Better Schools, Less Crime?

David Deming *

April 2010

Abstract

I estimate the effect of attending a first-choice middle or high school on young adult criminal activity, using data from public school choice lotteries in Charlotte-Mecklenburg school district (CMS). Seven years after random assignment, lottery winners have been arrested for fewer serious crimes, and have spent fewer days incarcerated. Lottery winners attended schools that were higher quality according to measures of peer and teacher inputs, as well as revealed preference, and the gain was roughly equivalent to switching from one of the lowest ranked schools to one at the district average. The reduction in crime is concentrated largely in the years after enrollment in the preferred school is complete. The effects are concentrated among African-American males whose *ex ante* characteristics define them as "high risk." As a result the CMS lottery assignment system, which gave priority to disadvantaged applicants, may have reduced crime relative to a simple lottery like those implemented by many U.S. charter schools.

*Harvard Kennedy School, 79 JFK St., Cambridge MA 02139 (email: demingd@nber.org). I would like to thank Lawrence Katz, Susan Dynarski, Brian Jacob, and Sandy Jencks for reading drafts of this paper and providing essential guidance and feedback. I benefited from the helpful comments of Josh Angrist, Amitabh Chandra, Roland Fryer, Alex Gelber, Josh Goodman, Bridget Long, Jens Ludwig, Erzo Luttmer, Juan Saavedra, Bruce Western, Tristan Zajonc and seminar participants at the Center for Education Policy Research (CEPR) series at Harvard University, the American Education Finance Assocation (AEFA) meetings, the Center for the Developing Child at Harvard University, and the University of Michigan. Special thanks to Tom Kane, Justine Hastings and Doug Staiger for generously sharing their lottery data, and to Eric Taylor and Andrew Baxter for help with matching the student and arrest record files. I gratefully acknowledge funding from the Julius B. Richmond Fellowship at the Center for the Developing Child and the Multidisciplinary Program on Inequality and Social Policy at Harvard. Web: `http://www.people.fas.harvard.edu/~deming/`.

1 Introduction

Can improvement in the quality of public schools be an effective crime prevention strategy? Criminal activity begins in early adolescence, and peaks when most youth should still be enrolled in secondary school (Wolfgang, Figlio and Sellin 1987; Farrington et al. 1986; Sampson and Laub 2003; Levitt and Lochner 2001). Crime is concentrated among minority males from high poverty neighborhoods (Freeman 1999; Pettit and Western 2004; Raphael and Sills 2006). An influential literature on "neighborhood effects" links criminal activity to neighborhood disadvantage through peer interaction models (Sah 1991; Glaeser, Sacerdote and Scheinkman 1996), or processes of socialization and collective efficacy (Sampson, Raudenbush and Earls 1997).

Schools may be a particularly important setting for the onset of criminal behavior.[1] Urban schools in high-poverty neighborhoods have high rates of violence and school dropout, and struggle to retain effective teachers (Lankford, Loeb and Wyckoff 2002; Murnane 2008; Cook, Gottfredson and Na 2009). Only 35 percent of inmates in U.S. correctional facilities earned a high school diploma or higher, compared to 82 percent of the general population (Bureau of Justice Statistics 2003). The best existing empirical evidence of the link between education and crime comes from Lochner and Moretti (2004), who use changes in compulsory schooling and child labor laws to estimate the effect of additional years of schooling on criminal activity. But the intensive margin of school quality is potentially more relevant for policy. In a human capital framework, low-skilled youth will engage in crime early in life because of low anticipated returns to schooling (Lochner 2004). If increased quality raises the return to investment in schooling, youth will stay in school longer, earn higher wages as adults, and commit fewer crimes.[2] Yet there is little evidence of the effect of school quality

[1] Since most public schools' assignment zones are defined by neighborhood, disentangling the separate influences of neighborhoods and schools is difficult. Jacob and Lefgren (2003) find that contemporaneous school enrollment leads to decreases in property crime but increases in violent crime, although their sample is not representative of large urban school districts.

[2] Additional compulsory schooling might accomplish the same goal, but the range of options for policy-makers is limited. The minimum school leaving age is already 18 in 18 states, and enforcement of truancy laws is sporadic (Oreopoulos 2006). Also, the population of "never takers" (i.e. youth who would drop out of school at the same age regardless of the law) might be particularly important.

on crime.[3]

In this paper I link a long and detailed panel of administrative data from Charlotte-Mecklenburg school district (CMS) to arrest and incarceration records from Mecklenburg County and the North Carolina Department of Corrections (NCDOC). In 2002, CMS implemented a district-wide open enrollment school choice plan. Slots at oversubscribed schools were allocated by random lottery. School choice in CMS was exceptionally broad-based. Ninety-five percent of students submitted at least one choice, and about forty percent chose a non-guaranteed school. Youth at higher *ex ante* risk for crime were actually more likely to choose a non-guaranteed school, allaying concerns about "cream-skimming" that might complicate the external validity of the findings (Epple and Romano 1998).

I estimate the causal effect of winning the lottery to attend a first-choice school on criminal activity through 2009, seven years after random assignment. Across various schools and for both middle and high school students, I find consistent evidence that winning the lottery reduces adult crime.[4] The effect is concentrated among African-American males and youth who are at highest risk for criminal involvement. Lottery winners also attend school longer and show modest improvements on school-based behavioral outcomes such as absences and suspensions. However, there is no detectable impact on test scores for any youth in the sample.

Nearly all of the reduction in crime occurs after enrollment in the preferred school is complete. Differences between lottery winners and losers persist to age 18 and beyond in both the middle and high school samples. The changes in peer and teacher quality experienced by lottery winners are roughly equivalent in magnitude to moving from one of the worst schools in the district to a school of average quality. Since nearly all of the lottery applicants stayed in CMS, winners and losers attended schools with similar budgets and governance

[3]Economic models of crime focus largely on changes in costs and benefits of crime for individuals on the margin of work and criminal activity (Becker 1968; Ehrlich 1973; Grogger 1998; Freeman 1999). A notable exception is Lochner (2004), who examines the onset of criminal behavior in a life-cycle model of schooling, crime and work. A recent paper by Weiner, Lutz and Ludwig (2009) finds a significant decline in homicide following school desegregation.

[4]Youth age 16 and above are considered "adult" by the criminal justice system in North Carolina. I do not observe juvenile crime.

structures. There were no additional community level interventions, such as in the Harlem Children's Zone (Dobbie and Fryer 2009). In sum, a treatment of between one and four years of enrollment in a higher quality public school led to large and persistent reductions in young adult criminal activity.

I also find strong evidence of heterogeneous treatment effects. I exploit the richness of pre-lottery administrative data and estimate the probability that a youth will be arrested in the future as a function of demographics, prior academic performance, behavior in school, and detailed neighborhood characteristics. The effect on crime of winning admission to a preferred school is strongly increasing in this *ex ante* prediction. Thus societal welfare gains from targeting resources to these youth might be substantial (Donohue and Siegelman 1998). Although random assignment of slots to oversubscribed schools is an ideal research design, it may be suboptimal from a welfare perspective if treatment effects can be predicted on the basis of observable characteristics (Bhattacharya and Dupas 2008). I simulate the effect of allocating slots based on *ex ante* crime risk rather than at random, and I find that this would reduce the social cost of crime by an additional 27 percent. While this allocation method is controversial (and in the case of race, illegal), it was executed at least in part by CMS, which gave a "priority boost" in the lottery to applicants who met an income standard based on eligibility for free or reduced price school lunches. I estimate that this priority boost lowered crime by 12 percent, relative to a lottery without priority groups such as the ones typically administered by US charter schools.

Several recent papers have found large positive impacts on test scores of winning admission to an oversubscribed public or charter school, using a lottery-based design (Hastings, Kane and Staiger 2008; Abdulkadiroglu et al. 2009; Dobbie and Fryer 2009; Hoxby and Murarka 2009; Angrist et al. 2010). Although these short-term test score gains are promising, data limitations have prohibited examination of longer-term outcomes measured outside the school setting.

There are at least two reasons why we might want to look beyond test scores and other school-based measures. First, there is an emerging literature on the unintended consequences of test-based accountability, which range from neglect of non-tested subjects to manipulation of the nutritional content of school lunches and outright teacher cheating (Jacob and Levitt 2003; Figlio and Winicki 2005; Jacob 2005). This leads to concerns that schools may raise student test scores through methods that do not translate to long-term improvements in skills or educational attainment. Second, even in the absence of distortionary incentives, the correlation between test score gains and improvements in long-term outcomes has not been conclusively established. Studies that relate test scores to earnings later in life, while suggestive, are not well-identified (Murnane, Willett and Levy 1995; Jencks and Phillips 1999; Currie and Thomas 2001). Furthermore, studies of early life interventions often find long-term impacts on outcomes such as educational attainment, earnings and criminal activity. despite fade out of test score gains in childhood (Krueger and Whitmore 2001; Belfield et al. 2006; Deming 2009). Thus programs can yield long-term benefits without raising test scores, and test score gains are no guarantee that impacts will persist over time.

This paper uses random assignment to examine the longer-term impact of school choice on crime, an important adult outcome measured outside the school setting. Studies of public school choice in Chicago and Tel Aviv examine high school graduation using school administrative data (Cullen, Jacob and Levitt 2006; Lavy 2009). Cullen, Jacob and Levitt (2006) find no impact of school choice on test scores or graduation but some benefits on behavioral outcomes, including self-reported criminal activity and during the years in which a student is enrolled. Taken together, the results here and in other studies suggest that looking only at test score gains may miss important benefits of interventions, particularly for disadvantaged youth. This paper also adds to the body of empirical evidence that links early education to future criminality (Garces, Thomas and Currie 2002; Lochner and Moretti 2004; Belfield et al. 2006; Weiner, Lutz and Ludwig 2009).

Although more research is needed to disentangle the relative contributions of neighborhoods and schools, this paper provides some evidence that schooling exerts a particularly strong influence on criminal behavior. The Moving to Opportunity (MTO) Demonstration found mixed impacts on crime (Ludwig, Duncan and Hirschfield 2001; Kling, Ludwig and

Katz 2005). MTO changed both neighborhoods and schools, although the change in measured school quality was not particularly large (Sanbonmatsu et al. 2006). Similarly, Jacob (2004) finds no independent impact on academic outcomes of moving out of high-density public housing. In contrast, the CMS open enrollment plan can be thought of as a pure school mobility experiment. Lottery winners continue to live in the same neighborhoods as lottery losers, and yet the reduction in crime persists even after schooling is largely complete.

The pattern of results is consistent with several possible explanations. Human capital theory predicts that offering youth admission to a better school would raise the return to investment in schooling, keeping them enrolled longer and increasing their opportunity cost of crime as adults (Lochner 2004). However, the results are also consistent with a model of peer influence where differential exposure to crime-prone youth exerts a long-lasting influence on adult crime. Without additional data on peer networks, these two hypotheses are difficult to disentangle, yet they could have very different policy implications. If the primary explanation for the results is an improvement in (non-peer) school inputs, then the estimates imply that investments in school quality will yield large reductions in the social cost of crime. However, any welfare calculation must account for shifting peer group composition due to school choice, including the possible negative externality imposed by lottery winners on their new peers. Estimates from the literature suggest that such spillovers are likely to be small in relation to the direct effect, and ambiguous in sign depending on the functional form of peer effects (Angrist and Lang 2004; Hoxby and Weingarth 2006; Carrell and Hoekstra 2008; Imberman, Kugler and Sacerdote 2009).[5] Still, because of the large one-year change in student assignments in CMS, extrapolation from the direct effect on lottery applicants is speculative and should be viewed with caution.

[5] Carrell and Hoekstra (2008) estimate the negative externality caused by children from families that are exposed to domestic violence and find that adding one of these children to a class of 20 causes each other child to commit 0.093 more infractions. Importantly, they find that the spillover effects on misbehavior are larger for low-income peers, which implies that concentrations of troubled students will generate more disruption. I show that the net effect of open enrollment in CMS was to distribute high-risk children across more schools than what would have happened in a pure neighborhood schools model. Thus, if the pattern of peer effects in Carrell and Hoekstra 2008 and other studies holds here, school choice would reduce overall crime (even if lottery winners' peers were negatively affected).

2 Data Description and Institutional Details

2.1 Data

With over 150,000 students enrolled in the 2008-2009 school year, Charlotte-Mecklenburg is the 20th largest school district in the nation. The CMS attendance area encompasses all of Mecklenburg County, including the entire city of Charlotte and several surrounding cities. Since the mid 1990s, the North Carolina Department of Public Instruction (NCDPI) has required all districts to submit a set of end-of-year (EOY) files that include demographic information, attendance and behavioral outcomes, yearly test scores in math and reading for grades 3 through 8, and subject-specific tests for higher grades. Internal CMS files obtained under a data use agreement also include identifying information such as name and date of birth, and students' exact addresses in every year, which I use to create detailed geographic identifiers. For more details on the nature and quality of the CMS administrative data, see the Data Appendix.

I match CMS administrative data to arrest records from the Mecklenburg County Sheriff (MCS).[6] I obtain these arrest records directly from the MCS website, which maintains an online searchable database that covers arrests in the county for the previous three years, counting from the day the website is accessed.[7] The data include all arrests of adults (age 16 and over in North Carolina) that occurred in the county, even if they were handled by another agency. Arrestees are tracked across incidents using a unique identifier that is established with fingerprinting. Critically, each observation includes the name and date of birth of the criminal.

The match was done using name and date of birth, and was exact in about 87 percent of cases. I obtained the remaining matches using an algorithm that assigns potential matches

[6] Since CMS is a "unified" school district, the geographic coverage of the school administrative data and the arrest records is identical.

[7] The web address is http://arrestinquiryweb.co.mecklenburg.nc.us/. I obtained the data by writing a script that loops over arrest numbers in consecutive order and copies the relevant information into a text file. See the Data Appendix for details.

a score based on the number and nature of differences.[8] I investigated match quality in several different ways, which are outlined in the Data Appendix.[9] Since the CMS open enrollment plan began in 2002, some older members of the sample could have been arrested prior to 2006, when the arrest data begin. To address this issue, I also obtained historical arrest records directly from MCS for members of the lottery sample only.[10] Finally, I add incarceration records from the MCS jail system and the North Carolina Department of Corrections (NCDOC). These county jail and state prison records are consistently available beginning only in 2006, and they were collected only for African-American male members of the lottery sample.[11] The data include number of days incarcerated, but probation and parole records are not included. See the Data Appendix for more details on the collection and coding of the arrest and incarceration data.

2.2 School Choice in Charlotte-Mecklenburg

From 1971 until 2001, CMS schools were forcibly desegregated under a court order. Students were bused all around the district to preserve racial balance in schools. After several years of legal challenges, the court order was overturned, and CMS was instructed that it could no longer determine student assignments based on race. In December of 2001 the CMS School Board voted on a policy of district-wide open enrollment for the 2002-2003 school year. School boundaries were redrawn as contiguous neighborhood zones, and children who lived in each zone received guaranteed access to their neighborhood school. The one-year change in student assignments was dramatic – about 40 percent of students at the middle and high school level were assigned to a different school than in the previous year. Because the inner

[8] As a specification check I ran the partial match algorithm a number of different ways, and I also estimated all the results in the paper using exact matches only. This made little difference. See the Data Appendix for details.

[9] These steps include verifying that there are no large time gaps in the data, that the age and demographic profile of arrests fits other studies, and that a high percentage of arrests among age-appropriate youth in Mecklenburg county are successfully matched to CMS data. See the Data Appendix for details.

