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

June 19, 2023

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 decrease adult criminality. Putting these results together, we show that school choice lotteries increase overall arrests and days incarcerated for young men.

We thank seminar participants at UNLV, 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, 2011; Abdulkadiroglu, Pathak, & Walters, 2018; Cullen, Jacob, & Levitt, 2006; Hastings, Kane, & Staiger, 2006; Hastings & Weinstein, 2008; 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). However, this literature has not examined the effect of winning a lottery on the subsequent school discipline and criminal justice outcomes of the students left behind, although in related work Lavy (2010) shows that the implementation of a school choice system in Tel-Aviv, Israel on net reduced the district level of violence and classroom disruption.

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 combine the indirect effects (on peers left behind) with

the direct effects (on lottery winners) to provide a more comprehensive assessment of the overall impact of school choice lotteries on young adult criminal justice outcomes.

While prior literature suggests 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). 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). Alternatively, a number of peer effects 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 living in the same neighborhood, are similar in age and attend the same middle school (Billings, Deming, and Ross, 2019). Our empirical analysis includes a sample of three cohorts of male 5th grade 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. To isolate the random portion of the lottery assignment, we condition on the share of the neighborhood group expected to win the lottery.[1] This estimation strategy isolates the effects of changes in the school attendance patterns of lottery winners that occur simply due to chance, and not due to variation in lottery priorities or application rates in each neighborhood. As a pseudo first stage, we confirm that lottery winners are substantially more likely to attend a non-neighborhood school than lottery losers and lottery losers are more likely to move to a different neighborhood.[2]

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 (4.4% of same grade male students in neighborhood) implies a 0.04 increase in the number of arrests from ages 16-22, which is an 8.3% increase over the sample average. The effects are more precisely estimated and

[1] Our estimation strategy is similar in concept to a control function approach (Wooldridge, 2015). We argue that the share of lottery winners from a CBG-School-Cohort is exogenous to adult crime once we condition on the expected win share. 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. We also include fixed effects for CBG, Middle School attendance zone and cohort to address spatial and temporal impacts on expectations about win shares.

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

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 24.7% increase in the probability of arrest, 59.5% more violent arrests and 84.6% more days incarcerated.[3]

Lottery applicants have substantially higher test scores than non-applicants, and thus the loss of these high-achieving, neighborhood peers from the school environment appears to negatively affect student's later life outcomes. This 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.[4] We also show that the effects on arrests persist and increase in absolute magnitude after age 19 when most youth are no longer in high school. 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). Consistent with a peer effects mechanism, we provide additional evidence that the effects are not related to underlying changes in neighborhood middle school characteristics and that increases in adult criminality are stronger for non-applicants who attended elementary school with lottery winners. Together these results support the importance of peer influences that 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.

[3] While the percentage changes are large for the low-risk subsample, absolute changes are more modest since the group affected is at lower risk of arrest based on observables.

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

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 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 that is based on cohort variation in unexpected win shares, all of which lead to 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 specific neighborhood trends may correlate with unexpected win probabilities.

Since lottery applicants tend to be positively selected on academics, the finding that most of our 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 substantial declines in both number of arrests and days 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 at least 72 fewer arrests and 2,497 fewer days incarcerated and generated costs to non-applicants of at least 100 more arrests and 5,434 more days incarcerated. In aggregate, we estimate at least a 0.6% net increase in total arrests and 3.9% more days incarcerated due to the lottery application process for our three cohorts of 5th grade students in our estimation sample (n=10,032). 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 sorting that would occur in the absence of school choice.[5]

[5] 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.

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 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 in North Carolina, while Mumma (2020) finds no effects in Massachusetts and North Carolina for students who attend schools near charter school openings.[6] 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 positive effects on test scores for lottery winners and no effects on test scores for students left behind, consistent with Gilraine, Petronijevic, & Singleton's (2021) study of North Carolina.

Our findings are especially important because school choice has proven popular among local voters and has helped to reduce the rates at which high-quality peers leave more disadvantaged school districts. More targeted school choice programs that promote school choice among lower-achieving students might change the type of selection into choice documented here

[6] 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.

as well as by Hastings, Kane, & Staiger (2006), Burgess et al. (2015) and Barseghyan et al. (2019). A more 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.[7] Nonguaranteed seats are assigned in three rounds, considering only first choices in the first round

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

of assignment and siblings of current students are guaranteed admission. 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 criteria based on geographic location and Title I choice status of the student's neighborhood school.[8] 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, he or she is assigned to their neighborhood school based on pre-specified school 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. The share of applications who won their first choice is about 35% over our sample period, which includes those entering sixth grade in the 2006-2007 to 2008-2009 school years.

[8] 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 individual winners as well as students who forego the lottery, 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. We also restrict to students who met the requirements for their first-choice program.[9] 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 (69%) 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 31% 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. We focus on the 6th grade lotteries for these cohorts, which allows us to observe the sample as adults. The data include student gender, race, an indicator for economic disadvantage, yearly

[9] 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.

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 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, and information on the number and nature of charges.[10] 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 between the ages of 16 and 22.[11] While we observe the future criminal behavior of CMS students, regardless of whether they transfer or drop out of school, they are limited to arrests in Mecklenburg County.[12]

[10] The Mecklenburg County Sheriff (MCS) tracks arrests across individuals using a unique identifier that is established with fingerprinting.
[11] Individuals arrested at age 16 or 17 were automatically charged in the adult criminal justice system in North Carolina until Dec 1, 2019.
[12] 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, it is unlikely that

Because we focus on criminal justice outcomes and arrests are overwhelmingly male (80%), we limit our sample to 13,493 boys observed in CMS in these cohorts, including 4,166 who specified some choice in the lottery. We exclude 2,025 students from our sample of applicants because they were either guaranteed admission, specified a choice in their neighborhood school, in lotteries 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.[13]

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, which leaves 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 2,141 6th grade lottery applicants and 7,891 non-applicants of which 5,433 had at least one applicant in their neighborhood and cohort.

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

there are significant numbers of arrests of CMS youth in the surrounding counties, as evidenced by the fact that there are very few young adult arrestees in Mecklenburg County who were not observed in CMS schools.

[13] 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.

students had at least one arrest and 8% were incarcerated at least one time from ages 16 to 22, which suggests a substantial amount of 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 in the sample includes around 20 students, which reflects a scale commonly associated with likely peer groups.

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 sample averages also highlight the relatively high test scores of lottery applicants, and show that applicants are more likely to be black rather than white or Hispanic. The higher achievement of lottery applicants is consistent with households that have strong preferences for school quality. Thus, we would expect 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. We focus on 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

by conditioning on the expected group level win share.[14] Because the lottery win probabilities depend on observables, we construct expected win shares based on student-specific win probabilities. We predict student-specific win probabilities using lottery-specific win shares for the student's first choice in their school choice application, as well as dummy variables for characteristics that determine lottery priority groups.[15] Let $\hat{P}_{ibst}$ 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.[16] We construct group level expected win shares (w_{bst}) using 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.[17] 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

[14] In our context, we can explicitly use factors that impact expected lottery wins for a specific program to predict win probabilities and thus we can add control variables (function thereof) to limit variation in our main independent variable to be conditionally random. Wooldridge (2015) and Petrin & Train (2010) provide a good discussion on control function approaches in econometrics.

[15] The prediction is based on estimates from probit regressions of a dummy variable for winning their first choice in the lottery on a vector of lottery observables. In all cases, the model includes the win ratio of applicants who applied to the same program as a predictor. For non-magnet applications, we also include year-specific dummies for the applicant scoring below grade level in reading and applying to an above average school in reading, applying to a non-Title I Choice school from a Title I Choice school, economically disadvantaged student applying to a non-Title I Choice school, and for living in the transportation zone of the application school. For magnet applications, we include year-specific dummy variables for applying to a non-Title I Choice school from a Title I Choice School, economically disadvantaged student from a Title I Choice school applying to a non-Title I Choice school, and for living in the transportation zone of the application school. In all cases, we also include the interaction term for each dummy with the lottery-specific ratio of applicants who won their first choice in the given year. See Appendix for a more detailed discussion of estimating expected win shares.

