

Can Parents Make Well-Informed School Choices?

Fang Lai

New York University, fang.lai@nyu.edu

Elisabeth Sadoulet

University of California at Berkeley, sadoulet@are.berkeley.edu,

Alain de Janvry

University of California at Berkeley, alain@are.berkeley.edu

September 2007

I. Introduction

Letting parents make school choices has been the subject of heated debates. The expectation is that devolving this role to parents will benefit both schools and children. As parents are not limited to their neighborhood school but are allowed to choose among a broader array of schools to maximize their children's academic achievement, the academic quality of these schools is expected to increase as they compete to attract students. In addition, as parents have private information about their children, they can achieve a better match between students and schools than a school district could do. Finally, school choice program with randomized lotteries might improve equity of education by random assignment of applicants to school of excessive demand. This major reform in the educational system has attracted much attention from scholars in the field of education. Studies have shown that the expected favorable outcomes require three conditions. First, parents' preferences for school have to be oriented at maximizing the academic success of their children. In a context of highly heterogeneous parental preferences, this may not obtain. If not all parents put high value on the academic quality of schools, introducing school choice might not lead to improved academic performance across schools. Second, parents have to be able to translate preferences into the choice of a sequence of schools. This requires good information not only about schools but also

about the expected choices of other parents in order to make strategic decisions about the sequence of schools from which to seek admission. Better schools will be filled up early, and requesting these schools at a later stage is not strategically efficient. More specifically, even if parents do seek academic improvement via school choice and know how to translate preferences as school choices, implementation of the system requires that parents understand the rules and avoid making unsound and eventually irrational choices. Unsound choices make them forfeit an important turn in the assignment process during which good schools they may have desired are being filled. Unsound choices can be very costly in that the child ends up in a worst school and may have a lower educational achievement.

Most existing school choice programs do not encourage revelation of the true parental school preferences (Adbdulkadiroglu and Sonmez, 2003). The final school choices are combinations of authentic preferences and strategic behavior in sequencing the selected school (e.g., to hedge the risk of ending up in a least desirable school). In this context, which is much more complex than a truthful preference revelation mechanism, factors such as incomplete information about schools and the preferences of others and imperfect understanding of the rules of the game are all likely to lead to unsound school choices. whether parents can make sound choices even if well intended is a big concern, with potentially high costs on children. A further concern is that parents of different socioeconomic backgrounds might have different preferences about schools, different abilities in making well-informed and strategic choice, and possibly uneven understanding of rules. Thus, giving parents the freedom of school choice might lead to unintended social stratification in school populations and educational achievements.

Mixed evidence about the effects of school choice on academic performance (e.g., Angrist, et al., 2002¹; Howell and Peterson, 2002; Teske and Schneider, 2001; Hsieh and Urquiola, 2006; Rouse, 1998; Cullen, Jacob and Levitt, 2005, 2006; Mayer, et. al., 2002; Krueger and Zhu, 2002) have also justified concerns about school choice programs. Moreover, there have been many studies on the heterogeneity of parental preferences

¹ They found a beneficial effect of winning a private secondary school lottery on student academic performance.

about schools and the resulting stratification (e.g., Hastings, et al., 2005a; Elacqua, Schneider, and Buckley, 2006; Hsieh and Urquiola, 2006). For example, Elacqua, Schneider, and Buckley (2006) compared parents' reported school choice sets and their responses to survey questions about what school characteristics they cared about in school choice, and found that demographics significantly affect school choice of parents in the Metropolitan Region of Santiago, Chile, despite the fact that all parents claimed to value academic quality in choosing a school. Still further, many papers have examined parents' strategies during the school choice process and worked on the optimal mechanism designs of school choice programs to balance efficiency and equity (e.g., Abdulkadioglu, Pathak, and Roth 2006; Pathak, Sonmez, 2006).

Regardless of parental preferences and mechanism design considerations, a fundamental question remains: Given available information, are parents willing and do they have adequate capacity to understand the school admission system and comply with the decision rules that have been set up by the school choice program designers? Furthermore, are they equally capable of making sound school choices? Even though many studies have expressed concern about parents' ability in making well-informed school choices (e.g., Ladd, 2002; Teske and Schneider, 2001), the existing evidence is based on thin data (e.g., responses to survey questions only) that are susceptible to various survey response biases, and only provide general insights such as that parents of disadvantaged status are less aware of the availability of a voucher program. There have been few rigorous empirical studies using both survey responses and parents' actual school choices of the willingness and capacity of parents to make sound school choices, and whether people with low socioeconomic status are especially likely to make unsound choices regardless of their preferences.

This paper provides rigorous empirical evidence on the prevalence of unsound school choices and the educational cost on children of unsound parent school choices. It uses a unique data set that combines both parental survey responses and the administrative records on their actual school choice for 4,948 students who were admitted to 28 public middle schools in Beijing's Eastern City District in 1999 under a reformed school admission mechanism. Under the new mechanism, which was first implemented

in 1998, parents were offered to make up to seven sequential school choices. Schools that were oversubscribed by applicants when their turn came would randomly choose among applicants regardless of their individual characteristics. Using administrative records of parental school choices, we identify several evident patterns of unsound school choices regardless of parents' actual preferences and strategies. We examine how these choices affect the student's status in the school system and their educational performance. We relate the occurrence of parents' unsound choices to their socioeconomic status and to their child's characteristics. We further explore the reasons for these unsound choices by using parents' responses to census data regarding their knowledge of the school system and their considerations in making school choices for their child. Results suggest that, regardless of parents' preferences and strategies, unsound school choices due to either misinformation or irrational decisions are very frequent. Students whose parents made unsound choices were more likely to end up in inferior schools and to have lower academic performance. Parents with lower income or education level, or whose child had a low primary school performance, are more likely to make unsound school choices. In addition, unsound choices are related to the level of knowledge parents have of the school system, the nature of their preferences in making school choices, and the complexity of the school choice mechanism.

The importance of these results for educational policy is evident. Even though the data set is for the Eastern City District of Beijing, the school choice program there resembles similar programs in other parts of the world. If in the context of China, where children are the only child of a family and people have a long tradition of valuing schooling, parents are not able to make sound school choices, this problem is likely to exist at a larger scale in school choice programs in other countries. The Chinese experiment and information thus deserves close attention. Results from that case indicate that offering parents school choice options is not enough to achieve favorable student outcomes. Helping parents fully understand the school allocation mechanism, systematically disclosing important school information, giving consideration to various patterns of parental behavior that deviate from rational choice theory, and directing people to effectively exploit the mechanism to maximize their child's benefits should be

incorporated into school choice programs. These are particularly important to people of disadvantaged socioeconomic status.

II. Background and Data

II.A. Middle School Choice and Admission in China

In order to equalize access to school resources across students of different socioeconomic status and ability, the government forfeited the previous merit-based middle school admission mechanism in 1998. It introduced instead a preference-based random assignment of primary school graduates to middle schools.

The Beijing Eastern City district was divided into fifteen school neighborhoods based on primary school enrollment in 1999. Students in each neighborhood could apply to four to seven middle schools, and some middle schools were available to more than one school neighborhood. Good schools were usually available to more than one school neighborhood, while low quality schools were only available to the school neighborhood of proximity. All schools were given neighborhood specific enrollment quota by the Education Bureau.

