019) who find in Israel that being separated from elementary school friends in middle school leads to large declines in social and school satisfaction and increases in violent behaviors for boys, while effects for girls are smaller in magnitude and focused on feelings of safety and social satisfaction. Together, our results support the importance of peer influences among boys, which are often solidified in middle school, and show that the loss of positive role models and friendships may lead to an increase in anti-social behavior that lasts beyond participation in K-12 education.

By controlling explicitly for the proportion of the peer group expected to win the lottery, we use identifying variation based on the random selection of lottery winners from the population

[8] The persistence of these effects contrasts with findings on high school peer effects on academic outcomes by Bifulco, Fletcher & Ross (2011), which are shown to fade over time (Bifulco et al., 2014).

of applicants. To confirm the quasi-random nature of this variation, we show that conditional on the expected lottery win share, actual win shares are uncorrelated with observable student attributes. We also provide a series of robustness tests in which we alter the combinations of neighborhood and school fixed effects, including a specification based solely on across cohort variation within attendance zone neighborhoods, all of which produce similar results. Finally, we show that using win shares of future cohorts for the neighborhood produces insignificant estimates that are substantially smaller in magnitude than our main estimates. This falsification test limits concerns that neighborhood-specific trends are correlated with unexpected wins.

Since lottery applicants tend to be positively selected on academics, the finding that effects are concentrated among low crime risk (and higher achievement) students is consistent with peer effects that operate between similar students (Billings, Deming & Ross, 2019; Fletcher, Ross & Zhang, 2020). Therefore, these findings speak to a broader literature on school segregation and criminal activity among youth; most notably, Billings, Deming & Rockoff (2014), Weiner, Lutz & Ludwig (2009) and Johnson (2011) show that racial segregation contributes to African-American youth involvement in the criminal justice system. However, unlike Billings, Deming & Ross (2019), our estimated effects are concentrated among low-risk students, and we do not find evidence of larger effects when the winner and the low-risk students left behind are the same race. This suggests that the social processes involved in the exit of positively selected peers may be quite different than, for example, rezoning, which could lead to increased concentrations of at-risk students (Billings, Deming and Rockoff, 2014).

To examine the role of benefits to lottery winners, we estimate the direct effects of winning the lottery on adult criminality in the sample of lottery applicants to oversubscribed schools. While this analysis is similar to Deming (2011), who also estimates the effects of winning the lottery on criminality in Mecklenburg County, our sample time period and even the mix of schools are different given our focus on more recent cohorts. Like Deming (2011), we find that winning the lottery leads to declines in the probability of being incarcerated for students in the top quintile of predicted arrest risk. However, we also show that winning the lottery reduced criminality among students at median and below median risk of arrest in our sample.

In aggregate, our results provide insight into the net effects of lotteries on young adult criminality. Specifically, we use the population of students in neighborhoods that are exposed to lottery applicants to calculate the number of additional arrests and incarceration days accumulated by non-applicants, as well as the reductions in arrests and incarceration for lottery winners. We find that the current constrained school choice system generated benefits to winners of around 101 fewer arrests and at least 2,070 fewer days incarcerated and generated costs to same neighborhood, non-applicants of at least 211 more arrests and 4,131 more days incarcerated. In aggregate, focusing on aggregates from significant estimates, we find at least a 4.8% net increase in total arrests and 1.3% more days incarcerated due to the lottery application process for our three cohorts of 5th grade students in our estimation sample (n=9,685). These net negative effects could be an understatement if lottery winners have broader effects on students assigned to the same middle school beyond peers residing in their local neighborhood. In the end, our net estimates are partial equilibrium in nature and thus cannot account for student residential sorting that would occur in the absence of school choice.[9]

While school choice in the U.S. has grown dramatically over the last few years and many papers have examined the impacts on lottery winners, the literature on the broader impacts of

[9] Altonji, Huang & Tabor (2015) develop an analytical approach for assessing the general equilibrium effects of introducing a system of school choice on students left behind based on estimating structural models of both student preferences over school attributes and the effect of school peer composition on student outcomes.

school choice is limited. Our results of worse net outcomes are consistent with recent theoretical work by Barseghyan et al. (2019) showing that in the presence of strong peer preferences, school choice can be welfare decreasing since aggregate peer quality is fixed. However, our findings on arrest and incarceration run contrary to several existing empirical papers that study the broader impacts of school choice on academic outcomes outside of the lottery setting. Gilraine, Petronijevic, & Singleton (2021) find positive aggregate effects of charter school expansion on math test scores in North Carolina, while Mumma (2020) finds no test score effects in Massachusetts and North Carolina for students who attend schools near charter school openings.[10] In Israel, Lavy (2010, 2021) documents cognitive and long term labor market benefits for children in school districts that adopt or expand school choice, and Hsieh and Urquiola (2006) find little to no aggregate impact from expanded school choice in Chile. Muralidharan & Sundararaman (2015) examine a lottery that provides private school vouchers in India, finding limited benefits in test scores to winners with no adverse effects on non-lottery participants in the same village. In supplementary analyses, we do find limited evidence of small positive effects on test scores for lottery winners and marginally significant decreases in test scores for students left behind.

Our findings are especially important because school choice has proven popular among local voters and often are implemented to reduce the exit of high-achieving peers from disadvantaged school districts (Cookson, Peter and Schneider, 2014; Hochschild and Scovronick, 2003). School choice programs that more strongly promote school choice among lower-achieving students may lead to a different type of selection into choice than documented here, as well as by Hastings, Kane, & Staiger (2006), Burgess et al. (2015) and Barseghyan et al. (2019). A more

[10] In related work, several papers examine the effects of school choice on racial segregation (Mumma 2022; Monarrez, Kisida and Chingos, 2022; Bifulco, Ladd and Ross, 2009a, 2009b) finding at most modest effects of choice on racial segregation.

representative and less positively selected composition of lottery winners would likely generate lower costs on non-applicants. For example, Rucinski & Goodman (2022) recommend providing lower-achieving students higher priorities in school choice lotteries, as well as removing academic admissions exams for certain choice schools. Notably, many state financing systems shift funding away from residentially assigned schools when students opt out of their assigned school, which might limit the ability of schools to address the negative effects experienced by students left behind. Ultimately, our findings suggest that the gains of school choice from retaining high-achieving students in public school districts likely comes at a significant cost to non-school choice students in the form of increased adult criminality.

School Choice In Charlotte-Mecklenburg Schools

We use data from Charlotte-Mecklenburg Schools (CMS) which is a large and diverse, urban school district that currently serves about 150,000 students. The CMS school choice system started in 2002 and shares a number of attributes with other commonly studied US schools districts (e.g. Boston, NYC, Chicago, etc.). By default, all CMS students are assigned to a neighborhood (home) school based on residential address. Through the school choice system, students can apply to several magnet school options or other neighborhood schools that are not based on their residential assignment. CMS uses a centralized lottery system to ration seats in oversubscribed programs. In the lottery, students can submit up to three program choices in order of preference.[11] Nonguaranteed seats are assigned in three rounds, considering only first choices in the first round

[11] We use the term program rather than school because students apply for specific grades as well as special magnet programs that in some cases encompass only a portion of classrooms in a school.

of assignment.[12] If there are more seats available than applicants for a given program, then all applicants to that program will be assigned to their first choice.

When a choice is oversubscribed, meaning that the number of applicants is greater than the number of available seats, assignment is quasi-random. Seat assignment is not completely random because the probability of winning for a particular student depends on that student's priority group. "Priority groups" refer to sets of students who meet some prespecified individual criteria such as economic disadvantage as well as geographic criteria such as the Title I choice status of the student's neighborhood school.[13] If a student is not assigned to their first choice, they remain in the unassigned pool and may win a seat to another choice in the following rounds. If a student does not win any of their lottery choices, they are assigned to their neighborhood school based on pre-specified school attendance zone boundaries. Because lottery choices are considered sequentially, students are most likely to win a choice by making it their first choice, and most seats are awarded in the first round.

In addition to the lottery rules, some magnet programs have specific entrance requirements, which are often based on end-of-grade exams in the prior year. For example, to enter one of the STEM programs in sixth grade, students must score at grade level in reading, math, and science on their fifth grade exams. When the requirements are based on test scores, we can observe whether a student met the stated requirements for their program of choice when determining the likelihood of admission. The share of applications who won their first choice is about 35% over our sample of applicants entering sixth grade in the 2006-2007 to 2008-2009 school years.

---

[12] Siblings of current students are guaranteed admission.

[13] Title I schools have a relatively high proportion of economically disadvantaged students. A Title I school is designated as a Title I choice school after failing to meet adequate yearly progress in the same subject for two consecutive years. No Child Left Behind (NCLB) mandated that students have the opportunity to attend a non–Title I choice school, but not necessarily to choose the school they attend.

To examine the effects of the lottery on both students who forego the lottery, as well as individual winners, our analysis focuses on two distinct samples: lottery applicants and non-applicants. To construct the sample of lottery applicants, we use any student who applied to a school (other than their residentially assigned neighborhood school) and who are neither guaranteed a seat nor have zero chance of admission, e.g., we drop sibling placements and limit to oversubscribed programs.[14] We also restrict to students who met the requirements for their first-choice program.[15] The sample of non-applicants includes students who did not specify any choice in the lottery and who are initially assigned to their neighborhood school – the school assigned based on their 5th grade residence – for their 6th grade year. Most of the students (64%) in the non-applicant sample have exposure to lottery applicants, meaning that would-be peers from their Census 2000 Block Group (CBG) and neighborhood school zone did apply to the lottery. The remaining 36% of non-applicants have no exposure to lottery applicants from their CBG, neighborhood school and cohort group, but they do share a CBG or neighborhood school with other students in the non-applicant sample.

Data

Given our focus on lottery applicants as well as non-applicants, our main data sample is comprised of administrative records from Charlotte-Mecklenburg Schools (CMS) for 24,883 5th grade students who attended public school in the county between the years of 2005-2006 to 2007-2008. Within the 28 middle schools in CMS, there are 32 oversubscribed programs across 27

[14] We also drop students with placements likely related to magnet continuation based on district documentation and data on current and application schools.

[15] We drop applicants to arts programs and leadership programs for our analysis because we do not observe whether the student met the entrance requirements. Arts schools require an audition or portfolio assessment, and leadership schools require an interview.

schools applied to by our lottery sample, which come from 59 distinct lotteries (program by cohort combinations). We focus on the 6th grade lotteries for these cohorts, which allows us to capture experiences when students are relatively young and still observe the sample as adults. The data include student gender, race, an indicator for economic disadvantage, yearly end-of-grade (EOG) test scores, days absent and days suspended from school. The EOG tests are standardized and administered across the state of North Carolina from 1993 to the present.

To measure adult criminal justice outcomes, the North Carolina Education Research Data Center (NCEDRC) linked CMS administrative data to arrest registry data for Mecklenburg County using first and last name as well as date of birth. The arrest data includes individual names and identifiers, information on the number and nature of charges and any spells of incarceration.[16] The incarceration data provides the time span in which an individual spent time in a county or state facility in North Carolina. The NCEDRC used a sequential matching algorithm that first matches individuals on exact full name (including middle name) and date of birth. After the first round, matching proceeds by excluding middle names, and finally by fuzzy matching on full name. Per NCEDRC rules, the resulting data only includes unique matches, thus not allowing probabilistic matching or multiple matches and limiting our analysis to unique matches.

The arrest rates in the matched dataset are comparable to other papers using administrative data in Mecklenburg County, which supports the validity of the matching process (Deming, 2011; Billings, Deming & Rockoff, 2014; Billings, Deming & Ross, 2019). An arrest rate of 13% for males over ages 16-22 in the merged data is in line with the 10-16% found in this literature, which varies some based on the age ranges and years examined. We define “offenders” as students who were arrested by Charlotte-Mecklenburg Police Department (CMPD) during our sample period

[16] The Mecklenburg County Sheriff (MCS) tracks arrests across individuals using a unique identifier that is established with fingerprinting.

between the ages of 16 and 22.[17] While we observe the future criminal behavior of CMS students, regardless of whether they transfer or drop out of school, our arrests outcomes are limited to Mecklenburg County.[18]

Because we focus on criminal justice outcomes and arrests are overwhelmingly male (80%), we limit our main sample to 13,493 boys observed in CMS in these cohorts, including 4,166 who specified some choice in the lottery. We exclude 2,372 students from our sample of applicants because they were either guaranteed admission, specified their first choice as their neighborhood school, were in lottery groups for which either all or no applicants won their first choice, applied to a magnet program with subjective admission criteria that we do not observe, or did not meet the specified admission requirements for the magnet program to which they applied.[19]

The non-applicant sample includes male students observed in CMS in 5th grade in the 2005-2006 through 2007-2008 cohorts who did not apply to the school choice lottery for their sixth-grade year and were assigned to attend their neighborhood school for 6th grade (during the school assignment process in the prior year). Of the 9,327 students who did not specify any choice in the lottery, we remove students who did not have at least one other same grade male student in their CBG-School-Cohort group. The treatment variables described below is undefined for these students because they have no same CBG-School-Cohort peers. This process leaves a sample of

[17] Individuals arrested at age 16 or 17 were automatically charged in the adult criminal justice system in North Carolina until Dec 1, 2019.

[18] Mecklenburg county contains Charlotte and the surrounding, relatively affluent suburbs. Most arrests are concentrated in and around the urban center. Further, the surrounding counties are lower density, have lower crime rates, and do not have any urban centers near the boundary with Mecklenburg County. As a result, the majority of young adult arrestees in Mecklenburg County are observed in CMS schools. As highlighted in Billings, Deming & Rockoff (2014) as well as Billings, Deming & Ross (2019), over 90% of age comparable arrestees can be matched to public student records.

[19] The admission process to Arts programs and Leadership programs includes subjective unobservable criteria, e.g., interviews or auditions. Aside from applying to an undersubscribed program, students may also have been guaranteed admission through a sibling placement, or by participating in a magnet program in 5th grade for which they received an automatic placement into the associated middle school magnet program.

7,903 students. Finally, 12 students are dropped, because they would be the only student left from their CBG in the sample after imposing the restrictions above. After these sample restrictions, the analysis samples include 1,794 6^{th} grade lottery applicants and 7,891 non-applicants of which 5,079 had at least one applicant in their neighborhood and cohort. Our main sample has applicants from 27 neighborhood schools, 297 CBGs, and 448 neighborhood school by CBG combinations.

Given our sampling requirements, the sample of non-applicants exhibit arrest and incarceration frequencies that are similar to the district averages, and most non-applicants have an applicant in the same cohort and neighborhood. Appendix Table A1 provides descriptive statistics of criminality measures, as well as other student outcomes. The first three columns include descriptive statistics for all CMS students, the lottery applicant sample, and the non-applicant sample, respectively. We find that 13% of CMS students had at least one arrest and 8% were incarcerated at least one time from ages 16 to 22, which suggests substantial overall involvement with the criminal justice system. Table 1 includes descriptive statistics for the number of applications and total number of same neighborhood male students, as well as student level attributes including race, ethnicity, and test scores. Importantly, the average neighborhood group (CBG-School-Cohort) in the sample of non-applicants includes over 20 students, which reflects a scale commonly associated with likely peer groups, and on average over 2 of these potential peers are lottery applicants.

Non-applicants have more peers in their local neighborhood and cohort, are more likely to be white and have lower test scores, relative to lottery applicants. Table 1 also highlights the relatively high test scores of lottery applicants, and shows that applicants are more likely to be black rather than white or Hispanic. The higher achievement of lottery applicants is consistent with applicants being from households that have strong preferences for school quality, and suggests that there is selection into lotteries for oversubscribed schools. The pattern of more black students selecting into the lottery may relate to an overrepresentation of lower-performing schools in neighborhoods with a higher proportion of minority students. Given this pattern of applicants, later analysis will also examine how results vary by race and academic achievement in elementary school. The last column of Table 1 presents balance tests that we discuss in more detail below.

Methodology

In our main analysis, we estimate the impact of lottery winners on the outcomes of non-applicants by isolating random variation in lottery winners for a given CBG-school-cohort. In short, we do this by estimating the effect of the CBG-School-Cohort specific win share for the 6^{th} grade lottery on arrest and incarceration outcomes between ages 16 – 22 for non-applicants. To address sorting on expected lottery outcomes, we adopt a control function approach for isolating the random variation created by the lottery by conditioning on the expected group level win share, similar to Abdulkadiroglu et al. (2017).

In comparison to Abdulkadiroglu et al. (2017), we create a more complex function, i.e., the share of a student's neighborhood peers who won the lottery. We adopt a more general methodology developed by Borusyak and Hull (2021), which shows that for general functions of a vector of random shocks that interact with individual and potentially endogenous observables, an exogenous instrument can be derived by recentering. That is, by subtracting the expected value of the instrument over the distribution of potential shocks. They also derive conditions under which the instrument can be used directly with the expected value included as a control function. We follow the second approach for consistency with Abdulkadiroglu et al.'s (2017) control function approach to estimating the effect of winning a lottery, and later show that results are robust to recentering the share of neighborhood peers who won the lottery.