[10] These data were recorded in almost exactly the same format as the more recent arrest records, although I cannot check their quality as easily.

[11] The data are limited to African-American males because I was unable to automate the collection process as well as for the arrest data. See the Data Appendix for details.

city of Charlotte is dense and highly segregated, African-American and poor students were even more likely to be reassigned.

The open enrollment lottery took place in the spring of 2002. CMS conducted an extensive outreach campaign to ensure that choice was broad-based, and 95 percent of parents submitted at least one choice (Hastings, Kane and Staiger 2008). Parents could submit up to three choices (not including their neighborhood school). Students were guaranteed access to their neighborhood school, and admission for all other students was subject to grade-specific capacity limits that were set by the district beforehand but were unknown to families at the time of the lottery (Hastings, Kane and Staiger 2008). When demand for slots among non-guaranteed applicants exceeded supply, admission was allocated by random lotteries according to the following strictly ordered priority groups:

1. Students that attended the school in the previous year and their siblings.
2. Free or reduced price lunch eligible (i.e. low income, "FRPL") students applying to schools where less than half of the previous year's school population was FRPL.
3. Students applying to a school within their own "choice zone".[12]

Applicants were sorted by priority group according to these rules and then assigned a random lottery number. Slots at each school were first filled by students with guaranteed access, and then remaining slots were offered to students within each priority group in order of their lottery numbers. CMS administered all of the lotteries centrally and applied an algorithm known as a "first choice maximizer" (Abdulkadiroglu and Somnez 2003). While this type of mechanism is not strategy-proof, Hastings, Kane and Staiger (2008) find little evidence of strategic choice by parents.

I begin with the full sample of middle and high school applicants. Since nearly all rising 12th graders received their first choice, I restrict the analysis sample to grades 6 through

[12] CMS divided schools into 4 "choice zones" and guaranteed transportation for students who applied to a school within their zone. This included magnet schools. The zones were constructed so that there was an even mix of mostly white "suburban" and mostly black "inner city" schools in each zone. In practice, this priority group was rarely used since very few students applied outside their choice zone.

11. Next I exclude the five percent of students who were not enrolled in any CMS school in the previous year. These students were much less likely to be enrolled in CMS in the following fall. Since previous enrollment was fixed at the time of the lottery, this restriction does not bias the results. The analysis sample consists of 21,132 high school students and 22,896 middle school students. The first column of Table 1 contains summary statistics for this sample. About sixty percent of the sample chose (and were automatically admitted to) their neighborhood school first. As shown in column 2 of Table 1, the remaining forty percent are more likely to be black and free lunch eligible, and they had lower test scores and higher rates of absence and out-of-school suspensions. About 75 percent of applicants to non-guaranteed schools were in lottery priority groups where the probability of admission was either zero or one. Even though these students chose a non-guaranteed school, there is no random variation in admission to exploit. In column 3 of Table 1 we see that the lottery subsample is similar to other applicants to non-guaranteed schools. The final lottery sample consists of 1,891 high school students and 2,320 middle school students.

Under busing schools were racially balanced, but the surrounding neighborhoods remained highly segregated. Thus the redrawing of school boundaries led to concentrations of minority students in some schools. Students who were assigned to these schools attempted to get out of them. Figure 1 displays the strong correlation between the racial composition of a school's neighborhood zone and the percent of students assigned to it who choose not to attend. Unlike many other studies of school choice, applicants to non-guaranteed schools are more disadvantaged than students who choose their neighborhood school.[13] Even within high-minority schools, from which most of the sample is drawn, lottery applicants are very similar in terms of race, socioeconomic status and average test scores to students who chose to remain in their neighborhood schools. Still, since lottery applicants had different preferences than their peers who chose to stay in the neighborhood school, they may differ on unobserved dimensions.

[13]See the Data Appendix for an analysis of selection into the lottery sample in a regression framework.

3 Empirical Strategy

If lottery numbers are randomly assigned, the winners and losers of each lottery will on average have identical observed and unobserved characteristics. Thus with sufficient sample size, a simple comparison of mean outcomes between winners and losers would identify the causal effect of winning each individual lottery. However, the sample here is not large enough to estimate the effect of winning each individual lottery. Instead, following Cullen, Jacob and Levitt (2006), I estimate ordinary least squares regressions of the form:

$$Y_{ij} = \delta \cdot W_{ij} + \beta X_{ij} + \Gamma_j + \varepsilon_{ij} \tag{1}$$

Y_{ij} is the outcome variable of interest for student i in lottery j. W_{ij} is an indicator variable equal to 1 if student i in lottery j had a winning randomly assigned lottery number, and zero if not. X_{ij} is a vector of covariates included for balance, Γ_j is a set of lottery (i.e. choice by grade by priority group) fixed effects, and ε_{ij} is a stochastic error term. I consider only first choices, so the number of observations is equal to the number of students in the lottery sample. In principle I could estimate a nested model that incorporates multiple choices. However, in practice nearly every student who did not receive their first choice was either automatically admitted to their second choice (if it was not oversubscribed) or automatically denied since all the slots were already filled.

Lottery fixed effects are necessary to ensure that the probability of admission to a first-choice school is uncorrelated with omitted variables in the error term. If, for example, savvy parents had some prior knowledge about the chance of admission, they might (all else equal) apply to schools where the probability of acceptance was higher. Thus comparing winners and losers across different lotteries might lead to a biased estimate. In the specification in equation (1), the δ coefficient gives the weighted average difference in outcomes between winners and losers across all lotteries, with weights equal to the number of students in the lottery times $p \cdot (1 - p)$ where p is the probability of admission (Cullen, Jacob and Levitt, 2006). Thus δ represents the intention-to-treat (ITT) effect of winning admission to a first-choice school for students in priority groups with non-degenerate lotteries. I cannot estimate the effect of attending a school for students with guaranteed access.

If the lotteries were conducted correctly, there should be no difference between winners and losers on any characteristic that is fixed at the time of application. I test this directly by estimating equation (1) with pre-treatment covariates such as race, gender and prior test scores as outcomes. The results, in the last column of Table 1, show that the lottery was balanced on observables and the randomization seems to have been conducted correctly. Even with proper randomization, however, the estimates could still be biased by selective attrition if leaving CMS or Mecklenburg County is correlated with winning the lottery. Since high school dropout rates are high for crime-prone youth, selective attrition is a serious concern for outcomes that come from the CMS administrative data. Students who drop out of school and are subsequently arrested in Mecklenburg County, however, are included in the data. Thus the main issue is selective migration. If lottery losers are more likely to leave the county, they may commit crimes in other jurisdictions. This would bias estimates downward. On the other hand, lottery winners may perform better in school and be more likely to leave the county to go to college, for example. This would bias the estimates upward. Still there are a few reasons to think that selective migration is not much of a concern here. First, the population of crime-prone youth is not very mobile. Attrition in grades K through 8 (where dropout is less of an issue) is negatively correlated with other predictors of crime and is much lower than average among future criminals.[14] Second, CMS assigns a withdrawal code to students who leave the district, and lottery status is uncorrelated with the code for out-of-county transfers. Additionally, the NCDOC state prison data includes information on county of arrest. Less than one percent of the sample spent time in state prison for offenses committed outside of Mecklenburg County, and there is no difference between lottery winners and losers.

[14] Ninety-one percent of future felons who were enrolled in CMS in 4th grade were still enrolled four years later (what would have been their 8th grade year). The overall average is eighty percent.

3.1 Predictors of Crime and Heterogeneous Treatment Effects

Most members of the lottery sample are probably not at high risk for criminal offending. Likewise, a small percentage of high-rate offenders are responsible for a large share of crimes (Wolfgang, Figlio and Sellin 1987; Freeman 1999). To test for heterogeneous treatment effects, I exploit the unusually long and rich panel of administrative data from CMS. Students with adult arrest records can be tracked all the way back to kindergarten in some cases, with yearly information on test scores and behavior and detailed neighborhood measures. I combine all of the individual correlates of criminal behavior into a single index and plot the treatment as a function of this *ex ante* crime risk. I estimate the probability that a student will have at least one arrest as a function of their history of test scores and behavior measures, demographic characteristics and neighborhood of residence. These measures are strong predictors of future criminality.[15] See the Appendix for more details on the estimation and for regression coefficients from this prediction.

In column 4 of Table 1, I present the average characteristics of youth who are in the top risk quintile according to this prediction. About ninety percent of the high risk sample is comprised of free lunch eligible African-American males. Their test scores are on average one standard deviation below the North Carolina state average, and they are absent and suspended many more days than the average student. Because the high risk students are overwhelmingly male, I exclude females from all subsequent analyses.[16].

To test for the possibility of heterogeneous treatment effects, I rank male youth according to their arrest risk and split the sample into five quintiles. I then estimate:

[15]The pseudo R-squared from the regression is about 0.23, compared to 0.24 when high school graduation is the dependent variable. Joint tests for the significance of each type of coefficient yield chi-squared values of 147 for test scores, 471 for behavior, and 249 for neighborhood fixed effects.

[16]I show in Appendix Table A4 that the number of arrests among females is extremely low, particularly for serious crimes. The crime prediction model greatly understates actual gender gaps in criminal offending. One way to show this is to regress a crime outcome such as felony arrests on the arrest prediction plus indicators for gender, race and free lunch status. The male coefficient comes in highly significant, while race and free lunch are insignificant, suggesting that the model does not do a good job accounting for gender differences. Results with females included are qualitatively similar, but do not identify "high risk" youth as accurately.

$$Y_{ij} = \sum_{q=1}^{5} \delta_q \cdot W_{ij} + \sum_{q=1}^{5} \phi_q (1 - W_{ij}) + \beta X_{ij} + \Gamma_j + \varepsilon_{ij} \tag{2}$$

where q indexes risk quintiles, and the rest of the notation is similar to equation (1). Separate coefficients by risk quintile for lottery winners (δ_q) and lottery losers (ϕ_q) allow me to test the hypotheses that lottery winners and losers are equal overall and within each quintile, and that the arrest risk quintiles are statistically different overall or within each group. I first estimate equation (2) for the main crime outcomes and plot the treatment effects and associated confidence intervals against each risk quintile. I then estimate simpler models where the first through fourth quintiles are pooled but the lottery is allowed to have a different effect on the top quintile "high risk" youth.

3.2 The Effect of Winning the Lottery on Measures of Enrollment and School Quality

Table 2 presents the effect of winning the lottery on enrollment and school characteristics for male applicants. Columns 1 through 4 present results for high school lottery applicants; columns 5 through 8 show the same for middle school applicants. The coefficients come from a regression like equation (2), but with the lowest four risk quintiles pooled together and a separate estimate for the top risk quintile. The odd numbered columns present control means for the estimates in each row. Below each estimate, and in subsequent tables, I report standard errors that are clustered at the individual lottery (i.e. choice by grade by priority group) level. The first row shows the effect of winning the lottery on attendance at a student's first choice school on the 20th day of the 2002 school year. The first stage is strong - lottery winners in all groups are over 55 percentage points more likely than losers to attend their first choice school. The coefficient is less than one mainly because some lottery losers successfully enroll in their first choice anyway.[17] For the main results in the paper, I report

[17]Some students moved into the school's neighborhood zone in the summer of 2002, after losing the lottery. Some lotteries were for special programs within schools, so a student might have been denied admission to the special program but accepted to the regular school. Finally, some students may have been admitted at

ITT estimates of the effect of winning the lottery. Later I discuss results that use the lottery as an instrumental variable for several of the outcomes in Table 2. Because a non-trivial fraction of lottery losers still manage to enroll, these estimates are not generalizable to all lottery applicants. Instead, they are local average treatment effects (LATEs) for students who comply with their lottery status (Angrist, Imbens and Rubin 1996).

The second row shows the effect of winning the lottery on total years enrolled in the first choice school. The treatment consisted of 1 to 1.5 additional years of enrollment on average, although notably from a much lower baseline for the top risk quintile. This suggests that the treatment "dose" was proportionally much larger for high risk youth. The third row shows the effect of winning the lottery on attendance at the student's neighborhood school, which is highly negative for all groups. Rows four through six show the effect of winning the lottery on the racial and family income composition of the school and on distance to assigned school. High school lottery winners attend schools that are demographically very similar to the schools attended by lottery losers, while middle school winners attend schools that are less African-American and higher income on average. All lottery winners travel farther to attend their first choice school, but the distance is greater for high school students.

The next five rows of Table 2 show the effect of winning the lottery on four measures of school quality. I normalize each of these measures to have mean zero and standard deviation one (separately for the middle and high school samples), to make them comparable to each other. Overall, lottery winners attend schools that are better on every dimension. The gain in measured quality for high risk youth is modestly larger than for the overall sample and starts from a much lower baseline, as indicated by the control means in each odd-numbered column. Interestingly, for high risk youth in both samples, the gains in average peer behavioral outcomes are larger than gains in peer test scores. Finally, we can see that lottery winners are much more likely to be enrolled in magnet schools. Magnet school enrollment comprises a larger share of the treatment in the high school sample, mostly due to the opening of a new magnet high school (Philip Berry Academy of Technology, a "career

the beginning of the school year when lottery winners did not enroll.

academy" that focuses on vocational and technical education) in the 2002-2003 school year.

The last four rows of Table 2 show the effect of winning the middle school lottery on high school characteristics. The sample is by necessity limited to students who were still enrolled in CMS in 9th grade. Although middle school lottery winners appear to attend better schools initially, these gains do not extend beyond the initial treatment. There is no statistically significant impact of winning the middle school lottery on the demographic composition, average test scores or average absences and suspensions of a student's high school.

In sum, lottery winners initially attend schools that are significantly better on several observable dimensions of quality. If school quality were normally distributed through CMS, then winning the lottery leads to average quality gains of around 0.3 standard deviations, with larger effects for high risk youth. Based on the control means in Table 2, lottery losers from the first four risk quintiles attend schools that are slightly worse than the district average, and winners attend schools that are slightly better. Measured quality gains are larger for high risk youth, and winning the lottery gets them into schools that are closer to the district average.

4 Results

4.1 Crime

Not all crimes are equal. Serious violent crimes such as murder, rape and armed robbery exact a heavy burden on their victims, so any welfare calculation should weigh these crimes more heavily. I measure crime severity in two ways. First, I use estimates of the victimization cost of crimes produced by Miller, Cohen and Wiersema (1996). These estimates, which were also used in an analysis of the of the Moving to Opportunity Demonstration by Kling, Ludwig and Katz (2005), consider tangible costs such as lost productivity and medical care, as well as intangible costs such as impact on quality of life, and are extremely high for fatal crimes.[18] To avoid the estimates being driven entirely by a few murders, I also report results with the cost of murder trimmed to twice the cost of rape, following Kling, Ludwig and Katz (2005). The second measure of severity weighs crimes by the expected punishment resulting from a successful conviction. In 1994 the state of North Carolina enacted the Structured Sentencing Act. Under structured sentencing, felony convictions are grouped into classes based on severity. This information is combined with the offender's prior record and other circumstances to determine a range of possible sentence lengths available to the judge. I group felony charges according to their class and assign the midpoint of the range of sentences for each of them. While both measures place a very high weight on murder, for example, the sentence weighted measure is better able to capture criminal intent.[19] I also examine the effect of winning the lottery on total days incarcerated in the county jail and state prison systems. These data are only available for African-American male members of the sample, from 2006 to the present. Since most high school sample members were already age 20 or above by 2006, I am missing prison time served during the peak criminal offending ages of 18 to 19. Incarceration data is likely to be much more complete for the middle school sample, however.