A student could apply to all the middle schools available in his school neighborhood, ranking them in order of preference. These choices were incorporated into a randomizing procedure as follows. First, a computer-generated 10-digit number was randomly assigned to each student. Schools first admitted the students who had reported them as their first choice (i.e., first-choice applicants), enrolling students with the lowest numbers first until they filled their enrollment quota in the specific school neighborhood. Thus, schools that had more first-choice applicants than space in fact randomly chose from the pool of first-choice applicants. If they could not fill up their quota with these applicants, they went to the pool of applicants selecting the school as their second choice and so on, until they met the quota in that school neighborhood. Likewise, students who missed their preferred choice in a particular round proceeded to their next best choice in their reported ranked application. If a student missed all the schools he applied to, he was randomly assigned to any middle school available to his neighborhood that had seats left.

Despite the general randomization, schools allowed some students to be admitted without randomization if their parents were employed in the school, if the students had received at least a city-level prize in academic or special skill achievements, or if a considerable direct payment to the school was made. Randomization was thus incomplete, with a fraction of the students escaping the random drawing process. Yet, 4,948 out of the 7,000 students were admitted by middle schools through randomization.

This school assignment mechanism does not encourage truthful revelation of preferences about schools by design. For example, in a given neighborhood with four schools available, if, for simplicity, all people preferred school No. 1 most, school No. 2 second, then school No. 3, and least preferred school No. 4, reporting truthful preferences might not maximize the outcome of all applicants. As all schools first randomly chose from students who applied to them as their first choice, those who only slightly preferred school No. 1 to school No. 2 might have higher expected utility when reporting school No. 2 as their first choice, as this way they had a much higher chance of admission than competing with the whole neighborhood for entrance to school No. 1. People who were more risk averse were also more likely to deviate from their true preferences and chose school No. 2 as their first choice. As a result, school No. 2 might also be filled up in the first round. If people were fully rational and expected this outcome, even if they chose school No. 1 as their first choice, they would not choose school No. 2 as their second choice following their true preference, as this would result in an invalid choice of the second round, as school No. 2 did not have seats available then. An invalid choice in the second round might also increase their risk of ending up in school No. 4, the least preferred school. To summarize, under this mechanism design, parents' school choices reflected more strategic behaviors than authentic preferences for schools. Either lack of full information or wrong expectations of schools might result in unsound choices detrimental to the final school assignment of the child. Using the Boston Mechanism, Abdulkadiroglu et al. (2006) observed similar outcomes where sincere students lost their priority to sophisticated students and were disproportionately unassigned under a non strategy-proof mechanism. In particular, if parents with different socioeconomic status tend to have different access to the relevant information and different levels of rationality,

the school choice outcomes might be unbalanced across students with different family backgrounds, reinforcing problems of educational inequality.

In 1999, the Education Bureau only distributed to the parents of primary school graduates a brochure introducing the general procedure of the preference-based randomization of the middle school admission, as well as qualifications for direct enrollment without randomization. Middle schools available to the neighborhoods and the quota of each school for each neighborhood were also reported in that brochure. However, any further instruction was left to each primary school, and different primary schools might have different levels of involvement in assisting parents' decisions. There was no organized assistance to parental school choice at the district level.

II. B. Data

A census was conducted in early 2002 by the Education Bureau of Beijing's Eastern City District in its 28 public middle schools. The census covered all 7,102 students enrolled in the third and last year of middle school, their parents, and their teachers. Dropout and repetition is almost inexistent in middle schools of this district, and hence our survey population is the population of students who entered middle school in 1999. Students were asked to give their opinions about their study environment, and to answer questions about their attitudes toward school and society. A questionnaire directed to the parents collected information on households' wealth, parents' education levels, and, most importantly, their reports of important factors affecting their choice of school for their children, and how well-prepared they were about schools and procedures when making the school choice decisions. We also had administrative data on parents' actual school choice sequences and the students' primary school test scores and primary school affiliation. Combining the parental self-reported data and the administrative records of their actual choices provides unique information about parental behaviors in school choice.

For the purpose of this paper, we only include the 4,948 students who enrolled to the middle schools via the standard school-selection and randomization process, thus with

school application records available. Among these 4,948 students, 231 students chose as their first choice a school that could accommodate all first choice students, entered the school of their choice, and thus did not actually go through the randomization process. The majority of this paper focuses on the remaining 4,717 students who were randomly chosen at least by their first choice schools. 556 of them went through another step of randomization by some schools that could accommodate all first-choice students but were oversubscribed in a later round. The final outcomes show that no students went through more than two steps of randomization. For simplicity, within a certain school neighborhood, we call the schools that were oversubscribed in the first round Type A schools, schools that were undersubscribed in the first round but oversubscribed in a latter round as Type B schools, and schools that were undersubscribed in all rounds as Type C schools. A school neighborhood had 1 to 3 Type A schools, 1 to 2 Type B schools, and 1 to 2 Type C schools.

III. Evidence from the Parents' Actual School Choice

We first examine the actual school applications filled in by parents in 1999. They result from the parents' true preference over the schools and their strategy in ordering them to maximize the chance of their child being assigned in a school of their liking. Observationally, we cannot distinguish true preference from strategic choice, and it is not the purpose of this analysis to question the validity of the parents' preference. But some school choice patterns reveal errors, regardless of true preference and meaningful strategy. Such errors are detected in that they lead to invalid entries in the application sequence. These invalid entries can only be detrimental to the final outcome of school assignment as the students could miss the round in which their choices were invalid. Such errors may be due to inadequate or inaccurate information about the schools, misunderstanding of the school choice mechanism, or lack of rationality in decision. In this section, we identify some evident and important examples of these patterns in the actual parents' school choices that we call "unsound choices". The number and percentage of parents making unsound choice of each type described below are reported in Table I.

III.A. Simple Technical Mistakes

The first two patterns are simple technical mistakes in the application process, namely, repetition of a school (repeated choice) and applying to schools not available to the school neighborhood (wrong entry). The latter can only be explained as a misinformed choice. The repeated choice of a given school reflects misunderstanding of the randomization process of the middle school admission. If a child missed a school of his choice in a given round, the school can only be filled and hence will never be available in a later round. The second application to the same school results in an invalid choice. Both of these simple technical mistakes thus waste an application slot and forfeit the corresponding round of choice in vain.

Table I shows that only 59 and 303 out of the total 4717 applicants reported repeated school or selected schools that were not available to the neighborhood, respectively. Together, this is less than 10% of the applicants, indicating that parents understood the basics of the new middle school admission system.

III.B. Unsound Choices

Other patterns demonstrate some unsound decisions due to mistaken expectations or irrationality in school choice. As mentioned above, the admission process revealed three types of school: Type A schools that could not accommodate all applicants who chose them as their first choice and thus randomly selected their students among applicants in the first round; Type B schools that could accommodate all applicants who chose them as first choice but randomly selected among applicants in later rounds; Type C schools that could accommodate whoever applied to them no matter how late in the application sequence. In examining the school choice applications, we identify three patterns of unsound choice that reveal either a misjudgment of a school popularity or a lack of understanding of the school allocation process.

III.B.A. Choosing a Type A School as the 2nd or Later Choice

Had the parents correctly anticipated that a school of their choice would be of Type A, they should never have selected it as the second or later choice in the application sequence, regardless of their preferences. Since the school is over-subscribed in the first

round, selecting a Type A school for the second or later round results in nothing more than wasting an application slot and might exacerbate the final school placement of the student. For that reason, primary schools were told to advise parents not to report more than one good school in their application sequences. Yet, 4310 out of 4717, more than 91%, of the applicants reported a Type A school as the second or later choice in their application sequence (Table I). The prevalence of this pattern suggests that predicting the popularity of these schools was difficult. By the same token, these errors may have less negative consequences on the final outcome. In the extreme case of all the applications having an A school in 2nd rank, then the second choice is globally invalid, and the next round of randomization only occurs on the third choice. With a very large number of invalid second choices, many school remain open for the third choice, and hence many children may end up in the same school as they would have had they not wasted their second choice.