In our context, we observe the number of seats available in all oversubscribed lotteries, the specific rules determining lottery priority, and the number of applicants and winners within observed priorities. Therefore, unlike Abdulkadiroglu et al. (2017) and many of the examples in Borusyak and Hull (2021) who rely on simulations, we directly calculate lottery win probability estimates using observed win shares within priority group bins. We use these probabilities to construct expected win shares based on student-specific win probabilities. Similar to Deming (2011), we focus on the first-choice lottery wins for two main reasons. First, almost all oversubscribed schools are filled with first choice winners thus limiting random variation in later rounds.[20] Second, prior literature focused on first choice winners and thus adopting this same structure makes our results more comparable.[21] We later show that results are robust to including 2nd and 3rd choices in this process.

Specifically, we predict win probabilities using observed win shares for the student's first choice in their school choice application and priority group.[22] Let $\widehat{P}_{ibst}$ represent the predicted probability that student *i,* from census block group *b*, neighborhood school zone *s* and cohort *t* wins their first choice in the lottery. We construct group level expected win shares (w_{bst}) using

[20] Focusing on the first round applicant sample, 38% of applicants won in the first round, while only 9% won in the second round and 6% won in the third round.

[21] One additional limitation in using the 2nd and 3rd choices is that the expected likelihood of winning must be based on the ex-post probability of winning after the realization of the prior stage lottery, rather that the a-priori probabilities.

[22] The prediction is based on observed group level win shares. We group students by application choice, year of application, whether their neighborhood school or magnet continuation was a Title I Choie school, whether the student is economically disadvantaged, and whether the student scored at grade level in reading in 4th grade (with a separate group for missing 4th grade scores). These are characteristics that determine lottery priorities. We do not observe whether the student lived within the small radius (1/3 mile) of a first choice full magnet, which is also a priority factor. See Appendix for a more detailed discussion of estimating expected win shares.

the student-specific win probabilities: $w_{bst} = \frac{1}{(n_{bst} - 1)} \sum_{i=1}^{n_{bst}} \hat{P}_{ibst}$, where n_{bst} is the number of students in CBG *b*, neighborhood school zone *s* and cohort *t* whether or not they apply for the lottery, and the win probability, $\hat{P}_{ibst}$, is equal to zero for non-applicants.[23] Scaling by the total number of students captures the expected loss of neighborhood peers as a share total neighborhood peers. Now, w_{bst} represents the aggregated expected outcome of the lottery, while accounting for the lottery rules and neighborhood-school-cohort groupings. That is, the average lottery realization that we would expect if the randomization process were repeated many times. The variation that we use is based on the actual realization of the lottery, while conditioning on the expected outcome. The win share, i.e., the actual realization of the lottery, is our main independent variable and can be written in the following way: $Z_{bst} = \frac{1}{(n_{bst}-1)} \sum_{i=1}^{n_{bst}} 1[won\ lottery]_{ibst}$, which is the share of students in the same CBG, school zone, and cohort who won their first choice in the lottery, divided by the number of peers in the group.

Using variation in Z_{bst}, while conditioning on the expected win share, w_{bst}, means that identification is based on the lumpiness of aggregated lottery wins. Among groups with the same expected win share, some CBG-School-Cohort groups will have a higher-than-expected share of students win the lottery, while others will have a lower-than-expected share win the lottery. Because most lottery winners comply with their assignment to a non-neighborhood school, we consider the random shock of applicants winning the school choice lottery as imposing a treatment on the non-applicant students residing in same neighborhood and assigned to the same school.

Specifically, we estimate Equation 1 where y_{ibst} is a measure of arrests or incarceration of student i, residing in block group b and attendance zone *s*, and in cohort t. The main right hand

[23] When calculating win ratios, we use the total number of male students in a neighborhood and cohort minus one, so this represents the number of would-be, or potential, male peers for each male.

side variable is the fraction of same cohort students in this small neighborhood who win the lottery (Z_{bst}). Since both the number of lottery applicants and the likelihood of winning may correlate with school and neighborhood unobservables, we condition on the expected fraction of winners (w_{bst}) from the same small neighborhood (*b,s*) and cohort, pre-determined student attributes (X_{ibst}), as well as CBG (δ_b), school attendance zone (η_s) and cohort (γ_t) fixed effects. We estimate cluster-robust standard errors at the CBG by school attendance zone by cohort level, which is the level of variation in win shares. Later, we show that results are robust to both the fixed effect structure and clustering at the CBG by school attendance zone level.

$$y_{ibst} = \beta_1 Z_{bst} + \beta_2 W_{bst} + \beta_3 X_{ibst} + \delta_b + \eta_s + \gamma_t + \varepsilon_{ibst} \qquad (1)$$

Note that W_{bst} is a vector of controls including, most importantly, (1) the expected fraction of winners (w_{bst}), but also including additional lottery related controls for (2) 2nd and 3rd choice wins in the lottery, (3) other applications, and (4) other wins.[24] For (3) and (4), "other" refers to applications that are not in the applicant sample, such as sibling placements, applications to undersubscribed lotteries, and applications to programs with subjective placement criteria. The vector of additional non-applicant control variables, X_{ibst}, includes a set of dummy variables for race / ethnicity, a dummy variable for economic disadvantage, lagged (5th grade) math and reading scores, and an indicator for English language learner.

This model is identified by randomness in the school choice lottery process where students in local neighborhoods with the same expected number of lottery winners experience different

[24] While we include the additional features to control for other lottery-related outcomes, the results are robust to excluding these controls.

treatments because one neighborhood has a higher than expected win rate and the other has fewer wins than expected. To support this identification strategy, we conduct balance tests by regressing the individual student attributes on the actual and expected win shares for each student attribute *k*.

$$X_{ibst}^{k} = \beta_{1k} Z_{bst} + \beta_{2k} W_{bst} + \delta_{bk} + \eta_{sk} + \gamma_{tk} + \varepsilon_{ibstk} \tag{2}$$

Returning to our descriptive statistics in Table 1, the final column presents estimates of β_{1k} from Equation 2 for the student attributes identified in each row. All student attributes appear to be uncorrelated with the share of lottery winners, as we fail to reject the null hypothesis of β_{1k} = 0 in every case. In addition, we report the p-value from the joint significance test of all of the attributes from a single regression of the win share, Z_{bst}, on the student attributes (X_{ibst}), while conditioning on the other lottery related variables and fixed effects. The p-value from the test for joint significance is 0.65, which highlights the insignificant explanatory power of the student level attributes in explaining win shares. Therefore, we find no evidence that the portion of students winning the lottery is related to student attributes once we properly control for neighborhood-cohort expected win share.[25]

Finally, we estimate heterogeneous effects across students by interacting Z_{bst} with a student attribute indicator variable (X_{ibst}^{k}). The most important source of heterogeneity in our analyses, presented with the main results, is on the a priori predicted arrest risk for our non-applicant sample. For this analysis, we start by predicting the probability of any arrest between

[25] The sample also passes balance tests in models that omit school and neighborhood effects (p-value = 0.59). This emphasizes the importance of the control function in our model, and without the inclusion of the control function sample indicates relatively severe imbalance (p-value = 0.00).

ages 16-22.[26] Using the predicted arrest risk, we construct dummy variables for high- and low-risk. HR_{ibst} is equal to one for individuals with predicted risk in the top half of the distribution and zero otherwise. LR_{ibst} is equal to one for those with predicted risk below the median, and zero otherwise. We then use the following specification to estimate the differential effects by predicted level of arrest risk.

$$y_{ibst} = \beta_1^{HR} Z_{bst} \times HR_{isbt} + \beta_1^{LR} Z_{bst} \times LR_{isbt} + \beta_2^{HR} W_{bst} \times HR_{isbt} + \beta_2^{LR} W_{bst} \times LR_{isbt} + \beta_3 X_{ibst} + \delta_b + \eta_s + \gamma_t + \varepsilon_{ibst} \quad (3)$$

As shown in equation (3), when estimating heterogeneous effects, we also interact the entire vector W_{bst} with the same attribute. Now, β_1^{HR} and β_1^{LR} represent risk-specific coefficients, which allows us to test whether effects differ between individuals with low- and high-risk of future arrest. We include heterogeneity estimates on several other dimensions, including by quintile of arrest risk using the analogous specifications. In each case, we interact the win share, Z_{bst}, and the expected win share and the rest of the vector of lottery-related variables, W_{bst}, with the dummy variables intended to capture heterogeneity.

[26] We estimate individual crime risk based on a set of student attributes prior to 6th grade and use the model to construct a composite measure of a student's likelihood of being arrested from age 16-22. To do this, we estimate a logistic regression using an indicator for any arrests between ages 16-22 on individual, CBG, and neighborhood school level covariates. The crime risk estimation results are provided in Appendix Table A2. The individual level predictors are a set of race / ethnicity dummy variables, 5th grade math and reading scores, an indicator for economically disadvantaged, as well as a continuous variable for age. For missing values, we use the mean value and include a dummy variable for missing. We also include means of each race/ethnicity, 5th grade test scores, and economically disadvantaged at both the CBG and neighborhood school levels.

Results

Since one of the main assumptions in our methodology is that lottery winners are less likely to attend their neighborhood school, relative to lottery losers, we formally test this assumption for the sample of lottery applicants. Table 2 presents estimates from regressing several attendance and movement related outcomes on a dummy variable for winning their first choice in the lottery.[27] Panel A provides results for all lottery applicants and Panel B provides risk-specific estimates, using interactions with indicators for above and below median crime risk. The outcome in column one is an indicator for having won the lottery for any of their choice schools yielding an estimate of about 0.75, which is less than one because some losers of the first lottery may win their second or third choice. Winning any choice in the lottery provides some opportunity for students to leave their neighborhood middle school without residential movement. Columns 2 and 3 indicate that first choice lottery winners are almost 70 percentage points more likely to attend their first-choice school in 6th grade and 35 percentage points less likely to attend their neighborhood school in 6th grade, relative to lottery losers.

The outcomes in columns 4 - 6 are dummy variables for whether the student changed assigned neighborhood school (*Change NS*), a sign of residential movement, and dummies for exiting the district (*Exit*) by grade 6 or grade 9, respectively. These results indicate lower attrition by lottery winners relative to losers in terms of residential relocation or exit from the school district, and this non-compliance would tend to weaken our reduced form effects biasing us away

[27] Consistent with Cullen, Jacob, & Levitt (2006). Deming (2011) and Abdulkadiroglu et al. (2017), we condition on the probability of winning the lottery (which is based on application choice by cohort and the observable priority characteristic specific lottery win rates) and a set of individual controls including dummies for race/ethnicity, economic disadvantage, and English language learner status, as well as 4th grade math and reading scores. When lagged test scores are missing, we use the overall sample mean value and include a dummy for missing any test score.

from finding effects on non-applicants.[28] Columns 7 and 8 provide evidence that lottery winners attend more desirable schools, ones in which students had higher test scores and lower predicted arrest rates in the prior year. Panel B indicates limited heterogeneity in these outcomes by crime risk with the exception of lower risk students generating most of the difference between winners and losers for residential movement and district exit.

Since our main analysis focuses on the impact of lottery winners on non-applicant peers, one may be concerned that an unexpectedly high number of lottery winners leaving the neighborhood school may impact the decision of non-applicants to attend their neighborhood school or lead to exit from the district. We formally test the effects of win shares (Z_{bst}) on changing neighborhood schools and exiting CMS by estimating equation 1 for our sample of non-applicants in Appendix Table A3. We find no evidence that non-applicants respond to lottery winning peers by exiting CMS both overall as well as for high and low arrest risk boys. However, we find modest evidence of effects on residential mobility. In the pooled sample, an increase in win share from no winners to the average proportion of winners in a neighborhood leads to a statistically significant 5.0% decrease in the probability of changing neighborhood schools.[29] Unlike exit, which may imply sample attrition, residential relocation is simply one aspect of treatment. Losing positive peers appears to increase the probability that non-applicants remain in the neighborhood schools and thus limits the concern that non-applicants with more peer lottery winners leave the neighborhood at a higher rate which could attenuate estimates.

[28] Students who lost their first choice are placed on a waitlist. Students may also be admitted through the first quarter of the school year as seats become available in their first-choice option. As highlighted in Bibler & Billings (2020), the presence of a negative effect on residential moves (Change NS) indicates that losers are moving away in response to the lottery results.

[29] Specifically, we find that a change from no winners to the average proportion of winners in a neighborhood (4.0% of same grade male students), indicates a 0.012 (-0.298*0.040) decrease in the probability of changing neighborhood schools, which is a 5.0% (0.012/0.239) decrease over our estimation sample average.

Table 3 provides our main estimates for the impact of lottery winners on those left behind (non-applicants). The first three columns present extensive margin outcomes for any arrest, arrest for violent crime and incarceration by age 22, and the remaining columns present the analogous intensive margin outcomes. Panel A includes estimates from the pooled sample of high- and low-risk non-applicants. Results in panel A are positive and significant for all arrest and violent arrest outcomes, and positive but insignificant for incarceration. To interpret coefficients, we use the change from no winners to the average proportion of winners in a neighborhood (4.0% of same grade male students). We find a 0.01 (0.260*0.040) increase in the probability of arrest and a 0.06 (1.454*0.04) increase in the number of arrests as a young adult, which represent increases of 8.3% (0.01/0.12) and 13% (0.06/0.46) over the estimation sample average for any arrests and number of arrests, respectively. Turning to Panel B, estimates for students with below median arrest risk are more precise and considerably larger. Among the low-risk non-applicants, a change from no winners to an average win share in their neighborhood generates a 1.3 percentage point increase in the probability of arrest, 0.024 more violent arrests and 1.2 more days incarcerated. Given lower rates of arrest and incarceration for this subsample, the estimates imply increases of 24.7%, 60.1% and 121% of the respective means.[30] All estimates for low-risk non-applicants are significant at the 1% level, except for the estimate for any violent arrest which is significant at the 5% level.

Finding strong effects among low-risk non-applicants is consistent with the positive selection of lottery applicants in terms of achievement, and negative selection in unobserved crime risk, which suggests that applicants are more similar in attributes to low-risk non-applicants. The presence of stronger peer effects between individuals with similar attributes is well established in

[30] We provide additional results in Appendix Tables A4 and A5 to show that results are consistent when using a variety of non-linear specifications that better address the intensive margin outcomes which are infrequent and, in the case of days incarcerated, can generate large values.

the literature (Billings, Deming & Ross, 2019; Black, Devereux & Salvanes, 2013; Lavy & Schlosser, 2011). The positive selection into the lottery is evident in Table 1, as applicants have higher average test scores in elementary school. Appendix Table A6 provides additional evidence for the positive selection of lottery applicants by highlighting that lottery losers who attend their home school score substantially higher than non-applicants on middle school tests, even after conditioning on lagged test scores and school fixed effects.

Robustness and Validation

We include additional results to highlight the validity of our identification strategy. First, we show that results are robust to different combinations of neighborhood fixed effects. Appendix Table A7 includes estimates for the effect of win shares on the probability of any arrest and days incarcerated from six different specifications. Columns 7 and 8 of Appendix Table A7 include school attendance zone by neighborhood fixed effects, which is equivalent to a cohort variation model across the three cohorts of students within a given geography. The cohort variation model provides almost identical results to our main models. Robustness to alternative vectors of fixed effects confirms the effectiveness of the Borusyak and Hull (2021) in eliminating endogeneity from treatment measure. Appendix Table A7 also includes a second set of standard errors from clustering at the block group by attendance zone level in square brackets. We observe only modest changes in standard errors especially in the low-risk subsample.

Next, we implement a falsification test for our main specification in which, for each student, we assign the win share and the expected win share from the cohort in the following year in the same small neighborhood. The results are displayed in Table 4. This test is effectively a replication of the main results in Table 3 where we assign treatment based on the one year later rising 6th graders in the same neighborhood. The maintained hypothesis is that unexpected wins for a future cohort of students in the school should not affect the outcomes of the current cohort. All estimates are statistically insignificant and, in most cases, substantially smaller in magnitude, relative to the main results. As a second falsification test, we re-estimate our main results in Table 3 using the win share of girl applicants to estimate the effect of same-neighborhood girls winning the lottery on non-applicant boys. To do this, we replace the win share and expected win share with the analogous values based on the sample of girls. As shown in Table A8, all coefficients are statistically insignificant and almost always smaller in magnitude.

We incorporate a number of specifications to test the sensitivity to how we model the lottery itself. Appendix Table A9 provides results for a version of the main results in Table 3 when including all rounds of the lottery rather than using only the first round. The estimated effects of a higher win share are similar to the main results in Table 3. Second, we re-estimate models using the re-centering approach used in Borusyak and Hull (2021), rather than including the expectation as a control function. Results using the re-centering approach is provided in Appendix Table A10 and are almost identical to Table 3.

Suspensions, Absences, and Peer Effects by Age

Given the strong evidence that unexpected lottery wins increase arrests and incarcerations among non-applicants in the same neighborhood, primarily for low crime risk individuals, we test whether these results extend to non-criminal justice outcomes and whether the behavioral effects emerge in the short-term, in addition to the long-term outcomes presented in Table 3.

Table 5 provides estimated effects of the win share on non-applicant school absences and suspensions. These estimates are based on Equation 1 using five outcomes related to absences and suspensions in the years following the lottery. The results suggest that low-risk non-applicants who are left behind by lottery winners experience substantially worse outcomes in these school-related measures, which corresponds to the estimated effects on arrest and incarceration. Panels A and B include up to 5 years post-lottery corresponding to grades 6 through 10. Panel A includes estimates for the pooled sample, for which only the estimate for absences is statistically significant at the 5% level.