[16] Results are robust to using the lottery-specific win ratios in place of the predicted win probabilities, where the win ratios are a key predictor in the regressions that we use to create win probabilities.

[17] 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.

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 number 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 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 cluster standard errors at the CBG by school attendance zone by cohort level, which is the level of randomization. Later, we show that results are robust to both the fixed effect structure and clustering at the CBG by school attendance zone level.

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

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. For (3) and (4), "other" refers to applications that are not in the applicant sample, like 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.[18]

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 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*.

$$X^k_{ibst} = \beta_{1k} Z_{bst} + \beta_{2k} W_{bst} + \delta_{bk} + \eta_{sk} + \gamma_{tk} + \varepsilon_{ibstk} \quad (2)$$

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} =

[18] For those missing a lagged test score, we use the mean value, and include a dummy variable for missing lagged score. 6% of the non-applicant sample is missing at least one lagged test score.

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 0.93, 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.[19]

Finally, we estimate heterogeneous effects across students by interacting Z_{bst} with a student attribute indicator variable (X^{k}_{ibst}). 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 ages 16-22.[20] 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.

$$y_{ibst} = \beta_1^{HR} Z_{bst} \times HR_{isbt} + \beta_1^{LR} Z_{bst} \times LR_{isbt} + \beta_2^{HR} W_{bst} \times HR_{isbt} + \beta_2^{LR} W_{bst} \times LR_{isbt} +$$

[19] 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).

[20] We estimate a logistic regression model for a dummy variable for any arrests between ages 16-22 on individual, CBG, and neighborhood school level covariates. The results of estimating predicted arrest are given in Appendix Table A2. The individual level predictors are a set of race / ethnicity dummy variables, 5th grade math and reading scores, and dummy variables for economically disadvantaged and exceptionality, 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.

$$\beta_3 X_{ibst} + \delta_b + \eta_s + \gamma_t + \varepsilon_{ibst} \quad (3)$$

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 vector of lottery-related variables, W_{bst}, with the set of dummy variables.

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.[21] 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.[22] 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

[21] Each regression follows the existing literature (e.g. Cullen, Jacob, & Levitt, 2006; Deming, 2011) and is conditional on lottery fixed effects (application choice by cohort), other lottery-related controls, 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.

[22] We estimate individual crime risk based on a set of student attributes prior to 6th grade and to construct a composite measure of a student's likelihood of being arrested age 16-22 based on demographics and elementary school performance/attributes or school attendance zone and CBG fixed effects. The crime risk estimation results are provided in Appendix Table A2.

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 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 and 5 are dummy variables for whether the student changed assigned neighborhood school (*Change NS*), a sign of residential movement, and a dummy for exiting the district (*Exit*), respectively. These results indicate limited attrition by lottery winners in terms of residential relocation or exit from the school district, which suggests that much of the gap between estimates for attending first choice school and not attending the neighborhood school arises from losers winning the lottery for non-first choice schools.[23] Columns 6 and 7 provide evidence that lottery winners are attending 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.

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 non-applicant decision to attend their neighborhood school. 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 limited evidence that non-applicants are sorting with respect to lottery winners both overall as well as for high and low arrest risk boys. Our only statically significant effect shows that an increase in win

[23] 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.

share from no winners to the average proportion of winners in a neighborhood leads to a 5.6% decrease in the probability of changing neighborhood schools.[24]

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 often sizable in magnitude, but imprecisely estimated with only number of arrests and number of violent arrests significant at the 10% level. To interpret coefficients, we use the change from no winners to the average proportion of winners in a neighborhood (4.4% of same grade male students). We find a 0.04 (0.872*0.044) increase in the number of arrests as a young adult, which is an 8.3% (0.04/0.46) increase over the estimation sample average. 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 1.4 percentage point increase in the probability of arrest, 0.02 more violent arrests and 1 more day incarcerated. Given lower rates of arrest and incarceration for this subsample, the estimates imply increases of 24.7, 59.5 and 84.6% of the respective means. All estimates are significant at a 1% 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

[24] Specifically, we find that a change from no winners to the average proportion of winners in a neighborhood (4.4% of same grade male students), indicates a 0.015 (-0.336*0.044) decrease in the probability of changing neighborhood schools, which is a 5.6% (0.015/0.263) decrease over our estimation sample average.

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. Further, in Appendix Table A4, we show 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.

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 in Appendix Table A5. The final two columns of Appendix Table A5 include school 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. Appendix Table A5 also includes standard errors clustered at the block group by attendance zone level. 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.

Heterogeneity

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 also based on Equation 1 using absences and suspensions in the years following the lottery as outcomes. The results suggest that some children left behind by lottery winners experience substantially worse outcomes in these dimensions. Panels A and B include up to 5 years post-lottery corresponding to grades 6 through 10. In the full sample estimates in Panel A, only the estimate for absences is statistically significant at the 10% 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 average number across our sample increases the number of absences by 0.33 (0.044*7.40) per year from 6th through 10th grade, which is an increase of 5.5% (0.32/5.87) over the mean absences among low-risk individuals. In addition, an increase in the win share from zero to the sample average increases the number of days in out of school suspension by 0.12 (0.044*2.74) per year from 6th through 10th grade, which is a 19% (0.12/0.63) increase over the sample average among low-risk non-applicants.

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 do not find significant effects on test scores or high school graduation, as shown in Appendix Table A6.

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, we estimate the main effects over two age ranges. Table 6 includes the estimated effects of the win share on low-risk non-applicant arrest and incarceration outcomes 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 28% ((0.044*0.17)/0.027)) 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, 36% ((0.044*0.307)/0.038)), 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 appear 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 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 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.

Finally, we recognize that two key correlates of likelihood of future arrest 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 A7. Panel A includes results for race and ethnicity and Panel B for student 5th grade test scores. The results for white students closely parallel the results for students with low risk of arrest in significance and magnitude. In fact, the estimated effect on days incarcerated is significant and implies an extra 1.73 (0.044*39.355) days incarcerated for white students exposed to the sample average win share, as opposed to no winners.[25] The estimated effects for students with above median test scores are all positive and statistically significant for likelihood of arrest and likelihood of violent arrest, but in several cases (number of arrests and days incarcerated) smaller in magnitude than estimates for

[25] We also examine whether the effects on those left behind are stronger among same race relationships controlling for share of wins among lottery applicants of the same race. We formally test this in Appendix Table A4 and do not find any evidence of additional effects on those left behind by lottery winners of the same race.

below median test score students. Appendix Table A8 expands the analysis in Panel 1 by examining effects by race and risk of arrest. We find effects for white students irrespective of their risk of arrest. In fact, in absolute terms, estimates 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.

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, we generate a series of figures that correlate win shares with school attributes 3 years post-lottery. Each dot represents the win share and characteristic combination for one neighborhood school-cohort observation. From the left-hand side of Figure 2, actual win share has a negative and statistically significant correlation with school by cohort average test scores and portion of students that are white, and a positive and significant correlation with economically disadvantaged student shares three years after a cohort's lottery. These results are consistent with a higher participation in the lottery for kids assigned to lower test score and higher economically disadvantaged neighborhood schools. The right-hand side of Figure

2 includes the analogous correlations between each characteristic and wins minus expected wins as a share of cohort size, highlighting that these correlations disappear or change sign when controlling for the expected win share. All three correlations are statistically insignificant. These results suggest that neighborhood schools are not changing along these dimensions in response to unexpected lottery wins as well as providing additional evidence that our control function approach address selection into the lottery.

The main result that having more same-neighborhood lottery winners increases arrests and incarcerations, primarily among low-risk non-applicants, supports a peer effect channel, as applicants are positively selected on academic performance which correlates with low arrest risk. Additionally, the response is primarily for behavioral outcomes, i.e., arrests, incarcerations, and middles 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 A9. 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. Panel B provides results for a model with more granular controls by including school 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. Coupling these two sets of results supports a primary role for peer influences in our main results where unexpected lottery wins increases 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 A10 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 resulting 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 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 white students and among students who are at low risk of being arrested.[26] Highlighting the complexity of peers groups in our context, Lavy and Sand (2019) find heterogenous effects on academic and behavioral outcomes for students based on the nature of friendships and number of friends that attend the same middle school.