III.B.B. Choosing a Top School as the 2nd or Later Choice

A more obvious error is the placement of the subset of A schools that are commonly acknowledged as top schools as second or later choice in an application sequence. The well-established reputation of these schools leaves less room for misinformation or wrong expectation on their popularity. Chances of getting in were expected to be small, even in the first round. Moreover, parents had been strongly warned against reporting these schools as second or later choice by the primary school teachers. Even so, 2041 out of 4717, or 57% of the applicants reported a top school as their second or later choice, resulting in a missed selection opportunity. These parents were either making a wild gamble over the behavior of their peers, or more likely did not fully understand the system.

III.B.C. Hedging Too Early

Had the applicants fully anticipated the final general equilibrium of the randomization process, they would report their ranked school choices in the order of $A \rightarrow B \rightarrow C$, and this regardless of the category of their preferred school. In particular,

parents that correctly anticipate that a school of their choice will fall under type C (that admit all applicants), should expect that the selection process will end there, and should not add more choices, notably of A and B schools behind this choice. Hence an A-C-B sequence is considered unsound choice, while A-C-0 is not, as it could simply reflect a true preference for school C over the available B schools. Similarly, parents should not select a Type A school after a type B school.

The reasons for early choice of a C school are multiple. Parents may genuinely prefer a school that turns out to be of type C for their children, and thus place it early in their sequence. Parents may also try hedging the risk of entering a least desirable school. However, as C schools never get fully subscribed, no matter how late the school comes in the sequence, the child will always be accepted if the process reaches this round. Hence hedging with a C school only requires the school to be put as the last choice, after any school that is preferred to it. Choosing a Type C school too early in the sequence might put the game to an end too early, thus missing opportunities to enter some better schools that might have positive probability of accepting non-first-choice applicants. In either case, preference or early hedging, if the parents understood the process well and made perfect expectations of the popularity of the schools, they should have left all subsequent applications blank as the choice of that type C school had ended the game instantly in that round.

We thus identify an application sequence as a “too early hedging” pattern using two conditions. First, the applicant chose the “hedging school” as early as the second choice. This is because different school neighborhoods have different numbers of middle schools available, ranging from 3 schools to 7 schools, and reporting a C school as early as the second choice would be too early a hedging even for school neighborhoods where only 3 schools were available for application. Second, the applicants chose other schools after the hedging choice, which indicated that they might not actually desire or expect to end the game as early as the second round. In fact, more than 80% of these applicants chose schools that were more popular than the school used for hedging. With these criteria, 817 (or 17%) out of all 4717 students were identified as early hedging.

The prevalence of these unsound choices, particularly the choice of a top school in 2nd or higher place and the early hedging, suggest that the understanding of the class assignment process was not fully understood by many parents. There is in effect an inherent complexity in the current design of school choice. The system request that parents make a strategic choice of schools rather than reveal their true preferences about the schools. And it may not be surprising that unsound choices would be made when the decision mechanism is complicated and the final outcomes difficult to predict. A mechanism encouraging truthful revelation of preferences might potentially decrease cases of misinformed school choices by simplifying the decision process. But unsound school choices will persist in any system not fully understood by parents. A successful school choice program should seriously consider these possibilities and find effective solutions to alleviate these problems.

IV. The Consequences of Unsound School Choices

Unsound school choices are invalid entries in the school choice sequence. For students, this may result in placement in inferior schools, which in turn may hurt future middle school performance. To analyze this, we first relate the indicators of making unsound school choices to the characteristics of the schools students were assigned to, and then focus on how students' performances at the HSEE in 2002 were affected by unsound school choices via deteriorated middle school placements.

The model relating school characteristics to misinformed school choices is

$$(1) \quad x_{isn} = \alpha + \delta_j \cdot I_{\{J_i=j\}} + \eta_n + \varepsilon_{isn},$$

where x_{isn} is a characteristic of school s which admitted student i from school neighborhood n . δ_j indicates the correlation between the type j unsound school choice and the characteristics of the school the student was finally assigned to. $I_{\{J_i=j\}}$ is an indicator equal to 1 if student i 's parents made type j unsound school choice, and equal to 0 otherwise. Type j unsound school choice can be a single mistake, e.g. hedging too early, or a combination of several mistakes, e.g., both simple technical mistakes and hedging too early. η_n is a school neighborhood fixed effect, which is always included in the

regression as we only compare students within the same school neighborhood who faced the same school choice set. ε_{isn} are the error terms.

Table II reports the δ_j corresponding to each individual unsound school pattern when we include them into model (1) one at a time. The four individual patterns included are simple technical mistakes (including repeated choice of the same school and wrong entry in school choice), hedging too early, reporting a top-tier school as one's second or later school choice, and reporting a Type A school as one's second or later school choice. Each entry in the table reports the coefficient of the unsound choice from regressing the school characteristics indicated by the corresponding column title on the indicator of the unsound choice pattern indicated by the corresponding row title, controlling for school neighborhood fixed effects. Results show that hedging too early and reporting a top-tier school as one's second or later choice in most cases are related to lower teacher-student ratio, lower percentage of teacher with quality rank III or higher, lower school average performance on the HSEE in 1999, the year in which the students in the data set entered middle school, lower average school performance on the HSEE in 2002, which was performed by these students, and lower value-added school effects as estimated by Lai, Sadoulet, and de Janvry (2007)². Moreover, students making these two mistakes were also more likely to end up in schools with peers from lower socioeconomic and academic backgrounds, as shown by the negative conditional correlation coefficients between the relevant choice patterns and the school average parents' years of education and student elementary school graduation test scores, respectively. Students reporting a top-tier school are significantly more likely to end up in schools of lower reputation category. Similar patterns can also be observed for simple technical mistakes. Interestingly, reporting a Type A school as one's 2nd or later choice sometimes is related to better school performance and quality measures. As there is less than 10% of the sample parents that did not report at least one Type A school as their 2nd or later choice, this indicator might actually pick up unobserved characteristics such as strong ambition that made parents make this choice despite its irrationality. In particular, as it is difficult to precisely

² Lai, Sadoulet, and de Janvry (2007) used the random assignment to estimate the school fixed effects on student performance, controlling for student performance upon graduating from primary schools, and other individual and family characteristics. The estimates of the school fixed effects in that paper are used as "the value-added school effects" here for the relevant analysis.

predict all the Type A schools ex ante, reporting a Type A school as 2nd or later choice is a prevalent pattern that more than 91% of the parents did it. Thus, the coefficient estimates of the relevant indicators might become arbitrary when most students in the sample fall into one category.

We also include all four unsound patterns into model (1) simultaneously. The results are very similar to those in Table II and are not reported.

