Information Acquisition and Provision in School Choice: An Experimental Study*

Yan Chen Yinghua He

July 3, 2017

Abstract

When participating in the school choice process, students often spend substantial time and effort acquiring information about different schools. In this study, we compare how two popular school choice mechanisms, the (Boston) Immediate Acceptance and the (Gale-Shapley) Deferred Acceptance, incentivize students' information acquisition. Our results show that only the Immediate Acceptance mechanism incentivizes students to learn their own cardinal and others' preferences. While our lab experiment yields results directionally consistent with our theoretical predictions, we also find that students systematically over-invest in information acquisition, especially when they believe that others invest more and when they are more curious. Our counterfactual policy analyses suggest that it is welfare-enhancing for educational authorities to provide more information to help each student learn both her own and others' preferences, even under strategy-proof mechanisms. Doing so improves match efficiency while reducing the socially wasteful costs of information over-acquisition.

Keywords: information acquisition, information provision, school choice, experiment

JEL Classification Numbers: C78, C92, D82

*We thank Simon Anderson, Nick Arnosti, Georgy Artemov, Larry Ausuble, Estelle Cantillon, Ting-Ting Ding, John Duffy, Federico Echenique, Maxim Engers, Guillaume Fréchette, Guillaume Haeringer, Astrid Hopfensitz, Tanjim Hossain, Fuhito Kojima, Rachel Kranton, Erin Krupka, Dorothea Kübler, Steve Leider, George Loewenstein David Miller, Noah Myung, Yesim Orhun, Emel and Erkut Ozbay, Scott Page, Szilvia Pápai, Tanya Rosenblat, Andy Schotter, Vernon Smith, Barry Sopher, Olivier Tercieux, Séverine Toussaert, Peter Troyan, Dan Vincent, Georg Weizsäcker, Roberto Weber, Huan Xie, and seminar participants at the University of Cologne, Concordia, Leicester, Maryland, Michigan, NYU, Rutgers, Toulouse, Tsinghua, UC-Irvine, Virginia, the 2015 ESA World (Sydney, Australia) and North American (Dallas, TX) meetings, the 2016 NSF/NBER Decentralization Conference (Providence, RI), and the 2017 European Behavioral Economics Meeting (Budapest, Hungary) for helpful discussions and comments, and Ming Jiang and Yingzhi Liang for excellent research assistance. Financial support from IFREE and the National Science Foundation through grant no. SES-0962492 to Chen and from the French National Research Agency (l'Agence Nationale de la Recherche) through Projects DesignEdu (ANR-14-CE30-0004) and FDA (ANR-14-FRAL-0005) to He is gratefully acknowledged. This research was approved by the University of Michigan IRB (HUM 00109086). Chen: School of Information, University of Michigan, 105 South State Street, Ann Arbor, MI 48109-2112. Email: yanchen@umich.edu. He: Baker Hall 246, Department of Economics MS-22, Rice University P.O. Box 1892, Houston, TX 77251. Email: yinghua.he@gmail.com.

1 Introduction

> "It was very hard, and very time-consuming," one New Orleans resident said of trying to find a school for her daughter, who entered kindergarten last fall. "I'm educated, I have a bachelor's degree, ... and I do have time to read articles online and research things." - Arianna Prothero. 2015. "Parents Confront Obstacles as School Choice Expands," *Education Week*.

School choice is now part of the education landscape in the US. However, when choosing a school, students often have imperfect information about their own preferences regarding candidate schools, partly because it is difficult to assess the potential educational outcomes each school provides (Dustan, de Janvry and Sadoulet 2015). Unfortunately, acquiring this information can be costly if a student faces too many choices, or must acquire information about a number of factors, such as academic performance, teacher quality, school facilities, extra-curricular activities offered, and peer quality. In New York City, for example, the 600-page *Directory of NYC Public High Schools* covers nearly 700 programs at more than 400 schools citywide. Given this extensive information, students incur substantial cost in processing the information to rank up to 12 high school programs they would like to attend (Nathanson, Corcoran and Baker-Smith 2013). These information frictions can lead to substantial welfare loss (Narita 2016).

The cost of information acquisition is particularly harmful to low-income students. Indeed, research has shown that, due to limited information, low-income high achievers in the U.S. tend not to apply to selective colleges, in spite of the fact that generous financial aid makes these colleges more financially accessible than the colleges these students end up choosing (Hoxby and Avery 2013, Hoxby and Turner 2015). Informational intervention can therefore substantially raise the number of applications from these students to selective colleges (Hoxby and Turner 2013). Furthermore, this application tendency is not limited to college selection. Indeed, similar undermatching phenomena are observed among low-income families in public school choice plans (Hastings and Weinstein 2008).

While research shows the existence of information disparity, the literature on matching and school choice instead typically assumes that all students have perfect knowledge about their own preferences, at least their ordinal ones. Our study differs in that we do not assume perfect knowledge. We thus contribute to the literature by examining how different school choice mechanisms incentivize student information acquisition and how information provision by educational authorities can promote student welfare. Specifically, we focus on two widely-used mechanisms, the (Boston) Immediate-Acceptance (hereafter IA) and the (Gale-Shapley) Deferred-Acceptance (hereafter DA). By taking into account both the benefits and costs of information acquisition, this study thus provides a more comprehensive evaluation of the mechanisms and more tailored guidance for the design of school choice and other matching markets.

Our results show that, in a school choice setting with unknown preferences and costly information acquisition, both the strategy-proof DA and the non-strategy-proof IA incentivize students to acquire information about their own ordinal preferences. However, we further find that only the non-strategy-proof mechanism induces students to learn their own cardinal preferences. Knowing these preferences can improve the efficiency of IA compared to DA (Abdulkadiroğlu, Che and Yasuda 2011). On the other hand, the lack of strategy-proofness also implies that information about others' preferences, or more generally others' strategies and actions, maybe used to compete with other students. As such, the game may become closer to zero-sum; the acquisition of information about others' preferences can be individually rational but socially wasteful, a disadvantage of a non-strategy-proof mechanism.

To better search for behavioral regularities as well as to quantify the welfare effects of information acquisition and provision, we conduct a laboratory experiment. Specifically, we use the Becker-DeGroot-Marschak mechanism to elicit willingness-to-pay (WTP) for information (Becker, DeGroot and Marschak 1964). Our experimental results show that students' WTP for their own and others' preferences under the non-strategy-proof IA is significantly greater than it is under the strategy-proof DA. These results are consistent with our theoretical predictions. However, we also find that their WTP is systematically higher than the theoretical prediction. Examining this result further, we decompose the WTP and find that both conformity and curiosity explain students' over-investment in information acquisition in our experiment. That is, a student tends to have a high WTP when she expects a high WTP by others (conformity)[1] and when she has a high WTP for non-instrumental information (curiosity).

In addition to documenting student information acquisition incentives, we provide insight for

[1]In the equilibrium of the game, a student is less willing to pay for information when others pay more. In other words, investments in information acquisition are strategic substitutes. We will elaborate this point further in Section 5.

policymakers by studying the welfare effects of three policies of information provision, relative to the benchmark case where everyone only knows the distribution of preferences. The first is a laissez-faire policy in which the education authority does not interfere with students' information acquisition. This policy corresponds to the real-life scenario in which educational authorities make information available but neither easily accessible nor easily understandable. Our experimental results show that this policy is always welfare-decreasing under DA and can be welfare-improving under IA only if the cost of information acquisition is very low.

The second policy freely provides information about a student's own cardinal preferences but is designed such that students acquire information about others' preferences at a cost. This policy reflects the recent trend of making information about schools more accessible and more understandable (Hastings and Weinstein 2008). Consistent with our theoretical prediction, our evidence shows that, compared to the laissez-faire policy, this policy achieves higher welfare under either mechanism, with a greater effect under IA.

The third policy offers free information about both a student's own and others' preferences, which increases the predictability of others' actions compared to the second policy. Due to the excessive WTP we observe in the laboratory, we find that providing information about others' preferences improves welfare by saving students from paying wasteful costs. This result may be expected under the non-strategy-proof IA, but not under the strategy-proof DA. Interestingly though, we show that free provision of others' preferences results in an even higher welfare gain under DA than under IA.

Together, our findings imply that school districts using either DA or IA can improve student welfare by providing information about both a student's own and others' preferences. The information related to a student's own preferences might be provided through accessible presentation materials on school offerings and performance (Hastings and Weinstein 2008), by knowledgeable guidance counselors and teachers (Sattin-Bajaj 2014). The information about others' preferences, which is about increasing the predictability of others' application behaviors, might be provided through data on student application statistics in past years as well as publishing applicants' actions and allowing students to revise their own applications upon observing others' actions. This latter approach has been implemented in the school choice context in Amsterdam (De Haan, Gautier, Oosterbeek and Van der Klaauw 2015) and Wake County, North Carolina (Dur, Hammond and Morrill 2015), as well as in the college admissions context in Inner Mongolia (Gong and Liang 2017).

Lastly, on the methodological front, we create and implement a simple measure of curiosity and show that it accounts for a significant fraction of the willingness to pay for information.

2 Literature Review

This study examines information acquisition incentives across school choice mechanisms. It thus contributes to the matching literature in which it is typically assumed that agents perfectly know their preferences (Gale and Shapley 1962, Roth and Sotomayor 1990, Abdulkadiroğlu and Sönmez 2003). One exception to this assumption is Chade, Lewis and Smith (2014), who consider the case where colleges observe signals of students' ability but cannot acquire more information. Allowing this possibility, both Lee and Schwarz (2012) and Rastegari, Condon and Immorlica (2013) study settings where firm preferences over workers are not completely known and are revealed only through interviews.

To our knowledge, the only theoretical papers that study the effects of matching mechanisms on information acquisition are those of Bade (2015), Harless and Manjunath (2015), and Artemov (2016). In the context of house allocations, Bade finds that serial dictatorship is the unique *ex ante* Pareto optimal, strategy-proof, and non-bossy allocation mechanism. However, in their study, Harless and Manjunath (2015) prove that the top-trading-cycles mechanism dominates the serial dictatorship mechanism under progressive measures of social welfare. Both papers focus on ordinal mechanisms. As we show in our paper, in any strategy-proof ordinal mechanism, students have no incentive to learn their cardinal preferences, while information about cardinal preferences can be welfare-improving, especially when students have similar ordinal preferences (Abdulkadiroğlu et al. 2011). Lastly, in an ongoing study, Artemov (2016) theoretically investigates information acquisition under DA in an environment similar to our experimental setting and finds that DA provides inadequate incentives for information acquisition.

Another unique feature of our study is that we examine incentives to acquire information regarding others' preferences. Other studies on information acquisition examine only the acquisition of information about one's own preferences. One exception in this body of literature is Kim (2008), who considers a common-value first-price auction with two bidders, one of whom learns her opponent's signal.

In addition to the matching literature, information acquisition is examined in various other fields, e.g., bargaining (Dang 2008), committee decisions (Persico 2004, Gerardi and Yariv 2008), contract theory (Crémer, Khalil and Rochet 1998, Crémer and Khalil 1992), finance (Barlevy and Veronesi 2000, Hauswald and Marquez 2006, Van Nieuwerburgh and Veldkamp 2010), and law and economics (Lester, Persico and Visschers 2009). In particular, there is a large body of theoretical literature on the role of information acquisition in mechanism design, especially in auction design, e.g., Persico (2000), Compte and Jehiel (2007), Crémer, Spiegel and Zheng (2009), Shi (2012), surveyed in Bergemann and Valimaki (2006). Within this stream of research, Bergemann and Valimaki (2002) show that the Vickrey-Clark-Groves mechanism provides the efficient incentives for information acquisition *ex ante* and implements the efficient allocation conditional on the private information *ex post*.

While the literature on information acquisition is mostly theoretical, there are a few experimental investigations. For example, Gabaix, Laibson, Moloche and Weinberg (2006) conduct two experiments with costly information acquisition to compare the directed cognition and fully rational models. In several other studies, Choi, Guerra and Kim (2015) compare the second-price (sealed-bid) auction with the English auction when bidders have independent values and are heterogeneously informed, while Gretschko and Rajko (2015) compare the two auctions in relation to information acquisition and bidding behavior in an independent and private value environment. Finally, Bhattacharya, Duffy and Kim (2015) study endogenous information acquisition in the context of voting behavior.

Finally, our study contributes to the experimental literature on school choice. This research has focused on strategy, stability and welfare comparisons across mechanisms and has typically assumed that students know their own preferences (Chen and Sönmez 2006, Featherstone and Niederle 2008, Calsamiglia, Haeringer and Klijn 2010, Klijn, Pais and Vorsatz 2012). Two experimental studies (Pais and Pinter 2008, Pais, Pinter and Veszteg 2011) examine matching mechanisms in different information settings. However, since they treat information settings as exogenous, their results focus more on the robustness of the mechanisms to information availability. More recently, several experimental studies of school choice have examined peer information sharing within networks (Ding and Schotter 2016) as well as intergenerational (Ding and Schotter 2015) and top-down advice (Guillen and Hing 2014). Our paper contributes to the experimental literature by providing the first experimental evidence on information acquisition in the school choice context.

3 Theoretical Analysis

In this section, we outline our theoretical model regarding the endogenous acquisition of information about one's own and others' preferences under both the Immediate and Deferred Acceptance mechanisms. We relegate the technical details and proofs to Appendix A.

3.1 The Model Setup

Our model begins with a finite set of students, I, to be assigned to a finite set of schools, S, through a centralized school choice mechanism. S is supplemented by a "null school" or outside option, s^0, and $\overline{S} \equiv S \cup s^0$. For each $s \in S$, there is a finite supply of seats, $q_s \in \mathbb{N}$, with enough seats in total to accommodate all students, $\sum_{s \in S} q_s = |I|$. Moreover, in our model, schools rank students using a common and even lottery with a single tie-breaking rule the realization of which is unknown to students when they enter the mechanism.

In our model, student i's valuations of schools are an independent draw from a distribution, F, denoted by a vector $V_i = [v_{i,s}]_{s \in S} \in \mathcal{V}$, where $\mathcal{V} = [\underline{v}, \overline{v}]^S$, $0 < \underline{v} < \overline{v}$, and $v_{i,s}$ is i's von Neumann-Morgenstern utility of school s. For notational convenience, we assume that $v_{i,s^0} = 0$ for all i, connoting that every school in S is acceptable to every student. Given our assumptions, our model is thus an independent-private-value model. In Section 3.5, we discuss how our results generalize to common- and interdependent-value models.

Note that our model treats student preferences as strict: for any pair of distinct schools, s and t in S, $v_{i,s} \neq v_{i,t}$ for all i. We therefore define strict ordinal preferences P on S such that sP_it if and only if $v_{i,s} > v_{i,t}$. We also augment the set of all possible strict ordinal preferences $\mathcal{P}$ with a "null preference" $P^{\phi} \equiv \emptyset$ denoting that a student has no information about her ordinal preference, expressed as $\bar{\mathcal{P}} = \mathcal{P} \cup \emptyset$. V^{ϕ} and $\overline{\mathcal{V}}$ are similarly defined. Furthermore, the distribution of cardinal preferences V conditional on P is denoted by $F\left(V|P\right)$, while the distribution of P implied by F is

$G(P|F)$. Finally, we impose a full-support assumption on $G(P|F)$, i.e., $G(P|F) > 0, \forall P \in \mathcal{P}$, indicating that every strict ordinal preference ranking is possible given the distribution of cardinal preferences. Necessarily, $G(P^{\phi}|F) = 0$.

While the value of the outside option and the distribution of preferences, $F(V)$ and thus $G(P|F)$, are always common knowledge, we introduce an information-acquisition stage for each i to learn her own preferences (P_i and/or V_i) or others' preferences (V_{-i}) before entering the mechanism. Because of the independent-private-value nature of our model, any acquisition of information about others' preferences aids a student only for the purpose of gaming or competing with other students in the process.

3.2 School Choice Mechanisms

As mentioned, we focus on two mechanisms popular in both the research literature and actual practice: the (Boston) Immediate Acceptance and the (Gale-Shapley) Deferred Acceptance mechanism.

The (Boston) **Immediate Acceptance** mechanism (IA) asks students to submit rank-ordered lists (ROL) of schools. Together with the pre-announced capacity of each school, IA uses predefined rules to determine the school priority rankings for students and executives the following procedure:

Round 1. Each school considers all students who rank it first and assigns its seats in order of their priority at that school until either there is no seat left at that school or no such student left.

Generally, in:

Round ($k > 1$). The kth choice of the students who have not yet been assigned is considered. Each school that still has available seats assigns the remaining seats to students who rank it as their kth choice in order of their priority at that school until either there is no seat left at that school or no such student left.

The process terminates after any Round k when either every student is assigned a seat at some school, or the only students who remain unassigned have listed no more than k choices.

The (Gale-Shapley) **Deferred Acceptance** mechanism (DA) can be either student-proposing or school-proposing. We focus on the student-proposing DA mechanism in this study. The DA mechanism uses information about the capacity of every school as well as students' ROLs of the schools in allocating seats. It utilizes strict rankings of schools over students using a pre-specified set of rules and proceeds as follows:

Round 1. Every student applies to her first choice. Each school rejects the least ranked students in excess of its capacity and temporarily holds the others.

Generally, in:

Round $(k > 1)$. Every student who is rejected in Round $(k-1)$ applies to the next choice on her list. Each school pools together new applicants and those on hold from Round $(k-1)$. It then rejects the least ranked students in excess of its capacity. Those who are not rejected are temporarily held.

The process terminates after any Round k when no rejections are issued. Each school is then matched with those students it is currently holding.

3.3 Information Acquisition in School Choice Games

Our study focuses on the role of information acquisition under the two mechanisms. We first investigate the incentive to acquire information regarding one's own preference. The timing of the game and the corresponding information structure are described as follows and also in Figure 1:

(i) Nature draws individual valuation V_i, and thus ordinal preferences P_i, from $F(V)$ for each i, but i knows only the value distribution $F(V)$. That is, $F(V)$ is common knowledge, while V_i and P_i are unknown to every student.

(ii) Each student i decides whether to acquire a signal about her ordinal preferences; if she decides to acquire a signal, she then decides how much she is willing to invest in information acquisition, denoted by $\alpha \in [0, \bar{\alpha}]$.

(iii) If ordinal preferences are learned, she then chooses an investment, $\beta \in [0, \bar{\beta}]$, to acquire a signal about her cardinal preferences.

(iv) Regardless of the information acquisition decision or outcome, every student plays the school choice game under either IA or DA.

Our model differentiates between learning about ordinal preferences and learning about cardinal preferences, as the former represents the case where a student acquires coarse information about the schools, whereas the latter represents the case where a student acquires more detailed in-

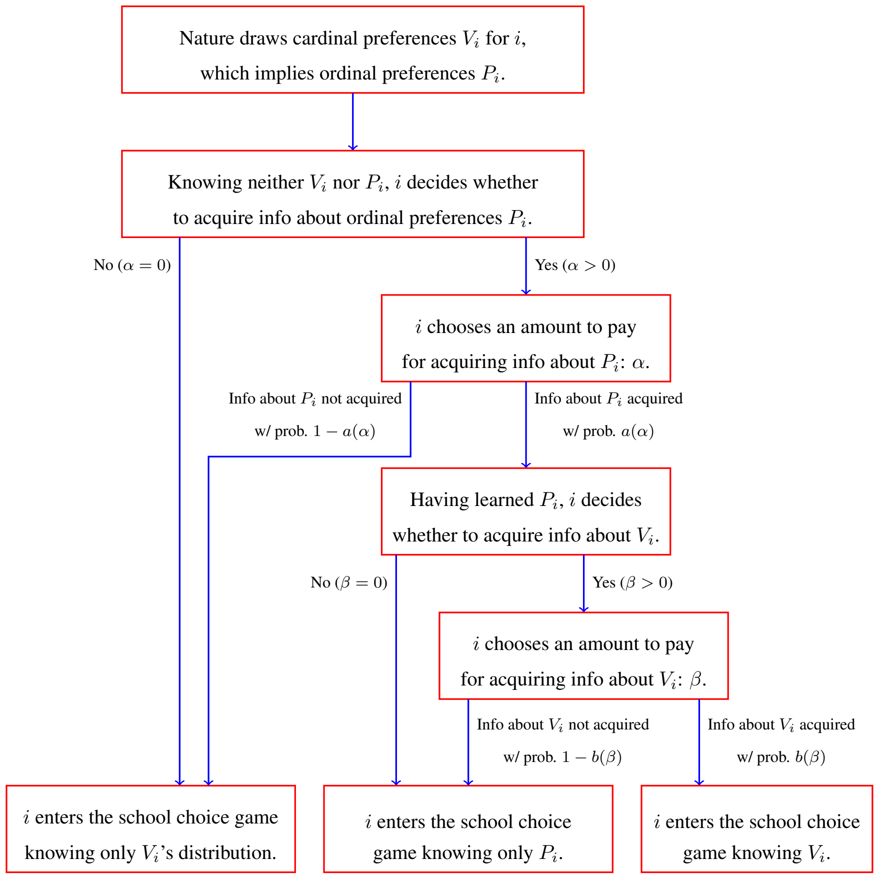

Figure 1: Acquiring Information about One's Own Preferences.

formation but at a greater cost.[2] As we shall see below, our model highlights that the mechanisms provide starkly different incentives for inquisition acquisition on these two dimensions.

Furthermore, in our model, i knows that others are engaging in information acquisition, but does not know what information they have acquired. Our signals, denoted by ω, are set up as "hard news" that reveals either the truth or nothing. The technology for information acquisition, especially how investments determine the probability of obtaining the information ($a(\alpha)$ and $b(\beta)$), is presented in Appendix A.

In our model, after observing the signal, students enter the school choice game under either DA or IA. Each student i submits an ROL denoted by $L_i \in \mathcal{P}$ such that sL_it if and only if s is ranked above t.[3] It is well-established that the student-proposing DA is strategy-proof (Dubins and

[2] The literature on both one-sided and two-sided matching usually assumes that agents know their own ordinal preferences (Roth and Sotomayor 1990, Bogomolnaia and Moulin 2001), but that cardinal preferences may be unknown due to "limited rationality" (Bogomolnaia and Moulin 2001).

[3] We restrict the set of actions to the set of possible ordinal preferences, $\mathcal{P}$. In other words, students are required to rank all schools in S. This analysis can be extended to allow ROLs of any length.

Freedman 1981, Roth 1982), while IA is not (Abdulkadiroğlu and Sönmez 2003). Given the above setup, we can now state our first proposition.

Proposition 1 (Information Acquisition Incentives: Own Preferences)**.** *In any symmetric Bayesian Nash equilibrium* $(\alpha^*, \beta^*(P, \alpha^*), \sigma^*(\omega))$ *under DA or IA, the following is true:*

(i) $\alpha^* > 0$*, i.e., students always have an incentive to learn their ordinal preferences;*

(ii) under DA, $\beta^*(P, \alpha^*) = 0 \; \forall P, \alpha^*$*, i.e., there is no incentive to learn cardinal preferences;*

(iii) under IA, there exists a preference distribution F *such that* $\beta^*(P, \alpha^*) > 0$ *for some* P*. That is, given some prior* F *on preferences, students under IA have an incentive to learn their cardinal preferences after observing their own ordinal preferences* (P)*.*

Remark 1. *Similar to the results for DA, students have no incentive to learn their own cardinal preferences when a strategy-proof mechanism elicits ordinal preferences.*[4]

For part (iii) in the above proposition, such distribution F is likely to exist. For example, as indicated in the proof, the distribution of ordinal preferences is common across students with a high probability; the second-most-popular school is likely worse than the best school. However, but with a small probability these two schools may be similar, if, for instance, some students live close to the second-best school.