Panel B presents separate estimates for the students with above and below median arrest risk. As with arrests and incarceration, we observe significant effects across the board for the low-risk subsample with substantial increases in the number of days absent and the number of days suspended. The suspension effects are concentrated on out of school suspensions. Among the low-risk non-applicants, we estimate that an increase from no winners in their neighborhood (CBG-school-cohort group) to the sample average win share increases the number of absences by 0.24 (0.040*5.93) per year and days in out of school suspension by 0.14 (0.040*3.599) per year from 6th through 10th grade. Both estimates are statistically significant at the 5% or 1% level and they represent increases over the low-risk sample means of 4% (0.24/5.87) in absences and 23% (0.14/0.61) in days in out of school suspension.

Finally, Panel C presents estimated effects on absences and suspensions for the low-risk subsample when restricting to observations during middle school, i.e., the three years post-lottery. We observe the same patterns in absences and suspensions for below median risk middle school students. The negative behavioral effects of losing peers begin in middle school soon after the peers are assigned to a different school through the lottery. However, we find small and mostly insignificant effects on test scores and high school graduation, as shown in Appendix Table A11.[31]

[31] The estimates for low-risk non-applicants in Table A11 imply that a change from no winners in the same neighborhood to the sample average win share reduces math scores by 0.016 (0.04*-0.388) standard deviations,

The main estimates for arrests and incarceration use outcomes aggregated over ages 16 to 22. To test how early the arrest and incarceration effects appear and how they evolve over time, Table 6 includes the estimated effects of the win share on low-risk non-applicant arrest and incarceration outcomes separately over ages 16 to 18 and ages 19 to 22. Coefficients indicate that going from no winners to the sample average win share increases the probability of any arrest at ages 16-18 by 25% ((0.040*0.173)/0.028)) over the sample average among low-risk individuals. Despite the higher incidence of arrest at ages 19-22, we estimate an even larger effect on the probability of any arrest, 38% ((0.040*0.352)/0.037)), at age 19-22. We find a similar pattern across the other arrest and incarceration outcomes in Table 6, which suggests that effects persist into early adulthood after affected students have left school.

Not only do we find a persistent behavioral response that continues into adulthood, the absolute effects of losing lottery applicants on arrest and incarceration are substantially larger for post-high school ages across all outcomes. Figure 1 displays estimated effects on several behavioral outcomes at different ages for the sample of low crime risk individuals. Each point corresponds to an estimated effect of going from zero to the sample average win share on some outcome as a percentage of the average outcome. We include estimated effects on suspensions and absences over three approximate age ranges following the lottery, as well as arrests, violent arrests, and days incarcerated for three different age ranges. This figure highlights a broader trend in the results, which is that the magnitude of the estimates, in percentage terms, grows significantly in young adulthood. The estimated effects on suspensions are relatively consistent in magnitude through age 16 and comparable to the effects on arrest and incarceration at ages 16-18. However, we observe substantially larger effects in adulthood as captured by arrests and incarceration. This

reading scores by 0.009 (0.04*-0.221) standard deviations and the probability of on-time graduation by 0.3 (0.04*-0.067) percentage points.

age trend provides evidence consistent with several papers in the education literature (Deming, 2009; Jacob, Lefgren, and Sims, 2010; Carrell and West, 2010; Chetty et al., 2014) that positive educational treatments (e.g., Head Start, high-quality teachers and in our case peers) can have benefits that fade-out initially but grow in young adulthood.

Heterogeneity

Finally, we recognize that two key correlates of arrest rates are race and academic performance. Therefore, we estimate heterogenous effects for young men using race/ethnicity dummy variable interactions with win share and lottery controls including expected win share, as well as dummies for above and below median test scores interacted with win share and the lottery controls. The results are shown in Appendix Table A12. Panel A includes results for race and ethnicity. The results for white students closely parallel the results for students with low risk of arrest in significance and magnitude. While the estimates for black students display a similar pattern and three of the estimates are statistically significant at the 10% level or better, the corresponding standard errors are generally larger.

Panel B of Table A12 includes separate estimates for students with above and below median 5^{th} grade test scores. In percentage terms, effect sizes across all outcomes are larger for above median test score students. All extensive margin effects for students with above median test scores are positive and statistically significant at the 5% level or better, and imply an approximate 25 to 28% increase in the probability of arrest and incarceration over the high test score sample mean. While the below median test score sample extensive margin estimates are all statistically insignificant, they suggest a 2 to 5% increase in the probability of arrest and incarceration over the below median sample means.

Intensive margin results suggest that the change from no winners to an average win share increases arrests and incarceration for above median test score students by over 20% and for below median test scores students by about 10%. Unlike the extensive margin effects for below median test scores students, the estimated effects on all arrests and violent arrests are significant at the 10 and 5% levels, respectively. Appendix Table A13 expands the analysis in Panel A of Table A12 by examining effects by race and risk of arrest. Estimated effects for white students are positive an frequently statistically significant, irrespective of their risk of arrest. In fact, in absolute terms, effects are larger for above median risk of arrest white students, likely due to the higher base of arrests and incarceration. We also find statistically significant effects for Hispanic and other race students with below median risk of arrest, but effects for black students are concentrated among high-risk students.

To this point, we have focused only on male students. A natural question is whether the same dynamics occur for female students? Appendix Table A14 tests for our main results on a sample of female students. Across all outcomes we find smaller and insignificant results. These same dynamics of greater criminal justice contact for non-applicants from the loss of lottery winners in the neighborhood does not appear to occur among female students, perhaps due to the nature of peer effects which provide stronger influences on male criminality. This result is consistent with Lavy & Sand (2019) where the loss of a peer entering middle school only negatively impacts misbehavior in school for boys and has a larger negative impact on school and social satisfaction for boys than for girls.

Mechanisms

Given the pattern of results highlighting larger effects on low-risk non-applicants that increase in magnitude with age, as well as the positive selection on academic performance by applicants, some potential mechanisms come forward. While we cannot rule out all possible channels, two main mechanisms are consistent with our results so far: 1) a decline in neighborhood school quality and 2) loss of positive peer influences when same-neighborhood students win the lottery. One would expect a limited impact of neighborhood level unexpected lottery wins on the broader neighborhood school environment, but it is plausible if lottery applicants provide positive inputs (e.g., student performance and behavior, as well as parental involvement) into the school.

To explore impacts on neighborhood schools, Figure 2 includes a series of figures that show the correlation between school-cohort win shares with school attributes 3 years post-lottery: average test scores, proportion of student who are white, and proportion of students who are economically disadvantaged. Each dot in the figure represents the win share and characteristic for one neighborhood school-cohort observation. The figure also presents the estimated slope with the standard error in parentheses in the upper right-hand corner. The panels on the left-hand side of Figure 2 measure actual win shares on the x-axis, which shows that win share has a negative and statistically significant correlation with average test scores and portion of students that are white, and a positive and significant correlation with the proportion of economically disadvantaged students. All characteristics are measured three years post-lottery. The correlations are consistent with higher lottery participation from neighborhood schools with lower test scores and a higher proportion of economically disadvantaged students. The right-hand side of Figure 2 includes the analogous correlation between each characteristic and wins minus expected wins as a share of cohort size. All three correlations change sign and become statistically insignificant when controlling for the expected win share. These results suggest that observable characteristics for neighborhood schools are not changing in response to unexpected lottery wins and provides additional evidence that controlling for expected win share addresses selection into the lottery.

To further demonstrate that the effect of win shares on non-applicant outcomes is not driven by school-level responses to lottery wins, we estimate the effect of win shares on the probability of arrest and days incarcerated from a specification with school by cohort fixed effects. By including school by cohort fixed effects, we compare students in the same school and cohort with different neighborhood group win shares. The estimates displayed in columns 9 and 10 of Appendix Table A7 are very similar to the main results in Table 3, which limits any concerns that results are driven by school level changes due to lottery wins. These findings above argue against changes in school quality as a mechanism behind our findings.

Further, several pieces of evidence are consistent with a peer effects channel. In particular, the increase in arrest and incarcerations among non-applicants are primarily driven by low-risk non-applicants, which corresponds to the positive selection of applicants. Additionally, the response is primarily for behavioral outcomes, i.e., arrests, incarcerations, and middle school absences and suspensions, rather than academic outcomes, which is consistent with research on school-based peer effects (Billings & Hoekstra, 2022, Kim & Fletcher, 2018, Billings, Deming & Rockoff, 2014).

To further investigate the peer effects mechanism, we test whether the response of non-applicants is related to the proportion of same-neighborhood lottery winners who also attended their same elementary school and present the results in Appendix Table A15. In panel A, we interact the main independent variable (win share) with the share of winners that attended the same elementary school as the focal student and interact our control function with the share of same elementary school applicants. Panel B provides results from a specification with more granular controls by including school attendance zone by CBG by cohort fixed effects. Results suggest that effects for low-risk non-applicants are stronger for those who attended elementary school with a higher share of lottery winners, which is consistent with a peer mechanism. While results are noisy, the interaction estimates are consistently positive and large. The implied effects in panel A are between 40 and 140 percent larger for non-applicants losing same elementary school neighbors, relative to losing peers who attended other elementary schools. The estimated differences are even larger in Panel B when including school attendance zone by CBG by cohort fixed effects.

Second, we split our sample into groups with a large versus small number of same cohort peers in the neighborhood. A small peer group implies closer social connections and potentially a larger negative effect of losing peers.[32] These results are shown in Appendix Table A16, and as hypothesized, larger effects are observed for non-applicants in neighborhoods with fewer peers. Coupling these two sets of results supports the idea that disruption of peer relationships plays a primary role in explaining why unexpected lottery wins increase adult criminal outcomes for neighborhood kids that did not apply for the lottery.

Finally, we examine whether the effects on those left behind are concentrated among same race relationships using an additional interaction for share of wins among lottery applicants of the same race. Appendix Table A17 tests whether the number of same race students winning the lottery as a share of total same race, same cohort students in the neighborhood has an additional effect on arrests and incarcerations of non-applicants. The estimates are statistically insignificant, which suggests that the effects on non-applicants of losing peers are comparable whether the winning peers are of the same race or not. These results differ from Billings, Deming & Ross (2019) who

[32] The sample of non-applicants is split evenly between large and small groups sizes with the large group averaging 34 students and 3 applicants while the small group averages 8 students and 1 applicant.

found strong same race effects in predicting criminal partnership. These differences are likely driven by the fact that Billings, Deming & Ross (2019) focuses on students at high risk of arrest, but lottery applicants tend to be positive selected and the estimated effects of losing an applicant peer are concentrated primarily among students who are at low risk of arrest.[33]

Lottery Winners and Adult Crime

To assess the impact of school choice on adult crime, we examine the direct effect of winning the first choice in the lottery on the arrest and incarceration outcomes for the lottery winners themselves. In this section, we largely replicate existing work by estimating the direct effect of winning a school choice lottery on the adult crime of lottery winners. Our analysis includes the sample of lottery applicants, focusing on comparing winners and losers. We initially mirror the sample split in the previous analysis by using above versus below median arrest risk, but also provide splits by quintile of arrest risk consistent with Deming (2011).

We follow Abdulkadiroglu et al. (2017) and calculate a control function based on the expected probability of an applicant winning the lottery using application program by priority characteristic specific lottery win rates. This differs from the Cullen, Jacob, & Levitt's (2006) and Deming's (2011) papers on the effect on school choice lotteries on winner's future arrests which include fixed effects intended to subsume the likelihood of winning instead of calculating the control function.[34] We then model arrests and incarceration (y_{il}) as a function of whether the applicant won the lottery, estimated win probability (P_{il},), and additional controls (X_i) for race

[33] The complicated nature of peer dynamics is well demonstrated in Carrell, Sacerdote & West (2013) where randomly assigning half of entering freshmen in the US Air Force Academy to peer groups that would maximize the academic performance of the lowest ability students led to unexpected negative impacts on these students. This result was attributed to subgroup dynamics as well as the mix of peers.

[34] We include additional estimates for the effects of winning the lottery on probability of arrest and days incarcerated using alternate specifications in Table A23.

and ethnicity, economic disadvantage, 4th grade math and reading scores, and ELL status.[35] The specification is displayed in Equation 4.

$$y_{il} = \gamma_1 W_{il} + \gamma_2 X_i + P_{il} + \varepsilon_{il} \tag{4}$$

Where W_{il} is a dummy variable equal to one if the applicant won their first choice in the lottery, and zero otherwise. Now, γ_1, describes the difference in outcomes, y_{il}, between the lottery applicants who won and lost their first choice in the lottery. After conditioning on win probability, we argue that this comparison describes the causal effect of the lottery outcomes (for 6th grade lottery) on arrest and incarceration outcomes measured from ages 16 to 22.

Table 7 provides descriptive statistics for the applicant sample and standard balance tests to show that lottery winners and losers are comparable on observables in this sample. The first two columns contain means and standard deviations among lottery winners and losers. Balance tests are included in columns 3 and 4. Each balance test is from a linear regression of the dummy variable for winning the lottery on the student level covariates. We report the estimated coefficients on the covariates in the table. Column 3 reports the coefficients without conditioning on the win probability. The estimates in column 4 are conditional on win probability. Consistent with prior work, we find that winning the lottery is uncorrelated with student characteristics, after controlling for lottery priorities in the CMS school choice lottery (Deming, 2011; Deming et al., 2014; Bibler & Billings, 2020).

[35] For students who are missing a 4th grade test score, we use the mean value and include a dummy for missing any test scores. 7% of the applicant sample is missing at least one 4th grade test score. We use 4th grade test scores for this part of the analysis because the 5th grade testing occurs after the lottery.

The main results from estimating Equation 4 are included in Table 8.[36] While the estimated effects of winning the lottery on the arrest and incarceration outcomes are all negative, suggesting that lottery winners benefit, most estimates in the pooled sample are statistically insignificant. However, the pooled sample results in Panel A show that lottery winners are 2.7 percentage points less likely to be incarcerated anytime between ages 16 to 22, which is a 45% decrease off the mean incarceration rate for lottery applicants. In the sample of low-risk lottery applicants, we find significant declines in both the probability of any violent arrests and any incarceration. However, some differences between the above and below median sample estimates also arise in the standard errors. Results for the low-risk subsample are more precisely estimated and statistically significant, while estimates for the high-risk subsample are often sizable and negative, but much noisier.[37]

Like Deming (2011), we also provide the estimated effects of winning the lottery on crime and incarceration for those with the highest ex-ante crime risk (top quintile). Appendix Table A20 provides results for 5 groups of lottery applicants based on quintiles of ex-ante arrest risk. Similar to Deming (2011), we find the largest declines in the highest quintile of arrest risk on the intensive margin. However, most estimates are statistically insignificant. In addition, we estimate statistically significant declines on the extensive and intensive margin for violent arrests in the second quintile.[38]

Appendix Table A22 includes estimated effects of winning the lottery on test scores, absences, suspensions, and graduation for lottery applicants. In contrast to the estimated effects on non-applicants for the same outcomes in Tables 5 and A11, we do not find any statistically

[36] The analogous estimates in the sample of female students are included in Appendix Table A18. All but one estimate for the effects of winning the lottery on criminal justice outcomes among girls are statistically insignificant.

[37] Estimates incorporating all lottery choices in Appendix Table A19 are muted, which is consistent with the fact that the most desirable and oversubscribed schools are often no longer available in rounds two and three.

[38] Appendix Table A21 presents similar results for non-applicants exposed to a lottery winner by risk quintile. Statistically significant effects are mostly concentrated in the 1st through the 3rd risk quintile.

significant effects of winning the lottery on the winner's number of absences or suspensions in the pooled sample. However, winning the lottery increases reading test scores by 0.04 standard deviations in the pooled sample, which is marginally statistically significant. The estimated improvements in test scores are larger in magnitude for the high-risk sample, but the estimates are statistically insignificant.

Net Benefit Analysis

The main estimates for the effects of winning the lottery on applicants and their peers suggest that some lottery winners benefit from lower arrest and incarceration rates as adults, while some non-applicants experience increased adult arrests and incarcerations due to the loss of lottery winning peers. Given the heterogeneity in both benefits and costs, we use a more formal accounting of effects by different levels of crime risk to assess the net impact of these two conflicting effects of the constrained school choice system on criminal justice outcomes in Charlotte-Mecklenburg county. To facilitate this analysis, two sets of results are provided in Table 9. Panel A displays aggregated arrests and days incarcerated for lottery winners and their peers left behind by above and below median crime risk, and Panel B includes the analogous aggregations for each risk quintile. The aggregation produces estimates for the total benefits to lottery winners calculated as the number of arrests and days incarcerated avoided, and costs to non-applicants in terms of increased arrests and days incarcerated.