We then examine specifically whether unsound school choices in 1999 are related to poor HSEE performance in 2002. Instead of having a direct impact on the student's academic performance, parental school choices (including unsound choices) affect the "selection channel" (the three-step sequential random assignment of students to at most three schools summarized from their application sequences) through which students were assigned to a school (Lai, Sadoulet, and de Janvry, 2007)³. The large number of sequences of school choices made by all students were summarized in 137 selection channels. Within the same selection channel, students were faced with the same probability of being randomly selected by the same final set of at most three schools. Thus, an unsound school choice should directly affect the selection channel a student ended up in. The student school assignment was then affected by the selection channel, chance, and some other individual or family characteristics. Finally, school quality, together with the student's individual and family characteristics that are important factors in academic performance, would directly affect the student's school performance. We do not have enough information or valid instrumental variables to estimate this whole system. However, we can examine whether unsound school choices hurt school performance via weakened status in school assignment by comparing the estimates from two reduced-form models:

$$(2) \quad y_{isn} = \alpha + \delta_j \cdot I_{\{J_i=j\}} + \eta_n + \varepsilon_{isn}$$

$$(3) \quad y_{iscn} = \alpha + \delta_j \cdot I_{\{J_i=j\}} + \nu_c + \eta_n + \varepsilon_{iscn}$$

³ Although students could make 3 to 7 ordered school choices, Lai, Sadoulet, and de Janvry (2007) showed that, eventually, at most only three schools of each choice sequence are in fact valid. Thus, all students went through a three-step "selection channel", where they were first randomly chosen by their first-choice school if they reported a Type A school as their first choice, then a school that randomized in the 2nd or higher round (some students did not have a 2nd step), and finally, if they missed the first two steps, they ended up in a Type C school in their neighborhood which could accommodate all applicants. Based on all choice sequences, there were eventually 137 such "selection channels". Students within the same selection channel were faced with the same probability of entering the same final set of at most three schools.

In these equations, y_{iscn} is the overall HSEE score of student i in school s who was from neighborhood n and ended up in selection channel c , and ν_c is a selection channel fixed effect. In model (2), δ_j might capture both the indirect effect of unsound school choices on the HSEE scores and the effects of observed or unobserved individual or family characteristics that affect both the probability of making misinformed school choices and the student academic performance. In model (3), δ_j is the marginal effect of the aforementioned observed and unobserved individual or family characteristics on the student HSEE performance that are orthogonal to the indirect effects of unsound choices via lesser quality selection channel placement. By including the selection channel fixed effect, the genuine indirect effect of the unsound school choice on the HSEE scores has been controlled for. As a result, if δ_j is significant in model (2) but insignificant in model (3), it indicates that the observed correlations between unsound school choices and HSEE scores are not mainly driven by the individual characteristics that are correlated with both the HSEE scores and the probability of making the relevant unsound school choices. Therefore, unsound school choices would significantly affect the student's HSEE performance, which was an essential determinant to her high school admission and placement, via weakened status in randomized school assignment (i.e., by ending up in a disadvantageous selection channel).

Table III reports the estimates of models (2) and (3). Each column indicates a regression using the overall HSEE scores across all five subjects (Chinese, Math, English, Physics, and Chemistry) as dependent variable, and the patterns indicated by the row titles as independent variables, controlling for variables in the control panel that are included in that column. As the independent variables entered the regressions simultaneously, the coefficients indicate marginal correlations between the unsound school choices and the HSEE scores after controlling for the parent's tendency of making other unsound school choices and other control variables included in the regressions.

Columns (1) and (2) correspond to models (2) and (3), respectively, using “simple technical mistakes”, “too early hedging”, and “reporting a top school as 2nd or later

choice” simultaneously as the main independent variables⁴. Column (3) follows model (2) and additionally controls for a rich set of individual characteristics including the student’s gender, elementary school affiliation and performance, parent’s ages, income, education, and profession-based socioeconomic status. The regression in column (3) serves as another check of whether the significant coefficients corresponding to the unsound school choice indicators are mainly driven by individual characteristics that are relevant to both academic performance and school choice patterns. Table III shows that only “reporting a top school as 2nd or later choice” has a large negative effect on the HSEE performance.

Even model (3) (column (2)) does not report the highest R-squares only when controlling for selection channels, the originally significant coefficient of the relevant unsound school choice pattern becomes insignificant. Controlling for a rich set of individual characteristics in column (3) does not decrease either the significance or the magnitude of the coefficient corresponding to the unsound choice. We can thus conclude that, as expected, unsound school choices affect student performance via weakened status in school placement, i.e., via an inferior selection channel.

V. Which Parents Are More Likely to Make Unsound School Choices?

We have found that unsound school choices were quite common among parents, and that unsound choices can lead to undesirable consequences in terms of school placement and educational outcomes. An important question naturally arises: do all parents have a similar tendency of making unsound school choices? Does this tendency vary across family and academic backgrounds? Answers to these questions have important implications for equity in access to education. To explore these answers, we use several logit models to predict the risks of making an unsound school choice of each aforementioned type. These models relate the indicator of each pattern of unsound school application to the applicants’ academic and family backgrounds, including the student’s

⁴ We did not include or “reporting a Type A school as 2nd or later choice” here as 91% of the students belonged to this pattern, which makes it difficult to justify it as a “mistake”, and meaningless to compare 91% to the other 9% students. By contrast, “reporting a Top school as 2nd or later choice” is more representative.

standardized primary school test scores, gender, and parents' income and education. Primary school and school neighborhood fixed effects are also included in the regressions.

Results in Table IV show that the tendency to make simple technical mistakes in school application—such as reporting a school unavailable to the school neighborhood or repeated choices—does not significantly depend on a student's individual characteristics or family backgrounds. This indicates that all parents knew about the basics of the middle school choice and admission mechanism regardless of their socioeconomic status, and that simple technical mistakes were random errors.

As to the unsound school choice patterns due to misinformation or irrational choice such as “too early hedging” and “reporting a Type A school or top school” as the second or later choice in the application sequence, we find some significant relationship between the risks of reporting these patterns and students' individual and family backgrounds. Parents with higher income or a higher-achievement child in primary school are less likely to engage in “too early hedging”. Parents with a higher level of education were less likely to report a top-tier or a Type A school as their second or later choices. These results correspond to intuition: parents with higher socioeconomic status in terms of income and education are less likely to make unsound mistakes that can lead to undesirable school placements and lower educational performance.

Parents of a child with higher primary school performance are less likely to report a top school as their second or higher order choice. However, such parents are more likely to report a Type A school as their second or higher order choice. This is apparently counterintuitive as the opportunity cost of making an unsound choice should be higher for students with higher primary school performance and thus parents should be more careful. However, it might reflect the imperfect rationality of the decision, as parents considered their high-achieving child deserving better schools than the system could entitle them to, and were reluctant in putting only one Type A school in their application. Their earnest aspirations for their child made them exaggerate the slim chances that some Type A schools might still have a slot left open after the first round. This is similar to behavior analyzed in psychology and economics where people tend to overestimate their

chance of winning a huge lottery as the stake goes up. Moreover, as aforementioned, accurately identifying all Type A schools is very difficult given the complex decision process provided by the new school admission system, and more than 90% people reported a Type A school as their second or later choice. Therefore, people who reported Type A schools as their 2nd or later choice were not more irrational than other parents, perhaps only more caring about a promising child.

We also estimated probit and linear probability models, and added various additional controls such as parents' age, party affiliations, and years of residency in the city. Results are similar to those in Table IV. We thus conclude that, while making an unsound choice is quite prevalent among applicants, parents with low socioeconomic status or with a child with weaker academic background were more likely to make misinformed choices.