We next consider a student's incentive to acquire information about others' preferences. To simplify notation and highlight the strategic acquisition of information beyond learning one's own preferences, we assume that everyone knows exactly her own cardinal preferences (V_i) but not others' preferences (V_{-i}). We further assume that the distribution of V_i, $F(V_i)$, is common knowledge with the same properties as before. The process and technology for information acquisition in this case are depicted in Figure 2.

To acquire information, student i may pay δ to acquire $\omega_{i,3} \in \bar{\mathcal{V}}^{(|I|-1)}$, a signal regarding V_{-i}. Similar to our earlier specification, with probability $d(\delta)$, she learns perfectly, $\omega_{3,i} = V_{-i}$; with probability $1 - d(\delta)$, $\omega_{3,i} = V_{-i}^{\phi}$, i.e., she learns nothing. The properties of $d(\delta)$ are specified in Appendix A. We can now state our second proposition.

[4]This statement does not extend to mechanisms that use information about cardinal preferences directly, e.g., Hylland and Zeckhauser (1979), Budish, Che, Kojima and Milgrom (2013), and He, Miralles, Pycia and Yan (2015).

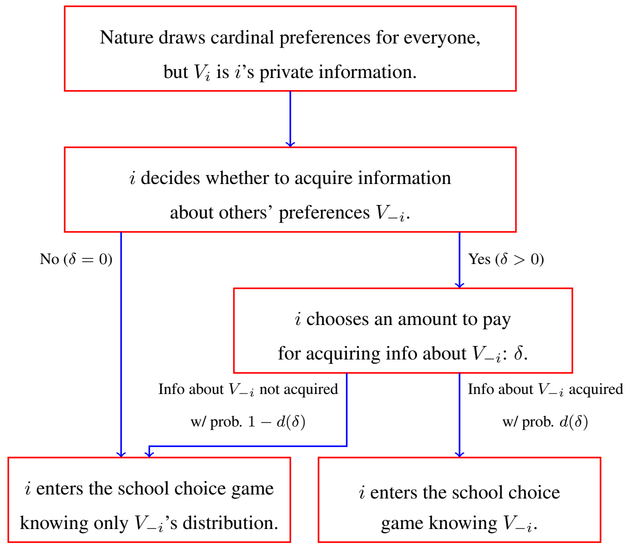

Figure 2: Acquiring Information about Others' Preferences.

Proposition 2 (Information Acquisition Incentives: Others' Preferences)**.** *Suppose* $(\delta^* (V), \sigma^* (\omega_3, V))$ *is an arbitrary symmetric Bayesian Nash equilibrium under a given mechanism. We then have:*

(i) $\delta^* (V) = 0$ *for all* V *under DA;*

(ii) There exists a preference distribution F *such that* $\delta^* (V) > 0$ *under IA for all* V *in some positive-measure set. That is, given some prior* F *regarding preferences, students of type* V *have an incentive to learn others' preferences.*

Remark 2. *Similar to the results for DA, students have no incentive to learn others' preferences when a strategy-proof mechanism elicits either ordinal or cardinal information from students.*

In short, this result provides another perspective on strategy-proofness as a desideratum in market design: a strategy-proof mechanism eliminates the incentive to acquire information about others' preferences.

We acknowledge that our analysis focuses on learning about others' preferences instead of their actions. One may argue that others' actions are the determinants of a student's outcomes. However, preferences and actions are equivalent in the equilibrium of our model; in other words, since the equilibrium strategy is "common knowledge," knowing others' preferences also allows a student to know others' actions.

3.4 Welfare Effects of Information Provision: The Case of Common Ordinal Preferences

Students always have an incentive to acquire information about their own preferences. In addition, they sometimes have an incentive to acquire information regarding others' preferences. However, despite the existence of an incentive, information is not always successfully acquired due to the cost of information acquisition. In this section, we examine the welfare implications that result when educational authorities provide information to participants.

To study whether providing information is welfare enhancing, we make the additional assumption that the provision of information decreases the cost of information acquisition to zero, while the lack of provision increases the cost to infinity. As the school choice game under IA is notoriously untractable without restrictions on preference domain, we focus on a special setting where everyone has the same ordinal (but different cardinal) preferences, similar to the setting examined in Abdulkadiroğlu et al. (2011) and Troyan (2012).[5] Note that we relax this assumption in our experiment (cf. Section 4).

In this setting, we start with a prior F and thus $G(P|F)$ such that after a P is drawn, it becomes everyone's ordinal preference. Again, every school is acceptable: $v_{i,s} > 0$ for all i and s; and F_{v_s} denotes the marginal distribution of the cardinal preference for school s.

We next represent the education authority's decision regarding how much information to release by sending a vector of signals to every i: $\bar{\omega}_i = (\bar{\omega}_{1,i}, \bar{\omega}_{2,i}, \bar{\omega}_{3,i}) \in \bar{\mathcal{P}} \times \bar{\mathcal{V}} \times \bar{\mathcal{V}}^{(|I|-1)}$, where $\bar{\omega}_{1,i}$ and $\bar{\omega}_{2,i}$ represent the signals of i's ordinal and cardinal preferences, respectively, and $\bar{\omega}_{3,i}$ is the signal of others' cardinal preferences. All signals are such that $\bar{\omega}_{1,i} \in \left\{P^{\phi}, P_i\right\}$, $\bar{\omega}_{2,i} \in \left\{V^{\phi}, V_i\right\}$, and $\bar{\omega}_{3,i} = \left\{V_{-i}^{\phi}, V_{-i}\right\}$, i.e., they are either perfectly informative or completely uninformative.

We can then study the *ex ante* welfare in equilibrium under each of the following information structures:

(i) Uninformed (**UI**): $\bar{\omega}_i = \left(P^{\phi}, V^{\phi}, V_{-i}^{\phi}\right), \forall i$;

(ii) Ordinally Informed (**OI**): $\bar{\omega}_i = \left(P_i, V^{\phi}, V_{-i}^{\phi}\right), \forall i$;

(iii) Cardinally Informed (**CI**): $\bar{\omega}_i = \left(P_i, V_i, V_{-i}^{\phi}\right), \forall i$;

(iv) Perfectly Informed (**PI**): $\bar{\omega}_i = (P_i, V_i, V_{-i}), \forall i$.

It should be noted that the identical ordinal preference is common knowledge under OI, CI,

[5] This setting is not a special case of the model in Section 3.3, because student preferences are correlated. However, it can be shown that the main results, Propositions 1 and 2, still hold true in this setting.

or PI. However, under UI, while everyone knows that the ordinal preference is the same across students, no one knows the realization of this preference order.

These four information structures reflect the four possible outcomes of different school choice policies. If an educational authority makes it difficult for students to acquire information about schools, the UI scenario holds. By contrast, if the educational authority makes some information easily accessible, then the OI scenario may emerge. If all information on a student's own preferences is readily available, then the CI scenario is more likely.

Our fourth possible structure, the PI scenario, relates to the gaming part of school choice under a non-strategy-proof mechanism. From Proposition 2, recall that individual students have an incentive to acquire information about others' preferences under IA. The literature has shown that this additional strategic behavior may exacerbate inequalities in access to public education. More precisely, if a student does not understand the game and therefore does not invest a sufficient effort to acquire information regarding others' preferences, she may be at a disadvantage when playing the school choice game. As a policy intervention, an educational authority can seek to mitigate this inequality by publishing students' actions and allowing students to revise their applications upon observing others' applications, as is the policy in Amsterdam (De Haan et al. 2015) and Wake County, North Carolina (Dur et al. 2015).

Note that, by the standard fixed point arguments, a symmetric Bayesian Nash equilibrium, possibly in mixed strategies, always exists under any of the four information structures . We now summarize the results regarding *ex ante* welfare under DA and IA in the following two propositions.

Proposition 3 (Ex Ante Welfare under DA)**.** *Under DA, the ex ante welfare of every student under any of the four information structures (UI, OI, CI, and PI) is* $\sum_{s \in S} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}\right)$ *in any symmetric equilibrium.*

This implies that there is no gain in *ex ante* student welfare when students receive more information under DA. Finally, we state our last proposition.

Proposition 4 (Ex Ante Welfare under IA)**.** *Under IA, we obtain the following ex ante student welfare comparisons in terms of Pareto dominance in any symmetric equilibrium:*

(i) Under the uninformed (UI) or ordinally informed (OI) structure, the student welfare is $\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s}dF_{v_s}(v_{i,s})$*;*

(ii) The welfare for cardinally informed (CI) students weakly dominates that for uninformed or ordinally informed students: CI $\geq$ *OI* $=$ *UI;*

(iii) The welfare for perfectly informed students (PI) weakly dominates that for uninformed or ordinally informed students: PI $\geq$ *OI* $=$ *UI;*

(iv) The welfare ranking in perfectly versus cardinally informed scenarios is ambiguous.

Proposition 4 suggests that it is always beneficial to provide more information about one's own cardinal preferences under IA, but that the effect of providing information about others' preferences is ambiguous. We provide the proof for part (iv) using two examples in Appendix A (sections A.9.4 and A.9.5). The intuition behind this proof is as follows: When perfectly informed, multiple high-type students at a school compete for the same school seats knowing the presence of other high-type students. Consequently, the strategy of always top-ranking that school can be sub-optimal; the school may end up being assigned to a low-type student, leading to a welfare loss. By contrast, when only cardinally informed, high-type students may always choose to top rank the school in a symmetric Bayesian Nash equilibrium.

3.5 Possible Extensions

Our results can be generalized to other school choice settings. For example, they can be generalized to the setting in which students have interdependent values over schools. In this setting, acquiring information about one's own values is achieved by learning more about the schools themselves as well as others' preferences. However, the information acquisition on others' values in the our model should be interpreted as information gathering for strategic purposes. In a setting of interdependent values, students decipher signals about others' preferences as both useful information about their own values and useful information about others' values. Our results allow us to identify the type of deciphering necessary under each mechanism.

Additionally, our model separately considers the acquisition of two types of information, but, in reality, students may acquire information about their own and others' preferences simultaneously.

Given the lack of strategy-proofness and the role of cardinal utility under IA, we expect our results to hold in the case of simultaneous information acquisition.

Furthermore, for our analysis of information provision, we assume that information acquisition cost can be set to infinity, which may not be feasible in reality. However, this limitation is not "binding" in terms of policy implications, because our experimental results show that educational authorities should increase information provision (Section 5) rather than increasing acquisition cost to infinite. We also abstract away from the cost of information provision. As a reference, a cost-benefit calculation of providing information can be found in Hoxby and Turner (2013); it shows that a cost of $6 per student can benefit low-income students significantly.

Finally, our model of information provision assumes students have the same ordinal preferences. This is not uncommon when comparing the welfare performance of the two mechanisms (Abdulkadiroğlu et al. 2011, Troyan 2012), with the justification that the common-ordinal-preference assumption is plausible in real life. Nonetheless, we realize that an extension to a more general preference domain would be fruitful and leave this endeavor for future work.

4 Experimental Design

To test our theoretical predictions when the common knowledge of rationality assumption is relaxed as well as to search for behavioral regularities in information acquisition in the school choice context, we conduct a laboratory experiment in the simplest possible environment. Our goal in this experiment is to compare student information acquisition behavior and its interaction with the incentives implied under IA or DA. We then use this observed behavior to estimate the welfare implications of various information-provision policies, using our theoretical predictions as a benchmark.

Compared to our theoretical framework, our experimental setting relaxes the common-ordinal-preference assumption and allow students to have different ordinal preferences. To simplify the game, we design the payoff distribution to reduce information acquisition to a single step. That is, upon learning one's own ordinal preferences, a student learns her cardinal preferences as well, since there is only one possible realization of cardinal preferences consistent with a given ordinal preference.

4.1 The Environment

We consider an environment with three students, $i \in \{1, 2, 3\}$, and three schools, $s \in \{a, b, c\}$. Each school has one available slot and ranks students using a lottery. Student cardinal preferences are i.i.d. random draws following the distribution in Table 1. The uncertainty in this setting comes from the value of school b, which can be either better or worse than that of school a, relaxing the common-ordinal-preference assumption. *Ex ante*, the expected payoff of being assigned to b is 0.3, which is less than 1/3 of the payoff from school a. In terms of welfare, there is one possible source of inefficiency, i.e., it is inefficient to assign a type-$(1, 0.1, 0)$ student to school b if there is at least one other student of type-$(1, 1.1, 0)$.

Table 1: Payoff Table for the Experiment

Students	$s = a$	$s = b$	$s = c$
$i \in$ {1,2,3}	1	0.1 with probability 4/5; 1.1 with probability 1/5	0

Notes: The above payoffs are in dollars. In the experiment, points are used to measure payoffs, with an exchange rate of 100 points = 1 USD.

Assuming that every student is an expected-utility maximizer, we solve all (symmetric) equilibria of the school choice game under either IA or DA for any given information structure. Derivations under the assumption that students are risk neutral or risk averse are relegated to our online appendix.[6] To measure the incentive to acquire information about own preferences (denoted as "OwnValue"), we endow every student with the common prior that everyone knows only the preference distribution. For each student, we then calculate the payoff difference between knowing or not knowing one's own preferences, taking into account that the other two students may or may not know their own preferences. This difference is our theoretical prediction related to student's willingness to pay (WTP, henceforth) for their OwnValue. Similarly, to measure student WTP for information about others' preferences (denoted as "OtherValue"), we treat preference realizations as private information. For any given student, we derive the payoff difference between knowing or not knowing others' preferences. These results are summarized in the last two columns of Table 3. While risk-averse students are often willing to pay less for information about either one's own or others' preferences, the same directional comparison between IA and DA maintains.

[6] Our online appendix can be found at http://yanchen.people.si.umich.edu/ under "Working Papers".

4.2 Treatments

In our experiment, we implement a 2 (mechanisms) $\times 2$ (information to be acquired) $\times$ 2 (information cost) factorial design to evaluate the performance of the two mechanisms $\{$IA, DA$\}$ under two information and cost conditions. The choice of the $2 \times 2 \times 2$ design is based on the following considerations.

(i) IA vs. DA (between-subject): While DA is dominant-strategy incentive compatible, IA is manipulable. Our theoretical analyses indicate that each mechanism provides different incentives for information acquisition.

(ii) Acquiring OwnValue vs. OtherValue (between-subject): Our theoretical analyses suggest that the incentive to acquire information depends on the type of information that may be acquired.

(iii) Free vs. costly information acquisition (within-subject): While a free information condition enables us to evaluate information provision policies, a costly information acquisition condition better reflects reality. As this variation is implemented at the within-subject level, we also take into account the order effect: For half of the sessions, subjects first experience 10 free information rounds and then experience 10 costly information rounds (denoted as "free-to-costly"); for the other half of the sessions, subjects experience costly-info rounds first and then free-info rounds (denoted as "costly-to-free").

In the free information treatments, participants are provided information about their own value (or others' values) at no cost. In comparison, in the costly information treatment, we use the Becker-Degroot-Marshak (BDM) mechanism (Becker et al. 1964) to elicit participant's WTP for their own or others' valuations of school b. Specifically, each subject is asked to enter her WTP for her own (or others') values in the interval, $[0, 15]$. The server then collects the WTP from each participant and generates a random number between $[0, 15]$ for each participant independently. If a subject's WTP is greater than the random number, she acquires the information and pays an amount equal to the random number; otherwise, she does not acquire the information and pays zero. The BDM procedure is incentive compatible under the assumption of monotonicity (Azrieli, Chambers and Healy 2017). To facilitate participant understanding of the BDM mechanism, we provide participants with numerical examples to illustrate and then test for their understanding by administering a quiz at the end of the instructions. Our instructions for the BDM mechanism are adapted from those in Benhabib, Bisin and Schotter (2010).

To elicit each participant's belief about the average WTP of the other two participants in her group, we use the binarized scoring rule (BSR) introduced in Hossain and Okui (2013). The BSR is incentive compatible under different risk attitudes and even when the decision maker is not an expected utility maximizer (Schotter and Trevino 2013). As such, it is more robust than the quadratic scoring rule. In our use of the BSR, each subject submits a guess about the average WTP of the other two participants. The server then computes the squared error between the guess and the actual average, i.e., $SE = (\text{guess} - \text{actual average})^2$. Next, the server randomly draws a number, R, uniformly from $[0, U]$. If $SE \leq R$, the subject receives a fixed prize of 5 points. Otherwise, she receives zero points. Based on our pilot sessions, we find that 90% of the squared errors fall at or below 49. Therefore, we use 49 as the upper bound in our BSR calculation, i.e., $U = 49$. The random number, R, is drawn independently for each subject, and for each round.

In the experimental instructions (Appendix B), we explain the DA (or IA) algorithm to participants in detail and include an example in the Review Questions to test participant understanding of the mechanisms. Following the convention in experimental economics, we do not inform the participants of their optimal strategies under either mechanism. Specifically, we do not tell participants that truthful ranking of schools is a dominant strategy under DA. This allows us to examine their naturally emerging strategies.

4.3 Experimental Procedures

Each experimental session consists of 20 rounds with costly (free) information for the first ten rounds, and free (costly) information for the next ten rounds. The order is counterbalanced for each treatment. Each session consists of 12 subjects.

At the beginning of each session, every subject is randomly assigned an ID number and is seated in front of a computer terminal. The experimenter then reads the instructions for the first ten rounds. After this, subjects have the opportunity to ask questions, and answers are provided to the full group. Subjects are then given 10 minutes to read the instructions at their own pace and to answer the review questions. After ten minutes, the experimenter distributes the answers and go over them with the group. Afterwards, subjects go through ten rounds of the experiment, randomly re-matched into groups of three at the beginning of each round. After the first ten rounds, the experimenter reads the instructions for the second ten rounds and answers any questions. Participants again complete a set of review questions, and then go through the second ten rounds of the experiment.

In the acquiring OwnValue treatments, each round consists of the following stages:

(i) Each participant is provided with the value distribution (Table 1) to induce the common prior.

(ii) Each participant is asked to rank the schools.

The server then collects the rankings, draws the school-b value for each subject, generates the tie-breaker, and allocates schools to participants. The allocation outcomes are shared with participants at the end of each round.

(iii) The server draws a new set of values for every participant. Everyone acquires her value for school b, either for free or by paying a cost:

  (a) For the *free* information treatment, everyone receives her own school-b value for free.

  (b) For the *costly* information treatment, we use the BDM mechanism to elicit each participant's willingness to pay. We tell the subjects that everyone will know the number of other subject(s) in her group who observe their value(s), regardless of whether she will observe her own value or not.

  The server collects the respective WTP assessments and generates a random number between $[0, 15]$ for each participant.

(iv) After step (iii), each subject receives the following feedback on her computer monitor:

  (a) *Free* information treatment: her school b value and the fact that every subject in her group is provided with his own value.

  (b) *Costly* information treatment: her WTP, her random number, and whether she observes her value,

    i. if she observes her value, she also receives the number of other subject(s) in her group who also observe their value(s);

    ii. if she does not observe her value, she receives only the number of other subject(s) in her group who observe their value(s).

(v) Each participant is then asked to rank the schools again.

The server again collects the rankings, generates a new tie-breaker for each participant, and allocates the schools.

(vi) At the end of each round, each participant receives the following feedback on her computer monitor:

   (a) For the game without OwnValue acquisition or provision: her ranking, her value, the tie-breaker, her allocation and her payoff.

   (b) For the game with free or costly OwnValue acquisition: her ranking, her value, the tie-breaker, her allocation and her payoff.

The OtherValue treatments proceed in a similar way, except that each subject always knows her own value for school b before ranking schools. In the OtherValue treatments, the information provided or acquired is the other two participants' values for school b.

At the end of 20 rounds, we implement the Holt and Laury lottery choice procedure to elicit subjects' risk attitudes (Holt and Laury 2002). In addition, after telling each subject her payoff from the risk elicitation task, we offer an opportunity for subjects to acquire information about the realization of the lottery, again using the BDM mechanism. Their WTP for this information is a measure of their *curiosity*, defined as an intrinsic preference for more information in the absence of any instrumental value associated with that information (Grant, Kajii and Polak 1998, Golman and Loewenstein 2015).

At the end of the experiment, each participant fills out a demographic and strategy survey on the computer and is then paid in private. Each experimental session lasts approximately 90 minutes. The average payment is $27.90, including a $5 show-up fee. The experiment is programmed in z-Tree (Fischbacher 2007).

Table 2: Features of Experimental Sessions

Information to Be Acquired	Immediate Acceptance		Deferred Acceptance	
OwnValue: Own Preferences	free-to-costly	3×12	free-to-costly	3×12
	costly-to-free	3×12	costly-to-free	3×12
OtherValue: Others' Preferences	free-to-costly	3×12	free-to-costly	3×12
	costly-to-free	3×12	costly-to-free	3×12

Notes: Each session, with 12 subjects, has 10 rounds with free information and another 10 rounds with costly information. For any given treatment, sessions with free information rounds first are denoted as "free-to-costly"; and the others with costly information first are denoted as "costly-to-free".

Table 2 summarizes the features of the experimental sessions. For each treatment, we conduct three independent sessions at the Behavioral Economics and Cognition Experimental Lab at the

University of Michigan. As mentioned, each session consists of 12 subjects. No subject participates in more than one session. This design gives us a total of 24 independent sessions and 288 distinct subjects. In addition, we conduct three sessions of our DA-OtherValue (free-to-costly) treatment using a z-Tree program with a coding error in the second ten rounds of the experiment, i.e., a participant's own value is not provided in the second ten rounds. In this case, we use the data from the first ten rounds for these sessions in our data analysis, since the instructions and program for the first half are both correct. With these additional sessions, we have a total of 27 independent sessions with 324 subjects. Our subjects are University of Michigan students, recruited using ORSEE (Greiner 2015). Experimental instructions are included in Appendix B; the data are available from the authors upon request.

5 Experimental Results

In the following discussion, we focus on participants' WTP for information and their welfare in the game under various information structures and information-provision policies. We provide our summary statistics for the key variables in Table C.3 of Appendix C. Since we find that our results on the rank-ordered lists of schools are largely consistent with our theoretical predictions, we relegate these to our online appendix.[7]

We introduce several shorthand notations in presenting the results. First, we let $x > y$ denote that a measure under treatment x is greater than the corresponding measure under treatment y at the 5% significance level or less. Second, we let $x \geq y$ denote that a measure under x is greater than the corresponding measure under y, but that this difference is not statistically significant at the 5% level.

5.1 Willingness to Pay for Information

Our theoretical model predicts that participants' WTP for their own valuation of school b should be greater under IA than that under DA. Figure 3 depicts the time series of the average WTP for a participant's own value for school b, with the theoretical predictions for risk neutral students represented by the horizontal dashed lines. The results in Figure 3 show that, while the average

[7] Our online appendix is posted under the "Working Papers" section on http://yanchen.people.si.umich.edu/.