To construct the aggregate effects of winning the lottery for each risk group, we use the risk-specific expected number of wins and risk-specific estimates of the effect of winning the lottery on each intensive margin outcome. For example, summing the predicted win probabilities across all high- and low-risk applicants suggests that the expected number of high- and low-risk winners is 278 and 373, respectively, which are displayed under *E[Wins]* in the first column of Panel A. From Table 8, the estimated effects of winning the lottery for high- and low-risk applicants on the number of all arrests are -0.35 and -0.01, respectively. Combining the risk-specific expected number of wins and estimates suggests that the lottery decreased all arrests by about 97 and 4 among, high- and low-risk applicants, respectively, which we show under the *All Arrests* column near the top of Panel A. Below the risk-specific effects, we sum up to calculate the *Total Effect on Applicants* of -101. We compute the results for each of the intensive margin outcomes in a similar manner and aggregate in two ways: all coefficients, regardless of statistical significance, and including only the statistically significant results. Benefits/costs arising only from significant estimates are indicated by bold numbers, which results in an estimated effect of 3,191 fewer days incarcerated among lottery applicants.

In the second half of Panel A, we estimate the corresponding aggregated change in the number of arrests and days incarcerated among the non-applicants with peers who applied to exit their school through the lottery. To calculate the aggregated effects, we use the group-specific expected win share along with the risk-specific estimated effects. For example, we sum the product of the expected win share (w_{bst}) and the estimated effect of peers winning the lottery on arrests for high-risk non-applicants, 1.42 from Table 3, across all high-risk students in the non-applicant sample. We find an aggregated increase in arrests of high-risk non-applicants of 252, and an estimated increase in arrests of 212 among the low-risk non-applicants. We replicate this for *Violent Arrests* and *Days Incarcerated* in the following columns. The analogous calculations for risk-quintile specific estimates and aggregates for each outcome are included in Panel B. Aggregating across applicants and non-applicants produces a total aggregated effect.

In all cases, whether using above and below median arrest risk, or using risk quintiles, and using all estimates or only significant estimates, the increases in arrests and days incarcerated of students residing in the local neighborhood of lottery winners substantially exceeds the declines in the analogous outcomes of the lottery winners themselves. The more conservative results, using above and below median crime risk, suggest that aggregate arrests increased by 362 when considering all estimates or 211 when aggregating significant coefficients only. Summing across the affected populations, we estimate 5,302 additional incarceration days when aggregating across all coefficients or 940 additional days incarcerated when aggregating significant coefficients only.

Conclusions

We estimate the effects of having peer lottery winners on arrest and incarceration using three cohorts of 5th grade students in Charlotte, NC. We find large negative effects of neighborhood peers winning the lottery on 5th grade boys who do not apply to the school choice lottery, including an increased likelihood of arrest, more total arrests, and increased days of incarceration between ages 16 and 22. The magnitude of the response is increasing in age with larger effects between age 19 and 22. The negative effects are concentrated among students with below median arrest risk., which appears consistent with effects for students who would have been likely to interact socially with the positively selected lottery applicants. Using measures of attendance and suspensions, we also observe evidence of negative behavioral effects in middle school among students at low risk of future arrest.

We examine two potential mechanisms for these effects: the effect of removing positively selected, same neighborhood, middle school peers, and broader effects of the lottery on middle school characteristics. We demonstrate that applicants have higher test scores in elementary school, and by examining lottery losers who remain at the assigned middle school show that they are positively selected experiencing test score gains in middle school conditional on elementary school test scores. We find that effects are larger when a non-applicant attended the same elementary school as the lottery school applicants, and find that effects are larger when the population of same neighborhood peers is smaller in size suggesting closer social relationships. On the other hand, we find no evidence of school quality effects as there is no relationship between lottery wins and middle school characteristics. In addition, estimates are robust to the inclusion of middle school attendance zone by cohort fixed effects to capture unobserved middle school changes across cohorts, further demonstrating that the results are not driven by observed or latent changes in school characteristics.

Estimating the direct effects on criminality of the applicants themselves, on the other hand, suggest that lottery winners experience lower arrest and incarceration rates than those who lost the same lotteries. Unfortunately, the benefits to lottery winners are not large enough to offset the negative effects on students left behind. Aggregating the costs and benefits of these two effects in terms of adult crime suggests that school choice increased total arrests and incarceration between ages 16 and 22 in our sample. This finding empirically validates the theoretical prediction of Barseghyan et al. (2019), that school choice can be welfare decreasing in the presence of strong peer preferences.

These results raise important questions about the overall impact of school choice programs in the key domain where school choice has consistently been shown to have positive impacts on lottery winners. The potential costs of school choice that arise from students who are left behind represents a significant cost that might be considered when deciding whether, or how, to expand school choice opportunities. Our results have important implications for public school systems as the popularity of school choice grows (Brunner et al., 2012; Tuttle, Gleason & Clark, 2012) and scholars continue to study and refine the lottery mechanisms used to implement school choice (Abdulkadiroğlu & Sönmez, 2003; Abdulkadiroglu & Andersson, 2022; Pathak, 2017).

References

Atila Abdulkadiroğlu, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane and Parag A. Pathak. "Accountability and Flexibility in Public Schools Evidence from Boston's Charters and Pilots," *The Quarterly Journal of Economics*, 126(2) (2011), 699-748

Abdulkadiroglu, Atila, Joshua D. Angrist, Yusuke Narita, and Parag A. Pathak. 2017. "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation," Econometrica 85:1373–1432.

Atila Abdulkadiroğlu, Parag A. Pathak, Christopher R. Walters. "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economics*, 10(1) (2018), 175-206.

Abdulkadiroglu, Atila, and Tommy Andersson. "School Choice," National Bureau of Economic Research Working Paper No. w29822 (2022).

Abdulkadiroğlu, Atila, and Tayfun Sönmez. "School Choice: A Mechanism Design Approach," *American Economic Review*, 93(3) (2003), 729-747.

Altonji, Joseph G., Ching-I. Huang, and Christopher R. Taber. "Estimating the cream skimming effect of school choice," *Journal of Political Economy* 123.2 (2015): 266-324.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. "Explaining the Effectiveness of Charter Schools?" *American Economic Journal: Applied Economics*, 5(4) (2013), 1-27

Bacher-Hicks, Andrew, Stephen B. Billings, David J. Deming. "The School to Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime," NBER Working Paper No. w26257 (2019).

Barseghyan, Levon, Damon Clark, and Stephen Coate. "Peer Preferences, School Competition, and the Effects of Public School Choice," *American Economic Journal: Economic Policy*, 11(4) (2019), 124-58.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen. "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections," *The Quarterly Journal of Economics*, 124 (2009), 105-147.

Bibler, Andrew, and Stephen B. Billings. "Win or Lose: Residential Sorting After a School Choice Lottery," *Review of Economics and Statistics*, 102(3) (2020), 457-472.

Bifulco, R. J. Fletcher, S. Oh, and S.L. Ross. "Do High School Peers Have Persistent Effects on College Attainment and Other Life Outcomes?" *Labour Economics*, 29 (2014), 83-90.

Bifulco, R. J. Fletcher and S.L. Ross. "The effect of classmate characteristics on individual outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy*, 3 (2011), 25-53.

Bifulco, R. H. Ladd and S.L. Ross. "The Effect of Public School Choice on Those Left Behind: Evidence from Durham, NC," *Peabody Journal of Education*, 84 (2009), 130-149.

Bifulco, R. H. Ladd and S.L. Ross. "Public School Choice and Integration: Evidence from Durham, NC," *Social Science Research*, 38 (2009), 78-85.

Billings, Stephen, Eric Brunner, and Stephen L Ross. "The Housing and Educational Consequences of the School Choice Provisions of NCLB: Evidence from Charlotte, NC," *Review of Economics and Statistics*, 100 (2018), 65-77.

Billings, Stephen B., David J. Deming, and Jonah Rockoff. "School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg," *The Quarterly Journal of Economics*, 129(1) (2014), 435-476.

Billings, Stephen B., David J. Deming, and Stephen L. Ross. "Partners in Crime," *American Economic Journal: Applied Economics*, 11(1) (2019), 126-150.

Billings, Stephen, and Mark Hoekstra. "The Effect of School and Neighborhood Peers on Achievement, Misbehavior, and Adult Crime," *Journal of Labor Economics*, Forthcoming (2022).

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. "Under pressure? The Effect of Peers on Outcomes of Young Adults," *Journal of Labor Economics*, 31(1) (2013), 119-153.

Borusyak, Kirill, and Peter Hull. "Non-random exposure to exogenous shocks: Theory and applications," In Press. *Econometrica*.

Brunner, Eric J., Sung-Woo Cho, and Randall Reback, "Mobility, Housing Markets, and Schools: Estimating the Effects of Inter-District Choice Programs," *Journal of Public Economics*, 96 (2012), 604–614.

Simon Burgess, Ellen Greaves, Anna Vignoles, Deborah Wilson. "What Parents Want: School Preferences and School Choice," *The Economic Journal*, 125(587) (2015), 1262-1289.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka. "The Long-run Effects of Disruptive Peers," *American Economic Review*, 108(11) (2018), 3377-3415.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West. "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica* 81.3 (2013): 855-882.

Carrell, Scott E. and James E. West. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy*, 118(3) (2010), 409–32.

Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt. "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, 74(5) (2006), 1191-1230.

Chakrabarti, R. "Vouchers, Public School Response and the Role of Incentives: Evidence from Florida," *Economic Inquiry*, 51(1) (2013), 500-526.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 104(9) (2014), 2633-79.

Cookson, Peter W., and Barbara Schneider. "Why School Choice?" Transforming Schools. Routledge, 2014. 557-577.

Damm, Anna Piil, and Cedric Gorinas. "Prison as a Criminal School: Peer Effects and Criminal Learning behind Bars," *Journal of Law and Economics*, 63(1) (2020), 149-180

Deming, David. "Early Childhood Intervention and Life-Cycle Development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 1(3) (2009), 111–34.

Deming, David J. "Better Schools, Less Crime?" *The Quarterly Journal of Economics*, 126(4) (2011), 2063-2115.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. "School Choice, School Quality, and Postsecondary Attainment," *American Economic Review*, 104(3) (2014), 991-1013.

Dills, Angela K and Rey Hernandez-Julian, "More choice, less crime," Education Finance and Policy, 6 (2) (2011), 246–266.

Dobbie, Will and Roland G. Fryer Jr. "The Medium-Term Impacts of High-Achieving Charter Schools," *Journal of Political Economy*, 123 (5) (2015).

Fabelo, Tony, Michael D. Thompson, Martha Plotkin, Dottie Carmichael, Miner P. Marchbanks, and Eric A. Booth. "Breaking Schools' Rules: A Statewide Study of How School Discipline Relates to Students' Success and Juvenile Justice Involvement," *New York: Council of State Governments Justice Center* (2011).

Fletcher, Jason M., Stephen L. Ross, and Yuxiu Zhang. "The Consequences of Friendship: Evidence on the Effect of Social Relationships in School on Academic Achievement," *Journal of Urban Economics,* 116 (2020), 103241.

Gilraine, Michael, Uros Petronijevic, and John D. Singleton. "Horizontal Differentiation and the Policy Effect of Charter Schools," *American Economic Journal: Economic Policy*, 13(3) (2021), 239-76.

Glaeser, Edward L., Bruce Sacerdote, and José A. Scheinkman. "Crime and Social Interactions," *The Quarterly Journal of Economics*, 111(2) (1996), 507–548.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. "Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery," NBER Working Paper No. 12145 (2006).

Hastings, Justine S., Jeffrey M. Weinstein. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *The Quarterly Journal of Economics*, 123(4) (2008), 1373–1414

Hsieh, Chang-Tai , Miguel Urquiola. "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program," *Journal of Public Economics*, 90 (2006), 1477 – 1503

Hess, Chris. "Residential Segregation by Race and Ethnicity and the Changing Geography of Neighborhood Poverty," *Spatial Demography*, 9(1) (2021), 57-106.

Hochschild, Jennifer L., and Nathan Scovronick. *American dream and public schools*. Oxford University Press, 2003.

Imberman, Scott. "The effect of charter schools on achievement and behavior of public school students," *Journal of Public Economics*, 95(7) (2011), 850-863

Jacob, Brian A., Lars Lefgren, and David P. Sims. “The Persistence of Teacher-Induced Learning Gains,” *Journal of Human Resources*, 45(4) (2010), 915–43.

Kim, Jinho, and Jason M. Fletcher. “The Influence of Classmates on Adolescent Criminal Activities in the United States,” *Deviant Behavior*, 39(3) (2018), 275-292.

Kling, Jeffrey R., Jens Ludwig and Lawrence F. Katz. “Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment,” *The Quarterly Journal of Economics*, 120(1) (2005), 87-130.

Johnson, Rucker C. “Long-run Impacts of School Desegregation and School Quality on Adult Attainments,” NBER Working Paper No. w16664 (2011).

Jordan, Reed and Megan Gallagher. "Does School Choice Affect Gentrification," *Posing the Question and Assessing the Evidence. Metropolitan Housing and Communities Policy Center Brief: Urban Institute* (2015).

Lavy, Victor. "Effects of Free Choice Among Public Schools," *The Review of Economic Studies*, 77(3) (2010), 1164-1191.

Lavy, Victor. "The Long-term Consequences of Free School Choice," *Journal of the European Economic Association*, 19(3) (2021), 1734-1781.

Lavy, Victor and Edith Sand. "The Effect of Social Networks on Students' Academic and Non-cognitive Behavioural Outcomes: Evidence from Conditional Random Assignment of Friends in School," *The Economic Journal*, 129 (2019), 439-480.

Lavy, Victor, and Analia Schlosser. "Mechanisms and Impacts of Gender Peer Effects at School," *American Economic Journal: Applied Economics*, 3(2) (2011), 1-33.

Ludwig, Jens; Duncan, Greg J.; and Hirschfield, Paul. “Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment,” *The Quarterly Journal of Economics*, 116 (2001), 655-80.

Ludwig, Jens, and Jeffrey R. Kling. “Is Crime Contagious?” *Journal of Law and Economics*, 50 (2007), 491–518.

Mills, Jonathan N., and Patrick J. Wolf, “Vouchers in the Bayou: The Effects of the Louisiana Scholarship Program on Student Achievement after 2 Years,” *Educational Evaluation and Policy Analysis*, 39(3) (2017), 464–484.

Monarrez, Tomás , Brian Kisida and Matthew Chingos. "The Effect of Charter Schools on School Segregation," *American Economic Journal: Economic Policy*, 14(1) (2002), 301-40.

Mumma, Kirsten Slungaard. "The Effect of Charter School Openings on Traditional Public Schools in Massachusetts and North Carolina," *American Economic Journal: Economic Policy*, 14(2) (2022), 454-74.

Muralidharan, Karthik, and Venkatesh Sundararaman. "The aggregate effect of school choice: Evidence from a two-stage experiment in India," *The Quarterly Journal of Economics*, 130(3) (2015), 1011-1066.

Patacchini, Eleanora and Yves Zenou. "Juvenile Delinquency and Conformism," *Journal of Law, Economics, and Organization*, 28 (2009), 1-31.

Pathak, Parag A. "What Really Matters in Designing School Choice Mechanisms," *Advances in Economics and Econometrics*, 1 (2017), 176-214.

Pop-Eleches, Cristian, and Miguel Urquiola. "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 103(4) (2013), 1289-1324.

Melanie Rucinski, Joshua Goodman. "Racial Diversity and Measuring Merit: Evidence from Boston's Exam School Admissions," *Education Finance and Policy*, 17(3) (2022), 408–431.

Schwartz, Amy Ellen, Ioan Voicu, and Keren Mertens Horn. "Do Choice Schools Break the Link Between Public Schools and Property Values? Evidence from House Prices in New York City," *Regional Science and Urban Economics*, 49 (2014), 1-10.

Sorensen, Lucy C., Shawn D. Bushway, and Elizabeth J. Gifford. "Getting tough? The effects of discretionary principal discipline on student outcomes," *Education Finance and Policy*, 17(2) (2022), 255-284.

Tuttle, Christina Clark, Philip Gleason, and Melissa Clark. "Using Lotteries to Evaluate Schools of Choice: Evidence from a National Study of Charter Schools," *Economics of Education Review*, 31(2) (2012), 237-253.

Weiner, David A., Byron F. Lutz, and Jens Ludwig. "The Effects of School Desegregation on Crime," NBER Working Paper #15380 (2009).