Lottery Winners and Adult Crime

To provide an overall assessment of the impact of school choice on adult crime, we incorporate 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. We limit our analysis to the sample of lottery applicants and focus on comparing winners and

[26] 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.

losers. We initially mirror the sample split in the earlier analysis by using above versus below median arrest risk, but also provide splits by quintile of arrest risk consistent with Deming (2011).

Specifically, we model arrests and incarceration (y_{ibst}) of lottery applicant *i* who applied in lottery *l*, which represents an application choice by year combination. As discussed earlier, we include controls for lottery fixed effects (ρ_l) and control for other characteristics that affect priority groups, which are captured in student attributes (X_{il}).[27] We include additional controls in X_{il} for race and ethnicity, economic disadvantage, 4th grade math and reading scores, and ELL status.[28] The specification is displayed in Equation 4.

$$y_{il} = \gamma_1 W_{il} + \gamma_2 X_{il} + \rho_l + \varepsilon_{il} \quad (4)$$

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 lottery fixed effects and other controls that affect lottery priority, 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.

[27] Lottery fixed effects are application choice by year indicators. The additional controls in X_{il} include year-specific dummy variables for economic disadvantage, Title I choice school, interactions between economically disadvantaged and Title I choice, below grade level in reading and applying to a nonmagnet with above average reading scores, scoring at grade level in both math and reading in 4th grade, academically gifted status, interactions between grade level achievement and academically gifted status with applying to an IB program, We also include single dummies for grade level in 4th grade math and reading, a dummy for missing grade level info, and two dummies indicating geographic proximity to a full magnet school that the student applied to.

[28] 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.

Table 7 provides descriptive statistics for the applicant sample and standard balancing test to show that lottery variation is random in this sample of students. The first two columns contain means and standard deviations among lottery winners and losers in our sample. Balance tests are included in columns 3 through 5. 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 when no lottery fixed effects or other controls are included. In columns 4 and 5, we include lottery fixed effects and other controls that affect priority groupings. 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). As with prior work, our variation is limited to oversubscribed programs.

The main results from estimating Equation 4 are included in Table 8. 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, in the pooled sample results shown in Panel A, we find that lottery winners are 2.5 percentage points less likely to be incarcerated anytime between ages 16 to 22, which is a 23% decrease off the mean incarceration rate for lottery applicants. In the sample of low-risk lottery applicants, we find significant declines in both arrests and incarcerations with the most precise estimates on the probabilities 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 more noisy.

Since prior work (Deming, 2011) focused on those with the highest ex-ante crime risk (top quintile), we replicate his analysis using our sample of lottery applicants and focusing on middle school students. Appendix Table A11 provides results for 5 groups of lottery applicants based on quintiles of ex-ante arrest risk for each outcome. Comparable to Deming (2011), we find statistically significant declines in the highest quintile of arrest risk on the intensive margin for both arrests and incarceration. However, we also estimate sizable and statistically significant declines on several outcomes in the second and third quintiles.[29]

Appendix Table A13 includes estimated effects of winning the lottery on test scores, absences, suspensions, and graduation for lottery applicants. In contrast to the earlier estimated effects on non-applicants for the same outcomes in Tables 5 and A6, we do not find any statistically significant effects of winning the lottery on the winner's number of absences or suspensions. However, we observe positive effects of 0.07 to 0.10 standard deviations on test scores for high-risk lottery winners, which we do not find for winner's non-applicant peers. Therefore, while not the focus of our paper, the school choice system appears to lead to slightly higher test scores for winners with no offsetting test score losses on the lottery winner's local neighborhood peers, similar to Gilraine, Petronijevic, & Singleton (2021) in North Carolina and Muralidharan & Sundararaman (2015) in India.

Net Benefit Analysis

To this point, one of the main takeaways is that some lottery winners benefit from lower arrest and incarceration rates as adults, while the departure of lottery winners increases adult arrests

[29] Appendix Table A12 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.

and incarcerations among some non-applicants. Given the heterogeneity in both benefits and costs, a formal accounting of effects by different levels of crime risk is required to assess the net benefit of constrained school choice on criminal justice outcomes. To implement this analysis, Table 9 provides two sets of results. Panel A is based on a simple split by above and below median crime risk and Panel B uses risk quintiles. This Table breaks down 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 from the previous section. 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 343 and 397, respectively, which are displayed under *E[Wins]* in the first column of Panel A. From Table 8, the estimated effects of winning the lottery among high- and low-risk applicants on the number of all arrests are -0.16 and -0.13, respectively. Using the risk-specific expected number of wins and estimates, we estimate that the lottery decreased all arrests by about 53 and 51 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 -104. 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. Significant estimates and the benefits/costs arising only from significant estimates are always indicated by bold numbers, which results in an estimated effect of 51 fewer arrests among lottery applicants.

In the second half of Panel A, we estimate the corresponding aggregated change in the number of arrests 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, 0.61 from Table 3, across all high-risk students in the non-applicant sample, which suggests that the lottery increased arrests by 126 among this group. Similarly, we find an estimated increase in arrests of 194 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 quintiles of crime risk, suggest that aggregate arrests increased by 170 when considering all estimates or 28 when aggregating significant coefficients only. Summing across the affected populations, we estimate 3,539 additional incarceration days when aggregating across all coefficients or 2,937 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 who are at below median arrest risk, and to some extent, among students who are white and scored higher on elementary school tests. The larger effects for students at low risk of future arrest appears consistent with effects for students who would have been more likely to interact socially with the positively selected lottery applicants. We also observe evidence of negative behavioral effects in middle school in terms of attendance and suspension, also among students at low risk of future arrest.

Estimating the direct effects on criminality on the applicants themselves, lottery winners, on the other hand, 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. We show through an accounting of costs and benefits to adult crime that school choice increased the total amount of 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

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.

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.

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.

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

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.

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.

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).

Reed, Jordan, 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.

Petrin, Amil, and Kenneth Train. "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47(1) (2010), 3-13.

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	2.81	4.55	2.60	
	(3.31)	(3.55)	(3.23)	
Group N	18.66	16.68	21.65	
	(16.85)	(15.18)	(17.52)	
Num. Wins	0.96	1.57	0.90	
	(1.42)	(1.76)	(1.36)	
Black	0.42	0.52	0.36	-0.02
	(0.49)	(0.50)	(0.48)	(0.12)
White	0.36	0.29	0.42	-0.08
	(0.48)	(0.45)	(0.49)	(0.09)
Hispanic	0.14	0.12	0.13	0.10
	(0.35)	(0.32)	(0.34)	(0.10)
Ec. Disadvantaged	0.48	0.47	0.45	0.03
	(0.50)	(0.50)	(0.50)	(0.10)
Math EOG, 5th Grade	0.13	0.33	0.13	-0.25
	(1.00)	(0.95)	(1.01)	(0.22)
Reading EOG, 5th Grade	-0.01	0.21	0.00	-0.21
	(1.00)	(0.95)	(1.01)	(0.23)
Missing Test Score	0.07	0.03	0.06	-0.00
	(0.26)	(0.17)	(0.24)	(0.07)
ELL	0.10	0.06	0.10	-0.01
	(0.31)	(0.24)	(0.30)	(0.09)
Observations	13,493	2,141	7,891	7,891
S.E. Clusters				1,134
P-value				0.929

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. Disadvantaged* 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 in 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 the 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				Att. School Mean	
	Won Any	App School	Neighb. Sch.	Change NS	Exit	Math	Pr(Arrest)
Panel A: Pooled							
Win	0.748***	0.680***	-0.355***	-0.054*	-0.014	0.254***	-0.012***
	(0.018)	(0.029)	(0.030)	(0.030)	(0.024)	(0.052)	(0.003)
Sample Mean	.512	.378	.267	.272	.121	.052	.128
N	2140	2140	2140	2140	2140	1975	1872
Panel B: By Risk							
$Win \times HR$	0.763***	0.704***	-0.329***	-0.005	0.005	0.273***	-0.021***
	(0.019)	(0.024)	(0.028)	(0.032)	(0.022)	(0.058)	(0.005)
$Win \times LR$	0.732***	0.660***	-0.376***	-0.097**	-0.029	0.236***	-0.004
	(0.029)	(0.043)	(0.040)	(0.038)	(0.035)	(0.059)	(0.003)
N	2140	2140	2140	2140	2140	1975	1872