WTP for OwnValue under IA is mostly within the theoretical bounds, it is substantially above the risk neutral prediction under DA. We also observe that the average WTP is lower (and closer to theoretical predictions) for those participants who receive their free information for the first ten rounds, indicating the importance of learning and cognitive load. We examine these empirical regularities further in subsequent subsections.

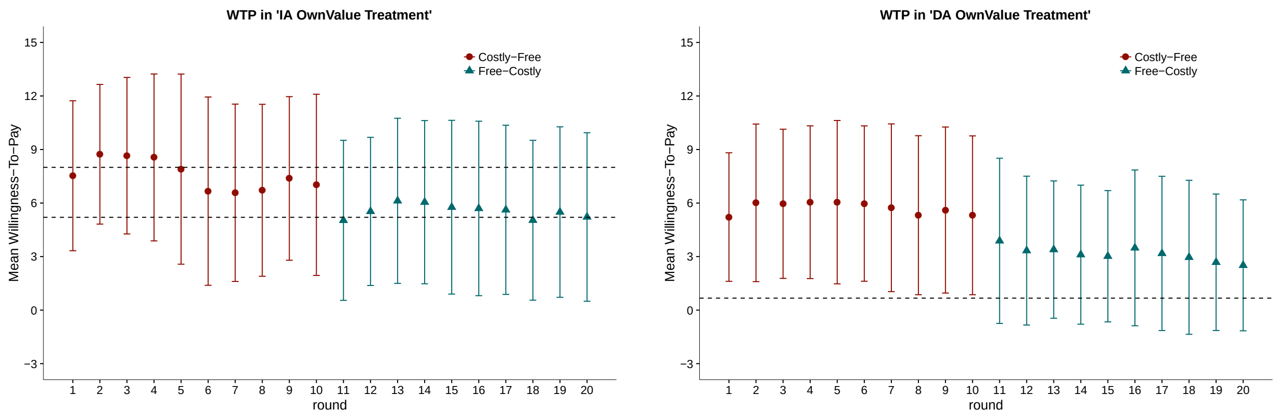

Figure 3: Average WTP for Own Value by Rounds

Notes: Horizontal dashed lines denote the theoretical predictions for the WTP of risk-neutral students. Error bars represent one standard deviation. The first ten rounds (costly first, i.e., costly information acquisition rounds played before the free information ones) and last ten rounds (free first, i.e., rounds with free information played before those with costly information acquisition) are from different sessions.

The last two columns of Table 3 provide the theoretical WTP for risk-neutral (averse) students for each treatment. In most cases, we find that risk aversion predicts a lower WTP in our environment. Note that our predictions are often intervals, as WTP depends on the number of others who successfully acquire information. As this number is uncertain *ex ante*, the prediction is an interval that takes into account all possibilities. The first four columns in Table 3 present the session average WTPs, with the corresponding standard deviations in parentheses.

We now state our first formal hypothesis regarding WTP to find out one's own value for the unknown school, and present the corresponding result.

Hypothesis 1 (WTP for own value)**.** *A subject's WTP to acquire information about her own preferences under IA is greater than it is under DA; both are positive. That is, IA > DA > 0.*

Result 1 (WTP for own value)**.** *A subject's WTP to acquire information about her own value under IA is significantly greater than it is under DA; both WTPs are positive.*

Table 3: Average Willingness-To-Pay for Information by Treatment

Treatment	All six Sessions	free-to-costly Sessions	costly-to-free Sessions	Theoretical Prediction Risk Neutral	Theoretical Prediction Risk Averse
IA OwnValue	6.56 (4.78)	5.56 (4.59)	7.57 (4.75)	[5.2, 8]	[4.9, 7.7]
IA OtherValue	4.51 (4.55)	4.00 (4.55)	5.02 (4.49)	[0, 0.24]	[0, 0.46]
DA OwnValue	4.44 (4.38)	3.16 (4.05)	5.72 (4.33)	0.67	[0.3, 0.4]
DA OtherValue	2.21 (3.16)	1.90 (3.25)	2.52 (3.04)	0	0

Notes: The WTPs are measured in experiment points. There are six sessions under each treatment. For any given treatment, three sessions have free information rounds first (denoted as free-to-costly), while the other three have costly information first (denoted as costly-to-free). Standard deviations are in parentheses and are calculated by treating each subject-round outcome as one observation. Therefore, for each treatment, there are 720 observations from the 10 costly rounds for these 72 subjects, half of which are from the costly-to-free treatments, while the other half are from the free-to-costly treatments.

Support: Table 3 presents the session average WTP for each treatment. Using each session as an independent observation, we reject the null of no difference in favor of Hypothesis 1 that IA > DA ($p = 0.03$, one-sided Wilcoxon rank-sum test). Furthermore, we find that the average WTP for one's own value under IA is 6.56 (s.d. 4.78), while that under DA is 4.44 (s.d. 4.38). Both are significantly different from zero at the 1% level. ∎

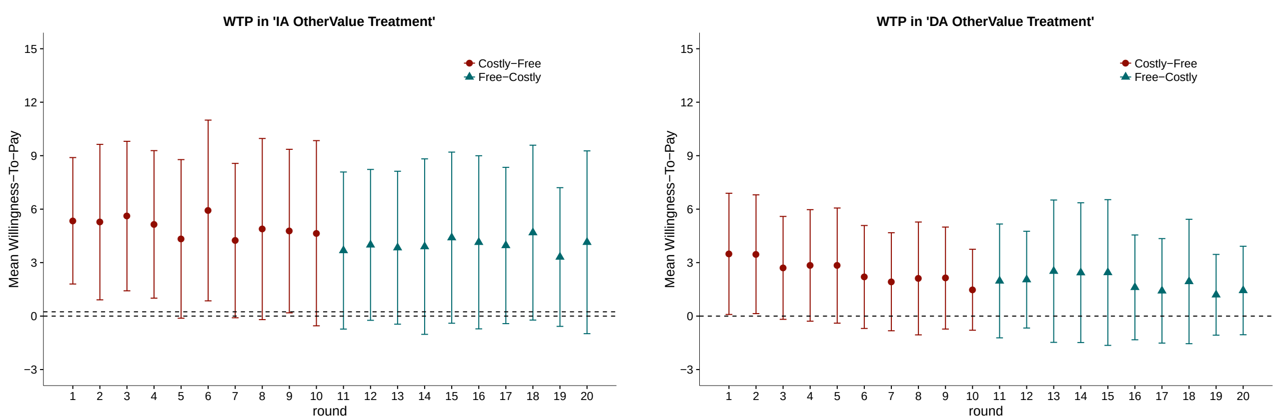

Figure 4: Average WTP for Others' Values by Rounds

Notes: Horizontal dashed lines denote the theoretical predictions for the WTP of risk-neutral students. Error bars represent one standard deviation. The first ten rounds (costly first, i.e., costly information acquisition rounds played before the free information aboutes) and last ten rounds (free first, i.e., rounds with free information played before those with costly information acquisition) are from different sessions.

We next examine our results regarding a subject's WTP for information about others' preferences. Figure 4 depicts the time series of the average WTP for OtherValue, again, with the theoretical predictions for risk neutral students represented by the horizontal lines. The results in Figure 4 show that our subjects exhibit a substantially greater WTP than our risk neutral (averse) predictions. We summarize our theoretical predictions in the following hypothesis and then present the corresponding result.

Hypothesis 2 (WTP for others' values)**.** *A subject's WTP for others' values is zero under DA regardless of risk attitude, whereas it is positive under IA. Therefore, IA > DA = 0.*

Result 2 (WTP for others' values)**.** *A subject's WTP for others' values under IA is significantly greater than it is under DA, but both are significantly different from zero: IA > DA >* 0.

Support: Table 3 presents the average WTP for others' values in each treatment. Treating each session as an independent observation, we reject the null of no difference in favor of Hypothesis 2 that a subject's WTP for others' values follows IA > DA, using a one-sided Wilcoxon rank-sum test ($p = 0.01$). Furthermore, we find that the average WTP for others' values under IA is 4.51 (s.d. 4.55), while that under DA is 2.21 (s.d. 3.16). Both are significantly different from zero at the 1% level. ∎

Our results further show that, under either IA or DA, a subject's WTP for her own value is significantly greater than it is for others' values ($p = 0.01$, one-sided Wilcoxon rank-sum test), consistent with our theoretical predictions.

In sum, we find that our WTP directional comparisons across mechanisms and across participants' own values and others' values are consistent with theory. However, we also find that participants indicate a greater WTP for information across treatments. This over-investment in information acquisition relative to equilibrium predictions has been observed in other endogenous information acquisition experiments in the context of jury or committee voting (Bhattacharya et al. 2015) as well as private value auctions (Gretschko and Rajko 2015). In particular, Bhattacharya et al. (2015) explain their observed over-investment in information acquisition as a result of a combination of poor strategic thinking and the quantal response equilibrium model. On the other hand, Gretschko and Rajko (2015) use regret avoidance to explain their observed overinvestment.

In our experimental design, we include a number of measurements of individual traits that may be relevant in examining information preferences, such as curiosity, risk preferences, and conformity. As excess WTP for information has important welfare implications (section 5.3), it is crucial to understand its determinants. In the following discussion, we first examine our results regarding the determinants of WTP for information at the subject level (section 5.1.1) and by using panel regressions (section 5.1.2). Based on these findings, we then decompose our observed WTP into the effects of behavioral and cognitive explanatory factors (section 5.2).

5.1.1 Determinants of WTP for Information: Subject-Average

We first investigate the determinants of WTP for information at the subject level. Because the elicited WTP is censored below at 0 and above at 15, we use the following Tobit model to analyze the determinants of the observed subject average WTP, denoted by $\overline{WTP}_i$:

$$\overline{WTP}_i^* = X_i'\beta + \varepsilon_i, \text{ and } \overline{WTP}_i = \max\{0, min\{\overline{WTP}_i^*, 15\}\},$$

where ε_i is normally distributed; X_i is a vector of independent variables and controls; and the latent variable $\overline{WTP}_i^*$ is not observed, while the subject-average WTP, $\overline{WTP}_i$, is the observed, censored version of $\overline{WTP}_i^*$. In the above specification, $\overline{WTP}_i$ is obtained by averaging all of subject i's WTPs from different rounds. Our Tobit model takes into account that $\overline{WTP}_i$ is censored for 18% of our main sample observations (43 out of 241 consistent subjects, defined below).[8] For our independent variables, we include the four treatment dummies and demographic variables. Furthermore, we consider the effects of the following factors:

(i) **% playing a dominated strategy with free info**: This variable measures the proportion of time a subject plays a dominated strategy in the free information rounds. As such, this variable is negatively correlated with a subject's understanding of the school choice game under DA. Note that there is no similar measure for IA, as it lacks a dominant strategy.

(ii) **Costly-to-free**: This dummy variable indicates that a session follows a costly-to-free information acquisition order. Playing the game with costly information acquisition in the first ten rounds imposes a higher cognitive load as participants must learn both the school choice game and the information acquisition game. This greater cognitive load, reflected in Table 3, may lead to a sub-optimal WTP.

[8] In a robustness check, we find quantitatively and qualitatively similar results in linear models (see Table C.4 in Appendix C).

(iii) **Curiosity**: This variable measures a subject's curiosity using her WTP for the lottery realization in the Holt-Laury risk elicitation task. As such information is non-instrumental, this WTP reflects a subject's "curiosity," or general preference for information.

(iv) **Risk aversion**: Risk aversion is measured as the switching point in the Holt-Laury lottery choice menu. Following the literature, we define a *consistent* subject as one who exhibits one switching point and chooses the right column in the last lottery choice. In our sample, we find that 241 of our 288 subjects (84%) are consistent; among these, we find that 78% are risk-averse, 16% risk-neutral, and 7% risk-loving. Theoretically, a greater degree of risk aversion suggests a lower WTP for information (see Table 3 in the online appendix).

Again, we first state our hypothesis and then discuss our results.

Hypothesis 3 (Understanding DA, Curiosity, and Cognitive Load)**.** *Participants who play dominated strategies under DA, who are in the costly-to-free information treatments, who are more curious or less risk averse, should exhibit a higher WTP for information about their own or others' values for school* b*.*

Table 4 presents the results for the four Tobit specifications investigating the determinants of subject-average WTP. Column (1) includes the results for our full sample, whereas columns (2)-(4) include only consistent subjects, progressively adding more controls.

While the treatment effects estimated from the Tobit model are largely consistent with Results 1 and 2, this set of analyses uncovers additional findings. First, we find that the variables, % playing a dominated strategy with free info and curiosity, are each positively correlated with WTP. Furthermore, we find that the timing of the information acquisition game in the experiment matters. That is, when subjects have to learn both the school choice mechanism and the information acquisition game in the first ten rounds, i.e., costly-to-free $= 1$, they exhibit a higher WTP. This is consistent with previous experimental findings that a higher cognitive load can cause sub-optimal play (Bednar, Chen, Liu and Page 2012). Lastly, consistent with our theoretical prediction, we find that subjects with a greater risk aversion exhibit a lower WTP (column 3); however, risk aversion becomes insignificant once we include demographic controls (column 4). We summarize our findings below.

Table 4: Determinants of Subject-Average WTP: Tobit Model

	(1) Full Sample	(2) Sub-sample	(3) Sub-sample	(4) Sub-sample
IA_OwnValue	6.45***	6.26***	5.22***	3.75
	(0.56)	(0.57)	(1.10)	(4.01)
IA_OtherValue	4.32***	4.05***	3.46***	1.69
	(0.62)	(0.72)	(1.21)	(4.15)
DA_OwnValue	4.13***	3.78***	2.94***	1.36
	(0.71)	(0.82)	(1.07)	(3.95)
DA_OtherValue	1.47***	1.01**	0.91	-0.53
	(0.45)	(0.47)	(1.13)	(4.02)
% playing a dominated strategy with free info[a]			6.85***	7.10***
			(2.02)	(2.62)
Curiosity			0.34***	0.34***
			(0.05)	(0.04)
Costly-to-free			1.88***	1.84***
			(0.45)	(0.35)
Risk Aversion			-0.28**	-0.23
			(0.13)	(0.15)
Female				-0.87**
				(0.43)
Graduate Student				0.44
				(1.96)
Black				-1.61
				(1.46)
Asian				-0.79*
				(0.45)
Hispanic				-2.97***
				(1.13)
$\#Obs.$	288	241	241	241

Notes: The outcome variable is the subject-level average WTP for information. There are 42 (out of 241, or 17%) subjects with an average WTP $= 0$ and one subject with WTP $= 15$. Columns 2-4 include only consistent subjects in the Holt-Laury lottery game. Column 4 also includes the following controls: age, ACT score, SAT score, dummy for other non-white ethnicities/races, dummy for ACT score missing, dummy for SAT score missing, dummy for degree missing, dummy for age missing, dummy for ethnicity missing, and dummy for gender missing. Standard errors clustered at the session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

a. "% playing a dominated strategy with free info" is defined as the percentage of times a subject plays a dominated strategy in the OwnValue or OtherValue treatment of DA in rounds without information acquisition. We find a mean $= 0.09$ (s.d. $= 0.14$) among all subjects ($n = 144$) who play the information acquisition game under DA. Their rounds without information acquisition, i.e., with no or free information provision, are considered in the construction of this variable. This variable equals zero for the IA treatments.

Result 3 (Understanding DA, Curiosity, and Cognitive Load). *Participants who play dominated strategies under DA, who are in the costly-to-free information treatments, and who are more curious exhibit a significantly higher WTP to acquire information about their own or others' values for school b.*

5.1.2 Determinants of Willingness to Pay for Information: Panel Data Analyses

We next explore within-subject time-series variations using panel data methods on the subject-round observations. Compared to our analysis of the subject average WTP in the previous subsection (5.1.1), panel regressions enable us to exploit the i.i.d. draws of subject values for school b in each round, and investigate factors affecting the dynamics of subject WTP within a session. Regarding the effect of the school choice mechanism on information acquisition incentives, we predict the following.

Hypothesis 4 (Mechanism Effect)**.** *(a) Participants should have no incentive to acquire others' values under DA, regardless of their school-b value; however, (b) those with with high values for school b should have an incentive to acquire others' values under IA.*

Furthermore, in our environment, investments in information acquisition under IA can be seen as strategic substitutes, as defined by Bulow, Geanakoplos and Klemperer (1985). Namely, participants' information-acquisition investments mutually offset one another in equilibrium: A participant has a lower WTP when she expects others having a higher WTP. The intuition behind this characterization is as follows: When a subject's own value is unknown, it is a dominant strategy to submit (a, b, c); the benefit of acquiring information on one's own value is to submit (b, a, c) if school b turns out to be the best; however, the magnitude of this benefit decreases with the number of other students playing (b, a, c). The latter is positively correlated with others' WTP for information. Similar arguments apply for the incentive to acquire information about others' values. This leads to our next hypotheses.

Hypothesis 5 (Information Acquisition as Strategic Substitutes)**.** *Participants' WTP to acquire information about others' values should decrease if they expect that others' WTP is greater.*

To test these hypotheses, we specify a panel regression model which considers WTP when a subject has a high value for school b under each mechanism, and a guess regarding others' WTP for information. Similar to our analysis of the subject-average WTP (Table 4), we again take into account that WTP is bounded within $[0, 15]$, and specify the following Tobit model:

$$\begin{aligned} WTP^*_{i,t} =& \alpha_i + \beta_1 High\_B \times IA\_OtherValue_{i,t} + \beta_2 High\_B \times DA\_OtherValue_{i,t} \\ & + \beta_3 WTP\_Guess_{i,t} + Controls_{i,t} + \varepsilon_{i,t}, \\ WTP_{i,t} =& \max\{0, min\{WTP^*_{i,t}, 15\}\}, \end{aligned} \tag{1}$$

where i indexes subjects, and t indexes rounds (within each session); and $WTP^*_{i,t}$ and $WTP_{i,t}$ respectively denote the true uncensored and the observed censored WTP. Given the non-linear nature of the Tobit model, we cannot consistently estimate α_i as subject fixed effects with a short panel (ten rounds). Consequently, we use a random effects Tobit model. For the above specification, we run each set of analyses for all four treatments both individually and pooled.[9]

Our explanatory variables include $High\_B \times IA\_OtherValue_{i,t}$, which equals one if in round t subject i has a high value for school b under the treatment IA OtherValue, and zero otherwise. We also include $High\_B \times DA\_OtherValue_{i,t}$, which we define similarly.[10] Our theoretical model predicts that the coefficient of $High\_B \times IA\_OtherValue_{i,t}$ should be positive, whereas that of $High\_B \times DA\_OtherValue_{i,t}$ should be zero. The results from our analysis support these predictions.

Another key explanatory variable is $WTP\_Guess_{i,t}$, which is subject i's estimate of her opponents' average WTP in round t. Here, we predict that a risk-neutral student i's own WTP and $WTP\_Guess_{i,t}$ should be independent of each other under DA and negatively correlated under IA.[11]

We also address the possibility that common shocks to i in round t may affect both $WTP_{i,t}$ and $WTP\_Guess_{i,t}$. We address this potential endogeneity concern with an instrumental-variable approach and present the results in Tables C.5 and C.6 in Appendix C. In these specifications, we use the realization of opponents' WTP in the previous round as an instrumental variable (IV) for $WTP\_Guess_{i,t}$. This variable is correlated with $WTP\_Guess_{i,t}$ as it provides information about how others play the game, but does not have a direct effect on $WTP_{i,t}$. Given the IV's validity, we

[9]As robustness checks, we present the respective fixed effects (Table C.5) and random effects linear panel regressions (Table C.6) in Appendix C.

[10]For the other treatments, IA or DA OwnValue, it is impossible to define a similar variable, as subjects do not know their own value for school b when deciding whether to acquire information.

[11]With risk-averse students, the WTP is still negatively correlated with $WTP\_Guess_{i,t}$ under IA, but the correlation is very weakly positive under DA for OwnValue. See the online appendix for more details.

fail to reject the null hypothesis that $WTP\_Guess_{i,t}$ is exogenous. Therefore, we conclude that endogeneity is not an issue for our analyses.

Finally, other controls include round (i.e., a linear time trend) and round in free-to-costly sessions. It should be emphasized that a large set of robustness checks reveals that inclusion of additional controls, such as a subject's accumulated wealth at the beginning of the round and whether a subject successfully acquired information in $t-1$, does not change our results significantly.

We present the results of our panel data analyses in Table 5, where the four columns differ in the inclusion or exclusion of demographics, wealth, and outcomes of information acquisition in the previous round. These results show that the coefficient on $WTP\_Guess_{i,t}$ is always positive and significant, contrary to our theoretical predictions. Our results also show that subjects lower their excess WTP over time, although this reduction is insignificant. Lastly, the coefficients on our four factors (% playing a dominated strategy with free info, Curiosity, Costly-to-free, and Risk Aversion) are similar to those in Table 4, although Costly-to-free and Risk Aversion are now sometimes insignificant. We summarize our results below.

Result 4 (Mechanism Effect)**.** *For participants with high values for school b, their WTP to acquire information about others' values is not significantly different from zero under DA, but is positive and significant under IA.*

By Result 4, we fail to reject Hypothesis 4(a) or (b), indicating that the mechanism effect on WTP for others' preferences is largely consistent with our theoretical predictions. By contrast, we obtain the following result regarding conformity.

Result 5 (Conformity)**.** *Participants who expect their opponents to pay more for others' values increase their own WTP to acquire information about their opponents' values for school b.*

By Result 5, we reject Hypothesis 5. We interpret this result as indicating that participants are willing to pay more to know others' preferences when they expect their opponents have high WTPs, a phenomenon we call *conformity*.