The main results of the paper are in Figures 2 and 3 and in Table 3. I first estimate equation (2) for selected crime outcomes and plot the point estimates and 90 percent confidence intervals by arrest risk quintile in Figures 2 and 3, for the middle and high school samples respectively. Each graph plots the coefficients from a model like equation (2), with a full set of lottery status by risk quintile interactions. The p-values from F-tests for equality of effects overall (and for each quintile, when statistically significant) and equality of quintiles

[18] The estimated social cost of murder is $4.3 million in 2009 dollars. The next costliest crime is rape, at about $125,000. Miller, Cohen and Wiersema (1996) do not include social cost estimates for drug crimes. Following Kling, Ludwig and Katz (2005), I assign costs to drug crimes according to felonies of equivalent standing. If instead I set the cost of drug crimes to zero, the estimates fall by about 25% in the high school sample but are unaffected for middle schools. This comes mostly from a large difference in the incidence of drug trafficking charges across treatment and control high school students (there were 16 drug trafficking charges in this sample, of which 14 occurred in the control group).

[19] For example, the difference between manslaughter and aggravated assault often comes down to luck (i.e. whether the bullet hit a critical organ or just missed it). The social cost measure would treat these two outcomes very differently, whereas the expected sentence length for these two crimes is very similar.

(in levels) are displayed on each graph. In Figure 2, we see that winning the lottery leads to fewer felony arrests overall (p=.078), and the effect is concentrated among the highest risk youth (0.76 felony arrests for lottery losers, 0.41 for winners, p=.013). Similarly, the trimmed social cost of crime is lower overall for lottery winners (p=.040), but the effect is concentrated among the top risk quintile youth ($11,000 for losers, $6,389 for winners, p=.036). The concentration of effects in the top risk quintile is even more pronounced for the middle school sample. The social cost of arrested crimes is $12,500 for middle school lottery losers and $4,643 for winners (p=.020), and the effect for days incarcerated is similarly large and concentrated among high risk youth (55.5 days for losers, 17.2 for winners, p=.003). For each of the eight outcomes in Figures 2 and 3, the level of crime committed by the top risk quintile is over twice that of the fourth quintile, and we can reject equality of quintiles at the 10 percent level for all eight outcomes.[20]

Table 3 shows regression results from a modified version of equation (2) where the first four risk quintiles are pooled, but the effect is allowed to vary for the top risk quintile.[21] In the first four columns I report estimates with the high and middle school samples pooled, with separate coefficients (from the same regression) for quintiles 1-4 and quintile 5. I first report results for the main outcomes of interest – number of felony arrests, social cost of arrested crimes, sentence-weighted crimes, and days incarcerated. In the last four rows I show results by type of felony charge. The odd numbered columns contain control means for each outcome, and the even-numbered columns show coefficients and standard errors, below in brackets.

Overall, winning the lottery led to an estimated reduction in the social cost of arrested crimes of over $30,000 for the top risk quintile, and over $11,000 for risk quintiles 1-4. Since more murders were committed by the control group than the treatment group (5 versus 1 in the combined high and middle school samples), the estimates are large and negative but

[20] Although I do not report the test statistics, equality of the 4th and 5th risk quintiles among lottery losers is rejected for all 8 outcomes in Figures 2 and 3.

[21] The models are estimated with the first through fourth risk quintile youth included, but I do not include the coefficients in the table.

relatively imprecise. When the cost of murder is trimmed, the effect becomes smaller but more precise. Winning the lottery led to a negative but insignificant drop of about $500 per male applicant in the first through fourth risk quintiles, but a decrease of over $6,000 per male applicant in the highest risk quintile. The effect for high risk males is large (over half of the control mean) and statistically significant at the one percent level. The results are of similar size and significance for the alternative measure of crime severity. High risk lottery winners commit crimes with a total expected sentence of about 26 months, relative to about 52 months for lottery losers. Finally, high risk lottery winners spend about 40 days in prison, compared to 70 days for lottery losers. Both the sentence-weighted and days incarcerated measures are statistically significant at the five percent level. The high overall level of incarceration among high risk youth is consistent with national trends - in 2006-2007, about 23 percent of black male high school dropouts in the U.S. were incarcerated on any given day (Sum et al. 2009).

Columns 5-6 and 7-8 show the top quintile results only, for the high and middle school samples respectively. Although the results for the main outcomes are similar, the pattern of effects by felony charges is different in each sample.[22] High school lottery winners are arrested for fewer of every type of charge, but the effect is largest for drug felonies (about two-thirds of the control mean). There is no overall effect on felony arrests or charges for high risk middle school lottery winners, but they commit many fewer index violent crimes (0.075 compared to 0.451 for losers). Since these crimes have the highest social cost and are punished most severely, the effects for social cost, sentence-weighted crimes, and days incarcerated are larger and more precisely estimated for the middle school sample. In Appendix Table A4 I present results separated by race and gender. I find statistically significant reductions in crime for African-American males overall, but nearly all of the results are statistically insignificant for other subgroups.

Winning the middle school lottery leads to substitution from more to less serious crimes,

[22] If someone is arrested once on seven counts of burglary, for example, this is seven charges but one arrest. Often there will be an outstanding warrant for an arrestee and they are processed at the same time on charges stemming from multiple incidents.

while winning the high school lottery leads to fewer (primarily drug) arrests overall.[23] Even though the effects are driven by high risk youth in both middle and high schools, the middle school sample appears more crime prone overall. The average number of felony arrests is about 0.7 in the top risk quintile for both samples, yet high school students have had many more years to accumulate arrests (and the average social cost of crimes is actually higher for the middle school sample). This is consistent with a developmental view of criminality, where delaying the onset of criminal offending among adolescents alters their future trajectory and prevents very serious crimes in the peak offending years (Moffitt 1993; Nagin and Tremblay 1999).

4.2 Pattern of Results Over Time

One possible explanation for the results is that winning the lottery entails longer bus rides to and from school, incapacitating youth during high crime hours. More generally, winning the lottery could prevent crime by removing high risk youth from "criminogenic" peers or neighborhoods (e.g. Sampson, Morenoff and Gannon-Rowley 2002; Kling, Ludwig and Katz 2005). Prominent models of criminal contagion treat individual crime as a function of contemporaneous exposure to crime-prone peers (Sah 1991; Glaeser, Sacerdote and Scheinkman 1996; Ludwig and Kling 2007). Both incapacitation and contagion explanations would predict a strong initial effect that fades over time. If, for example, drug market activity is concentrated within in a few schools, we might expect large differences in criminality in the high school years that diminish as enrollment in the treatment school ends and lottery winners and losers return to the same neighborhoods.

On the other hand, attending a better school might generate decreases in crime that persist long after enrollment is complete. In a human capital framework, increased school quality would raise the marginal productivity of investment in schooling. Youth who are given the opportunity to attend a better school would stay enrolled longer and acquire more

[23] This is supported by estimates where the dependent variable is dichotomous. High school lottery winners are less likely to ever be arrested but that is not true for middle school lottery winners.

skills, which would translate into a higher expected wage in the labor market. Higher wages raise the opportunity cost of crime and incarceration, lowering the optimal amount of crime committed (Lochner 2004). To the extent that skills acquired in school have a persistent effect on wages, reductions in crime would also be persistent. Alternatively, peer networks formed in middle or high school could have a persistent influence on adult criminality without affecting wages or employment directly. Although there is much evidence that social network formation is particularly important in the teenage years (e.g. Evans, Oates and Schwab 1992; Haynie 2001; Sacerdote 2001), there is little available evidence on the persistence into adulthood of criminal ties formed in adolescence. Finally, attending a better school might decrease the probability of arrest conditional on crime.[24]

Table 4 presents results by year since random assignment, for three of the main outcomes in Table 3. I present results separately for the high school and middle school samples, along with the median age of the sample at the beginning of each year. Standard errors are in brackets below the estimates, followed by control means for each period in curled brackets. Although I estimate models with the full sample, I only report the point estimates for high risk youth. For high school applicants, reductions in crime are concentrated in the fourth and fifth years following the lottery, when youth are around age 18-19 and no longer enrolled in their first choice school. The effects also come from post-treatment years in the middle school sample, although this is because data are only available beginning at age 16. High risk middle school lottery winners have a lower (but imprecisely estimated) social cost of arrested crimes in every period. The effect on felony arrests is negative and significant in year 5 but positive (though insignificant) in years 6 and 7. The effect on incarceration grows with time, which may help to explain the increase in felonies – the most serious offenders, who come disproportionately from the control group, are incapacitated and unable to commit further crimes. The patterns are similar for the other crime outcomes in Table 3. Data on

[24] Although I cannot provide any direct evidence on this, Lochner and Moretti (2004) find that the relationship between schooling and incarceration in the Census is similar to the relationship between schooling and self-reported crime, at least for white males. This suggests that higher levels of schooling do not greatly alter the probability of arrest conditional on crime.

incarceration are unfortunately unavailable for earlier periods in the high school sample.

4.3 Other Outcomes

A key limitation of this analysis is that I do not observe juvenile crime. This lack of early data could mask big differences in juvenile offending in the early years of the treatment. As an alternative, Table 5 shows the effect of winning the lottery on school disciplinary outcomes such as absences and suspensions, as well as test scores and course-taking. Because nearly all of the impacts on crime come from the highest risk youth, I report results for the highest risk quintile only, although the model is estimated with all male members of the sample. The first two rows show results for unexcused absences in the first two school years after the treatment, and the next two rows show the same thing but for out-of-school suspensions. Overall, lottery winners in both samples spend slightly more days in school. All four point estimates (2 samples, 2 years) for absences are negative, although only the 2003 middle school results are statistically significant. The effect for high school suspensions in 2003 is relatively large (a reduction of 3.7 from a baseline of 9.5 in the control group), but the other effects are small and statistically insignificant. Finally, I find that middle school lottery winners are less likely to be involved in a disciplinary incident where the punishment was long-term suspensions, expulsion or police involvement.[25]

In contrast to the results for crime and disciplinary outcomes, I find no evidence of test score gains.[26] Although results across various test subjects and grades are imprecisely

[25] I use a detailed disciplinary incident file maintained by CMS beginning in the 2006-2007 school year. Thus I cannot look at incidents for the high school sample at all or for any of the treatment years in the middle school sample. One difficulty with interpreting effects on absences and suspensions is that schools may differ in their discipline policies. If, for example, a higher-quality school maintains order by strictly enforcing rules, lottery winners might be more likely than losers to get suspended for equivalent behavior. Schools that succeed in keeping crime-prone youth in school longer may invest more resources in monitoring their behavior with low-level discipline, whereas "bad" schools might allow their behavior to escalate or not monitor them at all.

[26] For the middle school sample, the test score measures are results from standardized math and reading exams administered yearly for grades 3-8. High schools administer a set of end-of-course (EOC) exams in subjects such as Algebra I, Geometry, Biology and English. However, they are not taken by all students or even in the same grade in many cases, and so selection into test-taking may compromise interpretation of the results. The one exception is English I, which is taken in 9th grade by almost all students, so I include it as the only high school test score measure.

estimated, they are never distinguishable from zero, and I can rule out even modest (i.e. greater than 0.1 standard deviations) gains. Finally, I examine impacts on two measures of course-taking - whether a student was enrolled in remedial math (defined as less than Algebra I by 9th grade, which is the latest year a student can take the exam and graduate on time), and total math credits accumulated on EOC exams in 9th and 10th grade. High risk lottery winners in high school are much less likely to be enrolled in remedial math (19 percentage points from a control group baseline of 37 percent). However, there is no decrease in remedial math among lottery winners in the middle school sample.The impact on math credits is positive but imprecise in both samples.

Table 6 examines the effect of winning the lottery on enrollment, grade progression, and grade attainment for high risk youth. The school enrollment measures in the first four rows classify respondents as enrolled if they are present in CMS *in the year that they would have been in each grade if they progressed "on time."* For example, rising 6th grade lottery applicants would be enrolled in 9th grade in the 2005-2006 school year, so if they are still enrolled in CMS at the end of 2006 they are counted, even if they are not in grade 9. High risk middle school lottery winners are 18 percentage points more likely to be enrolled in CMS in their 10th grade year. The effect on 11th grade enrollment is about half the size (9 percentage points) but imprecisely estimated, and there is no impact on persistence into the 12th grade year. Is this difference in enrollment large enough to explain the impacts on crime? To test this, I estimate a regression of the trimmed social cost measure on similarly constructed grade enrollment dummies, a set of covariates and neighborhood fixed effects using high risk youth from the full sample. Then I multiply the estimated social cost of crime for each level of grade enrollment by the estimates in Table 6. If the cross-sectional relationship holds in the lottery sample, this rough calculation suggests that increased enrollment alone can explain about one-third of the total impact on crime for high risk middle school youth. Perhaps because 10th grade is around the time youth turn 16 and are legally permitted to leave school, enrollment beyond the grade 10 year is associated with a relatively large decline in crime. I find no impact on enrollment for high risk high school youth.

Next I measure grade progression by counting students as "on track" if they have advanced at least one grade for every year since the lottery and are not enrolled in an alternative school. The pattern here is exactly the opposite as the results for enrollment. High school lottery winners are more likely to be "on track" for 9th, 10th and 11th grade. The estimates are of similar size in absolute terms (between 12 and 14 percentage points) but grow in relative terms, as lottery losers increasingly fall behind or enroll in alternative schools. The effect fades to insignificance by 12th grade, however. In contrast, there is no effect on grade progression for high risk middle school lottery winners.

Despite the impacts on enrollment and progression, there is no detectable increase in high school graduationin either sample. Because I am limited to CMS administrative data, it is difficult to distinguish dropouts from subsequent GED recipients or transfers who may have graduated elsewhere.[27] Administrative records are particularly problematic for high risk youth, who are marginally attached to school and sometimes disappear from CMS well before the legal age of school leaving.[28] The graduation rate is only about 25 percent among high schoolers, and currently only about 10 percent among middle schoolers, although some who are still enrolled may subsequently graduate. Additionally, a bit less than 10 percent of the middle school sample never appears in any high school grade but subsequently appears in the arrest data. Because any intervention aimed at high school students would miss them altogether, this suggests that high school might be too late for the highest risk youth.

The effect of winning the lottery is largest at ages when most youth are mixing schooling, crime and work in some combination (Grogger 1998). If attending a better school increased the wages of lottery winners or their ability to find work, this might lead to a decrease in crime that persists after the treatment is complete. Still, I do not directly observe employ-

[27]Students who stop showing up for school are counted as either dropouts, transfers or no-shows, but there is considerable uncertainty across those categories. First, students are coded as dropouts only at age 16 and above. Second, transfers (even out-of-state) often show up subsequently in the Mecklenburg county arrest data.

[28]To illustrate the unreliability of coding, I calculate the average social cost of crimes for members of the sample who are recorded as transfers versus dropouts. Strikingly, despite the fact that some of the transfers are "real", the social cost of crime among them averages about $11,347, compared to $18,584 for verified dropouts.

ment or wages, and there are other explanations that are consistent with this pattern of results. Any explanation where exposure to peers early in life exerts a particularly strong influence on later criminality (either by raising the cost of legitimate activities, or through the formation of long-lasting peer groups) would also lead to the same pattern of results. In both samples combined, about 80 percent of students have already dropped out of school by the time they are arrested for their first felony. Furthermore, even among the remaining 20 percent, students with arrest records are often absent and/or suspended for long stretches of time before an arrest occurs. Thus it is plausible that keeping students enrolled longer, or maintaining a stronger attachment to school, reduces the overall amount of crime committed by delaying the onset of criminality through the peak period of offending (Moffitt 1993; Nagin and Tremblay 1999).