VI. Why Did Parents Make Unsound Choices?

VI. A. General Rationale

There are several reasons why parents might make an unsound choice. First, parents might have limited resources or connections to acquire the information about schools needed to make sound choices. In Beijing, important data such as the performance of middle schools in examinations is not publicly available. Data about school resources are not readily available either. Therefore, most parents have to base their school choices on vague information or long-established public opinions about schools. In particular, parents of lower socioeconomic status might have more limited resources to obtain information about the schools they consider applying to. With only one child, parents do not have a chance to learn by prior use of a school choice program.

Second, when data about schools are not readily available, parents may not all be willing to exert efforts to obtain such information. Parents' willingness to make efforts to obtain the relevant information might be different. Particularly, given the tight constraints

or high informational cost, such efforts might not be worthwhile for some parents of low socioeconomic status.

Third, even with enough information, parents might not make rational choices. Many studies in psychological economics have shown that people's behavior does not always follow rational expectations. For example, people tend to overestimate the slim chance of winning a multimillion lottery.

Finally, the school choice and admission mechanism does not encourage the revelation of parents' actual preferences as parents have to make strategic moves to optimize outcomes. This further complicates the decision-making process and makes the final outcome less predictable, because parents not only have to know about the quality of the match between each school and their child, but also need to make correct predictions about other people's choices. Moreover, the number of choices that parents were allowed to make was equal to the number of schools available in their school neighborhood, and this might have misled parents to fill in all the blanks in the application sequence. All school neighborhoods had access to three to seven middle schools. Therefore, making up to seven ordered school choices might be too difficult a task for parents under their given information set. In fact, *ex post*, as shown by the selection channels, students were sequentially randomly selected by not more than three schools, which indicated that around half or more of the choices made by parents were in fact useless. This reflects the complexity and difficulty in school choice posed by the mechanism design, which also contributed to the high frequency of unsound choices.

VI. B. Evidence Obtained by Combining Parents' Responses to Census Questions and Their Actual School Choices

In this section, we explore the various reasons mentioned above using both parent's actual school choices and their responses to census questions about knowledge and considerations in making school choices. Controlling for neighborhood fixed effects, we correlate unsound choices of parents and evaluations of various factors that might be important for their school choice, particularly their knowledge of school features and the

efforts they made to access the relevant information for their 2nd and 3rd school choices, and their particular educational concerns in making school choices.

We only look at parents' 2nd and 3rd school choices because the first choice is in most cases not relevant to the unsound choices. Ideally, we would look at parents' 2nd till last choices. However, parents' answers after the 3rd choice demonstrate significant carelessness, possibly as they wore out answering the same questions again and again for each of their school choices. Thus, we only use their answers regarding 2nd and 3rd school choices.⁵

Relevant results are reported in Table V. The upper panel reports parents' level of information about the schools of their 2nd or 3rd choices, and the lower panel reports the consideration they gave to factors corresponding to the upper panel. The questions asked in the census about parental considerations include: (1) matching student ability with school quality, (2) hedging the risk of entering the least desired school, (3) school admission quotas and chance of entry, (4) distance from home, (5) school spirit and discipline, (6) the quality of possible peers, (7) school facilities, (8) teacher quality, (9) the neighborhood surrounding the school, (10) school reputation and previous performance, and 11) other parents' school choice strategies. Parents were asked to indicate on a 1 (least important) to 5 (most important) scale the importance they attached to each factor. For the corresponding information and preparations, for each of their school choices, parents were asked whether they had the relevant information or did the relevant preparations when making school choices. We only use the average of their responses to the 2nd and 3rd school choices as the indicator of the knowledge or preparation for the aforementioned reasons.

First, from the upper panel of Table V, we find that parents had very limited information about the schools of their second choice, as only 30% of the parents in the sample claimed to know about the general conditions of the school, 32% knew the exact location of the school, and even less knew of each specific characteristics of the school of

⁵ We also included parents' first choice, and their 4th to 7th choices, respectively, into the model, and did not find significant inconsistencies with the results using information regarding the 2nd and 3rd choices only.

their 2nd choice. Moreover, they also demonstrated limited information and preparations about the game of school choice, as only 28% claimed they listened to teachers and other parents' opinions, which they should do, and only 30% had some sense about the chance of entry. The relevant knowledge and preparations for school of their third choice were even poorer.

The other entries in Table V report the coefficient of the choice pattern indicated by the row title when regressing the information/preparation or consideration indicated by the column title on that pattern, controlling for school neighborhood fixed effects. These coefficients indicate how the considerations and information/preparations of people who made each type of unsound choice differed from those of other people. For parents in each pattern of unsound choice, results tell us what considerations they had and whether they possessed the information they needed to play the game.

V.B.1. Simple Technical Mistakes

Parents who made these mistakes did not have significantly different preferences from other people, except that they weakly cared more about school discipline and spirit. Similarly, there are no significant differences in behavior except that they knew less about the school facilities (not significant at the 0.05 level). Therefore, simple technical mistakes appear as random errors that could be made by anyone.

VI. B.2. Too Early Hedging

According to statements about considerations in making choices in the lower panel of Table V, parents who hedged too early did not appear to care much about hedging the risk of entering a bad school or using the middle school admission system strategically as the coefficients regarding these two factors are not significant. Instead, they differentially cared more about distance to school and less about school performance and reputation. Actually, 52% of them reported to choose as their 2nd choice the nearest school that was not their first choice (i.e., the hedging choice). However, according to the upper panel, only 30% of the people in the whole sample knew about the exact location of their 2nd choice schools, but parents belonging to the “too early hedging” category did not know significantly better than other people in the sample. Thus, it is hard to claim

that they got what they really desired (distance advantage) if their choices were based on inadequate information about distance.

Moreover, distance considerations cannot justify why they chose other schools (80% of them chose other schools more popular than the hedging school) after the hedging schools, which demonstrates either inadequate understanding of the system and/or unwillingness to end the game as early as the 2nd round.

Choosing other schools after the hedging schools would not hurt them against their will if that school was actually their 2nd choice. It would hurt them if actually they preferred other schools but chose the hedging school as their 2nd choice because of misprediction of the popularity of the schools. In general, previous performance of a school is an important indicator of the popularity of that school, thus these parents' lack of attention to and knowledge about schools previous performance and reputation as shown in both panels might result in wrong expectations about popularity. Therefore, the "too early hedging" mainly demonstrates wrong expectations.

VI. B.3. Choosing a Top-tier School as Their 2nd or Later Choice

Unexpectedly, compared with other parents, these parents claimed to care more about the chance of entering the school of their choice and about other parents' strategy. They cared less about the actual school facilities and environments. Their revealed preferences correspond to those of a sophisticated player under the new school admission system. However, for various reasons that we cannot explicitly identify with the information we have, their behaviors showed the opposite. It seems that they had an inadequate understanding of the system and limited abilities to use the system to maximize their choices, as demonstrated in the lower panel by the negative coefficients of "teacher's and other parents' opinions" (to which they should have paid enough attention). In addition, even though they claimed to care more about the chance of school entry and parents' strategy, they had neither superior knowledge of the chance of being admitted to the schools of their 2nd or 3rd choices, nor superior knowledge of the various aspects of school characteristics that parents usually consider during school choice. In fact, they played against the rules and made more irrational choices than others despite their revealed preferences.

VI. B.4. Choosing a Type A school as Their 2nd or Later Choice

According to revealed preferences in the lower panel, these parents did not appear to be concerned with capacity constraints of the schools. Instead, they claimed to care more about teacher quality and school performance compared to people who did not make this kind of choice. More than 91% of the parents made this kind of choice, and parents' predominant preferences for higher school performance and teacher quality might have contributed to this pattern.