5.2 Decomposition of a Subject's WTP

In this section, we explore factors that may drive our observation that subjects over-invest in information acquisition. Specifically, we decompose our observed WTP treatment-by-treatment ac-

Table 5: Determinants of WTP: Random Effects Panel Tobit Analyses

	(1)	(2)	(3)	(4)
IA_OwnValue	2.77***	2.59***	2.72***	2.54***
	(0.75)	(0.80)	(0.76)	(0.81)
IA_OtherValue	0.99	0.59	0.95	0.55
	(0.97)	(0.85)	(0.99)	(0.86)
DA_OwnValue	2.31**	2.17**	2.30**	2.16**
	(1.13)	(0.95)	(1.04)	(0.99)
High_B $\times$ IA_OtherValue	3.16***	3.15***	3.16***	3.15***
	(1.14)	(0.86)	(1.10)	(0.90)
High_B $\times$ DA_OtherValue	-0.68	-0.66	-0.68	-0.66
	(1.07)	(1.04)	(1.04)	(1.16)
$WTP\_Guess_{i,t}$: Guess of Opponents' WTP in t	0.89***	0.88***	0.89***	0.88***
	(0.10)	(0.10)	(0.10)	(0.09)
% playing a dominated strategy with free info	7.79***	7.79**	7.95***	7.97**
	(2.57)	(3.59)	(2.39)	(3.43)
Curiosity	0.39***	0.39***	0.39***	0.39***
	(0.06)	(0.06)	(0.06)	(0.06)
Costly-to-free	1.31*	1.16*	1.69	1.59*
	(0.71)	(0.64)	(1.09)	(0.94)
Risk Aversion	-0.41**	-0.32	-0.40**	-0.32
	(0.19)	(0.21)	(0.18)	(0.20)
Round	-0.05	-0.05	-0.09	-0.10
	(0.07)	(0.08)	(0.09)	(0.09)
Round $\times$ costly-to-free	-0.10	-0.10	-0.10	-0.10
	(0.10)	(0.10)	(0.09)	(0.10)
Accumulated wealth up to $t-1$			0.00	0.00
			(0.00)	(0.00)
Successfully acquired info in $t-1$			0.29	0.29
			(0.25)	(0.24)
Other demographical controls	No	Yes	No	Yes
# of observations	2169	2169	2169	2169
# of subjects	241	241	241	241

Notes: The regression sample includes only consistent subjects in the Holt-Laury lottery game. There are 241 subjects in this sample each of whom has 9 observations from rounds 2-10. Columns (2) and (4) include additional demographical controls: dummy for female, dummy for graduate student, dummy for black, dummy for Asian, dummy for Hispanic, dummy for other non-white ethnicities/races, age, ACT score, SAT score, dummy for ACT score missing, dummy for SAT score missing, dummy for degree missing, dummy for age missing, dummy for ethnicity missing, and dummy for gender missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

cording to the following factors identified from our panel analysis: cognitive load, learning, % playing a dominated strategy with free info, conformity, curiosity, and risk aversion.

To perform our decomposition, we first estimate the Tobit model, as in equation system (1), for each separate treatment (columns 1-4 in Table 6). Doing so allows each factor to have a different effect in a given treatment. Indeed, we find that the coefficients for some of our key variables change significance levels (e.g., Round and Risk Aversion). As a comparison, we present our

Table 6: Determinants of WTP: Separate and Pooled Random Effects Panel Tobit Analyses

	IA		DA		
	OwnValue	OtherValue	OwnValue	OtherValue	Pooled
	(1)	(2)	(3)	(4)	(5)
IA_OwnValue					2.59***
					(0.80)
IA_OtherValue					0.59
					(0.85)
DA_OwnValue					2.17**
					(0.95)
High_B × IA_OtherValue		3.65***			3.15***
		(0.88)			(0.86)
High_B × DA_OtherValue				-0.63	-0.66
				(0.91)	(1.04)
Guess of Opponents' WTP in t	0.83***	0.77***	0.90***	1.06***	0.88***
	(0.16)	(0.17)	(0.07)	(0.19)	(0.10)
% playing a dominated strategy with free info	–	–	4.78*	6.16	7.79**
			(2.58)	(7.10)	(3.59)
Curiosity	0.39***	0.49***	0.35***	0.44***	0.39***
	(0.13)	(0.17)	(0.12)	(0.15)	(0.06)
Costly-to-free	1.70*	2.72***	-0.23	-0.41	1.16*
	(1.01)	(0.98)	(1.25)	(2.09)	(0.64)
Risk Aversion	-1.65***	-0.35	-0.37	0.31	-0.32
	(0.48)	(0.39)	(0.34)	(0.30)	(0.21)
Round	-0.11***	0.15	0.02	-0.24	-0.05
	(0.03)	(0.24)	(0.04)	(0.20)	(0.08)
Round × Costly-to-free	-0.20***	-0.19	-0.06	0.11	-0.10
	(0.07)	(0.26)	(0.06)	(0.22)	(0.10)
Other demographical controls	Yes	Yes	Yes	Yes	Yes
# of observations	567	513	576	513	2169
# of subjects	63	57	64	57	241

Notes: Estimates are from random effects panel Tobit models for each treatment separately and pooled, including only consistent subjects in the Holt-Laury lottery game. The sample includes 9 rounds of costly information acquisition for every subject, excluding the first round. Column 5 repeats column 2 in Table 5. All specifications include these additional controls: dummy for female, dummy for graduate student, dummy for black, dummy for Asian, dummy for Hispanic, age, ACT score, SAT score, dummy for ACT score missing, dummy for SAT score missing, dummy for degree missing, dummy for age missing, dummy for ethnicity missing, and dummy for gender missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

results based on the pooled regression in column (5), which reflects the same regression from column (2) of Table 5.

Based on these estimated coefficients, Table 7 presents the decomposition of subject WTP for information. Overall, this analysis shows that our six factors can explain the majority of the observed WTP.[12] We explore each factor below.

[12] As a robustness check, decompositions based on pooled regressions (column 5 in Table 6) are presented in Table C.7 in Appendix C, which show similar results.

Table 7: Decomposition of Subject WTP for Information

	IA OwnValue (1)	**IA OtherValue** (2)	**DA OwnValue** (3)	**DA OtherValue** (4)
WTP: data	**6.44**	**4.32**	**4.17**	**1.84**
	(4.87)	(4.68)	(4.30)	(2.86)
Model prediction[a]	6.37	4.20	4.13	1.73
	(3.23)	(2.95)	(2.90)	(2.10)
(i) Cognitive load[b]	1.04	1.02	0.01	-0.21
	(0.43)	(0.65)	(0.23)	(0.17)
(ii) Learning over rounds[b]	0.64	-0.09	0.06	0.31
	(0.59)	(0.27)	(0.13)	(0.36)
(iii) Conformity[b]	3.91	2.17	2.60	1.31
	(1.93)	(1.67)	(2.05)	(1.69)
(iv) % playing a dominated strategy with free info[b]			0.30	0.24
			(0.52)	(0.60)
(v) Curiosity[b]	1.63	1.34	0.95	0.52
	(1.69)	(1.64)	(1.30)	(1.11)
(vi) Risk aversion[b]	-1.64	-0.27	-0.44	0.27
	(1.46)	(0.32)	(0.44)	(0.31)
Total Explained by factors (i)-(vi)[c]	**5.13**	**3.46**	**3.22**	**1.62**
	(2.94)	(2.66)	(2.63)	(2.08)
Residual WTP[d]	**1.31**	**0.86**	**0.94**	**0.21**
	(3.75)	(3.78)	(2.99)	(2.12)
Theoretical prediction[e]	**[5.2, 8]**	**[0, 0.24]**	**0.67**	**0.00**
# of observations	567	513	576	513
# of subjects	63	57	64	57

Notes: Decompositions are based on the random effects Tobit model for each treatment (columns 1-4 in Table 6). The table reports the sample average, while standard deviations are in parentheses.
a. "Model prediction" is the predicted value of E(WTP) based on the corresponding estimated model, assuming that unobserved error terms are equal to zero. The predicted values are censored to $[0, 15]$.
b. The WTP explained by the corresponding factor is the difference between the model prediction with and without the factor. The former is predicted from the current values of all variables; the latter is calculated by setting the relevant variable value to zero (for factors "Cognitive load," "Conformity," "% playing a dominated strategy with free info," or "Curiosity") or setting the relevant variable to the counterfactual value (i.e., for "Risk aversion," we set the risk aversion measure to the risk-neutral value; for "Learning over round," we set "Round" to be the last round, i.e., "Round" $= 10$).
c. "Total Explained by factors (i)-(vi)" is the total WTP explained by the six factors above. Note that it is not the sum of the explained WTP of the six factors because of the censoring at 0 and 15.
d. "Residual WTP" is the difference between the observed WTP and the total WTP explained by the six factors.
e. The theoretical predictions are for risk neutral subjects.

(i) **Cognitive Load:** Our regression results indicate that the costly-to-free order is associated with an average of 1.16 points extra WTP in every round among all treatments (Table 6, column 5), but that this effect is not present in the DA treatments (Table 6, columns 3 and 4). Moreover, we find that the order also affects learning over rounds based on the coefficients on "Round" and "Round $\times$ Costly-to-free": in costly-to-free sessions, learning over rounds reduces the WTP faster, on average. Indeed, by round 10, we find that WTP in the costly-to-free sessions is reduced by 1.20 points relative to round 2, while this reduction is only 0.40

in the free-to-costly sessions.

To quantify the effect of cognitive load, we consider the counterfactual of replacing costly-to-free by free-to-costly. That is, we set all "Costly-to-free" and "Round $\times$ Costly-to-free" values to zero. We then measure the effect of cognitive load as the difference between the model prediction based on the true variable values and the prediction under the counterfactual. Both predictions are censored to guarantee that the predicted WTP falls between 0 and 15. The results in Table 7 show that the presence of cognitive load changes WTP by -0.21 to 1.04 points.

(ii) **Learning over rounds:** To assess learning over rounds, we consider the counterfactual of replacing a subject's behavior in rounds 2-9 with her behavior in round 10, as round 1 is omitted from our regression. Using the same censoring as above, we find that the estimated effect of learning, or the difference between the prediction with the observed variable values and the prediction under the counterfactual, accounts for between -0.09 to 0.64 points of WTP.

(iii) **Conformity:** Our conformity measure reflects the extent to which subjects positively respond to their beliefs about others' behavior, WTP_guess. Our results for conformity show that expecting that others will pay one extra point increases WTP by an average 0.88 points (column 5 in Table 6).

Although our theory predicts a negative correlation between a subject's own WTP and WTP_guess, we also consider the counterfactual where the correlation is zero. That is, in the counterfactual, there is no effect of WTP_guess on WTP. Our results show that the correlation between WTP and WTP_guess can explain 1.31 to 3.91 points of the observed WTP, or 50% to 71%, indicating that this is the single most important factor in explaining our observed WTP.

(iv) **% playing a dominated strategy with free info:** This measure is the proportion of time that a subject plays a dominated strategy in the DA treatment rounds without information acquisition. The results in Table 6 show that % playing a dominated strategy with free info increases a subject's WTP.

To quantify this effect, we consider the counterfactual without any misunderstanding of the mechanism (i.e., setting the variable to zero). We then calculate the difference between the model prediction with the observed variable values and the prediction under the counterfactual. The results in columns 3 and 4 in Table 7 show that the effect is 0.30 (0.24) under the DA OwnValue (OtherValue) treatments.

(v) **Curiosity:** From the regression in column 5 of Table 6, we see that a 1 point increase in WTP to pay for non-instrumental information is associated with 0.39 additional points of WTP in each round.

We thus consider the counterfactual where WTP for information in the school choice game is not associated with curiosity by setting the coefficient on Curiosity to zero. After similar calculations, we find that curiosity explains 0.52 to 1.63 points of our observed WTP.

(vi) **Risk Aversion:** To measure risk aversion, we use a subject's switching point in the Holt-Laury lottery choice game. In general, our results show that being more risk averse is correlated with a lower WTP, which is consistent with our theoretical predictions, albeit insignificantly so (Table 6, column 5). We further find that this correlation is heterogeneous across treatments, and that it becomes positive in the DA OtherValue treatment.

We thus consider the counterfactual where every subject is risk neutral (i.e., switching at the 5th choice in the Holt-Laury game), which requires us to change about 78% of our subjects from risk averse to risk neutral. Doing so, we find that risk aversion decreases WTP by 0.27 to 1.64 points, with the exception that it increases WTP in the DA OtherValue treatment.

Together, these findings can be summarized in the following result.

Result 6 (Decomposing WTP)**.** *The six factors combined explain 71-88% of the observed WTP for preferences about the unknown school. Of these, conformity explains 50-71% of the WTP, whereas curiosity explains 23-31% of the WTP.*

Overall, after accounting for the WTP explained by the six factors, the remaining WTP is similar to the level predicted by our theory for our two DA treatments. However, it is below the theoretical prediction level for the IA OwnValue treatment (1.31 versus [5.2, 8]), and slightly above the theoretical prediction for the IA OtherValue (0.86 versus [0, 0.24]).

Table 8: Effects of Information Acquisition & Provision: Policies Implemented in the Experiment

	Information Provision		Information Acquisition (Observed Effects)		
	Welfare effect^a		Welfare effect^a		Pr(Info Acquired)
	Theoretical	Observed	w/ low cost^b	w/ high cost^b	
IA OwnValue	9.60	8.16	2.30	0.22	0.44
IA OtherValue	0	-0.01	0.09	-1.23	0.28
DA OwnValue	5.34	4.26	-0.28	-1.57	0.30
DA OtherValue	0	0.42	-0.98	-1.45	0.14

Notes: More detailed estimates are shown in Tables C.8 and C.9 in Appendix C. a. Given a treatment, the welfare effect measures the difference between the average payoff (per round, per subject) with costly information acquisition (or free information provision) and the average payoff without it, while the costs of information acquisition are taken into account.
b. "Low cost" and "high cost" are two technologies for information acquisition. The former is the one used in the experiment, where a subject in expectation pays a half of her WTP when successfully acquiring the information; in the latter, the subject always pays her WTP if successfully acquiring information. With either of the technologies, subjects do not pay if they do not successfully obtain information.

5.3 Welfare Analysis

In our final set of analysis, we examine the welfare effects of both information provision by an educational authority and information acquisition by participants. To do so, we use two measures of welfare: the payoffs that subjects receive in the experiment and the efficiency of the allocation outcome. An allocation is deemed efficient if a subject, who values school b at 110, is matched with school b whenever such a subject exists. Similarly, when there are two or more subjects with high values for school b, an efficient allocation assigns one of them to school b.

5.3.1 Welfare Effects of Information Provision

Similar to our theoretical analyses provided in our online appendix, we first consider three information structures without information acquisition: (a) Uninformed (UI), (b) Cardinally Informed (CI), and (c) Perfectly Informed (PI).[13] Based on these analyses, we derive the following hypothesis:

Hypothesis 6 (Efficiency under Information Provision: Information Structures)**.** *With free information provision, the subject welfare and the fraction of efficient allocations under either mechanism follow the order of UI $<$ CI $=$ PI. Furthermore, the allocation is always (not always) efficient when subjects are either cardinally or perfectly informed under IA (DA).*

[13] By design, CI is equivalent to OI (everyone is informed of her ordinal preferences but not others' preferences). The theoretical *ex ante* welfare for risk-neutral subjects and the allocation efficiency are partially summarized in Table 8 and further detailed in Table C.8 (columns 2 and 4) in Appendix C.

Given our experimental design, we are able to perform both within- and between-treatment tests. For instance, to test the effect of information provision under IA (transforming UI to CI), we use the Wilcoxon matched-pairs signed-ranks test, since subjects first rank the schools knowing only the distribution (UI), and then rank the schools again after given free information about their own school b values (CI) under the free information provision rounds within the same treatment. To test if IA and DA reach the same level of efficiency under CI, we use a Wilcoxon rank-sum (or Mann-Whitney) test for two independent samples, as no individual subject experiences both treatments.

Result 7 (Efficiency under Information Provision)**.** *(i) There is no difference in subject payoffs or allocation efficiency between DA and IA for uninformed participants; (ii) Information provision that transforms uninformed subjects into cardinally informed ones improves both subject payoffs and allocation efficiency under both DA and IA, with a greater effect for IA; (iii) Information provision that transforms cardinally informed subjects to perfectly informed ones does not improve the performance of either mechanism; and (iv) IA outcomes with either cardinally or perfectly informed subjects are closer to the efficient outcome relative to the corresponding DA outcomes.*

Support: For parts (i)-(iii), columns 1 and 3 in Table C.8 report the observed payoff[14] and proportion of efficient allocations under each information condition, respectively. P-values for the Wilcoxon matched-pairs signed-ranks tests or the Wilcoxon rank-sum (or Mann-Whitney) tests are presented, treating each session as an independent observation. For Part (iv), we find that IA outcomes under CI or PI achieve 93-96% of the maximum payoffs and result in efficient allocations among 89-94% of all games; as a comparison, DA outcomes under CI or PI on average achieve only 87-89% of the maximum payoffs and result in efficient allocations among 81-84% of all games. ∎

In sum, Result 7 indicates that information provision about a participant's own preferences improves welfare under both mechanisms, with greater welfare gains under IA.

[14] Also available in Table C.8.

5.3.2 Welfare Effects of Information Acquisition

We now turn to the effects of costly information acquisition on welfare. As the information acquisition technology results in an endogenous probability of receiving the "hard news," there are likely both informed and uninformed subjects. If costs are not considered, we expect outcomes to fall between no information and free information provision.[15]

However, to evaluate the welfare effects of information acquisition, it is necessary to consider the costs. Our experiment can be regarded as a lower cost bound (labeled as "low cost"), as a participant pays half of her WTP in expectation only if information is successfully acquired. The statistics in Table C.9 show that the actual costs paid by subjects in each treatment are 22-34% of their WTP on average.[16] Our results in Table 8 indicate that for our low cost bound, costly information acquisition is welfare-improving in the IA OwnValue treatment as it increases the average payoff for each subject in each round by 2.3 points. However, in the IA OtherValue treatment, it is essentially welfare-neutral, yielding only 0.09 points in profits. In both DA treatments, costly information acquisition is welfare-decreasing.

We next use the same information acquisition technology as in the experiment except that subjects have to pay their WTP instead of the random number when successfully acquiring information (our high cost condition). Doing so, we find that the net loss for each subject in each round ranges from 1.23 in the IA OtherValue treatment to 1.57 in the DA OwnValue, while the only net gain is observed in the IA OwnValue treatment (0.22) (see Table 8).

5.3.3 Designing Information-Provision Policies: Counterfactual Analyses

Using the above results, we can now evaluate the welfare effect of various information provision policies. More specifically, we focus on the following three types of policies, measuring welfare relative to the uninformed baseline where no one knows her own preferences:

(i) **Laissez-Faire Policy**: The educational authority provides no information but lets students acquire information as they wish with either the low- or high-cost technology.

(ii) **Counterfactual 1: Free OwnValue and Costly OtherValue**: The educational authority makes all information relevant to OwnValue available but does not provide information about

[15] Appendix C.3 formalizes and tests additional hypotheses along these lines.

[16] Given the experimental design, when one's WTP is w, the expected cost of information acquisition is $w^2/30$.

Table 9: Welfare Effects of Counterfactual Policies (relative to UnInformed)

	Laissez-Faire Policy Costly Info Acquisition		*Counterfactual 1* Free OwnValue, Costly OtherValue		*Counterfactual 2* Free OwnValue
	w/ low cost[a]	w/ high cost[a]	w/ low cost[a]	w/ high cost[a]	& Free OtherValue
IA	2.34	-0.12	8.26	6.93	8.15
DA	-0.57	-2.00	3.99	2.69	4.68

Notes: This table presents the welfare effects of information acquisition and information provision relative to UI (i.e., nobody knows her own or others' preferences).
a. "Low cost" and "high cost" are two technologies for information acquisition. The former is the one used in the experiment, where a subject in expectation pays a half of her WTP when successfully acquiring the information; in the latter, the subject always pays her WTP if successfully acquiring information. Otherwise, subjects do not pay.

OtherValue. This policy corresponds to those employed by many school districts where information about school characteristics is readily available, but information about others' actions is not. In this setting, students rely on historical data to infer others' actions for the current year.

(iii) **Counterfactual 2: Free OwnValue and Free OtherValue**: The educational authority makes all relevant information freely available.

With the estimated effects of information acquisition and provision, we can now calculate the welfare, measured by the student average payoff in each round, created by each policy. Our results are summarized in Table 9. Taking the free-OwnValue-free-OtherValue policy as an example, we see that its welfare effect under IA is the sum of the welfare gain from providing OwnValue (8.16) and that from providing OtherValue (-0.01) as shown in Table 8.

To analyze the welfare effects under the laissez-faire policy, some additional assumptions are needed. For instance, under IA, we first take the net payoff gain of having OwnValue acquisition under either cost assumption (2.30 or 0.22 in Table 8). We then designate that only those who have successfully acquired OwnValue can engage in acquiring OtherValue. This comprises about 44% of our subjects. We further assume that this leads to 44% of the net payoff gain from acquiring OtherValue ($0.09 \times 44\%$ or $-1.23 \times 44\%$ as in Table 8, depending on cost level).[17] Similarly, for the free-OwnValue-costly-OtherValue policy, we take into account the effect of providing OwnValue as well as that of letting students acquire OtherValue (given that they know OwnValue already).

[17]Note that we ignore the fact that the game with this two-stage information acquisition will have both informed and uninformed players regarding their own values.

Overall, this analysis shows that making both OwnValue and OtherValue freely available leads to an additional 8.15 (4.68) points for every subject in each round under IA (DA). When the cost is low, the laissez-faire policy increases average payoffs only under IA. When the cost is high, this benefit disappears; in other words, engaging in information acquisition decreases social welfare, on average. In comparison, when the cost is high, the free-OwnValue-costly-OtherValue policy is always welfare-improving relative to UI, but is dominated by the free-OwnValue-costly-OtherValue policy.

Interestingly, the comparison between Counterfactuals 1 and 2 reveals that there is always a welfare gain for providing free information about others' preferences under DA. Theoretically, information about others' preferences should affect neither student strategies nor outcomes under DA. However, given our observed over-investment in information acquisition, we can interpret the free provision of information as a mechanism for reducing wasteful investment. This result has strong implications for school districts using DA, suggesting that educational authorities should actively provide information about others' preferences/actions even under DA.

6 Conclusion

This paper provides a number of important new insights for designing better school choice policies. Specifically, we present both theoretical and experimental evidence that the two most popular mechanisms, DA and IA, provide heterogeneous degrees of incentives for students to acquire information about participant preferences.

Our results first show that better information about a participant's own preferences improves student-school match quality, which is in line with recent calls for better information provision on school quality. However, our experimental results also show that students tend to over-invest in information acquisition. These results suggest that information provision may be effective in reducing this wasteful over-investments.

We further demonstrate that acquiring information about others' preferences is related to the gaming aspect of school choice, and that the incentive to do so is determined by a mechanism's strategy-proofness. Theoretically, only a non-strategy-proof mechanism should incentivize students to acquire information about others' preferences. However, our experimental results show that students seek to acquire information about others' preferences even under strategy-proof mechanisms. In our experiment, students still over-invest in learning others' preferences, even though the strategy-proofness of DA makes such information useless. This over-investment is more severe among those expecting that others are paying generously for information, as well as those who are more curious. We find that this wasteful investment can be avoided or reduced by making information about others' preferences freely available. These results thus suggest that educational authorities should consider information provision beyond what is currently provided.

References

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School Choice: A Mechanism Design Approach," *Amer. Econ. Rev.*, June 2003, *93* (3), 729–747.

__ **, Yeon-Koo Che, and Yosuke Yasuda**, "Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered," *The American Economic Review*, 2011, *101* (1), 399–410.

Artemov, Georgy, "School Choice Mechanisms, Information Acquisition and School Rankings," 2016. Unpublished Manuscript.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, 2017.

Bade, Sophie, "Serial dictatorship: The unique optimal allocation rule when information is endogenous," *Theoretical Economics*, 2015, *10* (2), 385–410.

Barlevy, Gadi and Pietro Veronesi, "Information Acquisition in Financial Markets," *The Review of Economic Studies*, 2000, *67* (1), 79–90.

Becker, Gary M., Morris H. DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9* (3), 226–232.

Bednar, Jenna, Yan Chen, Tracy Xiao Liu, and Scott Page, "Behavioral spillovers and cognitive load in multiple games: An experimental study," *Games and Economic Behavior*, 2012, *74* (1), 12 – 31.