Tables and Figures

Table 1: Characteristics Summary and Tests

	Summary of Characteristics			
	CMS	Lottery Apps	Non-Applicants	Tests
Num. Apps	2.36	4.15	2.17	
	(2.97)	(3.33)	(2.88)	
Group N	18.66	16.73	21.65	
	(16.85)	(15.07)	(17.52)	
Num. Wins	0.88	1.52	0.83	
	(1.32)	(1.67)	(1.27)	
Black	0.42	0.52	0.36	0.03
	(0.49)	(0.50)	(0.48)	(0.13)
White	0.36	0.29	0.42	-0.17
	(0.48)	(0.45)	(0.49)	(0.11)
Hispanic	0.14	0.11	0.13	0.13
	(0.35)	(0.31)	(0.34)	(0.11)
Ec. Disadvantage	0.48	0.46	0.45	0.06
	(0.50)	(0.50)	(0.50)	(0.12)
Math EOG, 5th Grade	0.13	0.37	0.13	-0.37
	(1.00)	(0.94)	(1.01)	(0.26)
Read EOG, 5th Grade	-0.01	0.27	0.00	-0.48*
	(1.00)	(0.93)	(1.01)	(0.26)
Missing Test Score	0.07	0.02	0.06	0.01
	(0.26)	(0.14)	(0.24)	(0.08)
ELL	0.10	0.05	0.10	0.04
	(0.31)	(0.22)	(0.30)	(0.10)
Observations	13,493	1,794	7,891	7,891
Clusters				1134
P-value				0.648

Notes: Summary of student and group level attributes. The first three columns show the mean and standard deviation for CMS, the sample of lottery applicants, and the non-applicant estimation sample, respectively. *Ec. Disadvantage* is an indicator for economically disadvantaged. *Num. Apps* = the number of other male applicants in student's cohort-school-CBG group. *Num. Wins* = the number of other male lottery winners from the applicant sample in the student's cohort-school-CBG group. *Group N* = the number of other male students in student's cohort-school-CBG group. *ELL* is an indicator for English Language Learner. Last column displays estimated coefficients from regressing the row variable on the win proportion, conditional on expected win proportion, covariates related to lottery applications and wins (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and school, CBG and year fixed effects. Standard errors are clustered at the CBG-Neighborhood School-Cohort level. The P-value refers to the joint test on the individual characteristics and test scores from regressing the win proportion on the characteristics, conditional on the other lottery related covariates and school, CBG, and year fixed effects.

Table 2: Effects of Winning Lottery on Own Attendance and Movement

		Attendance			Exit		Att. School Mean	
	Won Any	App School	Neighb. Sch.	Change NS	Grade 6	Grade 9	Math	Pr(Arrest)
Panel A: Pooled								
Win	0.755***	0.685***	-0.358***	-0.054*	-0.017	-0.049*	0.249***	-0.012***
	(0.018)	(0.028)	(0.029)	(0.032)	(0.027)	(0.029)	(0.045)	(0.004)
Dep Var Mean	.532	.4	.249	.272	.121	.234	.205	.127
Observations	1794	1794	1794	1794	1794	1794	1569	1569
Panel B: Effect Win, by Risk								
Win × HR	0.785***	0.712***	-0.341***	0.010	0.008	-0.027	0.261***	-0.021***
	(0.024)	(0.028)	(0.031)	(0.037)	(0.023)	(0.032)	(0.059)	(0.006)
Win × LR	0.730***	0.664***	-0.373***	-0.108***	-0.037	-0.068	0.235***	-0.005
	(0.032)	(0.043)	(0.044)	(0.037)	(0.038)	(0.045)	(0.056)	(0.004)
Observations	1794	1794	1794	1794	1794	1794	1569	1569

Notes: Estimated effect of winning first choice on attendance and movement of applicants. Each estimate is conditional on the set of student characteristics, and the estimated win probability. Panel A includes estimates for all applicants. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *Wins* × *HR* refers to estimates for the group with higher than median estimated risk, and *Wins* × *LR* shows estimates for the group with lower than median estimated risk. Each column in Panel B is estimated in a single regression using interaction terms. *Won Any* is a dummy for winning any choice in the lottery (first, second, or third). *App School* is an indicator variable for attending the middle school they applied to with their first choice in the lottery. *Neighb. Sch.* indicates whether the student attended their initially assigned neighborhood school in 6th grade. *Change NS* = Designated neighborhood school changes from year of lottery to the next year, indicating likely residential movement. *Exit* = Missing sixth grade school of attendance in CMS data, and missing 9th grade school in CMS data, each of which indicate likely district exit. The last two columns refer to average characteristics in the 6th grade school of attendance. *Math* is the average 6th grade standardized test score from the prior year, and *Pr(Arrest)* is the mean predicted arrest risk among the students attending that school in the same cohort. Standard errors are clustered at the application choice by year level.

Table 3: Effect of Peer Wins on Arrest Outcomes

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g)$	0.260**	0.155*	0.140	1.454**	0.612**	23.057
	(0.102)	(0.081)	(0.087)	(0.593)	(0.261)	(17.055)
Observations	7891	7891	7891	7891	7891	7891
Panel B: By Risk						
$(Wins/N_g) \times HR$	0.217	0.149	0.076	1.420	0.634	19.047
	(0.149)	(0.122)	(0.130)	(0.911)	(0.397)	(26.678)
$(Wins/N_g) \times LR$	0.333***	0.164**	0.243***	1.536***	0.601***	30.865***
	(0.094)	(0.071)	(0.072)	(0.415)	(0.183)	(9.340)
Dep Var Mean, HR	.208	.136	.135	.863	.361	15.51
Dep Var Mean, LR	.054	.022	.02	.124	.04	1.018
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the peer win proportion on outcomes. Panel A reports estimates for the full sample. Each estimate in Panel A is for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Neighborhood groups include male students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also includes interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table 4: Falsification Test, Lag Wins on Arrest Outcomes

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
Lag ($Wins/N_g$)	-0.033	-0.119	-0.062	-0.023	-0.118	18.591
	(0.131)	(0.857)	(0.100)	(0.402)	(0.100)	(21.652)
Observations	7311	7311	7311	7311	7311	7311
Panel B: By Risk						
Lag ($Wins/N_g$) $\times$ HR	-0.044	-0.100	-0.105	-0.029	-0.187	21.796
	(0.172)	(1.096)	(0.127)	(0.514)	(0.126)	(28.654)
Lag ($Wins/N_g$) $\times$ LR	-0.057	-0.138	0.029	-0.007	0.023	11.723
	(0.147)	(0.598)	(0.112)	(0.264)	(0.103)	(16.032)
Observations	7311	7311	7311	7311	7311	7311

Notes: Estimated effect of the win proportion from the same school-cbg in the following year on outcomes. Panel A reports estimates for the full sample. Each estimate in Panel A is for the effect of the proportion of students from their same neighborhood school and CBG in the *following* cohort (one year later) who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size from the following year, other lottery related group-level covariates from the following year (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *Lag (Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *Lag (Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also includes interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table 5: Peer Effects on Other Outcomes

		Suspensions			
	Absences	Total Days	Days ISS	Days OSS	Any Susp.
Panel A: Pooled					
$(Wins/N_g)$	4.517**	1.810	0.133	1.677	0.107
	(2.248)	(1.926)	(0.498)	(1.565)	(0.560)
Dep Var Mean	7.675	2.814	.711	2.103	.96
Observations	32974	31697	31697	31697	31697
Panel B: Effect Win, by Risk					
$(Wins/N_g) \times HR$	3.663	0.271	-0.226	0.497	-0.311
	(3.115)	(2.954)	(0.756)	(2.387)	(0.831)
$(Wins/N_g) \times LR$	5.929**	4.362***	0.763*	3.599***	0.792
	(2.359)	(1.540)	(0.458)	(1.207)	(0.538)
Dep Var Mean, HR	9.881	5.105	1.227	3.878	1.694
Dep Var Mean, LR	5.866	.891	.278	.613	.344
Observations	32974	31697	31697	31697	31697
Panel C: Grades 6-8, Low Risk					
$(Wins/N_g) \times LR$	6.156**	5.539***	1.286**	4.254***	1.136**
	(2.653)	(1.724)	(0.576)	(1.337)	(0.549)
Dep Var Mean, LR	5.495	.929	.298	.631	.338
LR Observations	11388	10451	10451	10451	10451
Observations	20747	19232	19232	19232	19232

Notes: Estimated effect of win proportion on non-applicant outcomes. Each is conditional on the set of student characteristics, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. ISS = In School Suspension; OSS = Out of School Suspension. Panels A and B include up to 5 years post lottery, or 6th through 10th grade years. Panel A includes estimates from the full sample. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. Panel C reports the estimated effects on the subsample with below median risk for the three years post-lottery, or 6th through 8th grade years. Panel B and Panel C are estimated using interaction terms between win proportion and high- and low-risk. Panels B and C regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table 6: Effect of Peer Wins on Arrest Outcomes (Low-Risk, by Age)

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Low-Risk, Age 16 to 18						
$(Wins/N_g) \times LR$	0.173**	0.077	0.117**	0.352**	0.138*	3.954
	(0.070)	(0.054)	(0.050)	(0.152)	(0.076)	(2.741)
Dep Var Mean, LR	.028	.011	.01	.048	.016	.369
Observations	7891	7891	7891	7891	7891	7891
Panel B: Age 18 to 22						
$(Wins/N_g) \times LR$	0.352***	0.173***	0.208***	1.184***	0.464***	26.911***
	(0.084)	(0.058)	(0.064)	(0.318)	(0.140)	(7.990)
Dep Var Mean, LR	.037	.014	.014	.076	.025	.649
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the peer win share on outcomes. Panel A reports estimates on the arrest outcomes from ages 16 to 18 for the low-risk sample. Each estimate in Panel A is for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for outcomes measured from ages 19 to 22 for low-risk students. Estimates are comparable to the low-risk estimates in Panel B of Table 3. The estimates in Table 3 use the outcomes measured over ages 16 to 22, whereas this table includes separate estimates for the outcomes measured at ages 16 to 18 and 19 to 22. These are estimated in a single regression using interaction terms, similar to Table 3. Each regression also includes interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table 7: Characteristics Summary and Tests (Applicant Sample)

			Tests	
	Won	Lost	Unconditional	Conditional
Made 2nd Application Choice	0.80	0.80	0.02	-0.03
	(0.40)	(0.40)	(0.05)	(0.03)
Made 3rd Application Choice	0.54	0.55	0.03	0.03
	(0.50)	(0.50)	(0.03)	(0.02)
Black	0.45	0.57	-0.07	-0.02
	(0.50)	(0.50)	(0.05)	(0.04)
White	0.37	0.24	0.05	0.07
	(0.48)	(0.43)	(0.06)	(0.05)
Hispanic	0.09	0.12	-0.06	-0.02
	(0.29)	(0.32)	(0.08)	(0.06)
Ec. Disadvantage	0.39	0.50	-0.04	0.03
	(0.49)	(0.50)	(0.03)	(0.02)
Math EOG, 4th Grade	0.46	0.23	0.03	0.00
	(0.97)	(0.95)	(0.02)	(0.01)
Read EOG, 4th Grade	0.35	0.16	0.00	-0.00
	(0.90)	(0.90)	(0.02)	(0.02)
Test Miss	0.05	0.05	0.03	-0.04
	(0.23)	(0.22)	(0.05)	(0.05)
ELL	0.04	0.06	-0.02	-0.03
	(0.20)	(0.23)	(0.05)	(0.05)
Observations	682	1,112	1,794	1,794
Clusters			59	59
P-value			0.015	0.511

Notes: Summary of characteristics in the lottery sample. *ELL* is an indicator for English Language Learner. *Ec. Disadvantage* is an indicator for economically disadvantaged. *Test Miss* is an indicator for missing either test score. The first two columns show the mean and standard deviation for lottery winners and losers. The last two columns display estimated coefficients from regressing an indicator for winning the first choice in the lottery on the listed characteristics. Standard errors are clustered at the application choice by year level. P-values refer to the joint test on the listed characteristics from the regression of winning on the characteristics. The column labeled *Unconditional* reports coefficients without conditioning. The *Conditional* column conditions on the estimated win probability.

Table 8: Effects of Winning Lottery on Own Arrest Outcomes

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
Win	-0.011	-0.011	-0.027**	-0.167	-0.072	-3.537
	(0.014)	(0.014)	(0.012)	(0.124)	(0.070)	(2.973)
Observations	1794	1794	1794	1794	1794	1794
Panel B: By Risk						
$Win \times HR$	-0.016	-0.001	-0.046**	-0.351	-0.149	-11.579*
	(0.028)	(0.027)	(0.023)	(0.267)	(0.154)	(6.676)
$Win \times LR$	-0.008	-0.020**	-0.012*	-0.011	-0.007	3.017
	(0.011)	(0.008)	(0.007)	(0.026)	(0.012)	(2.810)
Dep Var Mean, HR	.185	.112	.11	.734	.333	11.662
Dep Var Mean, LR	.041	.022	.019	.08	.029	1.364
Observations	1794	1794	1794	1794	1794	1794

Notes: Estimated effect of the winning first choice in the lottery on own arrests and incarceration. Each is conditional on the set of student characteristics, and the estimated win probability. Panel A includes estimates for all applicants. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *Wins* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *Wins* $\times$ *LR* shows estimates for the group with lower than median estimated risk. Each column in Panel B is estimated in a single regression using interaction terms. Standard errors are clustered at the application choice by year level.

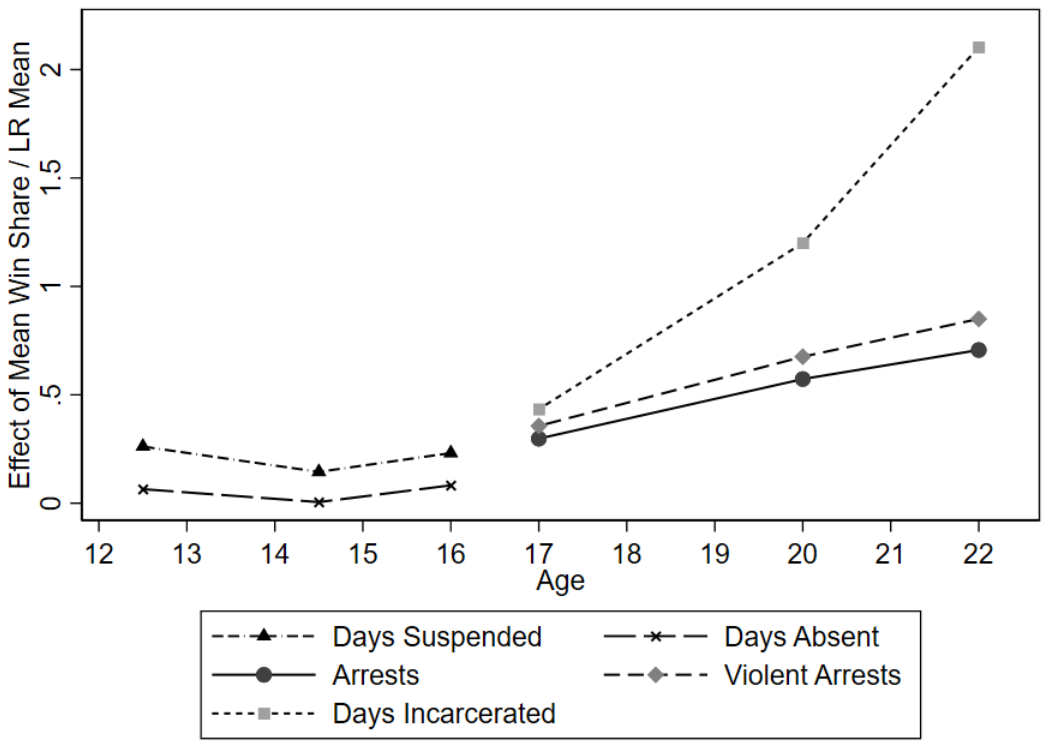
Table 9: Aggregated Effects

Panel A: By Above and Below Median Risk		Aggregated Effects: All Arrests	Aggregated Effects: Violent Arrests	Aggregated Effects: Days Incarcerated
	$E[Wins]$			
High-Risk Applicants	277.69	-96.84	-40.79	**-3191.48**
Low-Risk Applicants	373.41	-4.19	-2.97	1121.33
Total Effect on Applicants		-101.03	-48.34	-2070.15
Significant Effects				**-3191.48**
	$E[Displacements]$			
High-Risk Peers	2994.58	252.27	113.11	3240.91
Low-Risk Peers	3186.29	**211.59**	**83.80**	**4131.12**
Total Effect on Peers		463.86	196.91	7372.03
Significant Effects		**211.59**	**83.80**	**4131.12**

Panel B: By Risk Quintile		Aggregated Effects: All Arrests	Aggregated Effects: Violent Arrests	Aggregated Effects: Days Incarcerated
	$E[Wins]$			
Q5 (High)	90.47	-49.88	-16.59	-1498.39
Q4	118.47	-46.59	-19.79	-1579.13
Q3	159.08	4.19	-3.99	1138.92
Q2	135.29	-5.79	**-5.86**	-6.16
Q1 (Low)	147.79	2.47	0.10	-0.14
Total Effect on Applicants		-95.60	-46.13	-1944.90
Significant Effects			**-5.86**	
	$E[Displacements]$			
Q5 (High)	946.43	72.23	81.44	2901.55
Q4	1366.11	66.58	-1.12	-3170.39
Q3	1305.19	**137.80**	**50.86**	**5913.69**
Q2	1344.94	**81.67**	**34.05**	**1529.56**
Q1 (Low)	1218.21	**51.28**	**15.92**	**1099.00**
Total Effect on Peers		409.56	156.65	8273.41
Significant Effects		**270.75**	**100.83**	**8542.25**