5 Discussion and Policy Implications

Since criminal involvement can be predicted using information that is readily available to the school district, a lottery mechanism that gives priority to high risk youth could reduce crime more effectively. To quantify the benefits of targeting, I simulate the lottery and resulting distribution of students to schools under two alternative assignment rules. First, I assign open slots to the highest risk students (based on the prediction generated in Section 3.1) in descending order, for each lottery. While such an allocation system would be controversial, it is feasible since all the covariates are available to the school district. Second, I simulate a simple lottery with no priority groupings, similar to the decentralized lotteries conducted by many US charter schools. The CMS lottery system assigned a "priority boost" to free lunch-eligible (FRPL) students who applied to schools with a low fraction of FRPL students in the previous year. As a consequence, many poor (and high crime risk) students were automatically admitted to schools when other students had to win the lottery (or, in some cases, only FRPL students could be admitted, and no other students were admitted).

For both assignment rules, I simulate the lottery 500 times and calculate the new expected distribution of students to schools. In the last step, I use the original parameter values from the estimation of equation (2) for the social cost of crime outcome. This calculation makes some important assumptions. First, it assumes that students' choices were not strategic, and thus they would not have changed their preferences if the assignment rules changed. Second, it assumes that the relationship I estimate between crime risk and the social cost outcome is generalizable out of sample. Finally, it assumes that there are no differential spillover effects from lottery winners to their schoolmates under each scenario.

I estimate that if slots in oversubscribed schools were allocated to the highest risk students, the social cost of crime would fall by an additional 27 percent relative to the actual CMS assignment mechanism. A more realistic form of targeting is the method actually pursued by CMS – a "priority boost" for economically disadvantaged students. I estimate that this policy choice lowered the social cost of crime by about 12 percent, relative to a simple lottery with no preferential treatment. Most of the difference comes from changes in the middle school lottery, for two reasons. First, the effect is more strongly increasing in crime risk for the middle school lottery than for the high school lottery (see Figures 2 and 3). Second, there is much less sorting across choices at the middle school level, so there are many low and high risk students applying to the same schools.

CMS chose to implement an open enrollment school choice plan as an alternative to a traditional neighborhood schools model. They expanded capacity at schools where high demand was anticipated, including magnet schools that were located in the inner city. These schools increased yearly enrollment substantially and were in many cases still oversubscribed. Many low-performing schools, on the other hand, experienced large reductions in enrollment – by as much as 50 percent in some cases. Thus, relative to a pure neighborhood schools model, the net effect of open enrollment was to increase access to magnet and highly demanded schools for youth who would not otherwise be able to enroll. This strong demand response means that the treatment is not just a transfer from losers to winners, and could represent a real welfare gain.

While any welfare calculation would also have to include the possible negative externality imposed by these youth on their new peers, such an effect is likely to be a small fraction of the individual reduction in crime for two reasons. First, estimates from the peer effects literature are generally small (Angrist and Lang 2004; Hoxby and Weingarth 2006; Carrell and Hoekstra 2008; Imberman, Kugler and Sacerdote 2009). Second, lottery winners would need to have a much larger differential impact on their peers than they would have had in another school. Since they attended better schools on average (as did many other high risk youth who attended a non-guaranteed school but were not subject to randomization), disruptive students were less concentrated under open enrollment than they would have been in a neighborhood schools model. Depending on the nature of peer effects, the effect could go in either direction, but the available evidence suggests that concentrations of disruptive children increase overall misbehavior (Carrell and Hoekstra 2008; Imberman, Kugler and Sacerdote 2009).

All the results so far have been ITT estimates of the effect of winning the lottery. However, we can also calculate local average treatment effects (LATEs) for youth who comply with their lottery status, using the lottery as an instrument for enrollment.[29] Since the average "first stage" effect was around 0.55, the LATEs are a bit less than double the ITT estimates for each outcome. Following Hoxby and Murarka (2009) and Abdulkadiroglu et al. (2009), I can also calculate the per-year effect of enrollment in a first choice school. This is particularly large for high risk youth - each year of enrollment saves society over $55,000 in criminal victimization costs for arrested crimes. Finally, I use the lottery as an instrument for the quality of the school attended by applicants in the fall of 2002. I calculate the average of the four normalized school quality measures in Table 2. Assuming that all the treatment effect operates through measured school quality, a one standard deviation increase in school quality leads to a reduction in the social cost of arrested crimes of about $23,000 per applicant and

[29] The IV estimates are only valid if the monotonicity assumption ("no defiers" - i.e. no applicant would have enrolled if they lost or not enrolled if they won) holds (Angrist, Imbens and Rubin 1996). The group of compliers is a latent type, since we cannot directly observe who among the complier lottery losers would have enrolled if they had won (and vice versa for winners). Empirically, observed compliers are drawn from the middle of the distribution of arrest risk ($\widehat{\Pr}(arrest \mid X_{ij}) = .237$) relative to the lottery loser "always-takers" ($\widehat{\Pr}(arrest \mid X_{ij}) = .161$) and the lottery winner "never-takers" ($\widehat{\Pr}(arrest \mid X_{ij}) = .302$).

about $110,000 per high risk youth.

6 Conclusion

In this paper I estimate the longer-term effect of on adult crime of winning an admissions lottery to attend a better middle or high school. I find that winning the lottery greatly reduces crime, and the effect is concentrated among the highest risk youth in the sample. Importantly, the effects of winning the lottery persist beyond the treatment years into the peak ages of criminal offending and beyond. After enrollment in the first choice school is complete, youth attend similar schools and live in similar neighborhoods. Yet the impacts persist for seven years after random assignment. The findings suggest that schools may be a particularly important setting for the prevention of future crime. Most of the future criminals in the sample drop out of school at a very young age and are incarcerated for serious crimes prior to the age of high school graduation. For high risk youth on the margins of society, public schools may present the best opportunity to intervene.

The end of busing and the implementation of open enrollment in CMS was a significant policy change. The four lowest-ranked high schools lost over 20 percent of their enrollment from 2002 to 2003. In subsequent years, two of these schools were restructured as magnet schools that offered a series of specialized programs in a small school setting. Similarly, two of the lowest-ranked middle schools were subsequently closed. This suggests that open enrollment sent a strong demand signal to CMS that resulted in the shutting down or restructuring of low-performing schools. The No Child Left Behind Act of 2001 included a provision that allowed parents to transfer students from "persistently dangerous" public schools, but many states have set the legal threshold so high that very few schools qualify. The results here suggest that, to the extent that low quality schools are also persistently dangerous, allowing students to leave them for a better school might benefit individual students as well as society as a whole.

References

Abdulkadiroglu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane and Parag A. Pathak. 2009. "Accountability and flexibility in public schools: evidence from Boston's charters and pilots.".

Abdulkadiroglu, Atila. and Tayfun. Somnez. 2003. "School choice: A mechanism design approach." *American Economic Review* 93(3):729–747.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.

Angrist, Joshua D. and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *The American Economic Review* 94(5):1613–1634.

Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak and Christopher R. Walters. 2010. "Who Benefits from KIPP?".

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *The Journal of Political Economy* 76(2):169–217.

Belfield, Clive R, Milagros Nores, Steve Barnett and Lawrence Schweinhart. 2006. "The High/Scope Perry Preschool Program: Cost Benefit Analysis Using Data from the Age-40 Followup." *Journal of Human Resources* XLI(1):162–190.

Bhattacharya, Debopam and Pascaline Dupas. 2008. "Inferring Welfare Maximizing Treatment Assignment under Budget Constraints." NBER Working Paper No. 14447.

Carrell, Scott E. and Mark L. Hoekstra. 2008. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." NBER Working Paper No. 14246.

Cook, Philip J., Denise C. Gottfredson and Chongmin Na. 2009. "School crime control and prevention." Unpublished Working Paper.

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

Currie, Janet and Duncan Thomas. 2001. "Early Test Scores, School Quality and SES: Long Run Effects on Wage and Employment Outcomes." *Worker wellbeing in a changing labor market* 20:103–32.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1(3):111–134.

Dobbie, Will and Roland G. Fryer. 2009. "Are High-Quality Schools Enough to Close the Achievement Gap? Evidence from a Bold Social Experiment in Harlem." Unpublished Working Paper.

Donohue, John J. and Peter. Siegelman. 1998. "Allocating resources among prisons and social programs in the battle against crime." *The Journal of Legal Studies* 27(1):1–43.

Ehrlich, Isaac. 1973. "Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of political Economy* 81(3):521.

Epple, Dennis. and Richard E. Romano. 1998. "Competition between private and public schools, vouchers, and peer-group effects." *The American Economic Review* 88(1):33–62.

Evans, William., Wallace. Oates and Robert Schwab. 1992. "Measuring peer group effects: a model of teenage behavior." *Journal of Political Economy* 100:966–991.

Farrington, Daniel P., B. Gallagher, L. Morley, R.J. St Ledger and D.J. West. 1986. "Unemployment, school leaving, and crime." *British Journal of Criminology* 26(4):335.

Figlio, David N. and Joshua Winicki. 2005. "Food for thought: The effects of school accountability plans on school nutrition." *Journal of Public Economics* 89(2-3):381–394.

Freeman, Richard B. 1999. The economics of crime. In *Handbook of Labor Economics*, ed. Orley Ashenfelter and David Card. Vol. 3 Elsevier pp. 3529–3571.

Garces, Eliana, Duncan Thomas and Janet Currie. 2002. "Longer-term effects of Head Start." *American Economic Review* 92(4):999–1012.

Glaeser, Edward L., Bruce Sacerdote and Jose Scheinkman. 1996. "Crime and social interactions." *The Quarterly Journal of Economics* 111(2):507–548.

Grogger, Jeffrey. 1998. "Market wages and youth crime." *Journal of Labor Economics* 16(4):756–791.

Hastings, Justine S., Thomas J. Kane and Douglas O. Staiger. 2008. "Heterogeneous Preferences and the Efficacy of Public School Choice." Unpublished Working Paper.

Haynie, Dana L. 2001. "Delinquent Peers Revisited: Does Network Structure Matter?" *American Journal of Sociology* 106(4):1013–1057.

Hoxby, Caroline M. and Gretchen Weingarth. 2006. "Taking race out of the equation: School reassignment and the structure of peer effects." Unpublished working paper.

Hoxby, Caroline M. and Sonali. Murarka. 2009. "Charter Schools in New York City: Who Enrolls and How They Affect Their Students' Achievement.".

Imberman, Scott., Adriana D. Kugler and Bruce Sacerdote. 2009. "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees." NBER Working Paper No. 15291.

Jacob, Brian A. 2004. "Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago." *The American Economic Review* 94(1):233–258.

Jacob, Brian A. 2005. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89(5-6):761–796.

Jacob, Brian A. and Lars Lefgren. 2003. "Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime." *The American Economic Review* 93(5):1560–1577.

Jacob, Brian A. and Steven D. Levitt. 2003. "Rotten Apples: An Investigation of The Prevalence and Predictors of Teacher Cheating*." *Quarterly Journal of Economics* 118(3):843–877.

Jencks, Christopher and Meredith Phillips. 1999. "Aptitude or achievement: Why do test scores predict educational attainment and earnings." *Earning and learning: How schools matter* pp. 15–47.

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

Krueger, Alan B. and Diane M. Whitmore. 2001. "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR." *The Economic Journal* 111(468):1–28.

Lankford, Hamilton, Susanna Loeb and James Wyckoff. 2002. "Teacher sorting and the plight of urban schools: A descriptive analysis." *Educational Evaluation and Policy Analysis* 24(1):37.

Lavy, Victor. 2009. "Effects of Free Choice among Public Schools." *Manuscript, Hebrew University* .

Levitt, Steven D. and Lance Lochner. 2001. The determinants of juvenile crime. In *Risky Behavior among Youths: An Economic Analysis*, ed. Jonathan Gruber. University of Chicago Press pp. 327–73.

Lochner, Lance. 2004. "Education, work, and crime: A human capital approach." *International Economic Review* 45(3):811–843.

Lochner, Lance and Enrico Moretti. 2004. "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports." *The American Economic Review* 94(1):155–189.

Ludwig, Jens, Greg J. Duncan and Paul Hirschfield. 2001. "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment*." *Quarterly Journal of Economics* 116(2):655–679.

Ludwig, Jens and Jeffrey R. Kling. 2007. "Is Crime Contagious?" *The Journal of Law and Economics* 50(3):491–518.

Miller, Ted R., Mark A. Cohen and Brian Wiersema. 1996. "Victim costs and consequences: A new look." *Washington, DC National Institute of Justice, US Department of Justice* .

Moffitt, Terrie E. 1993. "Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy." *Psychological Review* 100:674–674.

Murnane, Richard. 2008. "Educating urban children." NBER Working Paper No. 13791.

Murnane, Richard J., John B. Willett and Frank Levy. 1995. "The growing importance of cognitive skills in wage determination." *The Review of Economics and Statistics* 77(2):251–266.

Nagin, Daniel. and Richard E. Tremblay. 1999. "Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency." *Child development* 70(5):1181–1196.

Oreopoulos, Philip. 2006. "Estimating average and local average treatment effects of education when compulsory schooling laws really matter." *The American Economic Review* pp. 152–175.

Pettit, Becky and Bruce Western. 2004. "Mass imprisonment and the life course: Race and class inequality in US incarceration." *American Sociological Review* 69(2):151–169.

Raphael, Steven and Melissa Sills. 2006. Urban Crime, Race, and the Criminal Justice System in the United States. In *Companion to Urban Economics*, ed. Daniel P. McMillen and Richard Arnott. Blackwell Publishing pp. 515–535.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates*." *Quarterly Journal of Economics* 116(2):681–704.

Sah, Raaj K. 1991. "Social osmosis and patterns of crime." *The Journal of Political Economy* 99(6):1272–1295.

Sampson, R.J., S.W. Raudenbush and F. Earls. 1997. "Neighborhoods and violent crime: A multilevel study of collective efficacy." *Science* 277(5328):918.

Sampson, Robert J., Jeffrey D. Morenoff and T. Gannon-Rowley. 2002. "Assessing "Neighborhood Effects": Social Processes and New Directions in Research." *Annual review of sociology* 28(1):443–478.

Sampson, Robert J. and John H. Laub. 2003. "Life-Course Desisters-Trajectories of Crime among Delinquent Boys followed to Age 70." *Criminology* 41(3):555–592.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan and Jeanne Brooks-Gunn. 2006. "Neighborhoods and academic achievement: Results from the Moving to Opportunity Experiment." *Journal of Human Resources* 41(4):649.

Sum, Andrew, Ishwar Khatiwada, Joseph McLaughlin and Shelia Palma. 2009. The Consequences of Dropping Out of High School. Technical report Center for Labor Market Studies, Northeastern University.

Weiner, David A., Byron Lutz and Jens Ludwig. 2009. "The effects of school desegregation on crime." NBER Working Paper No. 15380.

Wolfgang, Marvin E., Robert M. Figlio and Torstein Sellin. 1987. *Delinquency in a birth cohort*. University of Chicago Press.