Benhabib, Jess, Alberto Bisin, and Andrew Schotter, "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior*, 2010, *69* (2), 205 – 223.

Bergemann, Dirk and Juuso Valimaki, "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 2002, *70* (3), 1007–1033.

__ **and** __ , "Information in mechanism design," 2006.

Bhattacharya, Sourav, John Duffy, and Sun-Tak Kim, "Voting with Endogenous Information Acquisition: Theory and Evidence," 2015. Manuscript, Wuhan University.

Bogomolnaia, Anna and Hervé Moulin, “A new solution to the random assignment problem,” *Journal of Economic Theory*, 2001, *100* (2), 295–328.

Budish, Eric, Yeon-Koo Che, Fuhito Kojima, and Paul Milgrom, “Designing Random Allocation Mechanisms: Theory and Applications,” *American Economic Review*, 2013, *103* (2), 585–623.

Bulow, Jeremy I., John D. Geanakoplos, and Paul D. Klemperer, “Multimarket Oligopoly: Strategic Substitutes and Complements,” *Journal of Political Economy*, 1985, *93* (3), 488–511.

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn, “Constrained School Choice: An Experimental Study,” *American Economic Review*, September 2010, *100* (4), 1860–74.

Chade, Hector, Gregory Lewis, and Lones Smith, “Student Portfolios and the College Admissions Problem,” *The Review of Economic Studies*, 2014, *81* (3), 971–1002.

Chen, Yan and Tayfun Sönmez, “School Choice: An Experimental Study,” *Journal of Economic Theory*, 2006, *127*, 202–231.

Choi, Syngjoo, José-Alberto Guerra, and Jinwoo Kim, “Interdependent Value Auctions with Insider Information: Theory and Experiment,” 2015. Working Paper.

Compte, Olivier and Philippe Jehiel, “Auctions and information acquisition: sealed bid or dynamic formats?,” *The RAND Journal of Economics*, 2007, *38* (2), 355–372.

Crémer, Jacques and Fahad Khalil, “Gathering Information before Signing a Contract,” *The American Economic Review*, 1992, *82* (3), 566–578.

__ , __ , and Jean-Charles Rochet, “Strategic Information Gathering before a Contract Is Offered,” *Journal of Economic Theory*, 1998, *81* (1), 163–200.

Crémer, Jacques, Yossi Spiegel, and Charles Zheng, “Auctions with costly information acquisition,” *Economic Theory*, 2009, *38* (1), 41–72.

Dang, Tri Vi, “Bargaining with endogenous information,” *Journal of Economic Theory*, 2008, *140* (1), 339–354.

De Haan, Monique, Pieter A Gautier, Hessel Oosterbeek, and Bas Van der Klaauw, “The performance of school assignment mechanisms in practice,” Technical Report, CEPR Discussion Papers 2015.

Ding, Tingting and Andrew Schotter, “Learning and Mechanism Design: An Experimental Test of School Matching Mechanisms With Intergenerational Advice,” 2015. New York University Manuscript.

__ and __ , “Matching and chatting: An experimental study of the impact of network communication on school-matching mechanisms,” *Games and Economic Behavior*, 2016, pp. –.

Dubins, Lester E. and Deborah A. Freedman, “Machiavelli and the Gale-Shapley Algorithm,” *American Mathematical Monthly*, 1981, *88*, 485–494.

Dur, UM, Robert G Hammond, and Thayer Morrill, "Identifying the Harm of Manipulable School-Choice Mechanisms," Technical Report 2015.

Dustan, Andrew, Alain de Janvry, and Elisabeth Sadoulet, "Flourish or Fail? The Risky Reward of Elite High School Admission in Mexico City," *Unpublished manuscript*, 2015.

Featherstone, Clayton and Muriel Niederle, "Ex ante efficiency in school choice mechanisms: an experimental investigation," 2008. Stanford University Working Paper.

Fischbacher, Urs, "z-Tree: Zurich Toolbox for Ready-made Economic Experiment," *Experimental Economics*, 2007, *10* (2), 171–178.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg, "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *The American Economic Review*, 2006, *96* (4), 1043–1068.

Gale, David and Lloyd S. Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69*, 9–15.

Gerardi, Dino and Leeat Yariv, "Information acquisition in committees," *Games and Economic Behavior*, 2008, *62* (2), 436–459.

Golman, Russell and George Loewenstein, "The Demand For, and Avoidance of, Information," 2015. Carnegie Mellon University Working Paper.

Gong, Binglin and Yingzhi Liang, "A dynamic college admission mechanism in Inner Mongolia: Theory and experiment," 2017. University of Michigan Working Paper.

Grant, Simon, Atsushi Kajii, and Ben Polak, "Intrinsic Preference for Information," *Journal of Economic Theory*, 1998, *83* (2), 233 – 259.

Greiner, Ben, "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 2015, *1* (1), 114–125.

Gretschko, Vitali and Alexander Rajko, "Excess information acquisition in auctions," *Experimental Economics*, 2015, *18* (3), 335–355.

Guillen, Pablo and Alexander Hing, "Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism," *European Economic Review*, 2014, *70*, 178 – 185.

Harless, Patrick and Vikram Manjunath, "The Importance of Learning in Market Design," 2015. Working Paper.

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

Hauswald, Robert and Robert Marquez, "Competition and Strategic Information Acquisition in Credit Markets," *Review of Financial Studies*, 2006, *19* (3), 967–1000.

He, Yinghua, "Gaming the Boston School Choice Mechanism in Beijing," December 2014. Toulouse School of Economics Working Paper.

__ , **Antonio Miralles, Marek Pycia, and Jianye Yan**, "A Pseudo-Market Approach to Allocation Problems with Priorities," *Mimeo*, 2015.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *The American Economic Review*, 2002, *92* (5), 1644–1655.

Hossain, Tanjim and Ryo Okui, "The Binarized Scoring Rule," *The Review of Economic Studies*, 2013, *80* (3), 984–1001.

Hoxby, Caroline and Christopher Avery, "The Missing "One-offs": The hidden supply of high-achieving, low-income students," *Brookings papers on economic activity*, 2013, pp. 1–50.

__ **and Sarah Turner**, "Expanding college opportunities for high-achieving, low income students," *Stanford Institute for Economic Policy Research Discussion Paper*, 2013, (12-014).

__ **and** __ , "What High-Achieving Low-Income Students Know About College," Working Paper 20861, National Bureau of Economic Research January 2015.

Hylland, Aanund and Richard Zeckhauser, "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 1979, *87* (2), 293–314.

Kim, Jinwoo, "The value of an informed bidder in common value auctions," *Journal of Economic Theory*, 2008, *143* (1), 585–595.

Klijn, Flip, Joana Pais, and Marc Vorsatz, "Preference intensities and risk aversion in school choice: a laboratory experiment," *Experimental Economics*, 2012, pp. 1–22.

Lee, Robin S. and Michael Schwarz, "Interviewing in Two-Sided Matching Markets," *Mimeo*, 2012.

Lester, Benjamin, Nicola Persico, and Ludo Visschers, "Information Acquisition and the Exclusion of Evidence in Trials," *Journal of Law, Economics, and Organization*, December 2009, pp. 1–20.

Narita, Yusuke, "Match or Mismatch: Learning and Inertia in School Choice," Working Paper, Yale University 2016.

Nathanson, Lori, Sean Corcoran, and Christine Baker-Smith, "High School Choice in NYC: Assessing the Impact of Decision Supports of Low-Income Students," 2013. NYU Steinhardt Working Paper.

Nieuwerburgh, Stijn Van and Laura Veldkamp, "Information Acquisition and Under-Diversification," *The Review of Economic Studies*, 2010, *77* (2), 779–805.

Pais, Joana, Ágnes Pintér, and Róbert F Veszteg, "College Admissions and the Role of Information: An Experimental Study," *International Economic Review*, 2011, *52* (3), 713–737.

__ **and Ágnes Pintér**, "School choice and information: An experimental study on matching mechanisms," *Games and Economic Behavior*, 2008, *64* (1), 303 – 328.

Persico, Nicola, "Information Acquisition in Auctions," *Econometrica*, 2000, *68* (1), 135–148.

__ , "Committee Design with Endogenous Information," *The Review of Economic Studies*, 2004, *71* (1), 165–191.

Prothero, Arianna, “Parents Confront Obstacles as School Choice Expands,” *Education Week*, 2015, *34* (18), 1–16.

Rastegari, Baharak, Anne Condon, and Nicole Immorlica, “Two-sided matching with partial information,” in “In Proc. 14th ACM Conference on Electronic Commerce (EC13” 2013.

Roth, Alvin E., “The Economics of Matching: Stability and Incentives,” *Mathematics of Operations Research*, 1982, *7*, 617–628.

__ **and Marilda A. O. Sotomayor**, *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Cambridge: Cambridge University Press, 1990.

Sattin-Bajaj, Carolyn, “Two Roads Diverged: Exploring Variation in Students School Choice Experiences by Socioeconomic Status, Parental Nativity, and Ethnicity,” *Journal of School Choice*, 2014, *8* (3), 410–445.

Schotter, Andrew and Isabel Trevino, “Belief Elicitation in the Lab,” 2013. New York University Manuscript.

Shi, Xianwen, “Optimal auctions with information acquisition,” *Games and Economic Behavior*, 2012, *74* (2), 666–686.

Troyan, Peter, “Comparing school choice mechanisms by interim and ex-ante welfare,” *Games and Economic Behavior*, 2012, *75* (2), 936–947.

Appendix for

Information Acquisition and Provision in School Choice

Yan Chen Yinghua He

July 3, 2017

List of Appendices

A: Theoretical Setup and Omitted Proofs 48

B: Experimental Instructions: DA, OwnValue 64

C: Additional Analyses of Experimental Data 72

Appendix A Theoretical Setup and Omitted Proofs

A.1 Technology of Information Acquisition

Information acquisition in our model is covert. That is, i knows that others are engaging in information acquisition, but does not know what information they have acquired.

The information acquisition process consists of two stages (see Figure 1): i first pays a cost α to acquire a signal on the ordinal preference, $\omega_{1,i} \in \bar{\mathcal{P}}$. With probability $a(\alpha)$, she learns perfectly, $\omega_{1,i} = P_i$; by contrast, with probability $1-a(\alpha)$ she learns nothing, $\omega_{1,i} = P^{\phi}$. In the second stage, having learned ordinal preferences P_i, i may pay another cost, β, to learn her cardinal preferences by acquiring a signal $\omega_{2,i} \in \bar{\mathcal{V}}$, where $\bar{\mathcal{V}} \equiv [\underline{v}, \overline{v}]^{|S|} \cup V^{\phi}$. Here, with probability $b(\beta)$, she learns her cardinal preferences, $\omega_{2,i} = V_i$; by contrast, with probability $1 - b(\beta)$, she learns nothing, $\omega_{2,i} = V^{\phi}$, where V^{ϕ} denotes no cardinal preference information.

The technologies $a(\alpha)$ and $b(\beta)$ are such that $a(0) = b(0) = 0$, $\lim_{\alpha\to\infty} a(\alpha) = \lim_{\beta\to\infty} b(\beta) = 1$, $a', b' > 0$, $a'', b'' < 0$, and $a'(0) = b'(0) = +\infty$.[18] The cost of information acquisition is $c(\alpha, \beta)$, where $c(0,0) = 0$, $c_\alpha, c_\beta, c_{\alpha\beta}, c_{\alpha\alpha}, c_{\beta\beta} > 0$ for all (α, β) and $c_\alpha(0,0), c_\beta(\alpha, 0) < +\infty$ for all $\alpha > 0$. Given these restrictions, we limit our attention to $\alpha \in [0, \bar{\alpha}]$ and $\beta \in \left[0, \bar{\beta}\right]$, where $c(\bar{\alpha}, 0) = c(0, \bar{\beta}) = \overline{v}$, so that $c(\alpha, \beta)$ does not exceed the maximum possible payoff ($\overline{v}$).

After the two-stage information acquisition process, i observes signals $\omega_i = (\omega_{1,i}, \omega_{2,i}) \in \bar{\mathcal{P}} \times \bar{\mathcal{V}}$. If i pays (α, β), the distribution of signals is $H(\omega_i|\alpha, \beta)$, as outlined below:

$$\begin{array}{lll} H\left(\omega_i = \left(P^{\phi}, V^{\phi}\right)|\alpha, \beta\right) & = 1 - a(\alpha), & \text{(learning nothing)} \\ H\left(\omega_i = \left(P_i, V^{\phi}\right)|\alpha, \beta\right) & = a(\alpha)(1 - b(\beta)), & \text{(learning ordinal but not cardinal)} \\ H\left(\omega_i = (P_i, V_i)|\alpha, \beta\right) & = a(\alpha)\, b(\beta), & \text{(learning both ordinal and cardinal).} \end{array}$$

Together, these distributions imply that $H(\omega_i = (P, V)|\alpha, \beta) = 0$, if $(P, V) \notin \{(P^{\phi}, V^{\phi}), (P_i, V^{\phi}), (P_i, V_i)\}$. In other words, an agent cannot receive anything other than one of the three types of signals.

Upon observing signal ω_i, the posterior distributions of cardinal and ordinal preferences are:

$$F(V|\omega_i) = \begin{cases} F(V) & \text{if } \omega_i = \left(P^{\phi}, V^{\phi}\right), \\ F(V|P_i) & \text{if } \omega_i = \left(P_i, V^{\phi}\right), \\ 1_{V_i} & \text{if } \omega_i = (P_i, V_i); \end{cases} \quad G(P|\omega_i) = \begin{cases} G(P|F) & \text{if } \omega_i = \left(P^{\phi}, V^{\phi}\right), \\ 1_{P_i} & \text{if } \omega_i = \left(P_i, V^{\phi}\right), \\ 1_{P_i} & \text{if } \omega_i = (P_i, V_i); \end{cases}$$

where 1_{V_i} (or 1_{P_i}) is the probability distribution placing probability 1 on point V_i (or P_i).

A.2 Acquiring Information about Own Preferences

In our model, after observing the signal ω_i, students enter the school choice game under either DA or IA. Each student i submits an ROL denoted by $L_i \in \mathcal{P}$ such that sL_it if and only if s is ranked

[18] The infinite marginal productivity at zero input is consistent with, for example, the Cobb-Douglas function. When necessary, we define $0 \cdot \infty = 0$.

above t.[19] When i submits L_i and others submit L_{-i}, the payoff is represented by:

$$u\left(V_i, L_i, L_{-i}\right)=\sum_{s \in S} a_s\left(L_i, L_{-i}\right) v_{i,s} \equiv A\left(L_i, L_{-i}\right) \cdot V_i,$$

where $a_s\left(L_i, L_{-i}\right)$ is the probability that i is accepted by s, given $\left(L_i, L_{-i}\right)$, and $A\left(L_i, L_{-i}\right)$ is the vector of the probabilities determined by the mechanism. We further distinguish between two types of mechanisms: strategy-proof and non-strategy-proof. A mechanism is **strategy-proof** if:

$$u\left(V_i, P_i, L_{-i}\right) \geq u\left(V_i, L_i, L_{-i}\right), \forall L_i, L_{-i}, \text{ and } \forall V_i;$$

i.e., reporting true ordinal preferences is a dominant strategy. It is well-established that the student-proposing DA is strategy-proof (Dubins and Freedman 1981, Roth 1982), while IA is not (Abdulkadiroğlu and Sönmez 2003).

Under either mechanism, a **symmetric Bayesian Nash equilibrium** is defined by a tuple $\left(\alpha^*, \beta^*\left(P, \alpha^*\right), \sigma^*(\omega)\right)$ such that, for all i:

(i) A (possibly mixed) strategy $\sigma^*(\omega): \bar{\mathcal{P}} \times \bar{\mathcal{V}} \rightarrow \Delta(\mathcal{P})$,

$$\sigma^*(\omega) \in \arg\max_\sigma\left\{\int \int \int u\left(V, \sigma, \sigma^*\left(\omega_{-i}\right)\right) dF(V|\omega)\, dF\left(V_{-i}|\omega_{-i}\right) dH\left(\omega_{-i}|\alpha_{-i}^*, \beta_{-i}^*\right)\right\}.$$

With her own signal ω, everyone plays a best response, recognizing that others have paid $\left(\alpha_{-i}^*, \beta_{-i}^*\right)$ to acquire information. This leads to a value function given $\left(\omega, \alpha_{-i}^*, \beta_{-i}^*\right)$:

$$\Pi\left(\omega, \alpha_{-i}^*, \beta_{-i}^*\right) \equiv \max_\sigma\left\{\int \int \int u\left(V, \sigma, \sigma^*\left(\omega_{-i}\right)\right) dF(V|\omega)\, dF\left(V_{-i}|\omega_{-i}\right) dH\left(\omega_{-i}|\alpha_{-i}^*, \beta_{-i}^*\right)\right\}.$$

(ii) Acquisition of information about cardinal preferences $\beta^*\left(P, \alpha^*\right): \mathcal{P} \times[0, \bar{\alpha}] \rightarrow[0, \bar{\beta}], \forall P$,

$$\beta^*\left(P, \alpha^*\right) \in \arg\max_\beta\left\{\begin{array}{l} b(\beta) \int \Pi\left((P, V), \alpha_{-i}^*, \beta_{-i}^*\right) dF(V|P) \\ +(1-b(\beta)) \Pi\left(\left(P, V^\phi\right), \alpha_{-i}^*, \beta_{-i}^*\right)-c\left(\alpha^*, \beta\right) \end{array}\right\}.$$

Here, $\beta^*\left(P, \alpha^*\right)$ is the optimal decision given that one has learned her ordinal preference (P) after paying α^* to acquire P.

(iii) Acquisition of information about ordinal preferences $\alpha^* \in[0, \bar{\alpha}]$,

$$\alpha^* \in \arg\max_\alpha\left\{\begin{array}{l} a(\alpha) \displaystyle\int\left[\begin{array}{l} b\left(\beta^*(P, \alpha)\right) \int \Pi\left((P, V), \alpha_{-i}^*, \beta_{-i}^*\right) F(V|P) \\ +\left(1-b\left(\beta^*(P, \alpha)\right)\right) \Pi\left(\left(P, V^\phi\right), \alpha_{-i}^*, \beta_{-i}^*\right) \\ -c\left(\alpha, \beta^*(P, \alpha)\right) \end{array}\right] dG(P|F) \\ +(1-a(\alpha))\left[\Pi\left(\left(P^\phi, V^\phi\right), \alpha_{-i}^*, \beta_{-i}^*\right)-c(\alpha, 0)\right] \end{array}\right\}.$$

[19] We restrict the set of actions to the set of possible ordinal preferences, $\mathcal{P}$. In other words, students are required to rank all schools in S. However, the analysis can be extended to allow ROLs of any length.

The above expression already takes into account that the optimal β equals zero if one obtains a signal $\omega_1 = P^\phi$ in the first stage: $\beta^*(P^\phi, \alpha) = 0$ for all α.

A.3 Acquiring Information about Others' Preferences

To acquire information, student i may pay δ to acquire a signal of V_{-i}, $\omega_{i,3} \in \bar{\mathcal{V}}^{(|I|-1)}$. With probability $d(\delta)$, she learns perfectly, $\omega_{3,i} = V_{-i}$; with probability $1 - d(\delta)$, $\omega_{3,i} = V^\phi_{-i}$, she learns nothing. The distribution of signals and the posterior distribution of preferences are:

$$\begin{aligned} &K\left(\omega_{3,i} = V^\phi_{-i}|\delta\right) = 1 - d(\delta), \\ &K(\omega_{3,i} = V_{-i}|\delta) = d(\delta), \\ &K\left(\omega_{3,i} = V'_{-i}|\delta\right) = 0 \text{ if } V'_{-i} \notin \{V_{-i}, V^0_{-i}\}; \end{aligned} \qquad F(V_{-i}|\omega_{3,i}) = \begin{cases} F(V_{-i}) & \text{if } \omega_{3,i} = V^\phi_{-i}; \\ 1_{V_{-i}} & \text{if } \omega_{3,i} = V_{-i}. \end{cases}$$

The technology has the following properties: $d(0) = 0$, $\lim_{\delta \to \infty} d(\delta) = 1$, $d' > 0$, $d'' < 0$, and $d'(0) = \infty$. The cost for information acquisition is $e(\delta)$ such that $e(0) = 0$, $e', e'' > 0$ and $e'(0) < \infty$. Similarly, we restrict our attention to $\delta \in \left[0, \bar{\delta}\right]$, where $e\left(\bar{\delta}\right) = \bar{v}$.

Information acquisition is again covert. We focus on a symmetric Bayesian Nash equilibrium, $(\delta^*(V), \bar{\sigma}^*(\omega_3, V))$, where:

(i) A (possibly mixed) strategy $\bar{\sigma}^*(\omega_3, V) : \bar{\mathcal{V}}^{(|I|-1)} \times \mathcal{V} \to \Delta(\mathcal{P})$, such that

$$\bar{\sigma}^*(\omega_{3,i}, V_i) \in \arg\max_{\bar{\sigma}} \left\{ \int \int u(V_i, \bar{\sigma}, \bar{\sigma}^*(\omega_{3,-i}, V_{-i}))\, dF(V_{-i}|\omega_{3,i})\, dK\left(\omega_{3,-i}|\delta^*_{-i}\right) \right\}.$$

That is, given one's own signal $\omega_{3,i}$, each participant plays a best response, recognizing that everyone has paid δ^* to acquire information (denoted as δ^*_{-i}). We further define the value function, given $\left(\omega_{3,i}, \delta^*_{-i}\right)$ and V_i, as:

$$\Phi\left(V_i, \omega_{3,i}, \delta^*_{-i}\right) = \max_{\bar{\sigma}} \left\{ \int \int u(V_i, \bar{\sigma}, \bar{\sigma}^*(\omega_{3,-i}, V_{-i}))\, dF(V_{-i}|\omega_{3,i})\, dK\left(\omega_{3,-i}|\delta^*_{-i}\right) \right\}.$$

(ii) Acquisition of information about others' preferences $\delta^*(V) : \bar{\mathcal{V}} \to \left[0, \bar{\delta}\right]$, $\forall V$:

$$\delta^*(V_i) \in \arg\max_{\delta} \left\{ d(\delta) \int \Phi\left(V_i, V_{-i}, \delta^*_{-i}\right) dF(V_{-i}) + (1 - d(\delta))\, \Phi\left(V_i, V^\phi_{-i}, \delta^*_{-i}\right) - e(\delta) \right\}.$$

Here, $\delta^*(V_i)$ is the optimal information acquisition strategy.

The existence of such an equilibrium can be proven by similar arguments in the proof of Lemma 2, and the properties of information acquisition in equilibrium are summarized in Proposition 2.

A.4 Preamble to proofs

Before proving the propositions, let us summarize the properties of the two mechanisms. As the results can be easily verified by going through the mechanisms, we omit the formal proof.[20]

Lemma 1. *DA and IA (with single tie breaking) have the following properties:*

(i) Monotonicity: If the only difference between L_i and L_i' is that the position of s and t are swapped such that tL_is, $sL_i't$, and $\#\{s'' \in S|s'L_is''\} = \#\{s'' \in S|s'L_i's''\}$ for all $s' \in S\backslash\{s,t\}$, then:

$$a_s\left(L_i', L_{-i}\right) \geq a_s\left(L_i, L_{-i}\right), \forall L_{-i};$$

the inequality is strict when $L_j = L_i$, $\forall j \neq i$.