Notes: Estimated effect of winning first choice on attendance and movement of applicants. Each estimate is conditional on the set of student characteristics, a set of covariates related to the lottery priority groups, and lottery fixed effects. 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* refers to estimates for the group with higher than the 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. *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, likely indicating 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						
$(Wins/N_g)$	0.091	0.084	0.002	0.872*	0.432*	14.860
	(0.097)	(0.079)	(0.083)	(0.529)	(0.234)	(14.679)
Observations	7,891	7,891	7,891	7,891	7,891	7,891
Panel B: By Risk						
$(Wins/N_g) \times HR$	-0.037	0.028	-0.130	0.609	0.354	8.789
	(0.137)	(0.117)	(0.121)	(0.800)	(0.352)	(22.314)
$(Wins/N_g) \times LR$	0.309***	0.176***	0.223***	1.311***	0.568***	23.976***
	(0.089)	(0.066)	(0.070)	(0.415)	(0.205)	(8.965)
Dep Var Mean, HR	.207	.136	.135	.859	.361	15.284
Dep Var Mean, LR	.055	.022	.021	.129	.042	1.247
Observations	7,891	7,891	7,891	7,891	7,891	7,891

Notes: Estimated effects of the peer win proportion on outcomes. 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)* $\times$ *HR* shows estimates for the group with higher than the 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. In Panel B, we also include interaction terms between the high-risk dummy variable 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 $(Wins/N_g)$	0.007	-0.033	-0.037	-0.073	0.043	25.815
	(0.115)	(0.096)	(0.093)	(0.727)	(0.358)	(19.310)
Observations	7,636	7,636	7,636	7,636	7,636	7,636
Panel B: By Risk						
Lag $(Wins/N_g) \times HR$	0.052	-0.083	-0.049	0.035	0.081	35.122
	(0.156)	(0.124)	(0.122)	(0.959)	(0.469)	(25.741)
Lag $(Wins/N_g) \times LR$	-0.133	0.038	-0.045	-0.350	-0.099	7.375
	(0.149)	(0.105)	(0.100)	(0.592)	(0.287)	(18.147)
Observations	7,636	7,636	7,636	7,636	7,636	7,636

Notes: Estimated effect of the win proportion from the same school-cbg in the following year on outcomes. 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 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 the 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. In Panel B, we also include interaction terms between the high-risk dummy variable 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					
$(Wins/N_g)$	3.487*	0.475	-0.017	0.492	-0.185
	(2.060)	(1.786)	(0.432)	(1.476)	(0.513)
Dep Var Mean	7.674	2.814	.713	2.101	.96
Observations	32,993	31,717	31,717	31,717	31,717
Panel B: By Risk					
$(Wins/N_g) \times HR$	1.538	-1.161	-0.450	-0.710	-0.714
	(2.809)	(2.657)	(0.641)	(2.187)	(0.746)
$(Wins/N_g) \times LR$	7.401***	3.502**	0.766*	2.736**	0.780*
	(2.246)	(1.372)	(0.390)	(1.082)	(0.450)
Dep Var Mean, HR	9.961	5.181	1.249	3.932	1.72
Dep Var Mean, LR	5.874	.906	.28	.626	.348
Observations	32,993	31,717	31,717	31,717	31,717
Panel C: Grades 6-8, Low Risk					
$(Wins/N_g) \times LR$	6.106**	4.285***	1.074**	3.211***	0.973**
	(2.443)	(1.585)	(0.522)	(1.204)	(0.481)
Dep Var Mean, LR	5.496	.94	.301	.64	.339
LR Observations	11,606	10,654	10,654	10,654	10,654
Observations	20,763	19,249	19,249	19,249	19,249

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 Suspensions. 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, pooling across estimated risk. 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 the 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 each estimated using interaction terms with the win proportion and high- and low-risk. In Panels B and C, we also include interaction terms between the high-risk dummy variable 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						
$(Wins/N_g) \times LR$	0.170***	0.082*	0.119***	0.297**	0.111	2.949
	(0.062)	(0.048)	(0.042)	(0.135)	(0.073)	(2.596)
Dep Var Mean, LR	.027	.011	.01	.048	.016	.447
Observations	7,891	7,891	7,891	7,891	7,891	7,891
Panel B: Low Risk, Age 19 to 22						
$(Wins/N_g) \times LR$	0.307***	0.175***	0.175***	1.013***	0.457***	21.027***
	(0.082)	(0.057)	(0.063)	(0.333)	(0.166)	(7.708)
Dep Var Mean, LR	.038	.014	.014	.081	.026	.8
Observations	7,891	7,891	7,891	7,891	7,891	7,891

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 in this table we provide 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. We also include interaction terms between the high-risk dummy variable 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	Uncond.	Lottery FE	Lot FE w/cons
Made 2nd Application Choice	0.79	0.79	0.02	-0.02	-0.03
	(0.41)	(0.41)	(0.04)	(0.04)	(0.04)
Made 3rd Application Choice	0.55	0.53	0.04	0.04	0.03
	(0.50)	(0.50)	(0.03)	(0.02)	(0.02)
Black	0.46	0.55	-0.07	-0.03	-0.03
	(0.50)	(0.50)	(0.05)	(0.04)	(0.04)
White	0.35	0.25	0.02	0.05	0.05
	(0.48)	(0.43)	(0.06)	(0.04)	(0.05)
Hispanic	0.09	0.13	-0.08	-0.03	-0.03
	(0.29)	(0.33)	(0.07)	(0.06)	(0.05)
Ec. Disadvantaged	0.41	0.50	-0.03	-0.05*	
	(0.49)	(0.50)	(0.03)	(0.03)	
Math EOG, 4th Grade	0.40	0.20	0.03	0.00	0.00
	(0.99)	(0.94)	(0.02)	(0.01)	(0.02)
Reading EOG, 4th Grade	0.28	0.12	0.00	-0.01	-0.00
	(0.95)	(0.92)	(0.02)	(0.01)	(0.02)
Test Miss	0.06	0.09	-0.06	-0.05	-0.04
	(0.24)	(0.28)	(0.04)	(0.04)	(0.03)
ELL	0.05	0.07	-0.02	-0.03	-0.02
	(0.21)	(0.25)	(0.05)	(0.05)	(0.05)
Observations	753	1,388	2,141	2,140	2,140
S.E. Clusters			62	61	61
P-value			0.02	0.09	0.28

Notes: Summary of characteristics in the lottery sample. *ELL* is an indicator for English Language Learner. *Ec. Disadvantaged* is an indicator for economically disadvantaged. The first two columns show the mean and standard deviation for lottery winners and losers. The last three columns display estimated coefficients from regressing the dummy for winning the first choice in the lottery on the listed characteristics, conditional on other lottery-related control variables and lottery fixed effects (application by year). 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, conditional on other lottery features and fixed effects. The column labeled *Uncond.* reports coefficients without conditioning on any fixed effects or covariates. The *Lottery FE* column conditions on the application choice by year fixed effects only. The *Lottery FE w/cons* column reports estimates when conditioning on the lottery fixed effects plus other lottery-related control variables.

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	-0.012	-0.014	-0.025**	-0.148	-0.067	-3.282
	(0.015)	(0.014)	(0.012)	(0.129)	(0.072)	(3.179)
Observations	2,140	2,140	2,140	2,140	2,140	2,140
Panel B: By Risk						
$Win \times HR$	-0.002	0.003	-0.032	-0.155	-0.054	-5.759
	(0.030)	(0.027)	(0.022)	(0.238)	(0.136)	(6.397)
$Win \times LR$	-0.019	-0.029***	-0.019**	-0.129*	-0.073*	-1.082
	(0.013)	(0.009)	(0.008)	(0.069)	(0.037)	(3.047)
Observations	2,140	2,140	2,140	2,140	2,140	2,140

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, other covariates related to lottery outcomes, and lottery fixed effects. 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 the 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.