Table 1: Summary Statistics and Randomization Check

High School

	All Students	Chose Non-Home	Lottery	Top Risk Quintile	Randomization Check
	(1)	(2)	(3)	(4)	(5)
Male	0.50	0.50	0.54	0.86	0.021 [0.025]
Black	0.43	0.59	0.62	0.92	0.034 [0.024]
Free Lunch	0.47	0.63	0.64	0.93	0.016 [0.024]
Math (8th)	-0.06	-0.33	-0.28	-0.92	0.022 [0.046]
Reading (8th)	-0.02	-0.31	-0.26	-1.05	-0.019 [0.043]
Days Absent	9.6	11.7	11.1	18.7	0.49 [0.57]
Days Suspended	1.5	2.2	2.2	6.9	0.25 [0.39]
Sample Size	21,132	8,157	1,891	378	1,891

Middle School

	All Students	Chose Non-Home	Lottery	Top Risk Quintile	Randomization Check
Male	0.51	0.50	0.47	0.88	0.028 [0.021]
Black	0.46	0.61	0.62	0.93	0.033 [0.023]
Free Lunch	0.54	0.69	0.66	0.98	-0.027 [0.019]
Math (5th)	0.07	-0.17	-0.03	-0.89	0.006 [0.040]
Reading (5th)	-0.01	-0.23	-0.07	-0.98	-0.047 [0.042]
Days Absent	8.4	9.5	8.9	13.8	-0.32 [0.49]
Days Suspended	1.2	1.7	1.4	4.6	-0.15 [0.20]
Sample Size	22,896	9,397	2,320	464	2,320

Notes : Column 1 includes all high school (grades 9-11) and middle school (grades 6-8) students who were enrolled in CMS in the 2001-2002 school year. Column 2 restricts the sample students who listed as their first choice a school to which they were not guaranteed admission. Within that set of students, Column 3 includes only applicants to lotteries for which the probability of admission was neither zero nor one. Column 4 restricts the sample to lottery applicants who were in the top risk quintile according to the arrest prediction in Section 3.2. Column 5 reports point estimates from a regression like equation (1) with each row outcome as the dependent variable, with standard errors in brackets that are clustered at the lottery (i.e. school by grade by priority group) level. All covariates are from the 2001-2002 school year unless stated otherwise. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Table 2: Effect of Winning the Lottery on Enrollment and School Characteristics

	High Schools				Middle Schools			
	Risk Quintiles 1-4		Top Risk Quintile		Risk Quintiles 1-4		Top Risk Quintile	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Enrolled in 1st Choice	0.392	0.573***	0.174	0.692***	0.188	0.598***	0.068	0.580***
		[0.054]		[0.073]		[0.047]		[0.086]
Total Years Enrolled	1.06	1.49***	0.29	1.31***	0.45	1.13***	0.16	1.11***
		[0.17]		[0.22]		[0.09]		[0.14]
In Home School	0.399	-0.371***	0.489	-0.474***	0.555	-0.341***	0.406	-0.242***
		[0.028]		[0.065]		[0.053]		[0.069]
School Characteristics								
Percent Black	0.447	0.036	0.558	0.014	0.470	-0.054*	0.630	-0.061*
		[0.041]		[0.049]		[0.028]		[0.032]
Percent FRPL	0.488	0.011	0.621	-0.030	0.566	-0.071**	0.732	-0.087***
		[0.038]		[0.049]		[0.027]		[0.028]
Distance	6.63	2.01***	5.34	1.79***	6.03	0.48	5.19	0.49
(to assigned school)		[0.51]		[0.56]		[0.30]		[0.54]
School Quality Measures								
Academic	-0.076	0.183	-0.705	0.502***	-0.151	0.299***	-0.747	0.328**
(Test Scores)		[0.117]		[0.161]		[0.102]		[0.129]
Behavior	-0.041	0.449***	-0.706	0.870***	-0.126	0.289***	-0.836	0.452***
(Absent/Suspended)		[0.066]		[0.154]		[0.103]		[0.104]
Teacher Quality	-0.160	0.055	-0.772	0.435**	-0.155	0.382***	-0.455	0.472***
		[0.120]		[0.202]		[0.134]		[0.150]
Revealed Preference	-0.075	0.554***	-0.538	0.906***	0.073	0.329**	-0.538	0.368**
		[0.156]		[0.191]		[0.139]		[0.156]
Magnet School	0.165	0.331***	0.087	0.365***	0.090	0.181***	0.045	0.203***
		[0.113]		[0.122]		[0.051]		[0.049]
9th Grade School Characteristics								
Percent Black					0.478	-0.013	0.615	-0.025
						[0.021]		[0.038]
Percent FRPL					0.544	-0.014	0.675	-0.009
						[0.021]		[0.039]
Academic					-0.122	0.050	-0.754	0.053
(Test Scores)						[0.088]		[0.127]
Behavior					-0.097	0.035	-0.869	0.221
(Absent/Suspended)						[0.068]		[0.169]
Sample Size	1014				1081			

Notes: Each point estimate is from a regression like equation (2), where lottery status is fully interacted with indicators for whether an applicant is in the 1st-4th or 5th arrest risk quintiles. Results are for males only. Odd numbered columns present control means for each outcome, and standard errors are below each estimate in brackets and are clustered at the lottery (i.e. choice by priority group) level. Each peer input measure is calculated using data from the school year prior to the lottery and excludes sample members from the base rate calculation. Each school quality measure is normalized separately at the middle and high school level. Test scores are the average of prior year (or latest available) math and reading scores, and behavior is the same but for absences and out-of-school suspensions. Teacher quality is the average of the percentage of teachers with less than 3 years of experience, and a measure of undergraduate college competitiveness based on the Barron's rankings. Revealed preference is the school-level residual from a conditional logistic regression which predicts the probability that students will choose each school, conditional on a polynomial in distance and home school fixed effects. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Table 3: Effect of Winning the Lottery on Crime

	Full Sample				High School		Middle School	
	Risk Quintiles 1-4		Top Risk Quintile		Top Risk Quintile		Top Risk Quintile	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Felony Arrests	0.102	0.013	0.724	-0.123	0.761	-0.329***	0.699	0.105
		[0.035]		[0.097]		[0.126]		[0.175]
Total Social Cost	7,140	-12,185	36,464	-30,309	11,000	-14,106**	54,079	-42,799
		[7853]		[19,414]		[8,194]		[34,594]
Total Social Cost (murder trimmed)	1,350	-563	11,886	-5,948***	11,000	-3,916**	12,500	-7,843**
		[644]		[2,056]		[1,987]		[3,285]
Sentence-Weighted (in months)	3.8	3.1	52.5	-25.9**	58.6	-23.1*	48.3	-31.0**
		[2.5]		[10.6]		[11.7]		[14.5]
Total Days Incarcerated	7.8	5.2	70.0	-29.9***	91.4	-26.7	55.5	-36.2***
		[4.3]		[11.1]		[21.5]		[12.3]
Felony Charges								
Index Property	0.084	0.077*	0.404	0.018	0.435	-0.220	0.383	0.261
		[0.040]		[0.130]		[0.236]		[0.165]
Index Violent	0.023	0.019	0.378	-0.233*	0.272	-0.081	0.451	-0.379*
		[0.021]		[0.140]		[0.198]		[0.213]
Drug Felonies	0.035	-0.024	0.356	-0.089	0.478	-0.327**	0.271	0.174
		[0.024]		[0.091]		[0.148]		[0.139]
Other Felonies	0.053	0.049	0.387	-0.148	0.489	-0.279*	0.316	-0.056
		[0.040]		[0.093]		[0.143]		[0.118]
Sample Size	2095				1014		1081	

Notes: Each estimate is from a regression like equation (2), where the lottery treatment is interacted with indicators for whether an applicant is in the 1st-4th or 5th arrest risk quintiles. The sample is limited to males only. The Xij vector includes the prior year's math and reading test scores, absences and out of school suspensions, plus indicators for race and free lunch status. Odd numbered columns show control means for each outcome, and standard errors are below each estimate in brackets and are clustered at the lottery level. The first four columns show results for the middle and high school samples combined. Columns 5-6 and 7-8 show results for the top risk quintile only; quintiles 1-4 are included in the model but not shown. Social cost estimates are calculated using figures from Miller, Cohen and Wiersema (1996). The sentence-weighted estimates weigh crimes according to the expected time served from the NC Structured Sentencing Act. Index Property Crimes are larceny, burglary and auto theft. Index violent crimes are murder, aggravated assault, robbery and rape. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Table 4: Impact of Winning the Lottery on Crime over Time

Years since lottery	1-2	3	4	5	6	7
High School Sample						
Median age at beginning of year	15.5	17	18	19	20	21
Number of Felony Charges	-0.013	-0.328	-0.197***	-0.585**	0.070	-0.068
	[0.147]	[0.312]	[0.070]	[0.244]	[0.108]	[0.139]
Control Mean	{0.205}	{0.422}	{0.301}	{0.761}	{0.293}	{0.196}
Social Cost - Murder Trimmed	202	728	-2,626	-2,898**	169	-185
	[726]	[1,009]	[1,773]	[1,215]	[884]	[489]
Control Mean	{831}	{1,415}	{3,517}	{2,942}	[1,555]	{841}
Days in Prison				-9.18	-8.12	-0.44
				[6.94]	[12.22]	[15.68]
Control Mean				{24.28}	{30.73}	{27.61}
Middle School Sample						
Median age at beginning of year	13	14.5	15.5	16.5	17.5	18.5
Number of Felony Charges			0.032	-0.355*	0.042	0.246
			[0.112]	[0.188]	[0.176]	[0.168]
Control Mean			{0.163}	{0.549}	{0.429}	{0.338}
Social Cost - Murder Trimmed			-1,958	-2,282	-1,287	-2,383
			[2,197]	{1,412}	[978]	[1,780]
Control Mean			{2,475}	{2,598}	{1,972}	{5,151}
Days in Prison				-9.59***	-14.17***	-18.20**
				[2.67]	[4.49]	[7.62]
Control Mean				{11.31}	{21.23}	{24.97}

Notes: Each point estimate is from a regression like equation (2), where the lottery treatment variable is interacted with indicators for whether an applicant is in the 1st-4th or 5th arrest risk quintiles. Results are for males only. The Xij vector includes the prior year's math and reading test scores, absences and out of school suspensions, plus indicators for race and free lunch status. The effects are divided into years since random assignment, counting from June 1st of 2002. Standard errors are below each estimate in brackets and are clustered at the lottery (i.e. choice by priority group) level, and control means are below the standard errors in curled brackets. Social cost estimates are calculated using figures from Miller, Cohen and Wiersema (1996) and include victimization, but not justice system costs such as police or prisons. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Table 5: Effect of Winning the Lottery on Test Scores and Course-Taking

	High Schools		Middle Schools	
	Top Risk Quintile		Top Risk Quintile	
School Discipline	(1)	(2)	(3)	(4)
Unexcused Absences - 2003	11.10	-0.88	8.22	-2.30**
(in days)		[1.70]		[1.12]
Unexcused Absences - 2004	9.52	-0.96	8.00	-0.80
(in days)		[2.40]		[1.48]
Days Suspended - 2003	9.54	-3.73**	10.70	0.74
		[1.62]		[2.30]
Days Suspended - 2004	6.31	-0.24	10.90	-0.97
		[1.59]		[1.76]
Serious Incident - 2006-2007			0.158	-0.143***
(Police, Long Term Suspension, Expelled)				[0.042]
Test Scores and Course-Taking				
Math Score - 2003			-1.030	0.052
(in SD units)				[0.100]
Math Score - 2004			-0.927	-0.090
(in SD units)				[0.102]
Reading Score - 2003			-1.164	-0.076
(in SD units)				[0.172]
Reading Score - 2004			-1.190	-0.084
(in SD units)				[0.151]
9th Grade English Score	-1.195	-0.067	-1.033	-0.066
		[0.171]		[0.179]
Remedial Math	0.366	-0.191**	0.209	0.022
(<Algebra I, 9th Grade)		[0.078]		[0.090]
Math Credits - Grades 9-10	1.051	0.094	0.833	0.104
		[0.112]		[0.113]

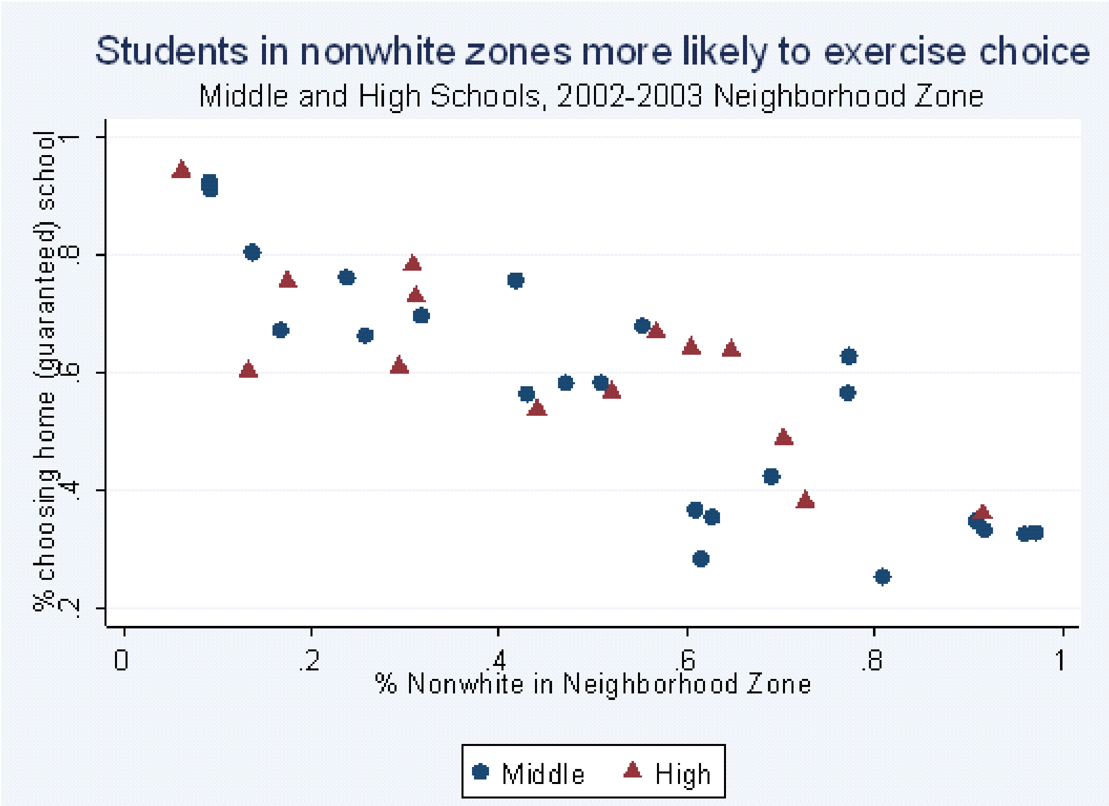
Notes: Each point estimate is from a regression like equation (2), where the lottery treatment variable is interacted with indicators for whether an applicant is in the 1st-4th or 5th arrest risk quintiles. Results are for males only. The Xij vector includes the prior year's math and reading test scores, absences and out of school suspensions, plus indicators for race and free lunch status. Odd numbered columns present control means for each outcome, and standard errors are below each estimate in brackets and are clustered at the lottery (i.e. choice by priority group) level. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Table 6: Effect of Winning the Lottery on School Enrollment