(ii) Guaranteed share in first choice: If school s is top ranked in L_i by i, $a_s\left(L_i, L_{-i}\right) \geq q_s/|I|$, for all L_{-i}.

(iii) Guaranteed assignment: $\sum_{s\in S} a_s\left(L_i, L_{-i}\right) = 1$ for all L_{-i}.

A.5 Proof of Lemma 2.

Lemma 2. *Under DA or IA, a symmetric Bayesian Nash equilibrium exists.*

This proof applies to either DA or IA. Note that, given any $(\alpha_{-i}, \beta_{-i})$ of other students, $\sigma^*(\omega)$ exists. This can be proven by the usual fixed point argument. Note that $\sigma^*(\omega)$ does not depend on a participant's own investments in information acquisition, but does depend on the signal received (ω).

Given ω, i's payoff function can be written as:

$$\int\int\int u_i\left(V, \sigma, \sigma^*\left(\omega_{-i}\right)\right) dF\left(V|\omega\right) dF\left(V_{-i}|\omega_{-i}\right) dH\left(\omega_{-i}|\alpha_{-i}, \beta_{-i}\right),$$

which is continuous in σ. Therefore, the value function $\Pi\left(\omega, \alpha_{-i}, \beta_{-i}\right)$ is continuous in $(\alpha_{-i}, \beta_{-i})$ by the maximum theorem.

For student i, the optimal information acquisition is solved by the first-order conditions (second-order conditions are satisfied by the assumptions on the functions $a(), b()$, and $c()$):

$$\begin{aligned} & a'\left(\alpha^*\right)\int\left[\begin{array}{l} b\left(\beta^*\left(P\right)\right)\int\Pi\left(\left(P,V\right),\alpha^*_{-i},\beta^*_{-i}\right)F\left(V|P\right) \\ +\left(1-b\left(\beta^*\left(P\right)\right)\right)\Pi\left(\left(P,V^{\phi}\right),\alpha^*_{-i},\beta^*_{-i}\right)-c\left(\alpha^*,\beta^*\left(P\right)\right)\end{array}\right]dG\left(P|F\right) \\ & -a'\left(\alpha^*\right)\left[\Pi\left(P^{\phi},\alpha^*_{-i},\beta^*_{-i}\right)-c\left(\alpha^*,0\right)\right] \\ & -a\left(\alpha^*\right)\int c_\alpha\left(\alpha^*,\beta^*\left(P\right)\right)dG\left(P|F\right)-\left(1-a\left(\alpha^*\right)\right)c_\alpha\left(\alpha^*,0\right)=0 \end{aligned}$$

$$b'\left(\beta^*\left(P\right)\right)\left[\int\Pi\left(V,\alpha^*_{-i},\beta^*_{-i}\right)dF\left(V|P\right)-\Pi\left(P,\alpha^*_{-i},\beta^*_{-i}\right)\right]-c_{\beta^*}\left(\alpha^*,\beta^*\left(P\right)\right)=0, \forall P\in\mathcal{P}.$$

[20] Similar results for IA and their respective proofs are available in He (2014).

Given the non-negative value of information and the properties of $a()$, $b()$, and $c()$, one can verify that there must exist α^* and $\beta^*(P)$ for all $P \in \mathcal{P}$ such that the first-order conditions are satisfied.

A.6 Proof of Proposition 1.

A.6.1 Proof of $\alpha^* > 0$

Given the existence of a symmetric equilibrium, let us suppose instead that $\alpha^* = 0$. This implies that $\beta^*(P) = 0$ for all $P \in \mathcal{P}$ and that the value function can be simplified as:

$$\begin{aligned}\Pi(\omega, \alpha^*, \beta^*) &= \Pi\left(\left(P^\phi, V^\phi\right), 0, \mathbf{0}\right) \\ &= \max_\sigma \left\{ \int \int u_i\left(V, \sigma, \sigma^*\left(\omega_{-i}\right)\right) dF(V)\, dF\left(V_{-i}\right) \right\}.\end{aligned}$$

Since $\alpha^* = 0$ and $\beta^* = \mathbf{0}$ (a $|\mathcal{P}|$-dimensional vector of zeros) is a best response for i, $\forall \alpha > 0$,

$$\begin{aligned}&\Pi\left(\left(P^\phi, V^\phi\right), 0, \mathbf{0}\right) \\ &\geq \left\{ a(\alpha) \int \Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right) dG(P|F) + (1 - a(\alpha))\, \Pi\left(\left(P^\phi, V^\phi\right), 0, \mathbf{0}\right) - c(\alpha, 0) \right\};\end{aligned}$$

or

$$c(\alpha, 0) \leq a(\alpha) \left[\int \Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right) dG(P|F) - \Pi\left(\left(P^\phi, V^\phi\right), 0, \mathbf{0}\right) \right], \forall \alpha > 0,$$

which can be satisfied if and only if $\Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right) = \Pi\left(\left(P^\phi, V^\phi\right), 0, \mathbf{0}\right)$ for all $P \in \mathcal{P}$, given that $\int \Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right) dG(P|F) \geq \Pi\left(\left(P^\phi, V^\phi\right), 0, \mathbf{0}\right)$ and $c'(0, 0) < a'(0) = \infty$.

In a given symmetric equilibrium σ^*, the finiteness of the strategy space implies that a finite set of lists $\left(L^{(1)}, ..., L^{(N)}\right)$ is played with positive probabilities $\left(p^{(1)}, ..., p^{(N)}\right)$ ($N \in \mathbb{N}$). Suppose that s_1 is bottom ranked in $L^{(1)}$ and s_2 is second to the bottom. Moreover, there exists an ordinal preference P^* such that $s_1 P^* s P^* s_2$ for all $s \neq s_1, s_2$. We also define $L^{(1)\prime}$ which switches only the rankings of the bottom two choices in $L^{(1)}$, s_1 and s_2.

Since $\Pi\left(\left(P^*, V^\phi\right), 0, \mathbf{0}\right) = \Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right)$, this implies that $L^{(1)}$ is also a best response to σ^* even if i has learned $P_i = P^*$. We then compare i's payoffs from submitting $L^{(1)}$ and $L^{(1)\prime}$.

By the monotonicity of the mechanism (Lemma 1), $a_{s_1}\left(L^{(1)\prime}, L_{-i}\right) \geq a_{s_1}\left(L^{(1)}, L_{-i}\right)$ and $a_{s_1}\left(L^{(1)\prime}, L_{-i}\right) \leq a_{s_1}\left(L^{(1)}, L_{-i}\right)$ for all L_{-i}. Moreover, $a_{s^*}\left(P^*, L_{-i}\right) > a_{s^*}\left(P, L_{-i}\right)$ when everyone else submits $L^{(1)}$ in L_{-i}.

Under either mechanisms, given a list, lower-ranked choices do not affect the admission probabilities at higher-ranked choices. Together with the guaranteed assignment (Lemma 1), this implies that $a_{s_1}\left(L^{(1)}, L_{-i}\right) + a_{s_2}\left(L^{(1)}, L_{-i}\right) = a_{s_1}\left(L^{(1)\prime}, L_{-i}\right) + a_{s_2}\left(L^{(1)\prime}, L_{-i}\right)$.

σ^* leads to a probability distribution over a finite number of possible profiles of others' actions (L_{-i}). With a positive probability, everyone else plays $L^{(1)}$. In this event, therefore, by submitting $L^{(1)\prime}$, i strictly increases the probability of being accepted by s_1 and decreases the probability of the least preferred school s_2, compared with that of submitting $L^{(1)}$. Furthermore, in any other possible profile of L_{-i}, the probability of being assigned to s^* is also always weakly higher when submitting $L^{(1)\prime}$. Hence, $L^{(1)}$ is not a best response to σ^* when $P_i = P^*$, and thus $\Pi\left(\left(P^*, V^\phi\right), 0, \mathbf{0}\right) \neq \Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right)$.

This contradiction proves that $\alpha^* = 0$ is not an equilibrium. Since an equilibrium always exists, it must be that $\alpha^* > 0$.

A.6.2 Proof of $\beta^*(P) = 0$ under DA

Suppose $\beta^*(P) > 0$ for some $P \in \mathcal{P}$ under DA or any strategy-proof ordinal mechanism. This implies that:

$$\begin{aligned}&\beta^*(P) \int \Pi\left((P, V), \alpha^*_{-i}, \beta^*_{-i}\right) dF(V|P) + (1-\beta^*(P))\, \Pi\left(\left(P, V^\phi\right), \alpha^*_{-i}, \beta^*_{-i}\right) - c\left(\alpha^*, \beta^*(P)\right)\\ &> \Pi\left(\left(P, V^\phi\right), \alpha^*_{-i}, \beta^*_{-i}\right),\end{aligned}$$

or,

$$\beta^*(P)\left[\int \Pi\left((P, V), \alpha^*_{-i}, \beta^*_{-i}\right) dF(V|P) - \Pi\left(\left(P, V^\phi\right), \alpha^*_{-i}, \beta^*_{-i}\right)\right] > c\left(\alpha^*, \beta^*(P)\right). \tag{2}$$

However, strategy-proofness implies that:

$$\int \Pi\left((P, V), \alpha^*_{-i}, \beta^*_{-i}\right) dF(V|P) = \Pi\left(\left(P, V^\phi\right), \alpha^*_{-i}, \beta^*_{-i}\right),$$

and thus Equation (2) cannot be satisfied. Therefore $\beta^*(P) = 0$ for all $P \in \mathcal{P}$.

A.6.3 Proof of $\beta^*(P) > 0$ for some P under IA

We construct an example where $\beta^*(P) > 0$ for some P given the distribution F under IA. For notational convenience, and in this proof only, we assume the upper bound of utility $\overline{v} = 1$ and the lower bound $\underline{v} = 0$, although we bear in mind that all schools are more preferable than an outside option. Suppose that F implies a distribution of ordinal preferences $G(P|F)$ such that for s_1 and s_2:

$$G(P|F) = \begin{cases} (1-\varepsilon) & \text{if } P = \bar{P}, \text{ s.t. } s_1 \bar{P} s_2 \bar{P} s_3 ... \bar{P} s_{|S|}; \\ \frac{\varepsilon}{|\mathcal{P}|-1} & \text{if } P \neq \bar{P}. \end{cases}$$

The distribution of cardinal preferences is:

$$F\left(V|\bar{P}\right) = \begin{cases} 1-\eta & \text{if } \left(v_{s_1}, v_{s_2}\right) = (1, \xi) \text{ and } v_s < \xi^2, \forall s \in S \backslash \{s_1, s_2\}; \\ \eta & \text{if } \left(v_{s_1}, v_{s_2}\right) = (1, 1-\xi) \text{ and } v_s < \xi^2, \forall s \in S \backslash \{s_1, s_2\}; \\ 0 & \text{otherwise.} \end{cases}$$

(ε, η, ξ) are all small positive numbers in $(0,1)$. Otherwise, there is no additional restriction on either $F(V|P)$ for $P \neq \bar{P}$ or v_s, $\forall s \in S \backslash \{s_1, s_2\}$.

Suppose that $\beta^*(P) = 0$ for all $P \in \mathcal{P}$. Section A.6.1 implies that $\alpha^* > 0$. If $\omega_i = \left(\bar{P}, V^\phi\right)$ (i.e., ordinal preferences are known but cardinal ones are not), then the expected payoff of being assigned to s_2 is:

$$E\left(v_{i,s_2}|\bar{P}\right) = (1-\eta)\,\xi + \eta\,(1-\xi)\,.$$

Furthermore, (η, ξ) are small enough such that $E\left(v_{i,s_2}|\bar{P}\right) < q_{s_1}/\,|I|$. Therefore, obtaining s_2 with certainty is less preferable than obtaining $q_{s_1}/\,|I|$ of s_1. In equilibrium, with a small enough (ε, η, ξ), it must be that:

$$\sigma^*\left(\left(\bar{P}, V^\phi\right), \alpha^*, \mathbf{0}\right) = \sigma^*\left(\left(P^\phi, V^\phi\right), \alpha^*, \mathbf{0}\right) = \bar{P}.$$

Therefore, from i's perspective, any other player, j, plays $\bar{P}$ with probability:

$$(1 - a(\alpha^*)) + a(\alpha^*)(1-\varepsilon) > 1 - \varepsilon.$$

It then suffices to show that student i has an incentive to deviate from such equilibrium strategies. Suppose that i has learned her ordinal preferences and $P_i = \bar{P}$. If furthermore she succeeds in acquiring information about V_i, there is a positive probability that $(v_{s_1}, v_{s_2}) = (1, 1-\xi)$. In this case, if she plays L_i s.t., $s_2 L_i s_1 L_i s_3 ... L_i s_{|S|}$ (or other payoff-equivalent strategies), her expected payoff is at least:

$$(1-\xi)\,(1-\varepsilon)^{(|I|-1)}\,,$$

while playing $P_i (= \bar{P})$ leads to an expected payoff less than:

$$(1-\varepsilon)^{(|I|-1)}\left[\frac{q_{s_1}}{|I|} + \left(1 - \frac{q_{s_1}}{|I|}\right)\xi\right] + \left(1 - (1-\varepsilon)^{(|I|-1)}\right).$$

This upper bound is obtained under the assumption that one is always assigned to s_1 if not everyone submits $\bar{P}$. When (ε, ξ) are close to zero, it is strictly profitable to submit L_i instead of $\bar{P}$:

$$\int \Pi\left(\left(\bar{P}, V\right), \alpha^*_{-i}, \mathbf{0}\right) dF\left(V|\bar{P}\right) > \Pi\left(\left(\bar{P}, V^\phi\right), \alpha^*_{-i}, \mathbf{0}\right),$$

because in other realizations of V, i cannot do worse than submitting $\bar{P}$. The marginal payoff of increasing $\beta\left(\bar{P}\right)$ from zero by Δ is then:

$$\Delta\left(b'(0)\left[\int \Pi\left(\left(\bar{P}, V\right), \alpha^*_{-i}, \mathbf{0}\right) dF\left(V|\bar{P}\right) - \Pi\left(\left(\bar{P}, V^\phi\right), \alpha^*_{-i}, \mathbf{0}\right)\right] - c_\beta\left(\alpha^*, 0\right)\right),$$

which is strictly positive given $c_\beta(\alpha^*, 0) < b'(0) = +\infty$. This proves that, under IA, $\beta^*(P) > 0$ for some $P \in \mathcal{P}$ given F.

A.7 Proof of Proposition 2.

For the first part, by the definition of strategy-proofness, information about others' types does not change one's best response. Therefore, $\delta^*(V) = 0$ for all V under any strategy-proof mechanism.

To prove the second part, we construct an example of $F(V)$ to show $\delta^*(V) > 0$ for some V under IA. For notational convenience and in this proof only, we assume the upper bound of utility $\overline{v} = 1$ and the lower bound $\underline{v} = 0$, although we bear in mind that all schools are more preferable than an outside option. The distribution of cardinal preferences is:

$$F(V) = \begin{cases} \frac{1}{2} - \varepsilon & \text{if } V = V^{(1)} \text{ s.t. } (v_{s_1}, v_{s_2}) = (1, 0), v_s \in (0, \xi) \ \forall s \notin \{s_1, s_2\}; \\ \frac{1}{2} - \varepsilon & \text{if } V = V^{(2)} \text{ s.t. } (v_{s_1}, v_{s_2}) = (0, 1), v_s \in (0, \xi) \ \forall s \notin \{s_1, s_2\}; \\ \varepsilon & \text{if } V = V^{(3)} \text{ s.t. } (v_{s_1}, v_{s_2}) = (1, 1-\eta), v_s \in (0, \xi) \ \forall s \notin \{s_1, s_2\}; \end{cases}$$

where (ε, ξ, η) are small positive values. In addition, $F\left(V \in [0,1]^{|S|} \setminus \{V^{(1)}, V^{(2)}, V^{(3)}\}\right) = \varepsilon$.

Suppose that, for student i, $V_i = V^{(3)}$. If $\delta^*(V) = 0$ for all V, the best response for i in equilibrium is to top rank either s_1 or s_2.

Given $F(V)$, there is a positive probability, $\left(\frac{1}{2} - \varepsilon\right)^{|I|-1}$, that every other student has $V^{(1)}$ and top ranks s_1. In this case, the payoff for i top-ranking s_1 is less than $q_{s_1}/|I| + \xi$, while top-ranking s_2 leads to $(1-\eta)$.

There is also a positive probability, $\left(\frac{1}{2} - \varepsilon\right)^{|I|-1}$, that every other student has $V^{(2)}$ and top ranks s_2. In this case, the payoff for i top-ranking s_1 is 1, while the payoff when top-ranking s_2 is at most $(1-\eta) q_{s_2}/|I| + \xi$.

Since $\int \Phi\left(V, V_{-i}, \delta^*_{-i}\right) dF(V_{-i}) \geq \Phi\left(V, V^{\phi}_{-i}, \delta^*_{-i}\right)$ and the above shows they are different for some realization of (V_i, V_{-i}), then:

$$\int \Phi\left(V, V_{-i}, \delta^*_{-i}\right) dF(V_{-i}) - \Phi\left(V, V^{\phi}_{-i}, \delta^*_{-i}\right) > 0.$$

The marginal payoff of acquiring information (increasing $\delta(V_i)$ from zero to Δ) is:

$$\Delta\left(d'(0)\left[\int \Phi\left(V, V_{-i}, \delta^*_{-i}\right) dF(V_{-i}) - \Phi\left(V, V^{\phi}_{-i}, \delta^*_{-i}\right)\right] - e'(0)\right),$$

which is positive for a small (ε, ξ, η) because $e'(0) < d'(0) = \infty$. This proves that $\delta^*(V) > 0$ for some V with a positive measure given F.

A.8 Proof of Proposition 3.

Under UI, the only information i has is that her preferences follow the distribution $F(V)$. Denote W^E_i as the expected (possibly weak) ordinal preferences of i such that $sW^E_i t$ if and only if $\int v_{i,s} dF_{v_s}(v_{i,s}) \geq \int v_{i,t} dF_{v_t}(v_{i,t})$. Given W^E_i, $\left(P^{E,1}_i, ..., P^{E,M}_i\right) \in \mathcal{P}$ are all the strict ordinal preferences that can be generated by randomly breaking ties in W^E_i if any exists. Therefore,

$M \geq 1$.

When others play L_{-i}, the expected payoff of i playing L_i is:

$$\int \sum_{s \in S} a_s \left(L_i, L_{-i}\right) v_{i,s} dF\left(V\right) = \sum_{s \in S} a_s \left(L_i, L_{-i}\right) \int v_{i,s} dF_{v_s}\left(v_{i,s}\right).$$

Since DA with single tie breaking is essentially a random serial dictatorship, it is therefore a dominant strategy that i submits any $P_i^{E,m}$ $m \in \{1, ..., M\}$. Moreover, a strategy that is not in $\left(P_i^{E,1}, ..., P_i^{E,M}\right)$ can never be played in any equilibrium, because there is a positive-measure set of realizations of the lottery that such a strategy leads to a strictly positive loss.

We claim that in equilibrium for any L_{-i}^* such that $L_j^* \in \left(P_i^{E,1}, ..., P_i^{E,M}\right)$, $j \neq i$, the payoff to i is:

$$\sum_{s \in S} a_s \left(P_i^{E,m}, L_{-i}^*\right) \int v_{i,s} dF_{v_s}\left(v_{i,s}\right) = \sum_{s \in S} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}\right), \forall m. \tag{3}$$

Note that, for any L_{-i}^*, $\sum_{s \in S} a_s \left(P_i^{E,m}, L_{-i}\right) \int v_{i,s} dF_{v_s}\left(v_{i,s}\right)$ does not vary across m given that any $P_i^{E,m}$ is a dominant strategy.

Since everyone has the same expected utility of being assigned to every school, the maximum utilitarian sum of expected utility is:

$$\sum_{s \in S} q_s \int v_{i,s} dF_{v_s}\left(v_{i,s}\right). \tag{4}$$

If Equation (3) is not satisfied and there exists i such that for some $\hat{L}_{-i}^*$:

$$\sum_{s \in S} a_s \left(P_i^{E,m}, \hat{L}_{-i}^*\right) \int v_{i,s} dF_{v_s}\left(v_{i,s}\right) > \sum_{s \in S} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}\right), \forall m. \tag{5}$$

The maximum utilitarian social welfare in (4) implies that there exists $j \in I \backslash \{i\}$ and $m \in \{1, ..., M\}$ such that:

$$\sum_{s \in S} a_s \left(P_j^{E,m}, \hat{L}_{-j}^*\right) \int v_{j,s} dF_{v_s}\left(v_{j,s}\right) < \sum_{s \in S} \frac{q_s}{|I|} \int v_{j,s} dF_{v_s}\left(v_{j,s}\right), \tag{6}$$

where $P_j^{E,m}$ is j's strategy in $\hat{L}_{-i}^*$ and $P_j^{E,m} = P_i^{E,m}$. We can always find such $P_i^{E,m}$ and $P_j^{E,m}$ because condition (5) is satisfied for all m. However, the even lottery implies that:

$$a_s \left(P_i^{E,m}, \left(L_{-(i,j)}^*, P_j^{E,m}\right)\right) = a_s \left(P_j^{E,m}, \left(L_{-(i,j)}^*, P_i^{E,m}\right)\right) \ \forall s \text{ if } P_i^{E,m} = P_j^{E,m},$$

and thus:

$$\sum_{s\in S} a_s \left(P_j^{E,m}, \left(L^*_{-(i,j)}, P_i^{E,m}\right)\right) \int v_{j,s} dF_{v_s}\left(v_{j,s}\right) = \sum_{s\in S} a_s \left(P_i^{E,m}, \left(L^*_{-(i,j)}, P_j^{E,m}\right)\right) \int v_{i,s} dF_{v_s}\left(v_{i,s}\right),$$

which contradicts the inequalities (5) and (6). This proves (3) is always satisfied.

Under OI, CI, or PI, the unique equilibrium is for everyone to report her true ordinal preferences, and thus the expected payoff (*ex ante*) is:

$$\begin{aligned}
&\int\int\sum_{s\in S} a_s\left(P, L_{-i}\left(P\right)\right) v_{i,s} dF\left(V|P\right) dG\left(P|F\right)\\
&=\int\int\sum_{s\in S}\frac{q_s}{|I|} v_{i,s} dF_{v_s}\left(v_{i,s}|P\right) dG\left(P|F\right)\\
&=\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}\left(v_{i,s}\right),
\end{aligned}$$

where $L_{-i}(P)$ is such that that $L_j = P, \forall j \in I\backslash\{i\}$.

A.9 Proof of Proposition 4.

A.9.1 Welfare under UI and OI

We first show UI = OI in any symmetric equilibrium in terms of *ex ante* student welfare.