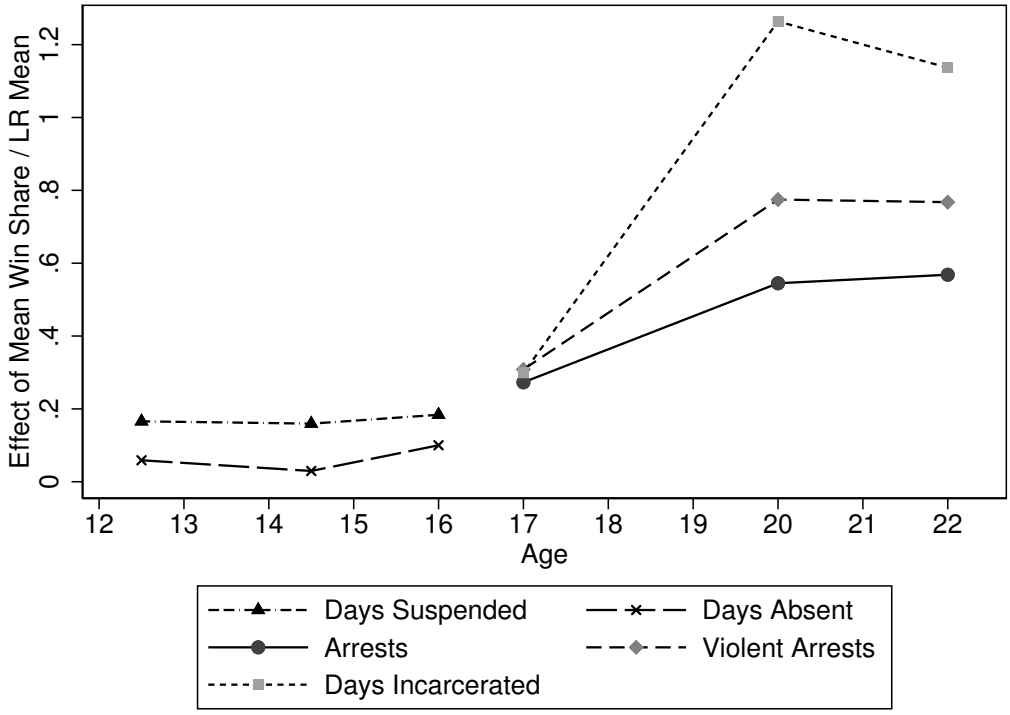
Table 9: Aggregated Effects

Panel A: By Above and Below Median Risk		Aggregated Effects		
		All Arrests	Violent Arrests	Days Incarcerated
	$E[Wins]$			
High-Risk Applicants	342.92	-53.15	-18.51	-1974.89
Low-Risk Applicants	396.71	**-51.17**	**-28.96**	-429.24
Total Effect on Applicants		-104.32	-47.47	-2404.13
Significant Effects		**-51.17**	**-28.96**	
	$E[Displacements]$			
High-Risk Peers	3393.25	126.24	73.38	1821.95
Low-Risk Peers	3608.23	**193.64**	**83.89**	**3541.5**
Total Effect on Peers		319.88	157.27	5363.45
Significant Effects		**193.64**	**83.89**	**3541.5**

Panel B: By Risk Quintile		Aggregated Effects		
		All Arrests	Violent Arrests	Days Incarcerated
	$E[Wins]$			
Q5 (High)	105.65	**-47.65**	-19.44	**-1846.18**
Q4	153.55	18.58	19.50	793.70
Q3	177.50	-42.42	**-30.00**	-196.50
Q2	141.81	**-24.53**	**-14.61**	**-651.05**
Q1 (Low)	161.13	-7.41	-3.38	-348.53
Total Effect on Applicants		-103.44	-47.93	-2248.56
Significant Effects		**-72.18**	**-44.60**	**-2497.23**
	$E[Displacements]$			
Q5 (High)	1115.74	18.46	55.54	2592.53
Q4	1572.08	85.30	-0.18	-2239.52
Q3	1444.51	69.11	**63.17**	**3663.49**
Q2	1522.98	**67.14**	**31.05**	**1138.37**
Q1 (Low)	1346.27	**33.27**	7.07	**632.41**
Total Effect on Peers		273.27	156.65	5787.28
Significant Effects		**100.41**	**94.22**	**5434.27**

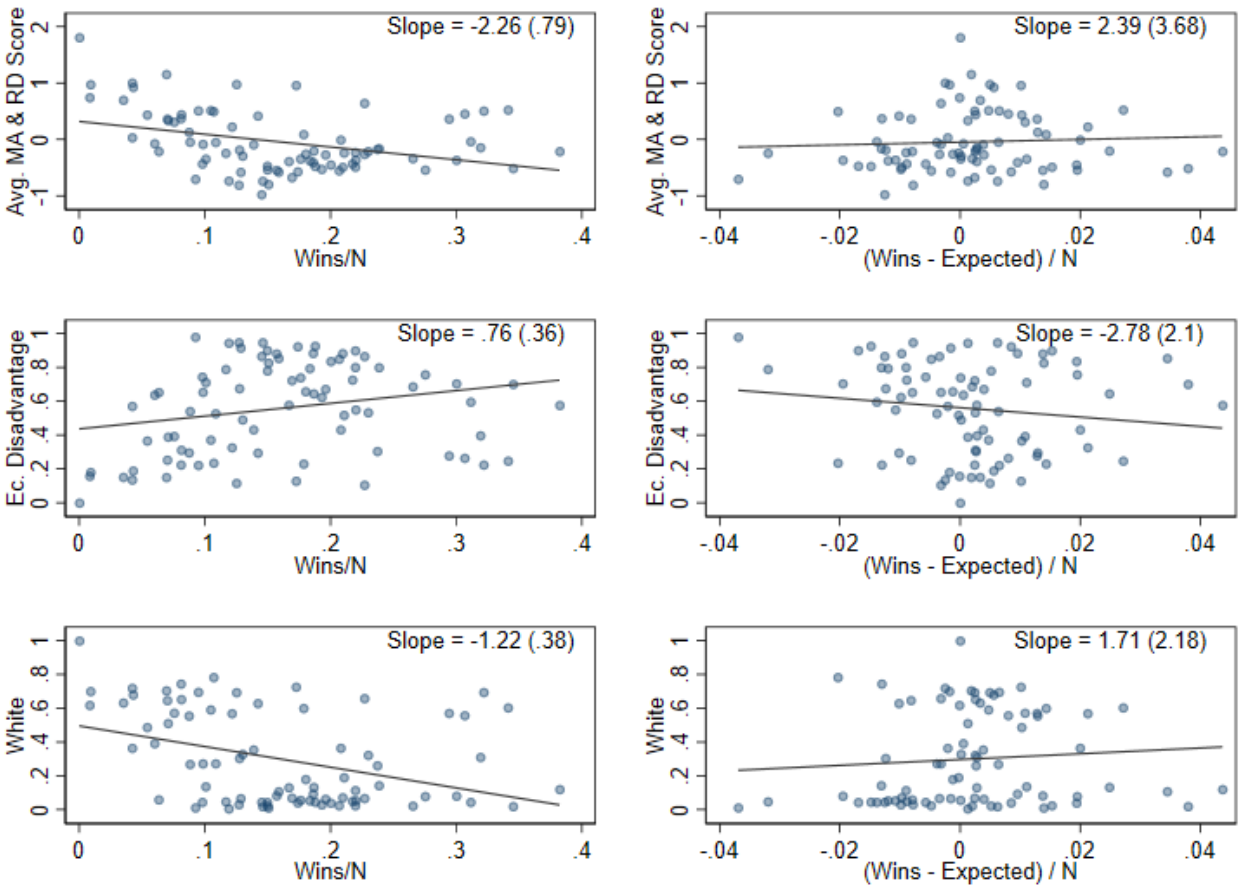
Notes: Aggregated effects on lottery applicants and their peers. The first column reports the expected number of wins among lottery applicants in each group *E[Wins]* or the expected number of displacements *E[displacements]*. *E[Wins]* are calulated as the sum of the estimated win probabilities for 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 displacement through the lottery. The aggregated effects on applicants are calculated using the expected number of wins and the relevant corresponding estimates for effects of winning on *All Arrests*, *Violent Arrests*, and *Days Incarcerated*, which come from Table 8 for Panel A and Table A11 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 relevant correspoding estimated effect on peer outcomes (Table 3 for Panel A and Table A12 for Panel B. Aggregates using only estimates with p-value<0.1 are bolded.