	High Schools		Middle Schools	
	Top Risk Quintile		Top Risk Quintile	
Enrollment	(1)	(2)	(3)	(4)
In CMS - Grade 9 Year	0.930	0.014	0.767	0.032
		[0.056]		[0.054]
In CMS - Grade 10 Year	0.673	-0.023	0.586	0.181***
		[0.082]		[0.068]
In CMS - Grade 11 Year	0.541	0.052	0.519	0.091
		[0.073]		[0.076]
In CMS - Grade 12 Year	0.348	0.008	0.376	-0.032
		[0.080]		[0.073]
Grade Progression				
"On Track" - Grade 9 Year	0.698	0.146**	0.534	0.032
		[0.056]		[0.054]
"On Track" - Grade 10 Year	0.345	0.133	0.271	0.055
		[0.084]		[0.065]
"On Track" - Grade 11 Year	0.207	0.121*	0.233	-0.079
		[0.071]		[0.054]
"On Track" - Grade 12 Year	0.163	0.030	0.173	-0.067
		[0.071		[0.047]
Final Status				
CMS Graduate	0.272	-0.029	0.105	-0.033
		[0.089]		[0.036]
Still Enrolled - 2009			0.143	0.031
				[0.064]
Verified Dropout (>9th Grade)	0.272	-0.064	0.226	0.103
		[0.054]		[0.065]
Transfer	0.207	0.098	0.278	-0.066
		[0.083]		[0.054]
No Show	0.250	-0.003	0.248	-0.035
		[0.052]		[0.058]

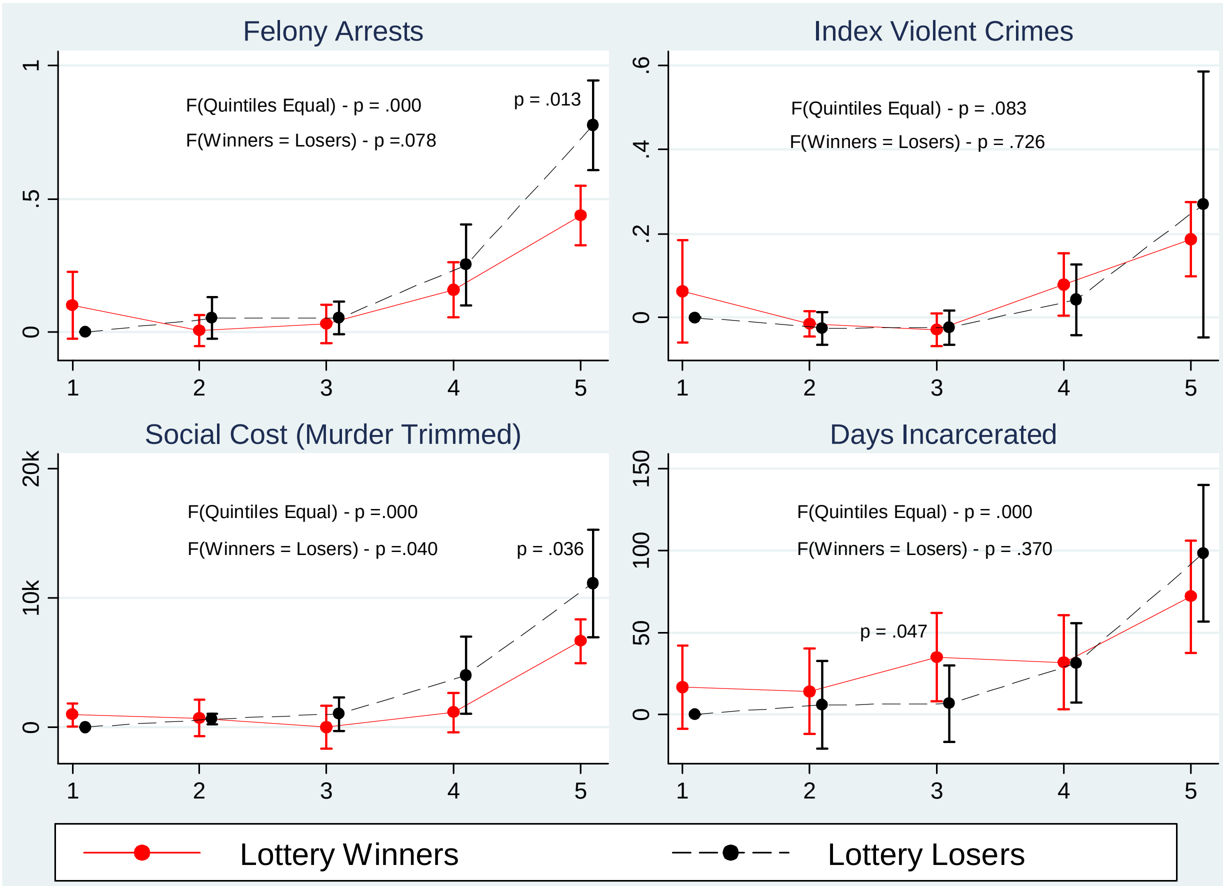
Notes: Each point estimate is from a regression like equation (2), where the lottery treatment variable is interacted with indicators for whether an applicant is in the 1st-4th or 5th arrest risk quintiles. Results are for males only. The Xij vector includes the prior year's math and reading test scores, absences and out of school suspensions, plus indicators for race and free lunch status. Odd numbered columns present control means for each outcome, and standard errors are below each estimate in brackets and are clustered at the lottery (i.e. choice by priority group) level. The enrollment variables track whether a student is enrolled in any CMS school in the year they would have been in each grade if they were progressing "on time". "On track" is defined as whether a student has advanced at least one grade per year since the lottery and is not enrolled in an alternative school. See the text for a discussion of the final status variables. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Figure 1



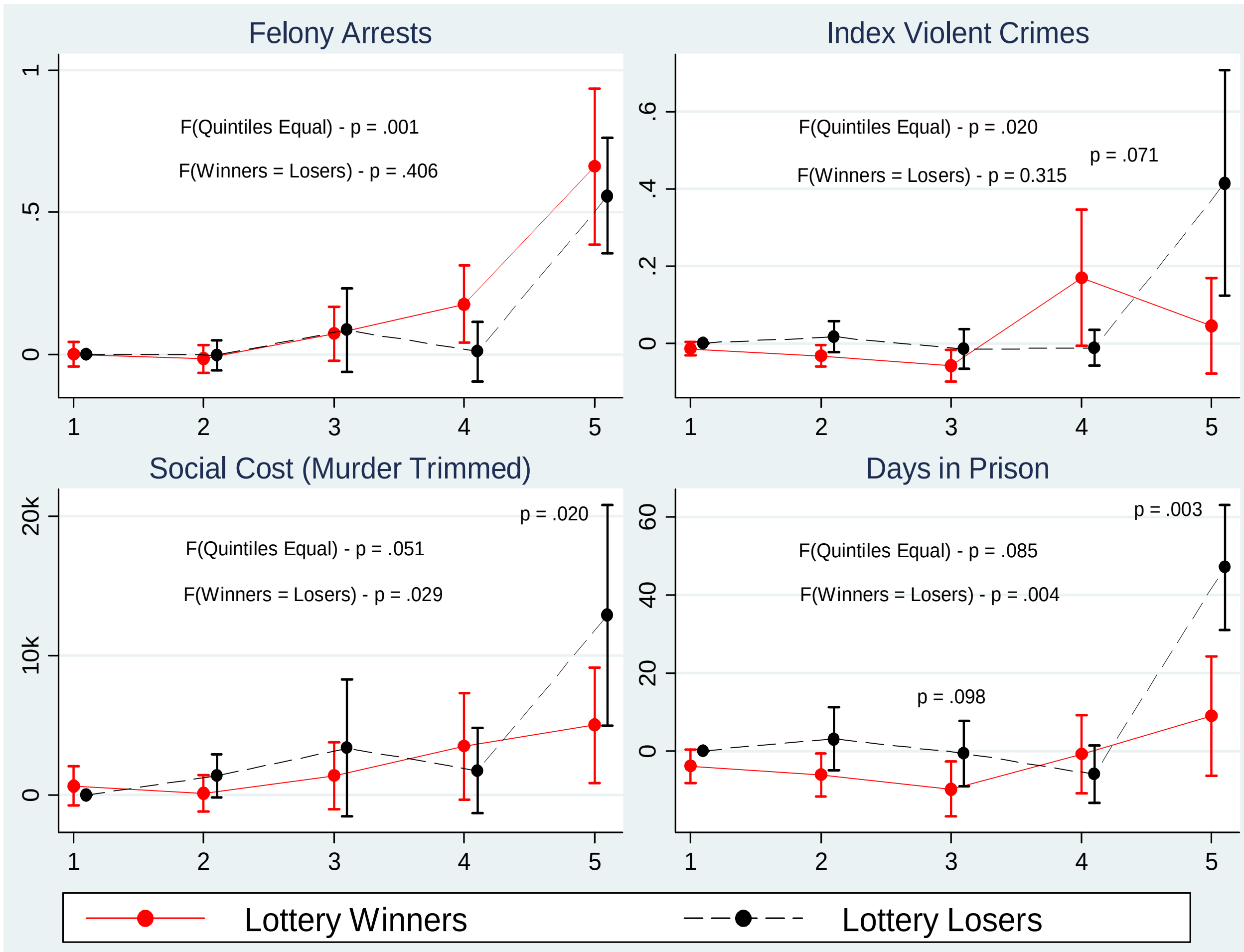
Students in nonwhite zones more likely to exercise choice
Middle and High Schools, 2002-2003 Neighborhood Zone
% choosing home (guaranteed) school
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
% Nonwhite in Neighborhood Zone
Middle
High

Figure 2 – Effect of Winning the Lottery on Crime, by Arrest Risk Quintile
High School Sample (N=1,014)



Notes: Each point estimate and 90 percent confidence interval are taken from a regression like equation (2) where the lottery treatment is fully interacted with indicators for whether a youth is in each risk quintile. F-tests for equality of treatment and control groups across all five quintiles and for equality of quintiles in levels are presented on each graph, as are test for equality within each quintile when statistically significant. The Days in Prison outcome is available for African-American males only (N=610).

Figure 3 – Effect of Winning the Lottery on Crime, by Arrest Risk Quintile
Middle School Sample (N=1,081)



Notes: Each point estimate and 90 percent confidence interval are taken from a regression like equation (2) where the lottery treatment is fully interacted with indicators for whether a youth is in each risk quintile. F-tests for equality of treatment and control groups across all five quintiles and for equality of quintiles in levels are presented on each graph, as are test for equality within each quintile when statistically significant. The Days in Prison outcome is available for African-American males only (N=649).

A Data Appendix

A.1 Sample and Data Sources

The analysis sample consists of 44,028 students in grades 6 through 11 who were enrolled in any CMS school in the previous year. These students listed as first choices 28 different middle schools and 17 different high schools. 26,474 students listed first a school to which they were guaranteed admission. Of the remaining 17,554 students, 5,033 were in lotteries where no students were offered admission, and 8,310 were in lotteries where all students were accepted. This left 4,211 students with admission to a first choice school that was subject to randomization (1,891 in high school and 2,320 in middle school). Nearly all schools had some applicants that were randomized (24 of the 28 middle schools, and 16 of the 17 high schools). Together with different priority groupings for grades and free lunch-eligible applicants, there were 72 lotteries in the middle school sample and 34 lotteries in the high school sample. About 46 percent of high school lottery applicants and 38 percent of middle school lottery applicants were admitted to their first choice school, although this varied tremendously by lottery.

The lottery file comes from **?** and includes students' individual choices, priority groupings, and lottery numbers. Within each priority group, lottery numbers were randomly assigned to students and slots were filled in ascending order by lottery number. I verified that the lottery numbers were accurate by plotting the probability of enrollment against within-priority-group lottery numbers and looking for evidence of a sharp break in enrollment at the minimum number cutoff. These graphs are available on request.

A.1.1 CMS Administrative Data

CMS maintains yearly student records that are linked longitudinally with a unique student identification number. The North Carolina Department of Public Instruction (NCDPI) requires CMS to report end-of-year (EOY) files for each school and grade with student enrollment, demographics, behavior measures and yearly test scores in a standard format. In addition to basic demographic information, these files include standardized math and reading End-of-Grade (EOG) tests for grades 3 through 8, End-of-Course (EOC) exams scores for specific subjects (such as Algebra I, Chemistry, and English I) taken mostly in high school, excused and unexcused absences, total days out-of-school suspended, special education classifications (with information about the nature and severity of the disability) and limited English proficiency status.

In addition to these EOY files, I have obtained more detailed information under a data use agreement with CMS and the Harvard Center for Education Policy Research (CEPR). The data are stored on secure computers with no internet connectivity in a room at CEPR. Access is restricted to identified researchers by means of a keycard system. The data include student's name, date of birth, and exact address. They also include yearly course enrollment information and grade received, which I can use to construct measures of grade point average and accumulated credits. I use address information to group students into census tract-by-school zone "neighborhoods", and I control for these neighborhood fixed effects in the crime prediction regression in Section 3.1. Following **?**, I also use address information to calculate straight-line distance from each student's home to each school, which I use in the revealed preference calculation in Table 1.

The CMS administrative data also contains dates of school enrollment and withdrawal. Each spell of enrollment has an associated withdrawal code. Withdrawal codes include high school graduation, transfer within CMS, transfer to private or charter schools, transfer to another public school in-state, out-of-state transfer, dropout, and no show, as well as other categories such as assignment to alternative schools, expulsion and death. CMS also provided a teacher information file, which includes courses taught, years of experience and information about the colleges attended and degrees obtained. I match each teacher's undergraduate institution to the Barron's Profile of American Colleges 2009, which groups schools into categories such as "competitive", "very competitive, and "most competitive", and use these classifications in the measure of teacher quality in Table 1.

A.2 Crime Data Collection and Match Process

Arrest data at the county level come from the Mecklenburg County Sheriff. The data include all arrests made in Mecklenburg county, including by arresting agencies with other jurisdictions (ex. Immigration and Naturalization Services, the US Marshals and other federal agencies, as well as city police from Charlotte and surrounding smaller cities). The data include all arrests made beginning on January 1st, 2006 through June 15th, 2009, with the exception of the approximately 3 percent of arrests that were expunged or missing. The data are collected at the arrest level, and include information on the classification (felony, misdemeanor, traffic), processing (bond amount, warrant, etc.) and exact description all associated charges at the time of arrest. Each arrest is assigned a unique 7 digit number in the order that it is processed, and first time arrestees are assigned a unique 6 digit identification number (established by fingerprinting) that links them across multiple arrests, if any. I have information on each arrestee's name and date of birth, which I use to match to the CMS administrative data, as well as home address at the time of arrest. MCS incarceration data cover the same period of time as the arrest data and are kept in a similar format. The unique 6 digit identification number links individuals to all spells of incarceration in MCS jails, and the associated charges. The data include name and date of birth and the first and last day of each incarceration spell.