According to the lower panel, they behaviors demonstrated a significant level of sophistication. Compared to the rest of parents, less than 9% of the sample, they claimed to know better about various school characteristics that were important to school choice decision, and to know better about the chance of school admission, and considered other people's choices and teacher's advice more seriously. However, making perfect prediction might be too difficult when so many people simultaneously act strategically, the information was limited, and the overall sophistication level was low as shown by the percentage of the sample possessing the relevant information about schools and chance of entry.

For these parents, one particular player might not lose much relative to the others, as 91% people made the same pattern of choice. However, overall efficiency could improve if the number of such invalid entries decreases. This suggests that the system could improve either by simplifying the design or by enriching the information set to improve the efficiency of school choice, and save significant amount of energy and efforts in making school choices that only result in invalid entries.

As mentioned above, the parents were asked to make three to seven ordered choices. In fact, three to seven ordered choices might be too many and significantly increase the difficulty of the decision-making process. Policy makers' intentions were to allow people to have a slot for every school they could choose in the school neighborhood. However, this was not beneficial to the efficiency of school choice decision, and parents were more likely to make mistakes when trying to report more ordered choices. All students were randomly selected by at most three schools. Thus, ex-post, it would be wiser to not fill in all the slots in the application sequences. The last

column of Table V shows the average number of vacant entries in the application sequences made by all parents, and the difference in the number of vacant entries in the sequences made between people who made each type of unsound choice, and the rest of the sample, controlling for the neighborhood fixed effects. The average number of vacant entries was only 0.86. Moreover, people who made unsound choices reported significantly less vacant entries in the sequence, especially for people who reported a Type A school as their 2nd or later choice. Therefore, trying to follow a misleading rule of the new school admission system, parents spent undue efforts filling in all the slots in the sequences, which only led to unsound choices and invalid entries in the application. To summarize, parents who reported a Type A school as their 2nd or later choice were hurt by the design of the school choice program.

VI. C. Reported Parental Possession of Information about the Schools They Chose and Efforts at Obtaining the Relevant Information

As parents' lack of essential knowledge about the schools they chose and lack of efforts in seeking the relevant information might be an important contributor to unsound school choices, it is interesting to look at how parents' information and preparations differ across different individual characteristics. In particular, parents with disadvantaged academic or socioeconomic backgrounds might have greater difficulties in accessing information about schools than parents with stronger backgrounds. The benefits of obtaining such information might also be different for different people. As a result, all parents might not be equally well-informed or well-prepared for making sound school choices.

To examine whether parents' knowledge about schools differs across students of different academic and socioeconomic backgrounds, we use an ordered-logit regression. The dependent variable is the number of schools among the schools in parents' first three choices for which they knew about or sought to obtain a certain piece of school information. The dependent variable thus varies from 0 to 3, and the relevant information includes all items of the information indicated by the column titles of Table VI. More

specifically, the dependent variables include whether parents knew about or inquired from reliable sources about the general conditions of the schools, whether they listened to advice from teachers or other parents, and whether they knew about the school admission quota and the chance of being admitted. These three items are the most basic knowledge about middle schools that most primary schools had claimed to have given to parents when helping them with their school choice decisions. Some pieces of information that were not readily available from the primary school teachers and staffs are also used as dependent variables in the ordered logit regressions. Parents usually had to exert extra discretion and efforts to acquire these types of information, including the middle schools' performance in the HSEE, school facilities, the neighborhood around the school (via an on-site visit), the location of the school and distance from their home, and the expected quality of student-school match. Independent variables include student and parental characteristics such as student gender, primary school test scores, and parents' income and education. Primary school fixed effects and school neighborhood fixed effects are included to control for systematic differences in information access across primary schools and school neighborhoods, as well as the difference in school choice options available across school neighborhoods.

Results in Table VI show that parents of girls and of students with better performance are significantly more informed or exerted more efforts in obtaining the relevant information about the middle schools they chose as their first three choices. Moreover, parents with higher income and education are in general better informed of the middle school they chose as their first three choices, knowing more or at least trying harder to know more about the school's academic performance, the exact school address and the distance from home to school, and the neighborhood around the school. They were also more likely to have paid a visit to the school before making a decision. Even for the basic preparations that primary schools expect all parents to have done before making a decision, parents were not uniformly well prepared: parents with higher income and more years of education were significantly more likely to inquire about the general conditions about the school they chose as their first choice, to listen to the advice of

primary school teachers and other parents, and to know about the school admission quota and the chance of entrance.

There might be several explanations for the universally positive significance (usually at least at 5% level) of primary school test scores and female dummy. First, both the benefit of entering a desirable school by extensive consultation and discreet choice and the opportunity cost of entering an undesirable school by careless choice are higher for students with better performance. Thus parents of students with better performance are more likely to do more research before making a final decision. Second, parents of better students might have more interactions with teachers, and thus teachers might, even unconsciously, provide them with better assistance in making decisions. Third, better primary school performance is positively correlated with stronger family backgrounds, including parents' education and income. Therefore, the primary school performance was highly correlated to stronger family backgrounds, which might be related to stronger and wider social connections enabling more access to the relevant information about the schools. As to the positive significance of the female dummy, this might reflect the fact that girls are in general considered to be more vulnerable to the external environment than boys, and thus parents are more careful in school choice for girls to better shield them from possible threats to their safety, behavior, and performance if they end up in a less desirable school. Moreover, as shown in the data, during primary school girls outperformed boys in both academic achievements and responsibilities assumed in student society. Thus, the female indicator might also pick up some unobserved performance effects, as students with superior performance might have better interactions with their teachers and benefit more from the teachers' assistance.

To disentangle the consequence of information constraints and parents' lack of preparation when making school choices from the consequence of parents' lack of attention to the importance of school choice and the relevant factors, we also ran the same regressions, adding additional controls including parental opinion about the importance of school to the students' development (ranging from 1 to 5, with 1 indicating "not important at all", and 5 indicating "very important"), parents' self-reported level of attention paid in their choice of schools (1 indicates "most attention" and 5 indicates "no

attention at all”), and, most importantly, parents’ self-declared importance attached to the relevant factor in making school choices. Indicators of parental self-declared importance of relevant factors include both the parents’ responses to the scale of importance (1 for “least important” and 5 for “highly important”) of the relevant factors and whether parents had reported the relevant factors as among the first three important considerations during their school choice. We obtain results very similar to those in Table VI. Thus, regardless of attention to school choice, parents with disadvantaged academic and socioeconomic backgrounds were more likely to make misinformed school choices, possibly due to limited access to information useful in making school choices.

VI. Conclusions

This paper uses parents’ responses to census questions regarding school choice in combination with the actual parental school choice from administrative records of 4717 parents who entered middle schools via a preference-based random assignment in Beijing’s Eastern City District in 1999 to examine parents’ capabilities in making sound school choice regardless of their preferences and strategies. Results show that many parents did not make rationally sound school choice, even though some of them claimed in the census responses that they did consider strategic choices and various important factors about schools in making their school choice. We show that unsound choices had a negative effect on the students’ final school assignment, which in turn indirectly affected their school performance. More importantly, parents of lower income or education level, or of a child with lower primary school performance, were more likely to possess less information about the schools they chose and to make an unsound school choice, which would lead to social and academic stratification across schools.

Insufficient information about schools, inadequate understanding of the randomization mechanism, decisions against the rule of rational choice, and inappropriate mechanism design of the school choice program all, to some extent, contribute to unsound school choices, with different reasons related to different types of unsound choices. All these call for more organized assistance during the school choice process and more publicly available information about schools. This is especially important to

disadvantaged parents and students, who are more likely to make unsound school choices, as shown in this study.