Under UI, the game can be transformed into one similar to that under PI but where everyone has the same cardinal preferences that are represented in terms of the expected utilities $\left[\int v_{i,s} dF_{v_s}\left(v_{i,s}\right)\right]_{s\in S}$. In a symmetric equilibrium, everyone must thus play exactly the same strategy, either pure or mixed, which further implies that everyone is assigned to each school with the same probability and has the same *ex ante* welfare:

$$\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}\left(v_{i,s}\right).$$

Under OI, everyone knows that everyone has the same ordinal preferences P. The game again can be considered as one under PI where everyone has the same cardinal preferences, $\left[\int v_{i,s} dF_{v_s}\left(v_{i,s}|P\right)\right]_{s\in S}$. Similar to the argument above, the payoff conditional on P is:

$$\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}\left(v_{i,s}|P\right),$$

which leads to an *ex ante* payoff:

$$\int\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}\left(v_{i,s}|P\right) dG\left(P|F\right) = \sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}\left(v_{i,s}\right).$$

A.9.2 Proof of CI $\geq$ UI = OI under IA

We then show CI $\geq$ OI = UI.

Under CI, a participant's cardinal preferences V_i are her private information, although her ordinal preferences P, which is common across i, are common knowledge. Suppose that $\sigma^{BN}(V) : [0,1]^{|S|} \to \Delta(\mathcal{P})$ is a symmetric Bayesian Nash equilibrium. We show that:

$$\begin{aligned}
&\int\int\left(\int A\left(\sigma^{BN}(V_i), \sigma^{BN}(V_{-i})\right) dF(V_{-i}|P)\cdot V_i\right) dF(V_i|P)\, dG(P|F) \\
&\geq \sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}(v_{i,s}).
\end{aligned}$$

The following is based on the same idea as that in the proof of Proposition 2 in (Troyan 2012). Note that $\int a_s\left(\sigma^{BN}(V_i), \sigma^{BN}(V_{-i})\right) dF(V_{-i}|P)$ is i's probability of being assigned to s in equilibrium when the realization of cardinal preferences is V_i. Furthermore, the *ex ante* assignment probability, i.e., the probability before the realization of P and V_i, is

$$\int\int\int a_s\left(\sigma^{BN}(V_i), \sigma^{BN}(V_{-i})\right) dF(V_{-i}|P)\, dF(V_i|P)\, dG(P|F),$$

which must be the same across students by symmetry. Therefore, we must have:

$$|I|\int\int\int a_s\left(\sigma^{BN}(V_i), \sigma^{BN}(V_{-i})\right) dF(V_{-i}|P)\, dF(V_i|P)\, dG(P|F) = q_s, \forall s\in S, \tag{7}$$

as in equilibrium all seats at all $s \in S$ must be assigned.

Suppose i plays an alternative strategy σ_i such that $\sigma_i = \int\int \sigma^{BN}(V_i)\, dF(V_i|P)\, dG(P|F) = \int \sigma^{BN}(V_i)\, dF(V_i)$. That is, i plays the "average" strategy of the equilibrium strategy regardless of her preferences. Her payoff given any realization of P is:

$$\begin{aligned}
&\int\left(\int A\left(\sigma_i, \sigma^{BN}(V_{-i})\right) dF(V_{-i}|P)\cdot V_i\right) dF(V_i|P) \\
&= \int\left(\int\left(\int\int A\left(\sigma^{BN}(V_i), \sigma^{BN}(V_{-i})\right) dF(V_i|P)\, dG(P|F)\right) dF(V_{-i}|P)\cdot V_i\right) dF(V_i|P) \\
&= \int\left(\textstyle\sum_{s\in S}\left(\int\int\int a_s\left(\sigma^{BN}(V_i), \sigma^{BN}(V_{-i})\right) dF(V_i|P)\, dG(P|F)\, dF(V_{-i}|P)\right) v_{i,s}\right) dF(V_i|P) \\
&= \int\left(\sum_{s\in S}\frac{q_s}{|I|} v_{i,s}\right) dF(V_i|P).
\end{aligned}$$

The last equation is due to (7). Since σ_i may not be optimal for i upon observing her preferences

V_i, we thus have *ex ante* welfare:

$$\begin{aligned}&\int\int\left(\int A\left(\sigma^{BN}\left(V_i\right),\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)\cdot V_i\right)dF\left(V_i|P\right)dG\left(P|F\right)\\ &\geq\int\int\left(\int A\left(\sigma_i,\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)\cdot V_i\right)dF\left(V_i|P\right)dG\left(P|F\right)\\ &=\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s}dF_{v_s}\left(v_{i,s}\right),\end{aligned}$$

which proves CI $\geq$ OI = UI in terms of Pareto dominance of *ex ante* student welfare.

A.9.3 Proof of PI $\geq$ OI = UI under IA

Under PI, everyone's cardinal preferences V_i are common knowledge. Given a symmetric equilibrium, by the same argument as above, PI must Pareto dominate OI and UI.

Suppose that $\sigma^{NE}\left(V_i,V_{-i}\right):[0,1]^{|S|\times|I|}\rightarrow\Delta\left(\mathcal{P}\right)$ is a symmetric Nash equilibrium. We show that:

$$\begin{aligned}&\int\int\int\left(A\left(\sigma^{NE}\left(V_i,V_{-i}\right),\left[\sigma^{NE}\left(V_j,V_{-j}\right)\right]_{j\in I\backslash\{i\}}\right)\cdot V_i\right)dF\left(V_{-i}|P\right)dF\left(V_i|P\right)dG\left(P|F\right)\\ &\geq\sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s}dF_{v_s}\left(v_{i,s}\right).\end{aligned}$$

Note that $a_s\left(\sigma^{NE}\left(V_i,V_{-i}\right),\left[\sigma^{NE}\left(V_j,V_{-j}\right)\right]_{j\in I\backslash\{i\}}\right)$ is i's probability of being assigned to s in equilibrium when the realization of cardinal preferences is (V_i,V_{-i}). Furthermore, the *ex ante* assignment probability, i.e., the probability before the realization of P and (V_i,V_{-i}), is

$$\int\int\int a_s\left(\sigma^{NE}\left(V_i,V_{-i}\right),\left[\sigma^{NE}\left(V_j,V_{-j}\right)\right]_{j\in I\backslash\{i\}}\right)dF\left(V_{-i}|P\right)dF\left(V_i|P\right)dG\left(P|F\right),$$

which must be the same across students by symmetry. Therefore, we must have, $\forall s\in S$:

$$|I|\int\int\int a_s\left(\sigma^{NE}\left(V_i,V_{-i}\right),\left[\sigma^{NE}\left(V_j,V_{-j}\right)\right]_{j\in I\backslash\{i\}}\right)dF\left(V_{-i}|P\right)dF\left(V_i|P\right)dG\left(P|F\right)=q_s,\tag{8}$$

as in equilibrium all seats at all $s\in S$ must be assigned.

Suppose i plays an alternative strategy σ_i such that

$$\sigma_i=\int\int\int\sigma^{NE}\left(V_i,V_{-i}\right)dF\left(V_{-i}|P\right)dF\left(V_i|P\right)dG\left(P|F\right).$$

That is, i plays the "average" strategy of the equilibrium strategy regardless of her and others' preferences. Her payoff given a realization of (V_i, V_{-i}) is:

$$
\begin{aligned}
&A\left(\sigma_i, \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right)\cdot V_i \\
&= \left(\int\int\int A\left(\sigma^{NE}(V_i, V_{-i}), \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right) dF(V_{-i}|P)\, dF(V_i|P)\, dG(P|F)\right)\cdot V_i \\
&= \sum\nolimits_{s\in S}\left(\int\int\int a_s\left(\sigma^{NE}(V_i, V_{-i}), \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right) dF(V_i|P)\, dG(P|F)\, dF(V_{-i}|P)\right) v_{i,s} \\
&= \sum_{s\in S}\frac{q_s}{|I|}v_{i,s}.
\end{aligned}
$$

The last equation is due to (8). Therefore, her payoff given a realization of P is:

$$
\begin{aligned}
&\int\int\left(A\left(\sigma_i, \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF(V_{-i}|P)\, dF(V_i|P) \\
&= \int\left(\sum_{s\in S}\frac{q_s}{|I|}v_{i,s}\right) dF(V_i|P).
\end{aligned}
$$

Since σ_i may not be optimal for i upon observing her and others' preferences (V_i, V_{-i}), we thus have:

$$
\begin{aligned}
&\int\int\int\left(A\left(\sigma^{NE}(V_i, V_{-i}), \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF(V_{-i}|P)\, dF(V_i|P)\, dG(P|F) \\
&\geq \int\int\int\left(A\left(\sigma_i, \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF(V_{-i}|P)\, dF(V_i|P)\, dG(P|F) \\
&= \sum_{s\in S}\frac{q_s}{|I|}\int v_{i,s} dF_{v_s}(v_{i,s}),
\end{aligned}
$$

which thus proves that $PI > OI = UI$ in terms of Pareto dominance.

We use two examples to show part (iii) in Proposition 4**:** Section A.9.4 shows that PI can dominate CI in symmetric equilibrium while the example in Section A.9.5 shows the opposite.

A.9.4 Example: PI dominates CI in symmetric equilibrium under IA

There are three schools (a, b, c) and three students whose cardinal preferences are i.i.d. draws from the following distribution:

$$
\begin{aligned}
\Pr\left((v_a, v_b, v_c) = (1, 0.1, 0)\right) = 1/2 \\
\Pr\left((v_a, v_b, v_c) = (1, 0.5, 0)\right) = 1/2
\end{aligned}
$$

Each school has one seat. For any realization of preference profile, we can find a symmetric Nash equilibrium as in Table A.1.

Table A.1: Symmetric Nash Equilibrium for Each Realization of the Game under PI

Realization of Preferences	Probability Realized	Strategy given realized type (1, 0.1, 0)	(1, 0.5, 0)	Payoff given realized type (1, 0.1, 0)	(1, 0.5, 0)
(1, 0.1, 0) (1, 0.1, 0) (1, 0.1, 0)	1/8	(a, b, c)	-	11/30	-
(1, 0.5, 0) (1, 0.1, 0) (1, 0.1, 0)	1/4	(a, b, c)	(b, a, c)	1/2	1/2
(1, 0.5, 0) (1, 0.5, 0) (1, 0.1, 0)	1/4	(a, b, c)	(a, b, c)	11/30	1/2
(1, 0.5, 0) (1, 0.5, 0) (1, 0.5, 0)	1/8	-	(a, b, c)	-	1/2

The above symmetric equilibrium leads to an *ex ante* student welfare:

$$\frac{1}{2}\left(\frac{1}{4}\frac{11}{30}+\frac{1}{2}\frac{1}{2}+\frac{1}{4}\frac{11}{30}\right)+\frac{1}{2}\left(\frac{1}{4}\frac{1}{2}+\frac{1}{2}\frac{1}{2}+\frac{1}{4}\frac{1}{2}\right)=\frac{14}{30}.$$

When everyone's preference is private information, we can verify that the unique symmetric Bayesian Nash equilibrium is:

$$\sigma^{BN}\left((1,0.1,0)\right)=\sigma^{BN}\left((1,0.5,0)\right)=(a,b,c)\,.$$

That is, everyone submits her true preference ranking. This leads to an *ex ante* welfare of:

$$\frac{1}{2}\frac{11}{30}+\frac{1}{2}\frac{15}{30}=\frac{13}{30}$$

which is lower than the above symmetric equilibrium under PI.

Also note that always playing (a,b,c) is also a symmetric Nash equilibrium under PI in all realizations of preference profile, which leads to the same ex ante student welfare as σ^{BN}.

A.9.5 Example: PI is dominated by CI in symmetric equilibrium under IA

There are three schools (a,b,c) and three students whose cardinal preferences are i.i.d. draws from the following distribution:

$$\begin{array}{l}\Pr\left((v_a,v_b,v_c)=(1,0.1,0)\right)=3/4\\ \Pr\left((v_a,v_b,v_c)=(1,0.9,0)\right)=1/4\end{array}.$$

Each school has one seat. For any realization of preference profile, we can find a symmetric Nash equilibrium as in Table A.2. The *ex ante* welfare under PI with the above symmetric equilibrium profile is:

$$\frac{3}{4}\left(\frac{9}{16}\frac{11}{30}+\frac{6}{16}\frac{1}{2}+\frac{1}{16}\frac{3073}{3610}\right)+\frac{1}{4}\left(\frac{1}{16}\frac{19}{30}+\frac{6}{16}\frac{99}{190}+\frac{9}{16}\frac{9}{10}\right)=\frac{22\,549}{43\,320}\approx 0.52052.$$

Table A.2: Symmetric Nash Equilibrium for Each Realization of the Game under PI

Realization of Preference	Probability Realized	Strategy given realized type (1, 0.1, 0)	Strategy given realized type (1, 0.9, 0)	Payoff given realized type (1, 0.1, 0)	Payoff given realized type (1, 0.9, 0)
(1, 0.1, 0) (1, 0.1, 0) (1, 0.1, 0)	27/64	(a, b, c)	-	11/30	-
(1, 0.9, 0) (1, 0.1, 0) (1, 0.1, 0)	27/64	(a, b, c)	(b, a, c)	1/2	9/10
(1, 0.9, 0) (1, 0.9, 0) (1, 0.1, 0)	9/64	(a, b, c)	(a, b, c) w/ prob 3/19 (b, a, c) w/ prob 16/19	3073/3610	99/190
(1, 0.9, 0) (1, 0.9, 0) (1, 0.9, 0)	1/64	-	(a, b, c) w/ prob 11/19 (b, a, c) w/ prob 8/19	-	19/30

Under CI, i.e., when a participant's own preferences are private information and the distribution of preferences is common knowledge, there is a symmetric Bayesian Nash equilibrium:

$$\sigma^{BN}\left((1,0.9,0)\right)=(b,a,c)\,;\sigma^{BN}\left((1,0.1,0)\right)=(a,b,c)\,.$$

For a type-$(1,0.1,0)$ student, it is a dominant strategy to play (a,b,c). Conditional on her type, her equilibrium payoff is:

$$\frac{9}{16}\left(\frac{1}{3}\left(1+\frac{1}{10}+0\right)\right)+\frac{6}{16}\frac{1}{2}+\frac{1}{16}=\frac{219}{480}.$$

For a type-$(1,0.9,0)$ student, given others follow σ^{BN}, playing (b,a,c) results in a payoff of:

$$\frac{9}{16}\frac{9}{10}+\frac{6}{16}\left(\frac{1}{2}\left(\frac{9}{10}+0\right)\right)+\frac{1}{16}\left(\frac{1}{3}\left(\frac{9}{10}+1+0\right)\right)=\frac{343}{480}.$$

If a type-$(1, 0.9, 0)$ student deviates to (a, b, c), she obtains:

$$\frac{9}{16}\left(\frac{1}{3}\left(\frac{9}{10}+1+0\right)\right)+\frac{6}{16}\left(\frac{1}{2}(1+0)\right)+\frac{1}{16}(1)=\frac{291}{480}.$$

It is therefore not a profitable deviation. Furthermore, she has no incentive to deviate to other rankings such as (c, a, b) or (c, b, a).

The *ex ante* payoff to every student in this equilibrium under CI is:

$$\frac{219}{480}\frac{3}{4}+\frac{343}{480}\frac{1}{4}=\frac{25}{48}\approx 0.52083,$$

which is higher than that under PI.

In this example, the reason that PI leads to lower welfare is because it sometimes leads to type-$(1, 0.9, 0)$ students to play mixed strategies in equilibrium. Therefore, sometimes school b is assigned to a type-(1,0.1,0) student, which never happens under CI in symmetric Bayesian Nash equilibrium.

Appendix B Experimental Instructions: DA, OwnValue

This is an experiment in the economics of decision making. In this experiment, we simulate a procedure to allocate students to schools. The procedure, payment rules, and student allocation method are described below. The amount of money you earn will depend upon the decisions you make and on the decisions other people make. Do not communicate with each other during the experiment. If you have questions at any point during the experiment, raise your hand and the experimenter will help you. At the end of the instructions, you will be asked to provide answers to a series of review questions. Once everyone has finished the review questions, we will go through the answers together.

Overview:

- There are 12 participants in this experiment.
- The experiment consists of three parts:
  - There will be 20 rounds of school ranking decisions and student allocations.
  - At the end of the 20 rounds, there will be a lottery experiment.
  - Finally, there will be a survey.
- At the beginning of each round, you will be randomly matched into four groups. Each group consists of three participants. Your payoff in a given round depends on your decisions and the decisions of the other two participants in your group.
- In this experiment, three schools are available for each group, school a, school b and school c. Each school has one slot. Each school slot will be allocated to one participant.
- **Your payoff** amount for each allocation depends on the school you are assigned to. These amounts reflect the quality and fit of the school for you.
  - If you are assigned to school a, your payoff is 100 points.
  - If you are assigned to school b, your payoff is either 110 points or 10 points, depending on a random draw. Specifically,
    - with 20% chance, your payoff is 110 points;
    - with 80% chance, your payoff is 10 points.
  - If you are assigned to school c, your payoff is 0.
- **Your total payoff** equals the sum of your payoffs in all 20 rounds, plus your payoff from the lottery experiment. Your earnings are given in points. At the end of the experiment you will be paid based on the exchange rate,

$1 = 100 points.

In addition, you will be paid $5 for participation, and up to $2.00 for answering the Review Questions correctly. Everyone will be paid in private and you are under no obligation to tell others how much you earn.

Are there any questions?

Procedure for the first 10 rounds:

- Every round, you will be asked to rank the schools twice:
  - Ranking without information (on your school b value): you will rank the schools without knowing the realization of your value for school b;
  - Ranking with information (on your school b value): the computer will first inform you of your school b value, and then ask you to rank the schools.
- **Ranking without information** consists of the following steps:
  - The computer will randomly draw the value of school b for each participant independently, but will not inform anyone of his or her value.
  - Without knowing the realization of school b value, every participant submits his or her school ranking.
  - The computer will then generate a lottery, and allocate the schools according to the Allocation Method described below.
  - The allocation results will not be revealed till the end of the round.
- **Ranking with information** consists of the following steps:
  - The computer will randomly draw the value of school b for each participant independently, and inform everyone of his or her school b value.
  - After knowing his or her school b value, every participant submits his or her school ranking.
  - After receiving the rankings, the computer will generate a lottery, and allocate the schools according to the Allocation Method described below.
- **Feedback**: At the end of each round, each participant receives the following feedback for each of the two rankings: your and your matches' school b values, rankings, lottery numbers, assigned schools, and earnings.
- At the beginning of each round, the computer randomly decides the order of the two rankings:

  - with 50% chance, you will rank the schools without information first;
  - with 50% chance, you will rank the schools with information first;

- The process repeats for 10 rounds.

Allocation Method

- **The lottery**: the priority of each student is determined by a lottery generated before each allocation. Every student is equally likely to be the first, second or third in the lottery.

- **The allocation of schools is described by the following method**:

  - An application to the first ranked school is sent for each participant.
  - Throughout the allocation process, a school can hold no more applications than its capacity.

    If a school receives more applications than its capacity, then it temporarily retains the student with the highest priority and rejects the remaining students.
  - Whenever an applicant is rejected at a school, his or her application is sent to the next choice.
  - Whenever a school receives new applications, these applications are considered together with the retained application for that school. Among the retained and new applications, the one with the highest priority is temporarily on hold.
  - The allocation is finalized when no more applications can be rejected.

    Each participant is assigned to the school that holds his or her application at the end of the process.

  Note that the allocation is temporary in each step until the last step.

Are there any questions?

An Example:

We will go through a simple example to illustrate how the allocation method works. This example has the same number of students and schools as the actual decisions you will make. You will be asked to work out the allocation of this example for Review Question 1.

Students and Schools: In this example, there are three students, 1-3, and three schools, A, B, and C.

Student ID Number: $1, 2, 3$	Schools: A, B, C

Slots: There is one slot at each school.

School	Slot
A	☐
B	☐
C	☐

Lottery: Suppose the lottery produces the following order:

$$\boxed{1 - 2 - 3}$$

Submitted School Rankings: The students submit the following school rankings:

	1st Choice	2nd Choice	3rd Choice
Student 1	A	B	C
Student 2	A	B	C
Student 3	B	A	C

The allocation method consists of the following steps: Please use this sheet to work out the allocation and enter it into the computer for Review Question 1.

Step 1 (temporary): Each student applies to his/her first choice. If a school receives more applications than its capacity, then it temporarily holds the application with the highest priority and rejects the remaining students.

Applicants		School		Hold	Reject
1, 2	$\longrightarrow$	A	$\longrightarrow$	☐	
3	$\longrightarrow$	B	$\longrightarrow$	☐	
	$\longrightarrow$	C	$\longrightarrow$	☐	

Step 2 (temporary): Each student rejected in Step 1 applies to his/her next choice. When a school receives new applications, these applications are considered together with the application on hold for that school. Among the new applications and those on hold, the one with the highest priority is on hold, while the rest are rejected.

Applicants		School		Hold	Reject
	$\longrightarrow$	A	$\longrightarrow$	$\square$	
	$\longrightarrow$	B	$\longrightarrow$	$\square$	
	$\longrightarrow$	C	$\longrightarrow$	$\square$	

Step 3 (temporary): Each student rejected in Step 2 applies to his/her next choice. Again, new applications are considered together with the application on hold for each school. Among the new applications and those on hold, the one with the highest priority is on hold, while the rest are rejected.

Applicants		School		Hold	Reject
	$\longrightarrow$	A	$\longrightarrow$	$\square$	
	$\longrightarrow$	B	$\longrightarrow$	$\square$	
	$\longrightarrow$	C	$\longrightarrow$	$\square$	

Step 4 (final): Each student rejected in Step 3 applies to his/her next choice. No one is rejected at this step. All students on hold are accepted.

Applicants		School		Accept	Reject
	$\longrightarrow$	A	$\longrightarrow$	$\square$	
	$\longrightarrow$	B	$\longrightarrow$	$\square$	
	$\longrightarrow$	C	$\longrightarrow$	$\square$	

The allocation ends at Step 4.

- Please enter your answer into the computer for Review Question 1.
- Afterwards, you will be asked to answer other review questions. When everyone is finished with them, we will go through the answers together.
- Feel free to refer to the experimental instructions before you answer any question. Each correct answer is worth 20 cents, and will be added to your total earnings.

Review Questions 2 - 7

2. How many participants are there in your group each round?
3. True or false: You will be matched with the same two participants each round.
4. Everyone has an equal chance of being the first, second or third in a lottery.

5. True or false: The lottery is fixed for the entire 20 rounds.
6. True or false: If you are not rejected at a step, then you are accepted into that school.
7. True or false: The allocation is final at the end of each step.

We are now ready to start the first 10 rounds. Feel free to earn as much as you can. Are there any questions?

Procedure for the second 10 rounds:

- Every round, you will again be asked to rank the schools twice.
- **Ranking without information** is identical to that in the first ten rounds.
- **Ranking with information**, however, will be different. We will elicit your willingness-to-pay for your school b value before you submit your ranking in each round. That is, the information about your school b value is no longer free. Specifically,
  - The computer will randomly draw the value of school b for each participant independently.
  - You will be asked your **willingness to pay** for this information. You can enter a number in the interval of $[0, 15]$ points, inclusive, to indicate your willingness to pay.
  - After everyone submits their willingness to pay, the computer will randomly draw a number for each participant independently. The number will be between 0 and 15, inclusive, with an increment of 0.01, with each number being chosen with equal probability.
    * If your willingness to pay is greater than the random number, you will pay the random number as your price to obtain your school b value. The computer will reveal your school b value and charge you a price which equals the random number.
    * If your willingness to pay is below the random number, the computer will not reveal your school b value and you will not be charged a price.