The original source for the 2006-2009 Mecklenburg county arrest and incarceration data is http://www.charmeck.org/Departments/MCSO/Inmate+Information/InmateLookup.htm. As the website states, "North Carolina Law makes this information public. The Mecklenburg County Sheriff's Office provides it via the internet for your convenience." The arrest data can be found at http://arrestinquiryweb.co.mecklenburg.nc.us/ and the incarceration data at http://mcsowebsvr.co.mecklenburg.nc.us/inmatesearch/inmate_search.asp. Both websites allow users to access information that is up to 3 years old, counting from the day the website is accessed (since I started collecting the data on January 1st, 2009, my data begin on January 1st, 2006). I collected the data by writing a script (also known as a macro) in an automation language called AutoIt. This program, which is similar to the more commonly used Perl, allows me to automate keystrokes, mouse clicks and other basic computer functions. MCS assigns arrest numbers consecutively in the order they are processed, so I wrote a script that entered arrest numbers in order into the website and copied all the relevant information into a text file. The websites both include name and date of birth, so I was able to connect arrests to individuals, and then individual arrestees (in some cases) to student records in CMS. Because of the format of the website, I was unable to fully automate collection of the incarceration data. Therefore, I collected incarceration data for African-American members of the lottery sample only.

I also obtain data from the North Carolina Department of Corrections (NCDOC). These data include spells of incarceration and associated charges and convictions for individuals who serve time in state prison. Members of the lottery sample can thus be linked to crimes committed outside of Mecklenburg county, but only if they spend time in state prison for those crimes. The NCDOC data include spells of incarceration prior to 2006, but only for individuals who are incarcerated or under the supervision of the justice system (i.e. on probation) as of 2009. Data from 2006 to the present do not have this limitation. Therefore, I also limit analysis of the NCDOC incarceration data to 2006 and later, for consistency. Like the MCS incarceration data, I was unable to fully automate collection of the NCDOC data, so I restrict to African-American members of the lottery sample only. See Appendix A.5 for example screen shots from the MCS and NCDOC websites.

Finally, I matched the crime data to CMS administrative data using first name, last name, and exact date of birth. To account for inconsistencies across data sources (i.e. hyphenated names, apostrophes, "Dave" vs. "David" etc.) I employed a partial matching algorithm. I used a STATA program written by Eric Taylor at CEPR called "lndmerge" that calculates the Levenshtein distance between two variables using optimal matching of sequences. The intuition is as follows: first the matching variables in each data source (i.e. name and date of birth) are combined into a unique string. Then all the observations in both datasets are combined into a matrix, and each combination is assigned a score (or distance) based on how many changes would need to be made to obtain an exact match. Longer strings are less likely to be exact matched, and so are penalized proportionately less for a change (i.e. David-Devid would count as a worse match than DavidDeming-DevidDeming). Using this method, about 87 percent of the matches were exact. I adopted various rules for accepting partial matches (a minimum score, minimum score plus exact match on first letter of last name, or on year of birth etc.) None of these made any difference in the main results, nor did restricting the analysis to exact matches only.

I conducted a number of tests to assess the quality of the match. First, since each arrest is given a unique identification number that is assigned consecutively in the order it was processed, I can calculate the fraction of arrest numbers that are missing from the data. Counting from the first day that the data were collected, this fraction is only 3.2 percent, and there are no large gaps. This suggests that nearly every arrest processed by MCS is present in the data.[1] Figure A1 plots the age profile of arrests in Mecklenburg County by type of offense. The Federal Bureau of Investigation (FBI) collects data on eight different "index" crimes for the Uniform Crime Reporting (UCR) Program, which covers law enforcement agencies across the country. Index property crimes are burglary, motor vehicle theft and felony larceny. Index violent crimes include murder/manslaughter, rape, robbery and aggravated assault.[2] The last category I include is felony drug offenses, which (based on weight) range from "possession with intent to distribute" all the way up to "trafficking." Index property and violent crimes peak at ages 17 and 18 respectively, which is consistent with other cohort studies of crime and delinquency (**???**). Interestingly, drug felony arrests do not peak until the early to mid twenties, and decline much more slowly with age than other categories of crime.

In the top panel of Table A1, I examine arrest rates of CMS attendees overall and by demographic group. I use six school cohorts of data, corresponding to students in grades 6 through 11 in 2002 and age 17 to 23 in 2009. The first and second rows show the fraction

[1] Most of the missing arrests have been expunged, and there is a slight increase in the number of expunged arrests in earlier years.

[2] The eighth crime is arson. The incidence of arson is very low in these data, so I do not include it here.

of CMS attendees who have a criminal record, and who have at least one felony arrest respectively, by race and gender. Not surprisingly, arrest rates vary dramatically, from about 34 percent for African-American males to about 3 percent for White or Asian females. Rows three through five show arrest rates by type of crime. African-American males are about six times more likely than white males to have at least one felony arrest, and about thirteen times more likely to be arrest for an index violent crime.

In the bottom panel of Table A1, I examine the percentage of arrests that are successfully matched to a CMS student by birth year and demographic group. Unmatched arrests could be students who were enrolled in private school, youth who travel to Mecklenburg County from elsewhere to commit crimes, or poor data quality. Match rates are highest for African-Americans (who are more likely than whites to attend public school) and for more recent birth years.[3] Since the CMS data only go back to the 1996-1997 school year, any student who left the district before that would not be matched. Since most criminals are high school dropouts, this is likely to result in fewer matches for the earliest birth cohorts. However, the weighted average match rate by birth year for the lottery sample exceeds 85 percent overall and 90 percent for African-American males. This high match rate is strong evidence of the quality of the data. It also highlights the important role that public school policies might play in city crime rates.

A.3 Selection into the Lottery Sample

Table A2 presents the average characteristics of lottery applicants compared to all CMS students. Column 1 shows control means and Column 2 shows coefficients from regressions of observable characteristics of students on an indicator for whether the student listed a non-guaranteed school as their first choice. Unlike many other instances of school choice, applicants to non-guaranteed schools are more disadvantaged than students who choose their

[3] Illegal aliens who are arrested by Immigration and Naturalization Services (INS) in North Carolina are often processed in Charlotte before they are sent to Atlanta and deported. This, along with the transient nature of the Charlotte's rising Latino population, accounts for the very low match rate among Latino arrestees.

neighborhood school. They are nearly twice as likely to be nonwhite and free or reduced price lunch eligible. Applicants to non-guaranteed schools also score about 0.4 standard deviations lower on both math and reading exams, and have been suspended and absent more days in the previous school year. Column 3 includes neighborhood school fixed effects, to assess the nature of within-school selection. Column 4 presents control means and Column 5 presents estimates where the sample is restricted to neighborhood schools where 60 percent or more of the assigned students are African-American or Latino.

Although applicants to non-guaranteed schools are more disadvantaged across schools, they are relatively similar on observables within the schools from which most of the lottery sample comes. Column 5 shows that, even with predominately minority schools, non-guaranteed applicants have test scores that are very similar to students who chose the neighborhood school. Furthermore, even within these high minority schools, applicants to non-guaranteed schools are absent and suspended more often. Column 6 looks only at students who were in non-degenerate lotteries (where the probability of admission was neither zero nor one). We see that applicants in the lottery sample have slightly higher test scores (about 0.1 standard deviations). However, this is largely because of the "priority boost" given to economically disadvantaged applicants, many of whom were automatically admitted and thus not subject to randomization.[4] Overall, the lottery sample is more disadvantaged than the average CMS student, but quite representative on observables of the students who attend high minority schools.

A.4 Arrest Prediction

I estimate the probability that a student will have at least one arrest as a function of yearly test scores in math and reading, absences and out-of-school suspensions, special education classifications, and neighborhood school zone by census tract fixed effects using each student's exact address in the year prior to open enrollment. For the high school sample I use data

[4]Because of the separate priority group assigned to FRPL students who apply to non-FRPL schools, most schools either had lotteries for them and denied everyone else, or automatically admitted them and had lotteries for non-FRPL students.

from grades 6 through 8, and grades 3 through 5 for the middle school sample.[5] I allow for second order polynomials in all of the continuous measures. The coefficients from the regression are listed in Table A3. In Columns 3 and 4, I reestimate the model with males only. These coefficients, which are the ones actually used in the crime prediction for the main results, differ very little from the prediction for the overall sample. Figure A2 plots the density of predicted criminality for all CMS students in grades 6 to 11, then for African-American males overall and from the seven lowest-performing schools (defined by average test scores) in the district. The distribution shifts rightward noticeably for these "high risk" subgroups.

A.5 Social Cost of Crime Calculations

The social cost of crime estimates from **?** include tangible costs such as lost productivity, medical and mental health care and other social services, and property damage. They also include estimates of intangible costs such as quality of life (based in part on the amount individuals are willing to pay to reduce the risk of death, and the compensatory component of jury damage awards - see **?** for details). Intangible costs make up most of the estimated cost of violent crimes, and are inherently difficult to monetize. Notably, the study does not include criminal justice system costs such as policing, crime and arrest processing, or incarceration. It also does not include the costs undertaken by individuals to avoid crime. Here I list the costs for the index property and index violent crimes, plus a few other notable crimes that drive the main estimates in the paper (all estimates are converted to 2009 dollars).

1. Murder - $4.38 million

2. Rape - $129,630

3. Aggravated Assault - $35,760

[5] In North Carolina, standardized End-of-Grade (EOG) tests in math and reading are administered for grades 3-8 only. While additional years of data would improve the precision of the estimates, it would also increase the percentage of respondents with missing data.

4. Domestic Assault - \$16,390
5. Simple Assault - \$2,980
6. Robbery - \$11,920
7. Motor Vehicle Theft - \$5,513
8. Burglary - \$2,086
9. Larceny - \$551

? do not monetize all crimes, and notably they exclude drug crimes from the estimation. One alternative is to impute a cost of zero for all drug crimes. This leaves the estimates for the middle school sample unchanged, but reduces the social cost estimates for the high school sample by approximately 25 percent. In the main estimates in the paper, I impute a cost of drug felonies that is equivalent to felonies of the same standing under the North Carolina Structured Sentencing Act. This varies by crime and the "schedule" of the controlled substance (for example, cocaine is schedule 2 and punished more severely than marijuana, which is schedule 6). The approximate classifications are below (for marijuana, crimes are roughly one step down in severity, so trafficking in marijuana = sell/deliver cocaine, roughly):

1. Drug Trafficking = Robbery = \$11,920
2. Sell/Deliver = Motor Vehicle Theft = \$5,513
3. Possession with Intent to Distribute = Burglary = \$2,086
4. Simple Possession (Felony) = Larceny = \$551

Appendix Tables and Figures

Table A1: Arrest Rates and Match Between School District and Arrest Data

Panel A: Arrest Rates by Race/Gender and Crime Type

	African-American		Hispanic		White/Asian		
	Male	Female	Male	Female	Male	Female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ever Arrested							
Any Arrest	0.34	0.13	0.16	0.04	0.10	0.03	
Any Felony	0.20	0.03	0.08	0.01	0.03	0.01	
Index Property	0.09	0.01	0.04	0.00	0.01	0.00	
Index Violent	0.07	0.00	0.02	0.00	0.01	0.00	
Drug Felony	0.08	0.01	0.01	0.00	0.01	0.00	
Sample Size	8,834	8,493	519	504	9,095	8,748	

Panel B: Percent of Arrests Matched to a CMS Attendee

	African-American		Hispanic		White/Asian		All
Year of Birth	Male	Female	Male	Female	Male	Female	Felonies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1980	0.26	0.20	0.01	0.00	0.11	0.04	0.19
1981	0.59	0.39	0.02	0.03	0.27	0.22	0.44
1982	0.65	0.56	0.03	0.08	0.34	0.25	0.53
1983	0.73	0.73	0.03	0.09	0.43	0.33	0.62
1984	0.72	0.66	0.04	0.09	0.48	0.42	0.64
1985	0.79	0.76	0.08	0.04	0.49	0.42	0.70
1986	0.83	0.74	0.12	0.24	0.53	0.43	0.75
1987	0.85	0.78	0.13	0.24	0.59	0.53	0.80
1988	0.90	0.86	0.23	0.31	0.72	0.67	0.85
1989	0.93	0.88	0.40	0.76	0.73	0.71	0.89
1990	0.93	0.91	0.57	0.75	0.82	0.68	0.90
1991	0.94	0.92	0.79	0.88	0.80	0.81	0.91
1992	0.95	0.94	0.74	0.83	0.81	0.80	0.91
1993	0.97	0.82	0.75	1.00	0.80	0.57	0.95
All Years	0.77	0.72	0.13	0.22	0.49	0.42	0.69
Sample Size	32,598	7,459	10,392	715	12,161	4,085	19,184

Notes: The sample in panel A consists of CMS attendees in grades K-5 in 1997 (ages 17-23 in 2009) that are still in CMS in grade 8 or higher. Index property crimes are felony larceny, burglary and motor vehicle theft. Index violent crimes are murder/manslaughter, aggravated assault, robbery and kidnapping. In Panel B the denominator is all arrests in Mecklenburg County.

Table A2: Selection into the Lottery Sample

	Outcome - Chose Non-Guaranteed School					
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.51	-0.01	-0.01	0.50	0.00	-0.00
		[0.01]	[0.01]		[0.01]	[0.01]
African-American or Latino	0.40	0.27***	0.13***	0.73	0.03	0.04
		[0.01]	[0.03]		[0.03]	[0.03]
Free / Reduced Lunch	0.40	0.26***	0.12***	0.71	0.04	0.01
		[0.01]	[0.02]		[0.02]	[0.03]
Math (standardized)	0.15	-0.41***	-0.16***	-0.36	-0.03	0.13***
		[0.01]	[0.01]		[0.04]	[0.04]
Reading (standardized)	0.15	-0.41***	-0.16***	-0.37	-0.04	0.11**
		[0.01]	[0.04]		[0.05]	[0.05]
Days Suspended	0.63	0.55***	0.36***	0.99	0.33***	0.04
		[0.04]	[0.08]		[0.13]	[0.08]
Days Absent	7.32	1.48***	1.02***	7.94	0.97***	0.37
		[0.09]	[0.19]		[0.31]	[0.30]
Home School FE			X	X	X	X
>60% Nonwhite Only				X	X	X
Non-Degenerate Lotteries Only						X
Sample Size	44,028			18,353		

Notes : The sample is all CMS students in rising grades 6-11 in the fall of 2002 who were enrolled in any CMS school in the previous year. The first column presents the control mean and the second column presents coefficients from a regression of the variable in each row on an indicator for whether the student listed a non-guaranteed school as their first choice. The third column adds neighborhood school fixed effects. Columns 4 shows the control mean and Column 5 shows estimates when the sample is restricted to schools where the assigned student population is 60% or more nonwhite. In Column 6 the independent variable of interest is an indicator for whether the student was in the lottery sample (i.e. they were in a priority group where the probability of admission was neither zero nor one.) Free or reduced price lunch is an indicator of socioeconomic status. Math and Reading are standardized scores administered in the years that students were in 5th grade (for middle school) and 8th grade (for high school). Standard errors are clustered at the neighborhood school level. * - sig. at 10% level. ** - sig. at 5% level. *** - sig. at 1% level.