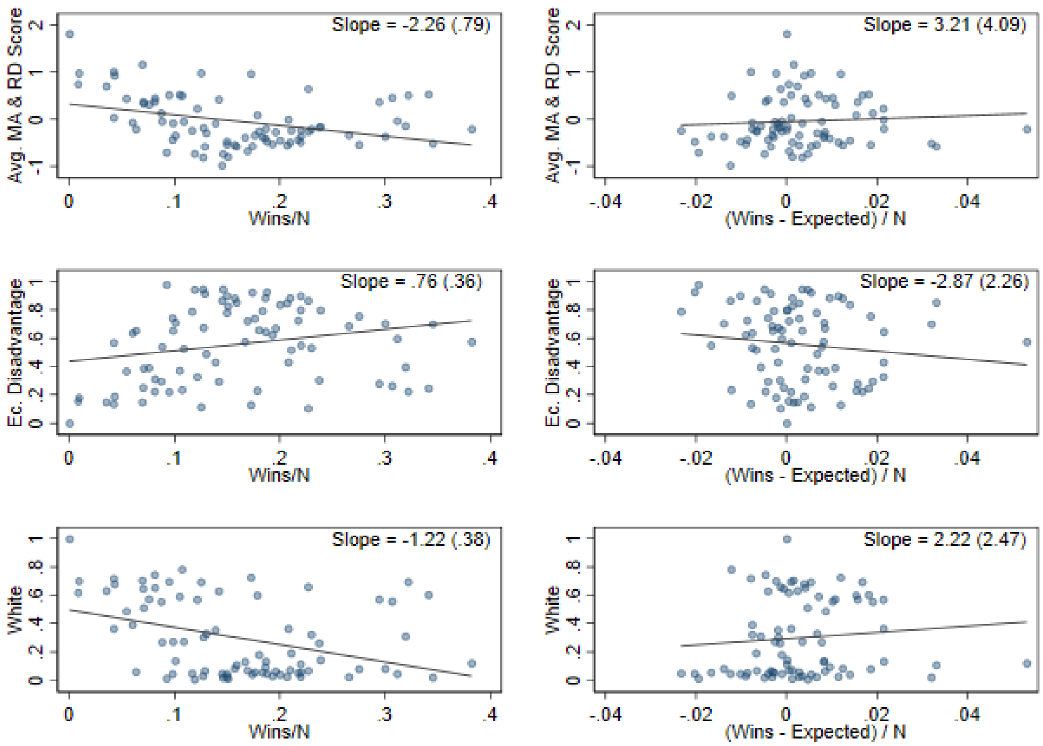
Notes: Aggregated effects on lottery applicants and their peers. The first column reports the expected number of wins among lottery applicants in each group *E[Wins]* or the expected number of peer wins *E[Displacements]*. *E[Wins]* are calculated as the sum of the estimated win probabilities for all applicants in the corresponding group. *E[Displacements]* are calculated as the sum of the expected number of wins for each non-applicant in each group. That is, each lottery applicant contributes $P_i \cdot N_g$ to the total number of displacements, where P_i refers to their win probability and N_g refers to the number of non-applicants in their group, i.e., the number of students exposed to the peer who may experience a peer displacement through the lottery. The aggregated effects on applicants are calculated using the expected number of wins and the corresponding estimates for effects of winning on *All Arrests*, *Violent Arrests*, and *Days Incarcerated*, which come from Table 8 for Panel A and Table A20 for Panel B. The estimated effects on peers are calculated using the sum of the expected lottery win share of each peer's group multiplied by the corresponding estimated effect on peer outcomes (Table 3 for Panel A and Table A21 for Panel B). Aggregates using only estimates with p-value<0.1 are in bold.

Figure 1: Estimated Effects on Low-Risk Peers by Age



Notes: Estimated effects of peer wins on in- and out-of-school outcomes for the low-risk non-applicants. Each point is a standardized estimated of the effect on low-risk non-applicants (β_1^{LR}) from a regression based on Equation (3). To obtain each point in the figure, we multiply the estimated effect on the outcome by the mean displacement (.04) and divide by the age- or grade-specific mean of the same outcome in the low-risk sample. Estimates are included for three distinct approximate age groups for each set of outcomes. For *Days Suspended* and *Days Absent*, one estimate includes the two years post-lottery (6th an 7th grade), one includes the third and fourth years post-lottery (8th and 9th grade), and one includes for the fifth year post-lottery. We do not include estimates for years that correspond to 11th and 12th grade, because students are eligible to drop out at age 17. For the other three outcomes *Arrests*, *Violent Arrests*, and *Days Incarcerated*, we include estimates for each of the three aggregations that we have in our data: Age 16 to 18, age 19 to 20, and age 21 to 22.

Figure 2: Lottery Wins and School Characteristics



Notes: Correlations between neighborhood school characteristics and lottery wins at the school level. Each panel shows observations and the slope estimate for the relation between some measure of win proportion represented on the x-axis and the school-cohort characteristic on the y-axis. Each dot represents one school-cohort observation. The left-hand column uses school-cohort level wins as a share of the school-cohort size and the right-hand column uses the number of wins minus expected wins as a share of the school-cohort size. The school-cohort size is measured as the number of students with that home school assignment at the time of the lottery. The school-cohort characteristics are the average of the school-cohort math and reading scores (top two panels), the proportion economically disadvantaged students (middle two panels), and the proportion of white students (bottom two panels). Each school-cohort characteristic measure includes 8th grade students in the designated school 3 years post-lottery, i.e., the 8th grade year.

Appendix A Tables and Figures

Table A1: Outcomes Summary

	CMS	Lottery Apps	Non-Applicants
A) Crime Outcomes			
Pr(Any Arrest)	0.13	0.11	0.12
	(0.33)	(0.32)	(0.33)
Pr(Violent Arrests)	0.08	0.07	0.07
	(0.27)	(0.25)	(0.26)
Pr(Incarceration)	0.08	0.06	0.07
	(0.26)	(0.24)	(0.26)
Num. Arrests	0.49	0.40	0.46
	(2.00)	(1.79)	(1.92)
Num. Violent Arrests	0.21	0.18	0.18
	(0.96)	(0.94)	(0.89)
Days Incarcerated	8.95	6.44	7.53
	(60.88)	(49.18)	(54.00)
B) Other Student Outcomes			
Reading Test Scores (6th-8th)	-0.05	0.18	-0.03
	(1.06)	(0.94)	(1.06)
Math Test Scores (6th-8th)	0.05	0.23	0.08
	(1.09)	(1.00)	(1.11)
Days Suspended	2.92	2.21	2.81
	(7.71)	(6.08)	(7.63)
Days Absent	6.57	5.59	6.60
	(7.83)	(6.31)	(7.92)
Graduated HS	0.91	0.96	0.91
	(0.28)	(0.21)	(0.29)
Dropped Out of HS	0.09	0.05	0.09
	(0.28)	(0.21)	(0.29)
Num. of Students in Sample	13,493	1,794	7,891

Notes: Panel A includes the summary statistics of each arrest and incarceration outcome from age 16 to 22 for our main sample of male 5th grade students in Charlotte-Mecklenburg Schools (CMS) between the years of 2005-2006 to 2007-2008. The columns show the mean and standard deviation for CMS, the sample of lottery applicants and the non-applicant sample, respectively. Panel B includes means and standard deviations for other student outcomes for the same three samples. The last row refers to the number of students in each sample. The number of observations contributing to each mean in Panel B varies.

Table A2: Arrest Prediction

	Fixed Effects	School and CBG Means
White	0.035	0.025
	(0.140)	(0.137)
Black	0.632***	0.599***
	(0.123)	(0.120)
Hispanic	-0.236*	-0.237*
	(0.143)	(0.140)
Ec. Disadvantage	0.702***	0.676***
	(0.077)	(0.075)
Math 5	-0.238***	-0.219***
	(0.044)	(0.043)
Read 5	-0.120***	-0.122***
	(0.044)	(0.042)
Math 5 Missing	-0.008	-0.035
	(0.272)	(0.258)
Read 5 Missing	0.322	0.290
	(0.259)	(0.246)
Age	0.024***	0.024***
	(0.005)	(0.004)
School Means		
White		-0.832
		(1.405)
Black		-1.890
		(1.432)
Hispanic		-2.144
		(1.531)
Ec. Disadvantage		0.914*
		(0.512)
Math 5		-0.210
		(0.764)
Read 5		-0.023
		(0.768)
CBG Means		
White		0.866
		(0.660)
Black		1.310**
		(0.592)
Hispanic		0.702
		(0.665)
Ec. Disadvantage		0.679*
		(0.356)
Math 5		0.451
		(0.280)
Read 5		-0.150
		(0.313)
Observations	12,716	13,493

Notes: Arrest prediction estimates from logistic regression with indicator for any arrest age 16-22. First column includes neighborhood school and CBG fixed effects. The Second uses neighborhood school and CBG mean characteristics. In addition to the listed covariates, we include a dummies for missing birthdate and 5th grade exceptionality.

Table A3: Effects of Peer Wins on Attendance and Movement

	Attend		Exit	
	Neighb. Sch.	Change NS	Grade 6	Grade 9
Panel A: Pooled				
$(Wins/N_g)$	0.228	-0.298**	-0.107	-0.061
	(0.144)	(0.151)	(0.092)	(0.127)
Dep Var Mean	.77	.239	.114	.261
Observations	7891	7891	7891	7891
Panel B: By Risk				
$(Wins/N_g) \times HR$	0.154	-0.120	-0.122	-0.088
	(0.178)	(0.185)	(0.124)	(0.173)
$(Wins/N_g) \times LR$	0.360*	-0.571***	-0.090	-0.020
	(0.206)	(0.207)	(0.131)	(0.175)
Observations	7891	7891	7891	7891

Notes: Estimated effect of win proportion on non-applicant outcomes. Each is conditional on the set of student characteristics, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel A includes estimates from the full sample. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. Panel B is estimated using interaction terms between win proportion and high- and low-risk. Panel B regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. *Attend Neighb. Sch.* indicates whether the student attended their initially assigned neighborhood school in 6th grade. *Change NS* = Designated neighborhood school changes from year of lottery to the next year, indicating likely residential movement. *Exit* = Missing sixth grade school of attendance in CMS data, or missing 9th grade school of attendance in CMS data, which indicate likely district exit. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A4: Effect of Peer Wins on Intensive Margin - Poisson

	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled			
$(Wins/N_g)$	1.498*	1.667*	2.519
	(0.798)	(0.863)	(1.535)
Observations	7027	6061	6056
Panel B: By Risk			
$(Wins/N_g) \times HR$	1.239	1.397	2.073
	(0.861)	(0.913)	(1.856)
$(Wins/N_g) \times LR$	4.240**	5.525*	5.939
	(2.082)	(2.865)	(4.699)
Dep Var Mean, HR	.914	.417	17.584
Dep Var Mean, LR	.148	.06	1.53
Observations	6990	6024	6019
Panel C: LR, Age 18-22			
$(Wins/N_g) \times LR$	6.976**	8.323**	9.674*
	(2.752)	(3.922)	(5.429)
Dep Var Mean, LR	.097	.04	1.043
Observations	6652	5598	5731

Notes: Estimated effects of the peer win proportion on intensive margin outcomes using Poisson regression. Panel A reports estimates across the sample, pooling by estimated risk. Each estimate in Panel A is for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Neighborhood groups include male students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* × *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* × *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel C reports estimates for the low risk group for outcomes measured over ages 18-22. Panel B and C regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A5: Days Incarcerated Robustness

	Main	Poisson	Topcode OLS	Topcode Poisson	ln(1+days)	Arcsin(days)
Panel A: Pooled						
$(Wins/N_g)$	23.057	2.519	13.510	1.423	0.534*	0.622*
	(17.055)	(1.535)	(9.276)	(1.157)	(0.298)	(0.344)
APE		24.709		9.831		
		(15.031)		(7.954)		
Observations	7891	6056	7891	6056	7891	7891
Panel B: By Risk						
$(Wins/N_g) \times HR$	19.047	2.139	8.774	0.782	0.302	0.350
	(26.678)	(1.860)	(14.348)	(1.428)	(0.455)	(0.522)
$(Wins/N_g) \times LR$	30.865***	5.888	21.618***	5.290	0.914***	1.066***
	(9.340)	(4.745)	(6.556)	(4.154)	(0.240)	(0.280)
Observations	7891	6056	7891	6056	7891	7891

Notes: Estimated effects of neighborhood win proportion on days incarcerated. Column labels refer to the method used for each result: (1) OLS using days incarcerated as outcome; (2) Poisson regression using days incarcerated as outcome; (3) and (4) are top coded versions of (1) and (2) with the top 1% values of days incarcerated replaced with the 99th percentile of the distribution of 250; (5) OLS using ln(1 + days incarcerated) as outcome; and (6) OLS using arcsin(days incarcerated) as outcome. Panel A reports estimates and average of the win proportion in the full sample of non-applicants. Each estimate in Panel A is for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery. Neighborhood groups include male students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports estimates for above and below median risk. Each column in panel B is estimated using interactions between risk indicators and the win proportion, expected win proportion, and other lottery covariates.

Table A6: Selection of Applicants on Test Scores

	Math	Reading	Math	Reading
Applicant, Lost Lottery	0.392***	0.378***	0.077***	0.061**
	(0.045)	(0.040)	(0.027)	(0.026)
Math EOG 5th			0.661***	0.230***
			(0.010)	(0.010)
Reading EOG 5th			0.181***	0.569***
			(0.010)	(0.011)
Observations	18085	18018	18085	18018

Notes: Comparison of applicants who lost the lottery and attended their initially assigned neighborhood school with non-applicants who attended their initially assigned neighborhood school. Each column reports estimates from a different regression of math or reading test score on an indicator for being in our lottery sample and losing the first choice in the lottery. Includes observations for up to 3 years post-lottery, i.e., 6th - 8th grade. Estimates in Columns 3 and 4 are conditional on 5th grade reading and math scores, with coefficients reported in the table. Each is also conditional on race/ethnicity, economic disadvantage, whether the student was ever ELL, indicators for missing lagged test scores, and CBG, neighborhood school, cohort fixed effects, and grade of test score. Sample limited to boys only. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A7: Effect of Peer Wins in Alternate Specifications

	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.
Panel A: Pooled												
$(Wins/N_g)$	0.093	15.482	0.231**	20.981	0.260**	23.057	0.248**	24.053	0.231**	22.245	0.177*	24.244
	(0.097)	(16.796)	(0.103)	(16.751)	(0.102)	(17.055)	(0.112)	(18.209)	(0.100)	(17.190)	(0.094)	(17.190)
	[.095]	[16.111]	[.11]	[17.763]	[.112]	[18.363]	[.126]	[20.645]	[.107]	[18.553]	[.104]	[19.001]
Observations	7897	7897	7891	7891	7891	7891	7856	7856	7891	7891	7888	7888
Panel B: By Risk												
$(Wins/N_g) \times HR$	0.062	17.329	0.201	19.885	0.217	19.047	0.258*	24.969	0.193	22.178	0.094	26.031
	(0.152)	(27.401)	(0.152)	(26.719)	(0.149)	(26.678)	(0.154)	(27.418)	(0.146)	(27.141)	(0.145)	(29.699)
	[.147]	[26.434]	[.163]	[28.214]	[.163]	[28.643]	[.172]	[30.159]	[.157]	[29.299]	[.161]	[32.462]
$(Wins/N_g) \times LR$	0.163**	15.723***	0.281***	23.737***	0.333***	30.865***	0.269**	26.960**	0.296***	24.757**	0.192**	18.732**
	(0.071)	(5.428)	(0.093)	(8.353)	(0.094)	(9.340)	(0.113)	(10.713)	(0.096)	(10.365)	(0.097)	(8.236)
	[.07]	[5.518]	[.094]	[7.962]	[.096]	[8.891]	[.114]	[9.898]	[.096]	[10.204]	[.101]	[8.266]
Observations	7897	7897	7891	7891	7891	7891	7856	7856	7891	7891	7886	7886
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓
School FE	✓	✓			✓	✓					✓	✓
CBG FE			✓	✓	✓	✓			✓	✓	✓	✓
School-CBG FE							✓	✓				
School-Cohort FE									✓	✓		
E[Win Prop.] Bins											✓	✓

Notes: Estimated effects of the win proportion on probability of arrest and days incarcerated using alternate specifications indicated in the bottom of the table. Each is conditional on the set of student characteristics, and other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications). Panel A includes estimates for the pooled sample of high- and low-risk non-applicants. Panel B includes separate estimates for above and below median risk, which are estimated using interaction terms. Panel A conditions on the expected win proportion (or expected win proportion bins in the last two columns) and Panel B conditions on the expected win proportion for high- and low-risk students (or expected win proportion by risk bins in the last two columns). There are two standard errors included for each estimate. Standard errors clustered at the CBG-Neighborhood School-Cohort level in parentheses, and standard errors clustered at the CBG-Neighborhood School level are included in brackets.