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 (.044) and divide by the age- or grade-specific mean of the same outcome in the low-risk sample. We include estimates at for three distinct approximate age groups for each set of outcomes. For *Days Suspended* and *Days Absent*, we estimate one for the two years post-lottery (6th an 7th grade for most), another for the third and fourth years post-lottery (8th and 9th grade for most), and one for the fith year post-lottery. We do not include estimates for years, which correspond to 11th and 12th grades, 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: Correlation between neighborhood school characteristics and lottery wins at the school level. Each panel shows observations and the slope estimate for the relation between a wins measure represented on the x-axis and the school-cohort measure on the y-asix. 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 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 proprotion economically disadvantaged students (middle two panels), and the proportion of white students (bottom two panels). Each school-cohort charactertic 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 Arrest)	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	0.42	0.46
	(2.00)	(1.85)	(1.92)
Num. Violent Arrests	0.21	0.19	0.18
	(0.96)	(1.01)	(0.89)
Days Incarcerated	8.95	7.38	7.53
	(60.88)	(53.61)	(54.00)
B) Other Student Outcomes			
Reading Test Scores (6th-8th)	-0.05	0.14	-0.03
	(1.06)	(0.96)	(1.06)
Math Test Scores (6th-8th)	0.05	0.20	0.08
	(1.09)	(1.00)	(1.11)
Days Suspended	2.93	2.23	2.82
	(7.72)	(6.10)	(7.64)
Days Absent	6.57	5.63	6.60
	(7.83)	(6.32)	(7.93)
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)
Observations	13,493	2,141	7,891

Notes: Summary of arrest and incarceration outcomes from age 16 to 22 for our main data sample of male 5th grade students that attended public school between the years of 2005-2006 to 2007-2008 in Charlotte-Mecklenburg Schools (CMS). The columns show the mean and standard deviation for CMS, the sample of lottery applicants and the non-applicant sample, respectively.

Table A2: Arrest Prediction

	Fixed Effects	School and CBG Means
White	0.0352	0.0255
	(0.140)	(0.138)
Black	0.632***	0.595***
	(0.123)	(0.120)
Hispanic	-0.236*	-0.232*
	(0.143)	(0.141)
Ec. Disadvantaged	0.702***	0.676***
	(0.0769)	(0.0749)
Math 5	-0.238***	-0.227***
	(0.0441)	(0.0430)
Read 5	-0.120***	-0.117***
	(0.0438)	(0.0426)
Math 5 Missing	-0.00821	-0.0363
	(0.272)	(0.257)
Read 5 Missing	0.321	0.290
	(0.259)	(0.245)
Age (Jan 1. Grade 5)	0.0246***	0.0242***
	(0.00460)	(0.00438)
School Means		
White		-0.0731
		(1.288)
Black		-1.120
		(1.304)
Hispanic		-1.162
		(1.389)
Ec. Disadvantaged		0.851*
		(0.471)
Math 5		-0.249
		(0.672)
Read 5		-0.00698
		(0.607)
CBG Means		
White		0.465
		(0.553)
Black		0.904*
		(0.496)
Hispanic		0.186
		(0.569)
Ec. Disadvantaged		0.614**
		(0.313)
Math 5		0.544**
		(0.233)
Read 5		-0.287
		(0.244)
Pseudo R-squared	0.137	0.111
Observations	12,716	13,493

Notes: Arrest prediction estimates from logistic regression with indicator for any arrrest 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 missig birthdate and 5th grade exceptionality.

Table A3: Effects of Peer Wins on Attendance and Movement

	Attend Neighb. Sch.	Change NS	Exit
Panel A: Pooled			
$(Wins/N_g)$	0.178	-0.336**	-0.075
	(0.127)	(0.135)	(0.087)
Sample Mean	.749	.263	.114
N	7891	7891	7891
Panel B: By Risk			
$(Wins/N_g) \times HR$	0.160	-0.254	-0.100
	(0.155)	(0.160)	(0.110)
$(Wins/N_g) \times LR$	0.229	-0.465**	-0.045
	(0.189)	(0.195)	(0.130)
N	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, pooling across estimated risk. 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 the 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 with the win proportion and high- and low-risk. In Panel B, we also include interaction terms between the high-risk dummy variable 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, likely indicating district exit. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A4: Selection of Applicants on Test Scores

	Math	Reading	Math	Reading
Panel A: Lottery Losers				
Applicant, Lost Lottery	0.35***	0.34***	0.06**	0.05**
	(0.04)	(0.04)	(0.02)	(0.02)
Math EOG 5th			0.66***	0.23***
			(0.01)	(0.01)
Reading EOG 5th			0.18***	0.57***
			(0.01)	(0.01)
Observations	18,390	18,323	18,390	18,323
Panel B: Lottery Winners				
Applicant, Won Lottery	0.27***	0.25***	0.03	-0.02
	(0.06)	(0.06)	(0.04)	(0.03)
Math EOG 5th			0.67***	0.22***
			(0.03)	(0.02)
Reading EOG 5th			0.16***	0.59***
			(0.02)	(0.03)
Observations	3,207	3,191	3,207	3,191

Notes: Panels A compares applicants who lost the lottery and attend their initially assigned neighborhood school with non-applicants who attend their initially assigned neighborhood school. Panel B compares applicants in our sample who won their first choice and attended their first choice school in 6th grade with other students who attended the same schools, but are not in our lottery sample due to gauranteed placements. Each column reports estimates from a different regression of a math or reading test score on a dummy for being in our lottery sample and losing (Panel A) or winning (Panel B) their first choice in the lottery. 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, dummy variables for missing lagged test scores, and CBG, neighborhood school (Panel A) or application school (Panel C), 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 A5: Effect of Peer Wins in Alternate Specifications

	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.	Pr(Arrest)	Days Incarc.
Panel A: Pooled								
$(Wins)/N_g$	-0.013	6.571	0.064	14.538	0.091	14.860	0.071	16.326
	(0.088)	(14.057)	(0.098)	(13.941)	(0.097)	(14.679)	(0.099)	(15.472)
	[0.085]	[13.719]	[0.102]	[14.911]	[0.102]	[15.791]	[0.113]	[17.496]
Observations	7903	7903	7891	7891	7891	7891	7856	7856
Panel B: By Risk								
$(Wins/N_g) \times HR$	-0.100	3.052	-0.051	11.329	-0.037	8.789	-0.011	11.065
	(0.137)	(22.914)	(0.142)	(21.474)	(0.137)	(22.314)	(0.138)	(22.837)
	[0.129]	[22.206]	[0.148]	[22.740]	[0.146]	[23.673]	[0.156]	[25.201]
$(Wins/N_g) \times LR$	0.157**	13.119**	0.253***	18.676**	0.309***	23.976***	0.245**	24.520**
	(0.063)	(5.810)	(0.087)	(8.394)	(0.089)	(8.965)	(0.101)	(10.909)
	[0.063]	[6.161]	[0.090]	[8.966]	[0.093]	[9.523]	[0.106]	[10.832]
Observations	7903	7903	7891	7891	7891	7891	7856	7856
Nbhd School FEs	✓	✓	-	-	✓	✓	-	-
CBG FEs	-	-	✓	✓	✓	✓	-	-
Nbhd School by CBG FEs	-	-	-	-	-	-	✓	✓

Notes: Estimated effect of the win proportion on the arrest probability and days incarcerated with different levels of fixed effects. 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). Both Panels include cohort fixed effects. Panel A includes estimates for the pooled sample of high- and low-risk non-applicants. Panel B includes separate estimates by above and below median risk. Panel A conditions on the expected win proportion and Panel B conditions on the expected win proportion for high- and low-risk students. There are two standard errors included for each estimate. Standard errors clustered at the CBG-Neighborhood School-Cohort level in parantheses, and standard errors clustered at the CBG-Neighborhood School level are inlcuded in brackets.