Table A3: Arrest Prediction

Dependent Variable: Ever Arrested (Logit Coefficients)

	All		Males Only	
	High (6-8 Xs)	Middle (3-5 Xs)	High (6-8 Xs)	Middle (3-5 Xs)
Demographics				
Male	**1.16 (0.05)**	**0.93 (0.05)**		
Black	**0.47 (0.07)**	**0.41 (0.07)**	**0.50 (0.08)**	**0.41 (0.08)**
Latino	**-0.70 (0.16)**	**-0.29 (0.11)**	**-0.60 (0.18)**	**-0.24 (0.13)**
FRPL	**0.32 (0.07)**	**0.47 (0.07)**	**0.31 (0.08)**	**0.37 (0.08)**
Math Scores				
6th / 3rd	-0.05 (0.07)	0.03 (0.06)	-0.03 (0.08)	0.03 (0.08)
squared	0.02 (0.03)	-0.02 (0.03)	0.02 (0.04)	-0.03 (0.03)
7th / 4th	-0.05 (0.07)	-0.01 (0.06)	-0.11 (0.09)	-0.05 (0.08)
squared	-0.00 (0.03)	0.01 (0.03)	-0.04 (0.04)	0.00 (0.03)
8th / 5th	-0.10 (0.07)	**-0.19 (0.06)**	-0.05 (0.08)	**-0.23 (0.07)**
squared	-0.05 (0.03)	-0.00 (0.02)	-0.04 (0.04)	0.03 (0.03)
Reading Scores				
6th / 3rd	**-0.14 (0.07)**	-0.09 (0.06)	-0.13 (0.08)	**-0.16 (0.08)**
squared	**-0.09 (0.03)**	-0.01 (0.03)	**-0.09 (0.04)**	-0.01 (0.04)
7th / 4th	**-0.14 (0.07)**	-0.05 (0.06)	**-0.13 (0.08)**	0.03 (0.08)
squared	-0.01 (0.03)	-0.04 (0.03)	0.01 (0.03)	-0.03 (0.04)
8th / 5th	-0.05 (0.06)	**-0.15 (0.06)**	-0.06 (0.07)	-0.12 (0.07)
squared	0.01 (0.02)	-0.04 (0.02)	0.02 (0.03)	-0.04 (0.03)
Special Education				
6th / 3rd	0.03 (0.09)	0.05 (0.07)	0.04 (0.09)	0.06 (0.08)
7th / 4th	-0.08 (0.11)	-0.06 (0.08)	-0.09 (0.12)	-0.08 (0.09)
8th / 5th	0.06 (0.09)	0.10 (0.06)	0.06 (0.10)	0.12 (0.07)
Days Absent				
6th / 3rd	0.002 (0.005)	0.001 (0.005)	-0.002 (0.006)	-0.005 (0.006)
7th / 4th	0.004 (0.004)	0.001 (0.005)	0.005 (0.005)	0.001 (0.001)
8th / 5th	**0.012 (0.003)**	**0.012 (0.004)**	**0.012 (0.004)**	**0.018 (0.006)**
Days Suspended				
6th / 3rd	0.015 (0.013)	**0.125 (0.039)**	0.018 (0.016)	**0.152 (0.045)**
7th / 4th	0.006 (0.011)	0.014 (0.034)	0.001 (0.013)	0.019 (0.039)
8th / 5th	0.008 (0.009)	0.028 (0.027)	0.005 (0.011)	0.037 (0.031)
Ever Suspended				
6th / 3rd	**0.29 (0.08)**	**0.31 (0.12)**	**0.29 (0.10)**	0.22 (0.14)
7th / 4th	**0.39 (0.08)**	**0.45 (0.11)**	**0.42 (0.09)**	**0.40 (0.12)**
8th / 5th	**0.60 (0.07)**	**0.54 (0.09)**	**0.53 (0.09)**	**0.51 (0.11)**
Sample Size	20,858	22,657	10,439	11,344
Pseudo R-squared	0.218	0.185	0.189	0.179
X^2 (Test Scores)	163.12	158.07	114.24	130.75
X^2 (Behavior)	538.77	390.92	357.54	270.57
X^2 (Geography)	260.51	259.28	228.5	288.3

Notes: Each row gives the logit coefficient from a regression that predicts the probability that a student will ever be arrested as a function of the covariates listed above, plus dummy variables for missing test scores in each year and census tract-by neighborhood school fixed effects. The density of these arrest predictions is graphed in Figure 3, and they are used to break students into the risk quintiles discussed in Section 3.1 The last 3 rows show test statistics for joint significance of the test score variables, the absence and suspension variables, and the geography fixed effects respectively. Values for missing data are imputed based on race and gender means, but only for students who were actually enrolled in CMS at the time. Coefficients in bold are sig. at the 5% level or greater.

Table A4: Effects of Winning the Lottery on Crime, by Race and Gender

	High School Sample				Middle School Sample			
	Male		Female		Male		Female	
	Black	Nonblack	Black	Nonblack	Black	Nonblack	Black	Nonblack
Felony Arrests	-0.148**	0.036	-0.043	-0.003	0.031	0.049	0.017	-0.023
	[0.064]	[0.047]	[0.037]	[0.003]	[0.091]	[0.051]	[0.024]	[0.017]
	{0.337}	{0.075}	{0.076}	{0.004}	{0.368}	{0.044}	{0.034}	{0.017}
Social Cost	-2,913**	375	-50	-20	-3,739**	489	-259	-44**
(murder trimmed)	[1,257]	[318]	[114]	[31]	[1,446]	[372]	[378]	[21]
	{5,399}	{607}	{336}	{44}	{5,887}	{570}	{727}	{50}
Sentence-Weighted	-7.41	2.91	-0.12	0	-9.91*	5.57	1.19	-0.16
	[6.10]	[2.14]	[0.52]	[0]	[5.71]	[4.19]	[1.83]	[0.12]
	{25.39}	{2.43}	{0.93}	{0}	{20.70}	{1.95}	{1.29}	{0.11}
Sample Size	610	404	559	318	649	432	797	442

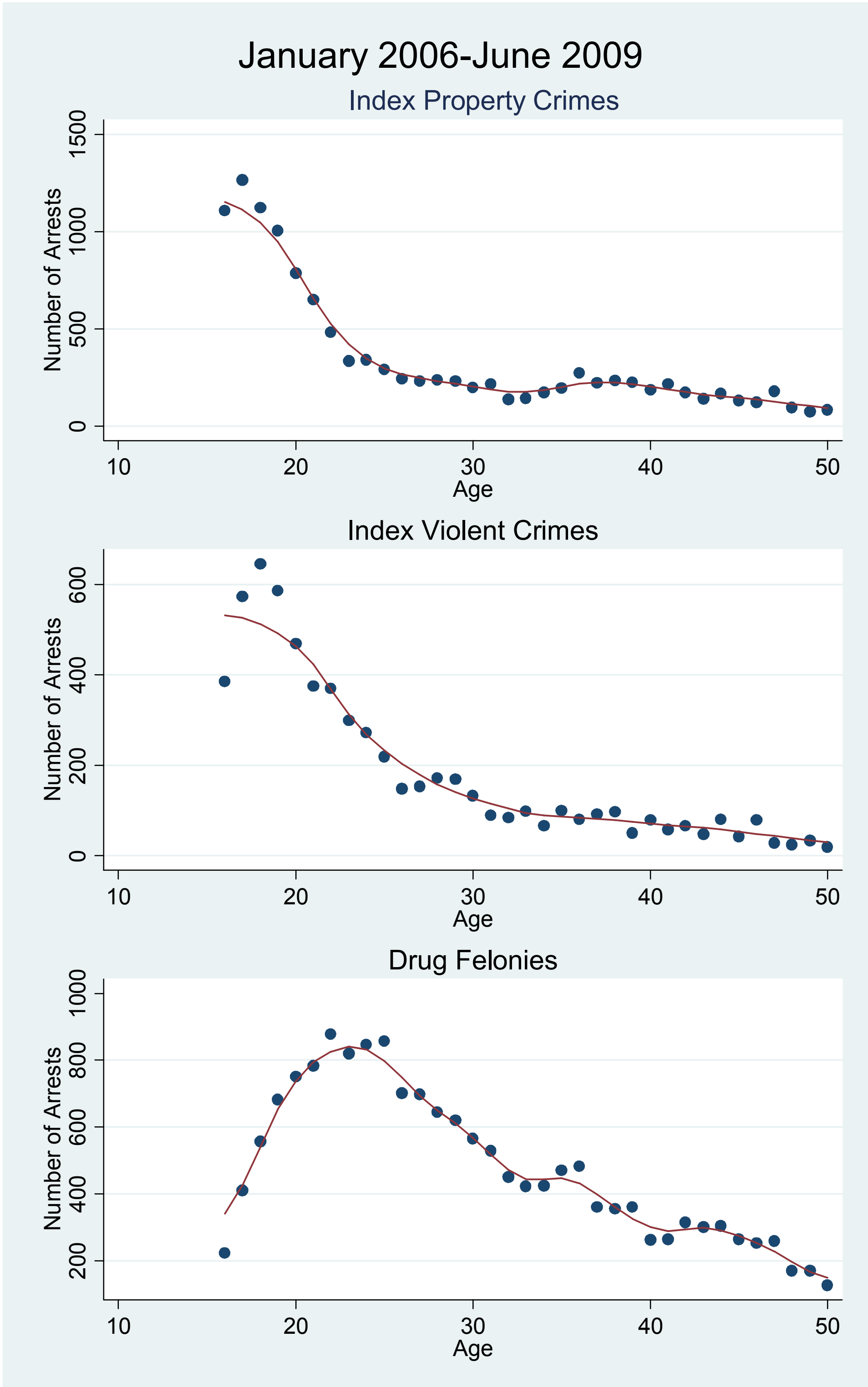
Notes : Each point estimate is from a regression like equation (1). The Xij vector includes free lunch status, prior math and reading scores, absences and out of school suspensions. Standard errors are below each estimate in brackets and are clustered at the lottery (i.e. choice by priority group) level. Control means are below the standard errors in curled brackets. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Table A5: Effect of Winning the Lottery on Crime - Alternate Specifications

Top Risk Quintile Only	High				Middle			
	OLS	Logit	Poisson	NBR	OLS	Logit	Poisson	NBR
Felony Arrests	-0.352**	-0.992***	-0.787***	-0.599***	0.101	0.226	0.020	0.069
	[0.126]	[0.317]	[0.243]	[0.228]	[0.180]	[0.405]	[0.268]	[0.236]
Total Days Incarcerated	-27.6	0.122	0.015	0.100	-38.3***	-0.39	-1.29***	-0.23
	[19.6]	[0.246]	[0.520]	[0.168]	[12.5]	[0.39]	[0.42]	[0.25]
Felony Charges								
Index Property	-0.239	-0.747	-0.697	-0.843*	0.261	0.648	0.430	0.286
	[0.250]	[0.539]	[0.544]	[0.477]	[0.173]	[0.565]	[0.328]	[0.399]
Index Violent	-0.089	0.384	-0.427	0.285	-0.376*	-0.690	-1.917**	-0.763*
	[0.199]	[0.719]	[0.878]	[0.595]	[0.201]	[0.457]	[0.773]	[0.453]
Drug Felonies	-0.342**	-1.680***	-1.454*	-0.996***	0.169	0.038	0.277	0.131
	[0.151]	[0.336]	[0.845]	[0.346]	[0.136]	[0.417]	[0.706]	[0.477]
Other Felonies	-0.287*	-0.708	-0.984	-0.285	-0.067	0.517	-0.336	0.091
	[0.145]	[0.702]	[0.668]	[0.619]	[0.123]	[0.361]	[0.412]	[0.350]
Sample Size	1014				1081			

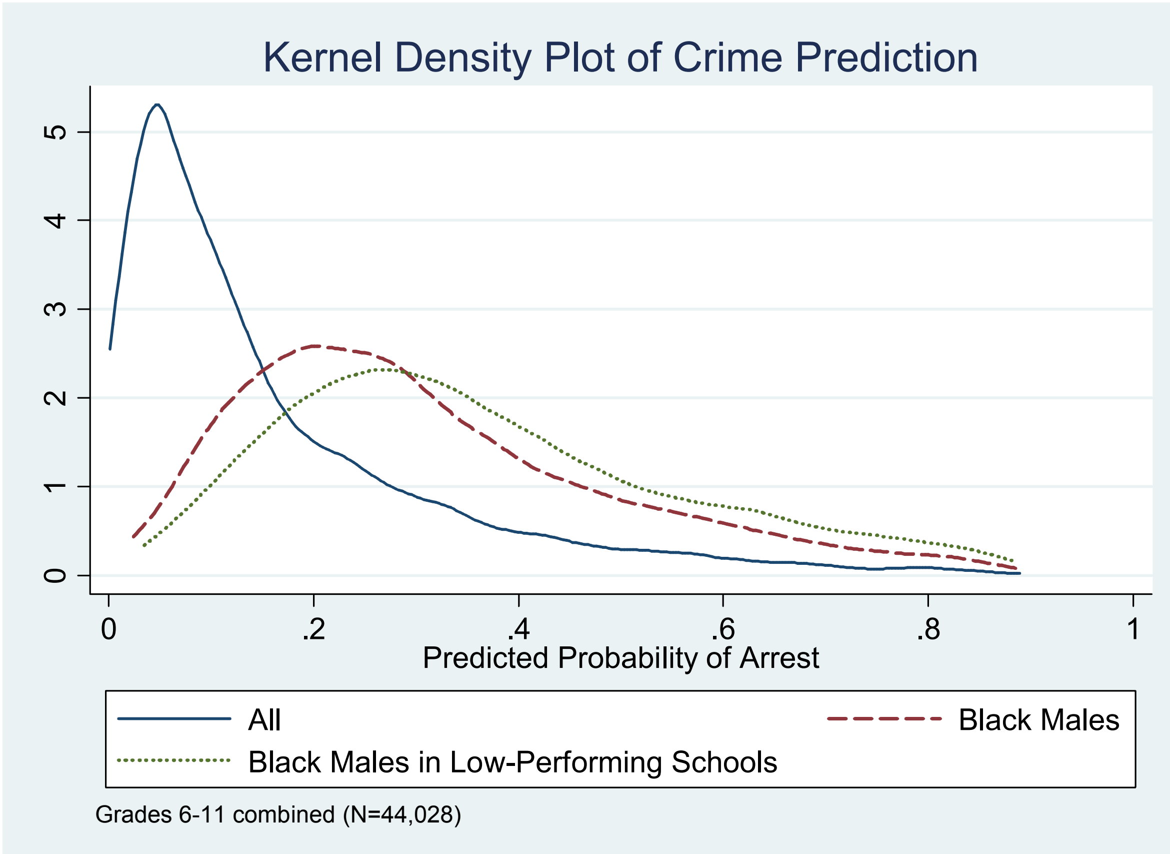
Notes: Each estimate is from a regression like equation (2), where the lottery treatment is interacted with indicators for whether an applicant is in the 1st-4th or 5th arrest risk quintiles. The Xij vector includes only the predicted probability of arrest estimated in Section 3.1. Block bootstrapped standard errors (with lotteries as clusters) are below each estimate in brackets. The first column contains OLS estimates, repeating the results in Table 4. The second column estimates a logit and converts each outcome into an indicator variable. Columns 3 and 4 present results using poisson and negative binomial count models. Index Property Crimes are larceny, burglary and auto theft. Index violent crimes are murder, aggravated assault, robbery and rape. * = sig. at 10% level; ** = sig. at 5% level; *** = sig. at 1% level.

Figure A1 – Age Profile of Crime in Mecklenburg County



Notes: Includes all arrests, not just those matched to CMS students. The data begin at age 16, when youths are treated as adults by the criminal justice system in North Carolina.

Figure A2


Kernel Density Plot of Crime Prediction
0
1
2
3
4
5
0
.2
.4
.6
.8
1
Predicted Probability of Arrest
All
Black Males
Black Males in Low-Performing Schools
Grades 6-11 combined (N=44,028)