As most existing school choice programs in the world are similar to the one in this paper, not providing a mechanism that encourages the truthful revelation of preferences about schools, a concern arises that even if all the arguments about benefits of introducing school choice are justified, parents might not be able to make school choices to their maximum benefit without sufficient information and assistance when they are faced with a complex decision process. One way of alleviating this might be to implement a mechanism forcing truthful revelation of preferences as suggested by Abdulkadiroglu and Sonmez (2003), which might make the decision process more straightforward. However, results from this paper suggest that, even under such a mechanism, benefits may not be fully realized as parental school choices can deviate from full rationality, for instance reporting a top-tier school as 2nd or later choice. Therefore, good mechanism design, disclosure of information, and organized assistance to parental choices, especially targeted at the more disadvantaged parents, should be indispensable components of school choice programs.

References

- Abdulkadiroglu, Atila, and Tayfun Sonmez. 2003. "School Choice: A Mechanism Design Approach." *The American Economic Review* 93 (3): 729-747.
- Abdulkadiroglu, Atila, Parag Pathak, and Alvin Roth. 2006. "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." Manuscript.
- Abdulkadiroglu, Atila, Parag Pathak, Alvin Roth, and Tayfun Sonmez. 2006. "Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access." Manuscript.
- Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. 2002. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *The American Economic Review* 92(5): 1535-1558.
- Cullen Berry J., Brian A. Jacob, and Steven D. Levitt. 2005. "The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools." *Journal of Public Economics* 89: 729-760.
- Cullen Berry J., Brian A. Jacob, and Steven D. Levitt. 2006. "The Effect of School Choice on Participants: Evidence from Randomized Lotteries." *Econometrica* 74(5): 1191-1230.

- Elacqua, Gregory, Mark Schneider, and Jack Buckley. 2006. "School Choice in Chile: Is It Class or the Classroom?" *Journal of Policy Analysis and Management* 25(3): 577-601.
- Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2005a. "Parental Preference and School Competition: Evidence from a Public School Choice Program," Yale University Department of Economics Working Paper.
- Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2005b. "Evaluating a School Choice Lottery: Preferences, Choices, and Heterogeneous Treatment Effects," Manuscript, Yale University.
- Howell, William. G., Patrick J. Wolf, David E. Campbell, and Paul E. Peterson. 2002. "School Vouchers and Academic Performance: Results from Three Randomized Field Trials." *Journal of Policy Analysis and Management* 21(2): 191-217.
- Hoxby, 2000a. "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review* 90, 1209-1238.
- Hsieh, Chang-Tai and Miguel Urquiola, 2006. "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program." *Journal of Public Economics*, 90(2006), 1477-1503.
- Krueger, Alan B. and Pei Zhu "Another Look at the New York City School Voucher Experiment." *National Bureau of Economic Research Working Paper* No. 9418, December 2002.
- Ladd, H., 2002. "School Vouchers: a Critical View." *Journal of Economic Perspectives* 16: 3-24.
- Lai, Fang, Elisabeth Sadoulet, and Alain de Janvry. 2007. "The Impact of Teacher Quality and School Quality on Educational Performance: Evidence from the Middle School Educational Reform in Beijing's Eastern City District." Manuscript.
- Mayer, Daniel P., Paul E. Peterson, David E. Myers, Christina Clark Tuttle, and William G. Howell. 2002. "School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program," Washington, DC: Mathematica Policy Research, Inc.
- Pathak, Parag, and Tayfun Sonmez. 2006. "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." Manuscript.
- Rouse, Cecilia. E. 1998. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *The Quarterly Journal of Economics* 113(2): 553-602.
- Teske, Paul and Mark Schneider. 2001. "What Research Can Tell Policymakers about School Choice?" *Journal of Policy Analysis and Management* 20(4): 609-631.

Table I. Number and Percentage of Parents Making Various Unsound Choices

| | <u>Simple technical mistakes</u> | | Choosing a type | Choosing a top | |
|--------------|----------------------------------|-------|-----------------|-------------------|-------------------|
| | Repeated | Wrong | A school as the | school as the 2nd | Too early hedging |
| | choice | entry | 2nd or later | or later choice | |
| Frequency | 59 | 303 | 4310 | 2676 | 817 |
| Percentage | 1 | 6 | 91 | 57 | 17 |
| Observations | 4717 | 4717 | 4717 | 4717 | 4717 |

The first two rows report the number (frequency) and percentage of parents making the type of unsound choice that is indicated by the corresponding column title.

Table II. Univariate analysis of the correlations between unsound school choices and the characteristics of the assigned schools

| | Teacher- student ratio | %rank III & IV teachers | Ave. years of teaching | School category | Value- added school effects | Ave. HSEE score 1999 | Ave. HSEE scores 2002 | Ave. peers' elementary school score | Ave. peers' parents' years of education |
|---|---------------------------|----------------------------|---------------------------|--------------------|--------------------------------------|----------------------------|--------------------------|--|--|
| Simple technical mistakes | 0.002** [0.02] | -1.34 [0.23] | -0.66** [0.02] | -0.07 [0.23] | 1.17 [0.23] | -3.43** [0.01] | 0 [1.00] | -0.01 [0.35] | -0.12** [0.04] |
| Reporting a type A school as the 2nd or later choice | 0.000 [0.84] | 2.16*** [0.01] | 0.59*** [0.00] | -0.07 [0.10] | 0.55 [0.41] | 1.37 [0.17] | 1.71** [0.03] | 0.02** [0.02] | 0.03 [0.42] |
| Reporting a top school as the 2nd or later choice | -0.001*** [0.00] | -1.97*** [0.00] | -0.69*** [0.00] | -0.16*** [0.00] | -3.92*** [0.00] | -3.38*** [0.00] | -5.69*** [0.00] | -0.08*** [0.00] | -0.22*** [0.00] |
| Too early hedging | -0.001** [0.08] | -8.44*** [0.00] | -2.07*** [0.00] | -0.04 [0.27] | -3.55*** [0.00] | -1.81** [0.03] | -2.78*** [0.00] | -0.07*** [0.00] | -0.08** [0.02] |

* significant at 10%; ** significant at 5%; *** significant at 1%

Each entry corresponds to a different regression: it reports the coefficient and the p-value (in bracket) of the unsound school choice indicated by the corresponding row title by regressing the school characteristics indicated by the corresponding column title on that individual school choice pattern, controlling for school neighborhood fixed effects

School category indicates the four-level reputation category, with 1 indicating the lowest reputation, and 4 indicating the highest reputation.

Value-added school effects are the school fixed effects estimates reported by Lai, Sadoulet, and de Janvry (2007), measuring the middle school effects on student performance controlling for primary school test scores, affiliations, and various individual characteristics.

Table III. The effect of unsound school choice on the HSEE average score

| Dependent variable: HSEE test score | | | |
|---|---------------------|----------------|--------------------|
| | (1) | (2) | (3) |
| Simple technical mistakes | 4.92 [0.40] | 4.26 [0.51] | 1.07 [0.86] |
| Too early hedging | 1.13 [0.73] | 0.35 [0.94] | 5.38 [0.11] |
| Reporting a top school as the 2nd or later choice | -11.96*** [0.00] | 2.22 [0.51] | -9.42*** [0.00] |
| Control | | | |
| Neighborhood fixed effect | Y | Y | Y |
| Selection channel fixed effect | | Y | |
| Individual & parental characteristics | | | Y |
| Observations | 2908 | 2892 | 2506 |
| R-squared | 0.03 | 0.12 | 0.17 |
| F-statistics of the unsound choices | 10.45 | 0.31 | 6.44 |
| p value of the joint F test | 0.00 | 0.82 | 0.00 |

p values in brackets

* significant at 10%; ** significant at 5%; *** significant at 1% choices.