Table A8: Effects of Same CBG-School-Cohort Girl Wins on Boy Outcomes

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g)$	-0.032	-0.053	0.061	0.209	-0.089	-6.914
	(0.106)	(0.087)	(0.094)	(0.580)	(0.267)	(17.331)
Observations	7458	7458	7458	7458	7458	7458
Panel B: By Risk						
$(Wins/N_g) \times HR$	0.019	-0.030	0.175	0.769	0.068	6.171
	(0.143)	(0.121)	(0.130)	(0.812)	(0.378)	(24.364)
$(Wins/N_g) \times LR$	-0.037	-0.088	-0.126	-0.351	-0.280	-19.002
	(0.119)	(0.094)	(0.087)	(0.515)	(0.204)	(11.761)
Dep Var Mean, HR	.201	.131	.129	.832	.346	14.839
Dep Var Mean, LR	.055	.022	.021	.128	.041	1.056
Observations	7458	7458	7458	7458	7458	7458

Notes: Estimated effects of the win proportion among same School-CBG-Cohort girls on outcomes in the sample of boys. Panel A reports estimates for the full sample. Each estimate in Panel A is for the effect of the proportion of same-neighborhood girls who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Neighborhood groups include students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A9: Effect of Peer Wins on Arrest Outcomes (All Rounds)

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g)$	0.206**	0.173**	0.122	1.164**	0.620**	31.730*
	(0.098)	(0.075)	(0.083)	(0.588)	(0.293)	(16.858)
Observations	7891	7891	7891	7891	7891	7891
Panel B: By Risk						
$(Wins/N_g) \times HR$	0.150	0.175*	0.066	1.115	0.685	38.527
	(0.137)	(0.105)	(0.119)	(0.857)	(0.427)	(24.604)
$(Wins/N_g) \times LR$	0.315***	0.173**	0.230***	1.287***	0.518***	20.411**
	(0.094)	(0.077)	(0.070)	(0.387)	(0.178)	(8.937)
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the proportion of peers who won a seat to a different school in any lottery round on outcomes. Panel A reports estimates across the sample. Each estimate in Panel A is for the effect of the proportion of students from their same initially assigned neighborhood school, CBG, and cohort who won any choice in the lottery to a different school on the arrest or incarceration outcome indicated by the column heading. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* × *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* × *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A10: Effect of Peer Wins on Arrest Outcomes (Using Wins-Exp)

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins - Expected)/N_g$	0.261**	0.156*	0.141	1.445**	0.616**	23.965
	(0.103)	(0.081)	(0.088)	(0.590)	(0.259)	(16.942)
Observations	7891	7891	7891	7891	7891	7891
Panel B: By Risk						
$[(Wins - Expected)/N_g] \times HR$	0.228	0.157	0.085	1.455	0.656*	20.047
	(0.151)	(0.121)	(0.131)	(0.908)	(0.392)	(26.646)
$[(Wins - Expected)/N_g] \times LR$	0.328***	0.163**	0.244***	1.502***	0.604***	30.442***
	(0.098)	(0.073)	(0.077)	(0.425)	(0.184)	(9.237)
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the peer wins minus expected wins as a proportion of groups size on outcomes. Estimates are comparable to Table 3. Panel A reports estimates for the full sample. Each estimate in Panel A is for the effect of the wins over expectation of students from their neighborhood group who won their first choice in the lottery as a proportion of the group size on the arrest or incarceration outcome indicated by the column heading. Neighborhood groups include male students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *[(Wins - Expected) / N]* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *[(Wins - Expected) / N]* $\times$ *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also include interaction terms between risk indicators and the movement related covariates. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A11: Peer Effects on Test Scores and Graduation

	Math	Reading	On Time Grad.	Dropout
Panel A: Pooled				
$(Wins/N_g)$	-0.098	-0.270*	0.156	-0.041
	(0.139)	(0.143)	(0.126)	(0.113)
Observations	19666	19598	5417	5417
Panel B: Effect Win, by Risk				
$(Wins/N_g) \times HR$	0.070	-0.318	0.340*	-0.185
	(0.174)	(0.194)	(0.181)	(0.168)
$(Wins/N_g) \times LR$	-0.388*	-0.221	-0.067	0.141
	(0.220)	(0.199)	(0.151)	(0.100)
Observations	19666	19598	5417	5417

Notes: Estimated effect of win proportion on non-applicant outcomes. Each is conditional on the set of student characteristics, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel A includes estimates from the full sample. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. Panel B is estimated using interaction terms between win proportion and high- and low-risk. Panel B regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. Test score results include up to three years post-lottery, grades 6-8 for most individuals. Rest score regressions also include indicators for years post-lottery. The last two columns show outcomes related to graduation and dropout. *On Time Grad.* is an indicator for graduating within 7 years post lottery (12th grade year for on-time progression). *Dropout* is an indicator equal to one if the student was ever observed as dropping out in the 7 years post lottery. The last two columns only include students who were observed graduating or dropping out in the NC data. These include up to one observation per student. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A12: Heterogeneity in Peer Effects

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Heterogeneity by Race / Ethnicity						
Boys, White	0.379***	0.266***	0.210**	2.268***	0.779***	49.791*
	(0.132)	(0.096)	(0.092)	(0.767)	(0.237)	(25.772)
Boys, Black	0.287*	0.245*	0.125	1.607	0.981**	24.965
	(0.167)	(0.143)	(0.143)	(1.043)	(0.484)	(21.880)
Boys, Hispanic	0.154	-0.082	0.212	0.669	0.009	31.986
	(0.217)	(0.170)	(0.198)	(1.037)	(0.441)	(28.165)
Boys, Other	-0.020	-0.318	-0.261	-1.442	-0.813	-141.671
	(0.309)	(0.264)	(0.260)	(2.328)	(0.983)	(139.440)
Observations	7891	7891	7891	7891	7891	7891
Panel B: Heterogeneity by Test Scores						
Above Median	0.493***	0.211**	0.221**	1.109**	0.374	16.614
	(0.115)	(0.096)	(0.094)	(0.549)	(0.238)	(12.605)
Below Median	0.115	0.129	0.095	1.727*	0.813**	28.861
	(0.150)	(0.119)	(0.130)	(0.924)	(0.400)	(27.346)
Dep Var Mean, High	.071	.033	.034	.206	.074	2.776
Dep Var Mean, Low	.18	.117	.113	.729	.306	12.727
Observations	7891	7891	7891	7891	7891	7891

Notes: Heterogeneity in estimated effects of the peer win proportion on outcomes. Panel A reports estimates by race / ethnicity and Panel B reports estimates by whether the average of the student's 5th grade math and reading test scores was above or below the median. Each column in each panel reports estimates from a single regression using interactions between group indicators and the proportion of students from their neighborhood group who won their first choice in the lottery. Each cell reports a group-specific estimate for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Each estimate is conditional on the set of student characteristics, interactions between the set of group indicators and expected wins as a proportion of the group size, interactions between the set of group indicators and other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A13: Heterogeneity in Peer Effects by Risk

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Heterogeneity by Race / Ethnicity						
White x HR	0.715	0.887**	0.717*	9.548**	2.875**	246.222
	(0.486)	(0.437)	(0.427)	(4.525)	(1.245)	(174.766)
White x LR	0.297**	0.145*	0.115	0.878*	0.408**	13.889
	(0.125)	(0.085)	(0.087)	(0.478)	(0.206)	(8.820)
Black x HR	0.303*	0.259*	0.109	1.814	1.093**	30.409
	(0.176)	(0.153)	(0.151)	(1.111)	(0.515)	(23.439)
Black x LR	0.087	0.039	0.289	-1.241	-0.491	-26.647
	(0.368)	(0.316)	(0.320)	(1.377)	(0.533)	(25.991)
Hispanic x HR	-0.077	-0.282	-0.100	-1.590	-0.991	-19.329
	(0.332)	(0.260)	(0.300)	(1.636)	(0.692)	(42.419)
Hispanic x LR	0.311	0.114	0.451***	2.829***	1.065**	87.349***
	(0.227)	(0.188)	(0.133)	(0.908)	(0.489)	(25.243)
Other x HR	-0.692	-0.943**	-0.853**	-6.147	-2.670	-328.041
	(0.543)	(0.421)	(0.424)	(3.750)	(1.652)	(245.045)
Other x LR	0.574***	0.296**	0.375**	3.497***	1.153**	44.308**
	(0.222)	(0.142)	(0.158)	(1.329)	(0.504)	(20.630)
Observations	7891	7891	7891	7891	7891	7891

Notes: Heterogeneity in the estimated effects of the peer win proportion on outcomes. Estimates by race / ethnicity and predicted risk. Each column reports estimates from a single regression using interaction terms between group indicators and the proportion of students from their neighborhood group who won their first choice in the lottery. Each cell reports a group-specific estimate for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Each estimate is conditional on the set of student characteristics, interactions between the set of group indicators and expected wins as a proportion of the group size, interactions between the set of group indicators and other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A14: Effect of Peer Wins on Arrest Outcomes (Girls)

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g)$	-0.034	0.023	0.032	0.001	0.058	-0.418
	(0.068)	(0.058)	(0.042)	(0.155)	(0.102)	(1.573)
Observations	7291	7291	7291	7291	7291	7291
Panel B: By Risk						
$(Wins/N_g) \times HR$	-0.013	0.084	0.072	0.184	0.172	0.770
	(0.091)	(0.078)	(0.058)	(0.210)	(0.135)	(1.870)
$(Wins/N_g) \times LR$	-0.028	-0.054	-0.032	-0.310*	-0.111	-2.512
	(0.051)	(0.035)	(0.032)	(0.164)	(0.075)	(2.240)
Dep Var Mean, HR	.107	.056	.043	.229	.103	1.094
Dep Var Mean, LR	.017	.006	.005	.037	.012	.308
Observations	7291	7291	7291	7291	7291	7291

Notes: Estimated effects of the peer win proportion on outcomes in the sample of girls. Panel A reports estimates for the full sample. Each estimate in Panel A is for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Neighborhood groups include female students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also include interaction terms between risk indicators and movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A15: Effect of Peer Wins Interacted with Elementary School Proxy

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Interaction with Mean of Winners from Same Elem School						
$(Wins/N_g) \times HR$	0.138	0.138	0.018	0.639	0.307	15.326
	(0.176)	(0.154)	(0.152)	(1.227)	(0.499)	(39.156)
$\times Mn(SameElem.)$	0.197	0.037	0.142	2.352	0.958	18.725
	(0.295)	(0.271)	(0.283)	(1.815)	(0.835)	(51.136)
$(Wins/N_g) \times LR$	0.231**	0.119	0.178**	1.364**	0.487**	26.629***
	(0.106)	(0.085)	(0.088)	(0.540)	(0.215)	(10.040)
$\times Mn(SameElem.)$	0.343*	0.150	0.209	0.606	0.377	12.601
	(0.201)	(0.149)	(0.158)	(0.927)	(0.410)	(20.710)
Observations	7891	7891	7891	7891	7891	7891
Panel B: Interactions with School-CBG-Cohort Fixed Effects						
$(Wins/N_g)$						
$\times Mn(SameElem.) \times HR$	0.421	-0.033	0.016	2.067	1.062	113.145
	(0.472)	(0.421)	(0.434)	(2.672)	(1.150)	(88.583)
$\times Mn(SameElem.) \times LR$	0.710**	0.336	0.374	1.868	1.185	72.456
	(0.348)	(0.284)	(0.290)	(1.718)	(0.747)	(56.388)
Observations	7716	7716	7716	7716	7716	7716

Notes: Estimated effect of the win proportion on outcomes including interaction terms with the mean of winners from the same elementary school. Panel A includes the main effect, win share interacted with high- or low-risk, and two additional interaction terms. The interactions $\times Mn(Same\ Elem.)$ include the share of winners from the group that the focal student attended school with in 5th grade. Panel A includes cohort, CBG, and School fixed effects in addition to the controls on student characteristics, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications, and interactions). Panel B includes cohort by school by CBG fixed effects (rather than cohort, CBG, and School fixed effects separately), and separate estimates of interactions between the win proportion and the fraction of winners who attended the same school in 5th grade for high- and low-risk non-applicants. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A16: Estimates by Group Size

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g) \times Large\ N$	0.082	-0.059	0.331*	1.050	0.168	37.894
	(0.232)	(0.167)	(0.178)	(1.491)	(0.561)	(32.809)
$(Wins/N_g) \times Small\ N$	0.319***	0.230***	0.154	1.620**	0.797***	26.441
	(0.112)	(0.089)	(0.098)	(0.645)	(0.288)	(18.269)
Dep Var Mean, Large N	.096	.051	.048	.303	.118	3.887
Dep Var Mean, Small N	.151	.096	.097	.617	.255	11.36
Observations	7891	7891	7891	7891	7891	7891
Panel B: By Group Size and Risk						
$Large\ N \times HR$	-0.102	-0.107	0.114	-0.454	-0.403	0.894
	(0.329)	(0.247)	(0.234)	(2.158)	(0.845)	(43.175)
$Large\ N \times LR$	0.338	-0.007	0.619***	2.940**	0.866	87.010***
	(0.252)	(0.216)	(0.178)	(1.301)	(0.613)	(32.473)
$Small\ N \times HR$	0.322**	0.240*	0.130	1.888*	0.925**	28.249
	(0.161)	(0.135)	(0.145)	(1.008)	(0.440)	(29.505)
$Small\ N \times LR$	0.290***	0.193***	0.161**	1.129***	0.521***	20.839**
	(0.105)	(0.074)	(0.081)	(0.416)	(0.182)	(9.188)
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the peer win proportion in large vs. small neighborhood groups. Panel A reports estimates by group sizee, pooling by estimated risk. Each estimate in Panel A is for the effect of the proportion of students from their neighborhood group who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading. Neighborhood groups include male students in the same cohort, initially assigned neighborhood school, and CBG. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. Rows with $\times$ *HR* show estimates for the group with higher than median estimated risk, and rows with $\times$ *LR* show estimates for the group with lower than median estimated risk. These are estimated in a single regression using interaction terms. Panel B regressions also include interaction terms between risk indicators and the movement related covariates, including the expected win share. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A17: Effect of Same Race Peer Wins on Arrest

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g)$	0.291**	0.174*	0.111	1.489**	0.440	17.017
	(0.123)	(0.101)	(0.103)	(0.703)	(0.311)	(23.437)
Same Race Win Share	-0.033	-0.016	0.038	-0.021	0.222	8.122
	(0.077)	(0.070)	(0.054)	(0.345)	(0.174)	(12.431)
Observations	7891	7891	7891	7891	7891	7891
Panel B: Low Risk						
$(Wins/N_g) \times HR$	0.258	0.170	0.023	1.394	0.292	8.365
	(0.185)	(0.153)	(0.154)	(1.092)	(0.468)	(36.966)
(Same Race Win Share) $\times HR$	-0.047	-0.019	0.060	0.009	0.407	12.663
	(0.112)	(0.102)	(0.079)	(0.528)	(0.268)	(18.881)
$(Wins/N_g) \times LR$	0.321***	0.148*	0.217***	1.483***	0.576**	25.596**
	(0.106)	(0.082)	(0.084)	(0.511)	(0.229)	(12.192)
(Same Race Win Share) $\times LR$	0.027	0.033	0.032	0.056	0.047	7.404
	(0.063)	(0.047)	(0.051)	(0.288)	(0.137)	(7.918)
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the peer win proportion on outcomes and additional effect from alternate peer groups defined at the school-cohort-CBG-race/ethnicity level. Panel A reports estimates across the sample, pooling by estimated risk. Each column in Panel A includes an estimate for the effect of the proportion of students from their school-CBG-cohort (*Wins / N*) who won their first choice in the lottery on the arrest or incarceration outcome indicated by the column heading, and another estimate for the proportion of students from their school-CBG-cohort-race/ethnicity group who won their first choice in the lottery. These are estimated in one regression. Each estimate is conditional on the set of student characteristics, expected wins as a proportion of the group size, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *(Wins / N)* $\times$ *HR* shows main estimates for the group with higher than median estimated risk, and *(Wins / N)* $\times$ *LR* shows estimates for the group with lower than median estimated risk. Each column in Panel B is estimated in a single regression using interaction terms. In Panel B, we also include interaction terms between risk indicators and the movement related covariates, including the expected wins variable. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A18: Effects of Winning Lottery on Own Arrest Outcomes (Girls)

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
Win	-0.006	-0.004	-0.002	-0.006	-0.005	0.069
	(0.009)	(0.008)	(0.006)	(0.035)	(0.013)	(0.109)
Observations	1925	1925	1925	1925	1925	1925
Panel B: By Risk						
$Win \times HR$	0.007	-0.002	0.001	-0.003	-0.008	-0.012
	(0.018)	(0.012)	(0.010)	(0.040)	(0.019)	(0.079)
$Win \times LR$	-0.018*	-0.006	-0.005	-0.011	-0.003	0.139
	(0.010)	(0.009)	(0.007)	(0.050)	(0.016)	(0.198)
Dep Var Mean, HR	.069	.029	.018	.122	.045	.13
Dep Var Mean, LR	.018	.01	.007	.052	.016	.129
Observations	1925	1925	1925	1925	1925	1925

Notes: Estimated effect of winning first choice in the lottery on own arrests and incarceration in the sample of girls. Each is conditional on the set of student characteristics, and estimated win probability. Panel A includes estimates for all applicants. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *Wins* × *HR* shows estimates for the group with higher than median estimated risk, and *Wins* × *LR* shows estimates for the group with lower than median estimated risk. Each column in Panel B is estimated in a single regression using interaction terms. Standard errors are clustered at the application choice by year level.