Table A6: Peer Effects on Test Scores and Graduation

	Math	Reading	On Time Grad.	Dropout
Panel A: Pooled				
$(Wins/N_g)$	0.080	-0.107	0.156	-0.041
	(0.132)	(0.138)	(0.126)	(0.114)
N	19,677	19,609	5,417	5,417
Panel B: By Risk				
$(Wins/N_g) \times HR$	0.244	-0.082	0.343*	-0.187
	(0.157)	(0.185)	(0.181)	(0.169)
$(Wins/N_g) \times LR$	-0.227	-0.167	-0.073	0.144
	(0.222)	(0.201)	(0.151)	(0.099)
N	19,677	19,609	5,417	5,417

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, pooling across estimated risk. 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 the 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 with the win proportion and high- and low-risk. In Panel B, we also include interaction terms between the high-risk dummy variable 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. We also include dummies for years post-lottery as controls in the test score regressions. The last two column show outcomes related to graduation and dropout. *On Time Grad.* is a dummy variable for graduating within 7 years post lottery (12th grade year for on-time progression). *Dropout* is a dummy variable equal to one if the students 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 A7: Heterogeneity in Peer Effects

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Heterogeneity by Race / Ethnicity						
White	0.361***	0.255***	0.192**	2.055***	0.736***	39.355**
	(0.116)	(0.087)	(0.079)	(0.672)	(0.225)	(19.994)
Black	-0.052	0.051	-0.156	0.457	0.518	10.455
	(0.157)	(0.141)	(0.138)	(0.936)	(0.427)	(20.773)
Hispanic	0.100	-0.007	0.167	0.364	0.040	25.803
	(0.179)	(0.147)	(0.167)	(0.949)	(0.441)	(27.924)
Other	0.150	-0.213	-0.185	-0.829	-0.458	-146.791
	(0.307)	(0.259)	(0.252)	(2.397)	(1.039)	(129.877)
Observations	7,891	7,891	7,891	7,891	7,891	7,891
Panel B: Heterogeneity by Test Scores						
Above Median	0.350***	0.170**	0.122	0.774	0.321	7.345
	(0.105)	(0.087)	(0.085)	(0.494)	(0.228)	(12.710)
Below Median	-0.073	0.034	-0.071	0.988	0.545	21.285
	(0.141)	(0.117)	(0.123)	(0.824)	(0.365)	(23.116)
Dep Var Mean, High	.071	.033	.034	.206	.074	2.777
Dep Var Mean, Low	.18	.117	.113	.729	.306	12.724
Observations	7,891	7,891	7,891	7,891	7,891	7,891

Notes: 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 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-speicific 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 A8: 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	0.829**	0.860**	0.581*	8.329**	2.724***	184.795
	(0.378)	(0.363)	(0.336)	(3.475)	(1.021)	(121.968)
White x LR	0.279**	0.144**	0.134*	0.910**	0.403**	12.838
	(0.109)	(0.069)	(0.071)	(0.395)	(0.174)	(8.031)
Black x HR	-0.050	0.061	-0.181	0.657	0.629	16.918
	(0.164)	(0.150)	(0.146)	(0.987)	(0.452)	(22.008)
Black x LR	-0.049	-0.064	0.199	-1.584	-0.906	-45.149
	(0.342)	(0.287)	(0.308)	(1.547)	(0.796)	(38.215)
Hispanic x HR	-0.149	-0.214	-0.064	-1.422	-0.912	-8.584
	(0.256)	(0.212)	(0.244)	(1.459)	(0.643)	(40.160)
Hispanic x LR	0.396*	0.267	0.424**	2.748**	1.352*	81.309**
	(0.222)	(0.178)	(0.177)	(1.204)	(0.707)	(37.703)
Other x HR	-0.430	-0.793*	-0.726	-4.848	-2.011	-337.253
	(0.565)	(0.449)	(0.446)	(4.383)	(2.076)	(241.755)
Other x LR	0.731***	0.344**	0.358**	3.392***	1.166**	33.983**
	(0.199)	(0.152)	(0.148)	(1.168)	(0.477)	(15.301)
N	7891	7891	7891	7891	7891	7891

Notes: Estimated effects of the peer win proportion on outcomes. Estimates by race / ethnicity and predicted risk. Each column in each panel 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-speicific 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 A9: 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						
$(Wins)/N_g \times HR$	-0.137	-0.034	-0.194	-0.192	0.030	5.942
	(0.159)	(0.146)	(0.145)	(1.047)	(0.437)	(30.957)
$\times Mn(Same\,Elem.)$	0.312	0.230	0.185	2.708	1.069	12.727
	(0.278)	(0.254)	(0.263)	(1.669)	(0.801)	(44.353)
$(Wins)/N_g \times LR$	0.217**	0.160**	0.183**	1.146**	0.505**	23.760**
	(0.093)	(0.075)	(0.079)	(0.495)	(0.223)	(10.440)
$\times Mn(Same\,Elem.)$	0.355*	0.062	0.150	0.664	0.240	2.458
	(0.206)	(0.152)	(0.143)	(0.874)	(0.388)	(18.878)
Observations	7,891	7,891	7,891	7,891	7,891	7,891
Panel B: Interactions with School-CBG-Cohort Fixed Effects						
$(Wins)/N_g$						
$\times Mn(Same\,Elem.) \times HR$	0.458	0.121	-0.024	2.386	1.060	75.313
	(0.439)	(0.412)	(0.416)	(2.529)	(1.175)	(78.950)
$\times Mn(Same\,Elem.) \times LR$	0.850**	0.365	0.461	3.370*	1.549*	75.167
	(0.341)	(0.285)	(0.289)	(1.880)	(0.891)	(55.112)
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, as well as two additional interaction terms. The interactions $\times Mn(Same\ \ Elem.)$ include the share of winners the 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 A10: Effect of Same Race Peer Wins on Arrest

	Pr(Arrest)	Pr(Violent)	Pr(Incarc.)	All Arrests	Violent Arrests	Days Incarc.
Panel A: Pooled						
$(Wins/N_g)$	0.139	0.135	0.003	1.048*	0.374	13.476
	(0.115)	(0.097)	(0.099)	(0.599)	(0.260)	(19.461)
(Same Race Win share)	-0.059	-0.063	-0.002	-0.218	0.071	1.701
	(0.066)	(0.061)	(0.051)	(0.306)	(0.147)	(11.537)
Observations	7891	7891	7891	7891	7891	7891
Panel B: By Risk						
$(Wins/N_g) \times HR$	0.028	0.098	-0.135	0.771	0.217	-0.193
	(0.170)	(0.147)	(0.147)	(0.926)	(0.398)	(29.257)
(Same Race Win Share) $\times HR$	-0.082	-0.086	0.006	-0.204	0.165	10.760
	(0.102)	(0.094)	(0.079)	(0.483)	(0.243)	(16.337)
$(Wins/N_g) \times LR$	0.315***	0.178**	0.219***	1.473***	0.570**	32.728*
	(0.101)	(0.074)	(0.080)	(0.570)	(0.233)	(18.427)
(Same Race Win Share) $\times LR$	-0.004	0.003	0.007	-0.198	0.010	-10.618
	(0.057)	(0.044)	(0.046)	(0.429)	(0.140)	(17.914)
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 inlcludes 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 together 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 the 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 the high-risk dummy variable and the movement related covariates, including the expected wins varable. Standard errors are clustered at the CBG-Neighborhood School-Cohort level.