Table IV. Heterogeneity in tendency of making unsound choices (Logit Regression)

| | Simple technical mistakes (1) | Reporting a type A school as the 2nd or later choice (2) | Reporting a top school as the 2nd or later choice (3) | Too early hedging (4) |
|-----------------------------------|----------------------------------|---|--|--------------------------|
| Log(parents' income) | 0.01 [0.83] | -0.02 [0.57] | 0 [0.82] | -0.05** [0.05] |
| Parents' years of education | -0.01 [0.87] | -0.05** [0.04] | -0.06*** [0.00] | -0.01 [0.76] |
| Female | 0.19 [0.26] | -0.15 [0.20] | 0.08 [0.22] | 0.08 [0.35] |
| Primary school test score | -0.03 [0.73] | 0.29*** [0.00] | -0.12*** [0.00] | -0.39*** [0.00] |
| Control | | | | |
| Primary school fixed effects | Y | Y | Y | Y |
| School neighborhood fixed effects | Y | Y | Y | Y |
| Observations | 3558 | 3764 | 4517 | 4375 |

p values in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Each column reports the result of a Logit regression. The dependent variable is the indicator of making the mistake indicated corresponding column title, and the independent variables of the regression in each column are indicated by the row titles.

Table V. Parental knowledge about the school of their 2nd and 3rd choices and their considerations in school choice

| Knowledge about schools of their 2nd and 3rd choice | General conditions | | | Teachers' or other parents' suggestions | School admission quota and chance of entry | School facilities | On-site visit of the neighborhood around the school | School location and distance from home | School reputation and performance record | The quality of school-student match | # of vacant entries in the application |
|---|---------------------------|-----------------|-------------------|---|--|-------------------|---|--|--|-------------------------------------|--|
| | Simple technical mistakes | -0.03 [0.39] | | | 0.01 [0.70] | 0.01 [0.62] | -0.04* [0.06] | 0.02 [0.30] | 0.01 [0.83] | -0.01 [0.62] | 0.03 [0.18] |
| Choosing a type A School as the 2nd or later choice | 0.05*** [0.01] | | | 0.05*** [0.01] | 0.03* [0.08] | 0.06*** [0.00] | 0.03* [0.09] | 0.03 [0.13] | 0.04** [0.04] | 0.03 [0.11] | -3.38*** [0.00] |
| Choosing a top-tier as the 2nd or later choice | -0.01 [0.41] | | | -0.02** [0.02] | -0.01 [0.17] | 0.01 [0.49] | 0 [0.80] | -0.03*** [0.01] | 0 [0.91] | -0.01 [0.29] | -1.41*** [0.00] |
| Hedging too early | -0.01 [0.58] | | | -0.01 [0.71] | 0.01 [0.70] | 0.01 [0.24] | 0.01 [0.34] | 0 [0.91] | -0.03** [0.04] | -0.02 [0.23] | -0.47*** [0.00] |
| % of the Sample knowing the info about the 2nd choice | 30 | | | 28 | 30 | 18 | 19 | 32 | 25 | 24 | Sample average |
| Observations | 4567 | | | 4567 | 4567 | 4567 | 4567 | 4567 | 4567 | 4567 | 0.86 |
| Evaluations of the importance of the factors during school choice | Spirit and discipline | Peer quality | Teacher quality | Other parents' strategies | Hedging the risk | Chance of entry | School facilities | Neighborhood around the school | Physical distance from home | School reputation and performance | Matching quality |
| Simple technical mistakes | 0.07* [0.09] | -0.01 [0.66] | 0.03 [0.32] | -0.01 [0.49] | 0.01 [0.84] | -0.04 [0.23] | 0.02 [0.29] | -0.01 [0.80] | 0.00 [0.90] | -0.04 [0.28] | -0.01 [0.79] |
| Choosing a type A school as the 2nd or later choice | -0.01 [0.60] | -0.01 [0.49] | 0.07*** [0.00] | 0.00 [0.66] | -0.01 [0.73] | -0.01 [0.75] | 0.00 [0.83] | -0.03 [0.14] | -0.02 [0.38] | 0.05* [0.06] | -0.01 [0.81] |
| Choosing a top-tier as the 2nd or later choice | 0.01 [0.68] | 0.00 [0.69] | -0.01 [0.58] | 0.01*** [0.00] | 0.00 [0.75] | 0.02** [0.05] | -0.02** [0.02] | -0.02* [0.07] | 0.01 [0.51] | -0.01 [0.44] | 0.01 [0.70] |
| Hedging too early | -0.02 [0.41] | -0.01 [0.54] | -0.03 [0.12] | 0 [0.28] | 0.01 [0.44] | 0.01 [0.48] | -0.01 [0.25] | 0.01 [0.36] | 0.05** [0.02] | -0.04* [0.05] | -0.01 [0.61] |
| Observations | 4405 | 4405 | 4405 | 4405 | 4405 | 4405 | 4405 | 4405 | 4405 | 4405 | 4405 |

p values in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Each entry reports the coefficient of the pattern indicated by the row title by regressing the information/preparation or considerations indicated by the column title on that pattern, controlling for the school neighborhood fixed effects.

Table VI. Heterogeneity in Parents' Knowledge about Schools of Their First Three Choices across Different Students and Parents

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|-----------------------------------|--------------------|---|--|-------------------|---|--|--|-------------------------------------|
| | General conditions | Teachers' or other parents' suggestions | School admission quota and chance of entry | School facilities | On-site visit of the neighborhood around the school | School location and distance from home | School reputation and performance record | The quality of school-student match |
| Log(parents' income) | 0.05*** [0.00] | 0.04* [0.05] | 0.04** [0.01] | -0.01 [0.54] | 0.03 [0.14] | 0.03* [0.06] | 0.04** [0.04] | 0 [0.36] |
| Parents' years of education | 0.04*** [0.00] | 0.03*** [0.01] | 0.06*** [0.00] | 0.02* [0.07] | 0.03** [0.03] | 0.04*** [0.00] | 0.06*** [0.00] | 0 [0.47] |
| Female | 0.25*** [0.00] | 0.16*** [0.00] | 0.21*** [0.00] | 0.15*** [0.01] | 0.12** [0.04] | 0.30*** [0.00] | 0.27*** [0.00] | 0.29*** [0.00] |
| Primary school test score | 0.14*** [0.00] | 0.08*** [0.01] | 0.15*** [0.00] | 0.01 [0.75] | 0.03 [0.30] | 0.15*** [0.00] | 0.16*** [0.00] | 0.13*** [0.00] |
| Control | | | | | | | | |
| Primary fixed effects | Y | Y | Y | Y | Y | Y | Y | Y |
| School neighborhood fixed effect: | Y | Y | Y | Y | Y | Y | Y | Y |
| Observations | 4517 | 4517 | 4517 | 4517 | 4517 | 4517 | 4517 | 4517 |

p values in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Each column reports an ordered-logit Regression: the dependent variable is the number of schools among the first three choices of the parents to which parents had reported to either possess or actively seek the information indicated by the column title. The independent variables are student and parent characteristics indicated by the row titles.