Table A19: Effects of Winning Lottery on Own Arrest Outcomes (All Rounds)

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
Win	0.000	-0.009	-0.018	-0.180	-0.084	-3.491
	(0.017)	(0.015)	(0.014)	(0.121)	(0.070)	(2.942)
Observations	1796	1796	1796	1796	1796	1796
Panel B: By Risk						
$Win \times HR$	0.012	0.000	-0.031	-0.391	-0.177	-10.659
	(0.037)	(0.033)	(0.031)	(0.262)	(0.154)	(6.825)
$Win \times LR$	-0.009	-0.017*	-0.008	-0.002	-0.005	2.703
	(0.013)	(0.009)	(0.007)	(0.029)	(0.011)	(2.695)
Dep Var Mean, HR	.188	.114	.111	.718	.317	11.019
Dep Var Mean, LR	.039	.021	.017	.075	.027	1.352
Observations	1796	1796	1796	1796	1796	1796

Notes: Estimated effect of winning any choice to a different school in the lottery on own arrests and incarceration in the sample of girls. Each is conditional on the set of student characteristics, and estimated win probability. Panel A includes estimates for all applicants. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *Wins* × *HR* shows estimates for the group with higher than median estimated risk, and *Wins* × *LR* shows estimates for the group with lower than median estimated risk. Each column in Panel B is estimated in a single regression using interaction terms. Standard errors are clustered at the application choice by year level.

Table A20: Effect of Winning on Own Arrest Outcomes by Quintile

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: By Risk Quintiles						
Q5 (High)	-0.035	-0.053	-0.085*	-0.551	-0.219	-16.563
	(0.056)	(0.059)	(0.046)	(0.380)	(0.256)	(11.606)
Q4	-0.001	0.032	-0.006	-0.393	-0.167	-13.329
	(0.036)	(0.029)	(0.031)	(0.459)	(0.245)	(10.228)
Q3	-0.023	0.000	-0.027	0.026	0.025	7.159
	(0.026)	(0.018)	(0.018)	(0.143)	(0.054)	(6.745)
Q2	-0.017	-0.040**	-0.022	-0.043	-0.043**	-0.046
	(0.022)	(0.017)	(0.016)	(0.037)	(0.021)	(0.439)
Q1 (Low)	0.016	0.003	-0.003	0.017	0.001	-0.001
	(0.011)	(0.009)	(0.003)	(0.017)	(0.011)	(0.305)
Observations	1794	1794	1794	1794	1794	1794
Panel B: $Q5$ and $Q1 - Q4$						
$Q5$	-0.035	-0.054	-0.085*	-0.560	-0.223	-16.694
	(0.056)	(0.059)	(0.046)	(0.375)	(0.253)	(11.468)
$(Q1 - Q4)$	-0.006	-0.003	-0.016	-0.089	-0.041	-0.967
	(0.013)	(0.011)	(0.010)	(0.110)	(0.056)	(2.215)
Dep Var Mean, Q5	.274	.182	.182	1.14	.532	20.277
Dep Var Mean, Q1-Q4	.076	.04	.037	.238	.1	3.338
Observations	1794	1794	1794	1794	1794	1794

Notes: Estimated effect of the winning first choice in the lottery on own arrests. Each is conditional on the set of student characteristics, and estimated win probability. Panel A includes separate estimates for each quintile of predicted arrest risk. Panel B includes a low-risk estimates, which pools quintiles 1-4, and high-risk estimate which includes quintile 5. Both panels are estimated using interaction terms between predicted risk indicators and the treatment. Standard errors are clustered at the application choice by year level.

Table A21: Effect of Peer Wins on Arrest Outcomes by Risk Quintile

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: By Risk Quintile						
Q5 (High)	0.063	0.186	-0.039	1.039	1.172	41.745
	(0.221)	(0.208)	(0.216)	(1.650)	(0.781)	(38.383)
Q4	0.307	0.054	0.124	0.880	-0.015	-41.909
	(0.226)	(0.178)	(0.187)	(1.181)	(0.536)	(47.247)
Q3	0.249	0.192	0.336**	2.205*	0.814*	94.640*
	(0.192)	(0.165)	(0.162)	(1.307)	(0.416)	(49.041)
Q2	0.335**	0.200**	0.197**	1.411**	0.588**	26.426**
	(0.130)	(0.086)	(0.093)	(0.557)	(0.243)	(10.433)
Q1 (Low)	0.314**	0.120*	0.161**	1.101***	0.342**	23.601***
	(0.153)	(0.072)	(0.078)	(0.419)	(0.169)	(8.039)
Observations	7891	7891	7891	7891	7891	7891
Panel B: $Q5$ and $Q1 - Q4$						
$(Wins/N_g) \times Q5$	0.075	0.195	-0.036	1.082	1.210	43.721
	(0.220)	(0.207)	(0.216)	(1.649)	(0.782)	(38.660)
$(Wins/N_g) \times (Q1 - Q4)$	0.324***	0.136*	0.207**	1.407**	0.413*	17.160
	(0.111)	(0.082)	(0.087)	(0.551)	(0.221)	(20.437)
Dep Var Mean, HR	.316	.233	.226	1.503	.655	29.28
Dep Var Mean, LR	.081	.039	.039	.231	.083	2.847
Observations	7891	7891	7891	7891	7891	7891

Notes: Estimated effect of the win proportion on outcomes. Each is conditional on the set of student characteristics, other lottery related group-level covariates (number of 2nd and 3rd choice wins, number of other wins, and number of other applications), and cohort, school, and CBG fixed effects. Panel A includes separate estimates for each quintile of predicted arrest risk. Panel B includes a low-risk estimate, which pools quintiles 1-4, and high-risk estimate which includes quintile 5. Both panels are estimated using interaction terms between predicted risk indicators and the treatment. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A22: Effects of Winning on Own Other Outcomes

	Math	Read	Absences	Suspensions Total Days	Days ISS	Days OSS	Any Susp.	On Time Grad.	Dropout
Panel A: Pooled									
Win	0.011	0.044*	-0.492	-0.294	-0.063	-0.231	-0.073	0.015	0.001
	(0.038)	(0.026)	(0.376)	(0.338)	(0.092)	(0.263)	(0.089)	(0.017)	(0.010)
Observations	4705	4697	7756	7234	7234	7234	7234	1313	1313
Panel B: Effect of Win, By Risk									
$Win \times HR$	0.059	0.051	-0.276	-0.379	-0.083	-0.295	-0.135	0.055	-0.013
	(0.046)	(0.038)	(0.598)	(0.708)	(0.182)	(0.556)	(0.185)	(0.036)	(0.020)
$Win \times LR$	-0.037	0.032	-0.654*	-0.179	-0.036	-0.143*	-0.009	-0.020*	0.015*
	(0.046)	(0.032)	(0.370)	(0.116)	(0.058)	(0.079)	(0.045)	(0.012)	(0.008)
Observations	4705	4697	7756	7234	7234	7234	7234	1313	1313

Notes: Estimated effect of winning first choice on lottery applicant outcomes. Each is conditional on the set of student characteristics, and estimated win probability. Standardized math and reading scores in columns 1 and 2 are from up to three years after the lottery, or 6th-8th grade for most students. Absences and Suspension include up to 5 years post lottery, or 6th through 10th grade. The last two columns show outcomes related to graduation and dropout. *On Time Grad.* is an indicator for graduating within 7 years post lottery (12th grade year for on-time progression). *Dropout* is an indicator equal to one if the student was ever observed as dropping out in the 7 years post lottery. The last two columns only include students who were observed graduating or dropping out in the NC data, and include up to one observation per student. Standard errors are clustered at the application choice by year level.

Table A23: Effect of Wining on Own Outcomes from Alternate Specifications

	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.
Panel A: Pooled								
Win	-0.011	-3.537	-0.012	-5.438*	-0.021	-2.920	-0.012	-3.727
	(0.014)	(2.973)	(0.017)	(2.856)	(0.013)	(2.858)	(0.014)	(2.988)
	[.014]	[2.947]	[.016]	[3.062]	[.014]	[2.807]	[.014]	[2.916]
Observations	1794	1794	1733	1733	1775	1775	1794	1794
Panel B: By Risk								
$Win \times HR$	-0.016	-11.579*	-0.002	-13.023*	-0.026	-9.336	-0.014	-9.843*
	(0.028)	(6.676)	(0.029)	(7.346)	(0.027)	(6.265)	(0.028)	(5.729)
	[.028]	[5.36]	[.03]	[6.003]	[.027]	[4.917]	[.027]	[4.566]
$Win \times LR$	-0.008	3.017	-0.020	0.849	-0.016	2.034	-0.009	1.225
	(0.011)	(2.810)	(0.017)	(2.569)	(0.013)	(3.104)	(0.011)	(2.912)
	[.011]	[2.838]	[.015]	[3.014]	[.012]	[2.969]	[.012]	[3.019]
Observations	1794	1794	1733	1733	1775	1775	1794	1794
Pr(Win)	✓	✓	✓	✓				
School FE			✓	✓				
CBG FE			✓	✓				
Lottery-Priority FE					✓	✓		
5pp Pr(Win) Bins							✓	✓

Notes: Estimated effect of winning the lottery on own probability of arrest and days incarcerated using alternate specifications as specified in the bottom of the table. Each is conditional on the set of student characteristics. Panel A includes estimates for all applicants. Panel B reports the same set of estimates for the high- and low-risk students, based on the estimated probability of arrest. *Wins* × *HR* shows estimates for the group with higher than median estimated risk, and *Wins* × *LR* shows estimates for the group with lower than median estimated risk. Each column in Panel B is estimated in a single regression using interaction terms. Two standard errors are included for each estimate. Standard errors clustered at the lottery level (application choice by year) are in parentheses, and standard errors clustered at the lottery by priority level are included in brackets.

Appendix B

Estimating Lottery Win Probabilities and Expected Win Shares

We start with the sample students who submitted at least one choice in the lottery over our sample period. The lottery proceeds in three rounds. In the first round, only the first choices of the applicants are considered. We focus on the first-round applications in our main analysis that meet the following criteria:

(1) Submitted at least one choice in the lottery.
(2) Their first choice was to a program that is not in their designated home school.
(3) Their first-choice program did not include subjective criteria in the admission process.[1]
(4) They did not receive a sibling placement, or other apparent magnet continuation placement.[2]

The resulting sample includes students for whom winning the lottery means that they will be assigned to a different school than their residentially zoned school based on their residence at the time of the lottery for 6th grade admissions, which takes place during their 5th grade school year. In some cases, students may apply to a program within their neighborhood school, but those applicants do not meet criteria (2), so they are not included in our main sample.

We group applicants based on application choice, year of application, and observable lottery criteria which include whether the student's neighborhood school or magnet continuation school is a Title I choice school (0 or 1), whether the student is economically disadvantaged (0 or 1), and whether the student was at grade level in reading on the fourth grade standardized reading exam (0, 1, or missing). We construct the predicted win probabilities, $\hat{P}_{ibst}$, using the proportion of applicants within the application choice, year, Title I School Status, economic disadvantage status, and reading level groups who won the first choice in the lottery. The lottery sample includes the set of students with $\hat{P}_{ibst}$ between zero and one who also met the entrance requirements for their first-choice program.

We use the predicted win probabilities to construct expected wins by summing win probabilities within CBG-School-Cohort to obtain an expected number of wins. We define the expected win proportion as the expected number of wins divided by group size (minus one). That is,

$$W_{bst} = \frac{1}{(n_{bst} - 1)} \sum_{i=1}^{n_{bst}} \hat{P}_{ibst}$$

Where n_{bst} refers to the number of students in the group, i.e., same census block group and school zone in the same 5th grade cohort. W_{bst} represents a CBG-School-Cohort specific

[1] Arts programs require an audition and leadership programs which require an interview. Both include admission criteria that are subjective and not observable.

[2] We filter out applicants that likely receive automatic placements which is inferred from the data and CMS documentation.

expected share of peers in the lottery applicant sample who will receive an assignment to a different school by winning their first choice in the lottery.

Constructing Group Level Lottery-Related Controls

The focal students in this analysis are non-applicants who have a win probability of zero. The students who contribute to the expected win share are lottery applicants who meet the lottery sample criteria. The sample generating the expected win share and the estimation sample are distinct. In addition, students who specify a lottery choice, but do not meet the criteria outlined above do not contribute to either sample. Effectively, we can view their assignments as non-random, and they would not contribute to identification. However, we do use this subsample to construct some of the lottery-specific conditioning variables. These additional conditioning variables account for sibling slots, non-first choices and other applicants not meeting the criteria for random lottery assignment, which may impact student expectations regarding lottery win probabilities and peer movement. However, conditioning on these additional lottery-related outcomes is inconsequential to our results.

The additional lottery-specific conditioning variables are:

1. *Number of students who won their second or third choice in the lottery*

The number of students who won their second or third choice in the lottery is constructed using the same sample of lottery applicants in oversubscribed lotteries who did not receive automatic placements and met the stated requirements for their program of application. We construct the total number of students in the group who won their second or third choice as the sum of an indicator for winning the second or third choice within the CBG-School-Cohort group.

2. *Number of other students submitting lottery applications*

As outlined in the previous section, we limit the lottery applicant sample to those who met several criteria, i.e., they are in an oversubscribed lottery and there is some random assignment that can be used for estimation. The remaining lottery applicants, e.g., who receive a sibling placement or apply to an undersubscribed lottery, are not used in either of the estimation samples. We construct two conditioning variables from this sample, the first of which is the number of applicants in the CBG-School-Cohort group who do not meet the criteria for the estimation sample.

3. *Number of other students assigned to their lottery choice*

Using the same set of students who applied but did not meet the criteria for the lottery estimation sample, we construct a variable for the number of applicants who were assigned to their lottery choice. This would include, for example, students who received sibling placements, lottery winners who applied to arts or leadership programs, or applicants to undersubscribed lotteries.

Estimating win probabilities in all rounds

In our robustness analysis we include main results using variation in wins from all three lottery rounds. Again, we focus on students who submitted at least one lottery application, did not receive a sibling placement in the first round, did not apply to their zoned neighborhood school in the first round, and otherwise did not have an inferred automatic continuation placement. We do include applicants to arts and leadership programs in this sample because they may win a seat a later round.

We calculate round specific win probabilities sequentially. We calculate win probabilities in round 1 using the same method we use in the main analysis, which focuses on the first round only. That is, we group students by application choice, year, title I default assignment status, economic disadvantage, and whether they were at grade level in reading = (0, 1, missing).[3] We calculate the win probability as the proportion of winners within these groups. Because admission to arts and leadership magnet programs include subjective criteria based on interviews and auditions, we treat these outcomes as deterministic outcomes.

In the next steps we calculate conditional win probabilities in the second and third round using the observed lottery-cohort specific win rates for previously unseated applicants.[4] Again, we remove students who received apparent automatic placements in this round (sibling placements, to their neighborhood school, or inferred continuations) and treat assignments to arts and leadership programs as deterministic. Among the students without automatic placements, we use the observed win rates to impute the conditional win probability for all others with that choice. That is, in each round we calculate the fraction of students with a given round specific lottery choice and year of application who won a seat in that round of the lottery and use that fraction as the probability of winning that choice, conditional on not winning any prior choice. We use this as the conditional win probability for all students who made that choice, regardless of whether they received an earlier placement in the lottery.

In rounds 2 and 3 we do not make use of the priority groups. In the CMS assignment, priorities are not absolute, such that students awarded a seat through a given priority are capped as a proportion of available seats in that program. We also impute conditional probabilities for students with applications to lotteries with no observed variation in that round. For example, if all students to a given second round application received a placement in the first round, we have no observed placements to generate the second-round conditional win probability. In these cases, we impute a conditional win probability using the average win rate in that round and year.

At this stage we have three probabilities: probability of winning in the first round (Pr(win1)); probability of winning in the second round conditional on not winning in the first round (Pr(win2 | win1 = 0)); and probability of winning in the third round conditional on not winning in round 1

[3] We use fourth grade reading level when available.
[4] We are limited in our computation because we do not observe published capacities for each program.

or 2 (Pr(win3 | win1=0, win2=0)). From these, we construct unconditional probabilities of winning each round, and compute the estimated probability that the student wins a seat to a program outside of their neighborhood school. This is the treatment of interest, which aligns with our analysis using first round choices only, in which we focus on lottery placements that lead to movement across schools.

In this case, our lottery sample includes all applicants who met the criteria who have a probability of receiving a placement to a different school than their residentially assigned neighborhood school between zero and one. In addition, we drop students who were ultimately assigned to a program with subjective admission criteria (arts and leadership) and students with a positive probability of winning a seat to a program for which they did not meet the entrance requirements. The resulting set is the lottery applicant sample for the analysis including all rounds. Like the analysis using first round results only, we use this sample to create the predicted and actual group level win proportions at the CBG-School-Cohort level. Additionally, we create two conditioning variables for the number of other lottery applications and wins.

Estimating Arrest Risk

We start with the full sample of students in the district from the three cohorts that we use (Column 1 of Table 1). We predict the probability of arrest conditional on a set of individual characteristics, and school and CBG level means. The dependent variable is an indicator for any arrest from ages 16 to 22. The explanatory variables are indicators for race/ethnicity, economic disadvantage, 5th grade math and reading scores, dummy variables for missing test scores, indicators for fifth grade exceptionality, age, a dummy for missing age, CBG level means of each race/ethnicity indicator, economic disadvantage, math and reading test scores, and school attendance zone level means of each race/ethnicity indicator, economic disadvantage, math and reading test scores. We estimate the parameters using a logistic regression, and predict the individual probability of arrest, $\hat{R}_{ibst}$. In our main heterogeneity analyses, e.g., Panel B of Table 3, we split the sample into groups with high and low arrest risk using the sample median $\hat{R}_{ibst}$ to create risk indicators.