Table A11: 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)	-0.005	-0.032	-0.040	-0.451**	-0.184	-17.475**
	(0.053)	(0.044)	(0.041)	(0.205)	(0.129)	(7.608)
Q4	0.033	0.053	-0.001	0.121	0.127	5.169
	(0.042)	(0.034)	(0.030)	(0.407)	(0.220)	(13.311)
Q3	-0.056**	-0.040*	-0.042**	-0.239	-0.169*	-1.107
	(0.026)	(0.021)	(0.021)	(0.191)	(0.085)	(6.733)
Q2	-0.031	-0.049**	-0.031**	-0.173**	-0.103**	-4.591*
	(0.026)	(0.019)	(0.014)	(0.085)	(0.050)	(2.469)
Q1 (Low)	0.009	-0.005	-0.010	-0.046	-0.021	-2.163
	(0.014)	(0.010)	(0.009)	(0.056)	(0.032)	(1.938)
Observations	2,140	2,140	2,140	2,140	2,140	2,140
Panel B: Effect Win, by Risk						
$Win \times Q5$	-0.007	-0.036	-0.041	-0.474**	-0.196	-18.202**
	(0.053)	(0.044)	(0.041)	(0.205)	(0.126)	(7.896)
$Win \times (Q1 - Q4)$	-0.012	-0.010	-0.021*	-0.087	-0.043	-0.618
	(0.016)	(0.014)	(0.012)	(0.145)	(0.075)	(3.893)
Dep Var Mean, Q5	.27	.175	.175	1.09	.505	17.714
Dep Var Mean, Q1-Q4	.078	.043	.039	.272	.124	5.165
Observations	2,140	2,140	2,140	2,140	2,140	2,140

Notes: Estimated effect of the winning first choice in the lottery on own arrests. Each is conditional on the set of student characteristics, covariates related to lottery outcomes, and lottery fixed effects. Panel A includes separate estimate for each quintile of predicted arrest risk. Panel B includes a low-risk esitmate, 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 A12: 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 Quntile						
Q5 (High)	-0.277	-0.069	-0.337*	0.223	0.671	31.321
	(0.200)	(0.191)	(0.192)	(1.447)	(0.670)	(34.330)
Q4	0.153	0.047	0.043	0.968	-0.002	-25.416
	(0.213)	(0.170)	(0.165)	(1.173)	(0.549)	(47.414)
Q3	0.212	0.346**	0.318**	0.989	0.904*	52.427*
	(0.176)	(0.149)	(0.152)	(1.016)	(0.467)	(26.941)
Q2	0.306**	0.162**	0.157*	1.053**	0.487**	17.854*
	(0.119)	(0.078)	(0.086)	(0.521)	(0.230)	(9.535)
Q1 (Low)	0.244*	0.055	0.088	0.659*	0.140	12.527*
	(0.131)	(0.059)	(0.071)	(0.361)	(0.153)	(7.097)
Observations	7,891	7,891	7,891	7,891	7,891	7,891
Panel B: Boys, By Risk						
$(Wins/N_g) \times Q5$	-0.265	-0.066	-0.333*	0.252	0.697	33.053
	(0.200)	(0.191)	(0.192)	(1.446)	(0.673)	(34.574)
$(Wins/N_g) \times (Q1 - Q4)$	0.240**	0.140*	0.146*	0.985**	0.345	8.333
	(0.102)	(0.075)	(0.076)	(0.489)	(0.216)	(17.559)
Dep Var Mean, Q5	.313	.231	.222	1.492	.65	28.553
Dep Var Mean, Q1-Q4	.082	.039	.039	.232	.084	2.988
Observations	7,891	7,891	7,891	7,891	7,891	7,891

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 quntile of predicted arrest risk. Panel B includes a low-risk esitmate, which pools quintiles 1-4 and high-risk estimate which includes quintile 5 only. 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 A13: Effects of Winning on Own Other Outcomes

				Suspensions					
	Math	Read	Absences	Total Days	Days ISS	Days OSS	Any Susp.	On Time Grad.	Dropout
Panel A: Pooled									
Win	0.017	0.061**	-0.510	-0.335	-0.042	-0.292	-0.063	0.002	0.004
	(0.040)	(0.024)	(0.383)	(0.356)	(0.087)	(0.288)	(0.084)	(0.017)	(0.010)
Observations	5,555	5,540	9,256	8,691	8,691	8,691	8,691	1,558	1,558
Panel B: Effect Win, by Risk									
$Win \times HR$	0.077**	0.096***	-0.213	-0.381	-0.047	-0.334	-0.122	0.020	0.000
	(0.035)	(0.030)	(0.580)	(0.661)	(0.162)	(0.530)	(0.161)	(0.031)	(0.017)
$Win \times LR$	-0.044	0.024	-0.697*	-0.249	-0.032	-0.217	-0.002	-0.015	0.009
	(0.048)	(0.030)	(0.367)	(0.184)	(0.062)	(0.142)	(0.054)	(0.014)	(0.010)
Observations	5,555	5,540	9,256	8,691	8,691	8,691	8,691	1,558	1,558

Notes: Estimated effect of winning choice on lottery applicant outcomes. Each is conditional on the set of student characteristics, characterstics related to lottery outcomes, and lottery fixed effects. Standardized math and reading scores in columns 1 and 2 are from up to three years after the lottery, or 6th, 7th, and 8th grade for most students. Absences and Suspension are from the 5 years post lottery, or 6th through 10th grade years. The last two column show outcomes related to graduation and dropout. *On Time Grad.* is a dummy variable for graduating within 7 years post lottery (12th grade year for on-time progression). *Dropout* is a dummy variable equal to one if the students 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 application choice by year level.

Appendix B
Estimating Lottery Win Probabilities and Expected Win Shares

We start with the sample students who 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) Met the entrance requirements for their first-choice program.
(4) Their first-choice program did not include subjective criteria in the admission process.[1]
(5) They were in an oversubscribed lottery. That is, they did not receive a sibling placement, and the program-cohort pool contains at least one lottery winner and at least one lottery loser.

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 calculate the proportion of applicants meeting these criteria who won their first choice in a given program for the lottery, $\overline{w}_l$. This is a baseline win probability for each student and accounts for the degree to which a given program is oversubscribed. In our main analysis, we estimate win probabilities using $\overline{w}_l$ and a vector of characteristics that can alter the probability of winning the lottery.[2]

There are two types of programs that students may have applied to: a magnet program and a neighborhood school other than their residentially assigned school. Because of minor differences in the lottery processes, we estimate admission probabilities separately for magnet and non-magnet lotteries. In each case, we estimate the parameters using a probit regression for the probability of winning the lottery and predict the individual win probabilities.

In all cases, our model for estimating the probability of an individual winning the lottery includes the proportion of lottery winners in that program ($\overline{w}_l$) as a predictor. We include other characteristics that determine lottery priorities, which may impact win probabilities. For non-magnet applications, we include year-specific dummies for the applicant scoring below grade level in reading and applying to an above average school in reading, applying to a non-Title I Choice school from a Title I Choice school, economically disadvantaged student applying to a non-Title I Choice school, and for living in the transportation zone of the application school. For magnet applications, we include year-specific dummy variables for applying to a non-Title I Choice school from a Title I Choice School, economically disadvantaged student from a Title I

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

[2] In an alternative approach, we simply use the lottery-specific win share, $\overline{w}_l$, in place of the estimated win probabilities. Our results are robust to that change.

Choice school applying to a non-Title I Choice school, and for living in the transportation zone of the application school. In all cases, we also include the interaction term for each dummy with the lottery-specific ratio of applicants who won their first choice in the given year.

We construct the predicted win probabilities, $\hat{P}_{ibst}$, using the estimated parameters and characteristics. Effectively, we rely on the lottery-specific win rates, allowing for adjustments based on observable student and school level characteristics to obtain a refined estimate of the win probabilities. The results are robust to the simplified process of using the lottery-specific win rates as the win probabilities. In that case, we use $\hat{P}_{ibst} = \overline{w}_l$.

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

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. 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 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, $\widehat{R}_{ibst}$. We use split the sample into those with $\widehat{R}_{ibst}$ above and below the sample median to create the indicators for high and low arrest risk.