    It can be demonstrated that, given the procedures we are using, it is best for you, in terms of maximizing your earnings, to report your willingness to pay for your school b value truthfully since doing anything else would reduce your welfare. So it pays to report your willingness to pay truthfully.
  - You will also be asked to **guess** the average willingness to pay of the other two participants in your group, again, in the interval of $[0, 15]$ points, inclusive.
  - You will be rewarded for guessing the average of your matches' willingness to pay correctly. Your payoff from guessing is determined by the squared error between your guess and the actual average, i.e., (your guess - the actual average)^2. Specifically, the computer will randomly choose a number between 0 and 49, with each number being

chosen with equal probability. You will earn 5 points, if your squared error is below the random number and zero otherwise. Therefore, you should try to guess as accurately as possible.

  - Regardless of whether you obtain your school b value, the computer will reveal the number of participant(s) in your group who have obtained their school b value(s).
  - Every participant submits his or her school ranking.
  - After everyone submits their rankings, the computer will generate a lottery, and allocate the schools according to the same Allocation Method used in the first ten rounds.

- **Feedback**: At the end of each round, each participant receives the same feedback for each of the two rankings as in the first ten rounds.

  In addition, for ranking with information, the computer will also tell you: your and your matches' willingness to pay, the actual prices paid, the random numbers, whether each participant in your group knows their school b values, the guesses, and guess earnings.

- The process repeats for 10 rounds.

Are there any questions? You can now proceed to answer review questions 8-10 on your computer. Recall each correct answer is worth 20 cents, and will be added to your total earnings. Again, feel free to refer to the instructions before you answer any question.

Review Questions 8 - 10

8. Suppose you submitted 1.12 as your willingness to pay to obtain your school b value, and the random number is 5.48. Do you get to know your school b value? What price do you pay?
9. Suppose you submitted 10.33 as your willingness to pay to obtain your school b value, and the random number is 8.37. Do you get to know your school b value? What price do you pay?
10. Suppose your guess for the average willingness to pay of the other two participants is 7, and the actual average is 10. The computer draws a random number, 14. What is your earning from your guess?

Lottery Experiment

Procedure

- **Making Ten Decisions**: On your screen, you will see a table with 10 decisions in 10 separate rows, and you choose by clicking on the buttons on the right, option A or option B, for each of the 10 rows. You may make these choices in any order and change them as much as you wish until you press the Submit button at the bottom.

- The money prizes are determined by the computer equivalent of throwing a ten-sided die. Each outcome, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, is equally likely. If you choose Option A in the row shown below, you will have a 1 in 10 chance of earning 200 points and a 9 in 10 chance of earning 160 points. Similarly, Option B offers a 1 in 10 chance of earning 385 points and a 9 in 10 chance of earning 10 points.

Decision	Option A	Option B	Your Choice
1	200 points if the die is 1 160 points if the die is 2-10	385 points if the die is 1 10 points if the die is 2-10	A or B

- **The Relevant Decision**: One of the rows is then selected at random, and the Option (A or B) that you chose in that row will be used to determine your earnings. Note: Please think about each decision carefully, since each row is equally likely to end up being the one that is used to determine payoffs.

  For example, suppose that you make all ten decisions and the throw of the die is 9, then your choice, A or B, for decision 9 below would be used and the other decisions would not be used.

Decision	Option A	Option B	Your Choice
9	200 points if the die is 1-9 160 points if the die is 10	385 points if the die is 1-9 10 points if the die is 10	A or B

- **Determining the Payoff**: After one of the decisions has been randomly selected, the computer will generate another random number that corresponds to the throw of a ten-sided die. The number is equally likely to be 1, 2, 3, ... 10. This random number determines your earnings for the Option (A or B) that you previously selected for the decision being used.

  For example, in Decision 9 below, a throw of 1, 2, 3, 4, 5, 6, 7, 8, or 9 will result in the higher payoff for the option you chose, and a throw of 10 will result in the lower payoff.

Decision	Option A	Option B	Your Choice
9	200 points if the die is 1-9 160 points if the die is 10	385 points if the die is 1-9 10 points if the die is 10	A or B
10	200 points if the die is 1-10	385 points if the die is 1-10	A or B

  For decision 10, the random die throw will not be needed, since the choice is between amounts of money that are fixed: 200 points for Option A and 385 points for Option B.

We encourage you to earn as much cash as you can. Are there any questions?

Appendix C Additional Analyses of Experimental Data

In this appendix, we first present the summary statistics of the experimental data (Table C.3) and then examine the robustness of our analyses in Section 5.1 regarding willingness to pay for information.

Table C.3: Summary Statistics

	Full Sample	Consistent Subjects in the Holt-Laury Lottery Choice Game[a]				
		All Treatments	IA		DA	
			OwnValue	OtherValue	OwnValue	OtherValue
	(1)	(2)	(3)	(4)	(5)	(6)
WTP for info	4.42	4.24	6.44	4.32	4.17	1.84
	(4.57)	(4.56)	(4.87)	(4.68)	(4.30)	(2.86)
Guess of others' WTP	5.09	5.01	7.03	5.25	4.39	3.22
	(2.90)	(2.90)	(2.44)	(2.89)	(2.63)	(2.17)
Info Acquired	0.29	0.28	0.43	0.26	0.28	0.11
	(0.46)	(0.45)	(0.50)	(0.44)	(0.45)	(0.31)
High_B×IA_OtherValue	0.05	0.05	-	0.20	-	-
	(0.22)	(0.21)	-	(0.40)	-	-
High_B×DA_OtherValue	0.05	0.04	-	-	-	0.19
	(0.21)	(0.21)	-	-	-	(0.39)
Wealth after all rounds	2063.19	2064.45	2036.24	2203.25	1947.82	2087.78
	(394.54)	(398.71)	(405.90)	(356.60)	(407.50)	(385.12)
% playing a dominated strategy with free info	0.05	0.04	-	-	0.09	0.07
	(0.11)	(0.11)	-	-	(0.13)	(0.14)
Costly-to-free	0.50	0.51	0.52	0.49	0.52	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Risk aversion	5.59	6.68	6.41	6.49	6.77	7.05
	(2.80)	(1.44)	(1.28)	(1.56)	(1.43)	(1.38)
Consistent in Holt-Laury	0.84	-	-	-	-	-
	(0.37)	-	-	-	-	-
Curiosity	4.20	3.87	5.47	4.35	3.51	2.01
	(4.95)	(4.86)	(5.46)	(5.06)	(4.55)	(3.39)
Demographics						
Female	0.54	0.54	0.44	0.56	0.55	0.60
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
Graduate student	0.14	0.15	0.21	0.18	0.14	0.05
	(0.35)	(0.35)	(0.41)	(0.38)	(0.35)	(0.22)
Black	0.05	0.03	0.08	0.02	0.00	0.02
	(0.21)	(0.17)	(0.27)	(0.13)	(0.00)	(0.13)
Asian	0.36	0.40	0.41	0.33	0.48	0.35
	(0.48)	(0.49)	(0.49)	(0.47)	(0.50)	(0.48)
Hispanic	0.02	0.02	0.03	0.00	0.05	0.00
	(0.14)	(0.14)	(0.18)	(0.00)	(0.21)	(0.00)
Age	20.83	20.80	21.70	19.82	21.19	20.37
	(3.47)	(3.05)	(3.57)	(3.48)	(2.54)	(1.97)
# Observations	2592	2169	567	513	576	513
# Subjects	288	241	63	57	64	57

Notes: This table reports means and standard deviations (in parentheses) for the variables used in the main analysis. Corresponding to the regressions controlling for lagged variables, every subject's 9 rounds (2nd to 10th) with costly information acquisition are included, whereas the rounds with free information and the first round with costly information are excluded from in the calculation of these staticstics.

a. A subject is not consistent in the Holt-Laury lottery choice game if she has more than one switching point or makes a dominated choice. Columns 2-6 exclude 47 inconsistent subjects.

C.1 Willingness to Pay for Information: Robustness of Results

In Section 5.1, we use a Tobit model to investigate the determinants of WTP for information. Here, we present the results from linear panel regressions that allow more flexible specifications and instrumental variables. In short, the following results are consistent with those in the main text, indicating that the endogeneity issue is not a concern.

Corresponding to Table 4 in Section 5.1, we present Table C.4 where the subject-average WTP is regressed on treatment types and other controls. The two sets of results are qualitatively the same.

In comparison with results from a random effect Tobit model in Tables 5 and 6, the next two tables provide the results from our analyses of the determinants of WTP in random and fixed effects panel regressions. In all specifications, our outcome variable is the subject-round WTP. Our specification is as follows:

$$\begin{aligned} WTP_{i,t} =& \alpha_i + \beta_1 High\_B \times IA\_OtherValue_{i,t} + \beta_2 High\_B \times DA\_OtherValue_{i,t} \\ &+ \beta_3 WTP\_Guess_{i,t} + Controls_{i,t} + \varepsilon_{i,t}, \end{aligned}$$

where i is the index for subjects and t for rounds (with each session); α_i is subject fixed effects; and all control variables are time-subject-specific. Other controls are the same as in Section 5.1. Depending on whether the model is random or fixed effects, we have different interpretations of α_i.

The endogeneity of $WTP\_Guess_{i,t}$ is plausible if there are common shocks in round t that increase everyone's $WTP_{i,t}$ and $WTP\_Guess_{i,t}$. We address this issue with an IV approach where the lagged $WTP\_Average_{-i,t-1}$ is the instrumental variable. Here, $WTP\_Average_{-i,t-1}$, i' opponents' WTP in round $t-1$, is correlated with $WTP\_Guess_{i,t}$, as a participant might rely on the opponents' WTP in the previous round to make her guess of others' WTP this round. Moreover, $WTP\_Average_{-i,t-1}$ should not affect her decision in round t directly, as opponents in each round are randomly drawn.

Our fixed-effect results are presented in Table C.5. The first three columns are from OLS regressions, while column 4 is from an IV regression, where the instrument for the potentially endogenous variable, $WTP\_Guess_{i,t}$, is $WTP\_Average_{-i,t-1}$. Column 5 shows the first-stage result.

When we consider $WTP\_Average_{-i,t-1}$ as an IV for use $WTP\_Guess_{i,t}$, column 5 presents the first-stage result which shows that $WTP\_Average_{-i,t-1}$ is positively correlated with $WTP\_Guess_{i,t}$ (significant at 1% level).

Column 4 is the IV regression result. Observationally, IV results are not very different from OLS results (column 3), although the coefficient on $WTP\_Guess_{i,t}$ is decreased. We next perform an endogeneity test. Under the null hypothesis that $WTP\_Guess_{i,t}$ can actually be treated as exogenous, the test statistic is distributed as chi-squared with degrees of freedom equal to one. It is defined as the difference of two Sargan-Hansen statistics: one for the IV regression, where the $WTP\_Guess_{i,t}$ is treated as endogenous, and one for the OLS regression, where $WTP\_Guess_{i,t}$

Table C.4: Determinants of Subject-Average WTP: Linear Regression

	(1) Full Sample	(2) Sub-sample	(3) Sub-sample	(4) Sub-sample
IA_OwnValue	6.56***	6.41***	5.70***	4.41
	(0.55)	(0.56)	(0.90)	(3.67)
IA_OtherValue	4.51***	4.31***	4.00***	2.48
	(0.48)	(0.54)	(0.91)	(3.78)
DA_OwnValue	4.44***	4.16***	3.66***	2.29
	(0.63)	(0.70)	(0.87)	(3.73)
DA_OtherValue	2.21***	1.92***	2.02**	0.74
	(0.30)	(0.27)	(0.91)	(3.72)
% playing a dominated strategy with free info[a]			5.44***	5.41**
			(1.85)	(2.32)
Curiosity			0.29***	0.29***
			(0.05)	(0.04)
Costly-to-free			1.61***	1.56***
			(0.36)	(0.30)
Risk Aversion			-0.27**	-0.23
			(0.11)	(0.14)
Female				-0.71*
				(0.38)
Graduate Student				0.52
				(1.76)
Black				-1.83
				(1.46)
Asian				-0.48
				(0.40)
Hispanic				-1.92**
				(0.78)
N	288	241	241	241
R^2	0.65	0.63	0.73	0.74

Notes: The outcome variable is subject-level average WTP for information. Columns 2-4 exclude participants with multiple switching points in the Holt-Laury lottery game or those who make inconsistent choices. Column 4 also includes the following controls: age, ACT score, SAT score, dummy for other non-white ethnicities/races, dummy for ACT score missing, dummy for SAT score missing, dummy for degree missing, dummy for age missing, dummy for ethnicity missing, and dummy for gender missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

a. "% playing a dominated strategy with free info" is defined as the percentage of time when the subject played a dominated strategy (i.e., non-truth-telling) in the OwnValue or OtherValue treatment of DA in rounds without information acquisition. Mean = 0.09, standard deviation = 0.14 among all subjects ($n = 144$) played the information acquisition game under DA. Only rounds without information acquisition, i.e., with no information or free information provision, are considered. This variable equals zero for both treatments of IA, because dominant strategies are not defined under IA.

is treated as exogenous. It turns out that the test statistic is 1.64 (p-value 0.20), which leads us to conclude that $WTP\_Guess_{i,t}$ is exogenous.

In summary, the results in Table C.5 are similar to those in Tables 5 and 6 from a random effect Tobit model. Moreover, the IV results in column 4 are qualitatively similar to other results in Table C.5.

When we repeat the same analyses with random effect panel regressions, we obtain similar results (Table C.6).

Table C.5: Determinants of WTP: Fixed Effects Model and IV Regression Results

	(1) FE	(2) FE	(3) FE	(4) FE	(5) IV	(6) 1st Stage
High_B × IA_OtherValue	2.25**	2.25**	2.24**	2.24**	2.24***	-0.05
	(0.81)	(0.81)	(0.80)	(0.80)	(0.79)	(0.26)
High_B × DA_OtherValue	-0.09	-0.09	-0.10	-0.10	-0.10	-0.02
	(0.43)	(0.43)	(0.43)	(0.43)	(0.40)	(0.24)
Round	-0.02	-0.02	-0.03	-0.03	-0.03	-0.12***
	(0.03)	(0.03)	(0.06)	(0.06)	(0.04)	(0.03)
Round × costly-to-free	-0.08	-0.08	-0.08	-0.08	-0.07	0.03
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)
Accumulated wealth up to $t-1$			0.00	0.00		
			(0.00)	(0.00)		
Successfully acquired info in $t-1$			-0.02	-0.02		
			(0.22)	(0.22)		
"**Endogenous**": Guess of Opponents' WTP in t	0.60***	0.60***	0.60***	0.60***	0.47***	
	(0.07)	(0.07)	(0.07)	(0.07)	(0.10)	
IV: Average WTP of Opponents in $t-1$	–	–	–	–	–	0.15***
						(0.01)
# of Observations	2169	2169	2169	2169	2169	2169
# of Subjects	241	241	241	241	241	241
R^2	0.18	0.18	0.18	0.18	0.17	0.12

Notes: The outcome variable is WTP for information of each subject in each round. Regressions exclude participants with multiple switching points or making dominated choices in the Holt-Laury lottery game and only include observations from rounds 2-10. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
"Average WTP of Opponents in $t-1$" is used as IV for "Guess of Opponents' WTP in t" (1st-stage results in the last column, i.e., dependent variable = "Guess of others' WTP in t"). A endogneity test for "Guess of Opponents' WTP in t" based on Sargan-Hansen statistics gives a p-value of 0.20. That is, we fail to reject the null hypothesis that the variable is exogenous.

C.2 Decomposition based on Pooled Regression

Table 7 in Section 5.1 presents the decomposition of excess WTP based on Tobit models for each treatment. As a robustness check, we also present results based on pooled regressions (Table C.7). Although results change to some extent, we still find "Conformity" explains the largest part of the excess WTP.

Table C.6: Determinants of WTP: Random Effects Model and IV Regression Results

	(1) RE	(2) RE	(3) RE	(4) RE	(5) IV	(6) 1st Stage
IA_OwnValue	1.52***	1.43***	1.24***	1.16***	2.09***	3.02***
	(0.43)	(0.44)	(0.41)	(0.44)	(0.57)	(0.30)
IA_OtherValue	0.34	0.15	0.22	0.06	0.48	1.52***
	(0.47)	(0.49)	(0.46)	(0.48)	(0.46)	(0.40)
DA_OwnValue	1.12***	1.01***	1.01***	0.90**	1.19***	0.66**
	(0.38)	(0.39)	(0.38)	(0.38)	(0.46)	(0.31)
High_B $\times$ IA_OtherValue	2.28***	2.27***	2.35***	2.31***	2.24***	-0.04
	(0.82)	(0.82)	(0.86)	(0.84)	(0.73)	(0.25)
High_B $\times$ DA_OtherValue	-0.08	-0.06	-0.02	0.01	-0.08	-0.03
	(0.43)	(0.43)	(0.43)	(0.43)	(0.44)	(0.24)
Round	-0.01	-0.01	0.02	0.02	-0.03	-0.12***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.03)
Round $\times$ costly-to-free	-0.08	-0.08	-0.07	-0.07	-0.07	0.03
	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.04)
% playing a dominated strategy with free info	3.69**	3.39*	3.24**	2.92	3.86*	2.57**
	(1.55)	(1.82)	(1.56)	(1.81)	(2.25)	(1.27)
Curiosity	0.22***	0.23***	0.21***	0.21***	0.25***	0.12***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Costly-to-free	0.86**	0.75**	0.47	0.31	1.08***	1.50***
	(0.36)	(0.36)	(0.52)	(0.54)	(0.37)	(0.33)
Risk Aversion	-0.29***	-0.26**	-0.27***	-0.25**	-0.25**	0.04
	(0.10)	(0.11)	(0.09)	(0.11)	(0.13)	(0.08)
Accumulated wealth up to $t-1$			-0.00	-0.00		
			(0.00)	(0.00)		
Successfully acquired info in $t-1$			0.66***	0.63***		
			(0.21)	(0.21)		
"**Endogenous**": Guess of Opponents' WTP in t	0.62***	0.62***	0.65***	0.65***	0.44***	
	(0.06)	(0.07)	(0.06)	(0.06)	(0.10)	
IV: Average WTP of opponents in $t-1$						0.16***
						(0.01)
Other Demographic Controls	No	Yes	No	Yes	Yes	Yes
# of Observations	2169	2169	2169	2169	2169	2169
# of Subjects	241	241	241	241	241	241

Notes: The regression sample is the same as that in column 5 in Table 6. Each of the 241 subjects has 9 observations from 9 rounds. Estimates are from random effects panel Tobit models. All specifications include additional controls: dummy for female, dummy for graduate student, dummy for black, dummy for Asian, dummy for Hispanic, age, ACT score, SAT score, dummy for ACT score missing, dummy for SAT score missing, dummy for degree missing,dummy for age missing, dummy for ethnicity missing, and dummy for gender missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

"Average WTP of Opponents in $t-1$" is used as IV for "Guess of others' WTP in t" (1st-stage results in the last column, i.e., dependent variable = "Guess of others' WTP in t"). Without clustered standard errors, a endogneity test for "Guess of Opponents' WTP in t" based on Hausman's specification test gives a p-value of 0.52. That is, we fail to reject the null hypothesis that the variable is exogenous.

Table C.7: Decomposition of Subject WTP for Information Based on the Pooled Regression

	IA OwnValue	**IA OtherValue**	**DA OwnValue**	**DA OtherValue**
WTP: data	**6.44**	**4.32**	**4.17**	**1.84**
	(4.87)	(4.68)	(4.30)	(2.86)
Model prediction[a]	6.29	4.17	4.15	1.78
	(2.79)	(3.01)	(2.65)	(1.74)
(i) Cognitive load[b]	0.68	0.53	0.57	0.32
	(0.29)	(0.31)	(0.31)	(0.22)
(ii) Learning over rounds[b]	0.31	0.25	0.27	0.16
	(0.29)	(0.26)	(0.27)	(0.19)
(iii) Conformity[b]	4.16	2.64	2.34	1.06
	(1.95)	(2.06)	(1.82)	(1.17)
(iv) % playing a dominated strategy with free info[b]			0.45	0.25
			(0.79)	(0.56)
(v) Curiosity[b]	1.70	1.21	0.98	0.40
	(1.73)	(1.49)	(1.35)	(0.87)
(vi) Risk aversion[b]	-0.33	-0.28	-0.35	-0.27
	(0.29)	(0.33)	(0.33)	(0.27)
Total Explained by factors (i)-(vi)[c]	**5.50**	**3.57**	**3.46**	**1.50**
	(2.70)	(2.78)	(2.59)	(1.70)
Residual WTP[d]	**0.93**	**0.75**	**0.71**	**0.34**
	(4.01)	(3.84)	(3.14)	(2.28)
Theoretical prediction[e]	**[5.2,8]**	**[0,0.24]**	**0.67**	**0.00**
# of observations	567	513	576	513
# of subjects	63	57	64	57

Notes: Decompositions are based on a random effects panel Tobit model that pools observations from all four treatments (columns 5 in Table 6). This is in contrast to Table 7 which uses a separate regression for each treatment. The table reports the sample average, while standard deviations are in parentheses.

a. "Model prediction" is the predicted value of E(WTP) based on the corresponding estimated model, assuming that unobserved error terms are equal to zero. The predicted values are censored to $[0, 15]$.

b. The WTP explained by the corresponding factor is the difference between the model prediction with and without the factor. The former is predicted from the current values of all variables; the latter is calculated by setting the relevant variable value to zero (for factors "Cognitive load," "Conformity," "% playing a dominated strategy with free info," or "Curiosity") or setting the relevant variable to the counterfactual value (for "Risk aversion," we set the risk aversion measure to the risk-neutral value; for "Learning over round," we set "Round" to be the last round, i.e., "Round" $= 10$).

c. "Total Explained by factors (i)-(vi)" is the total WTP explained by the six factors above. Note that it is not the sum of the explained WTP of the six factors because of the censoring at 0 and 15.

d. "Residual WTP" is the difference between the observed WTP and the total WTP explained by the six factors.

e. The theoretical predictions are for risk neutral subjects.

C.3 Welfare Effects of Information Provision and Acquisition: Additional Analyses

In this section, we present additional analyses related to our welfare effects. The results in Table C.8 show the effects of information provision on payoffs and allocation efficiency.

Regarding the effects of information acquisition, if costs are not taken into account, we expect outcomes to fall between no information and free information provision. This is because the information acquisition technology results in an endogenous probability of receiving the "hard news," with both informed and uninformed participants. This leads to our final set of hypotheses.

Hypothesis 7 (Efficiency under Information Acquisition: Information Structures)**.** *With costly information acquisition and not taking into account costs, the ex ante subject welfare and the fraction of efficient allocations under either mechanism follow the order of UI* $<$ *{Acquiring OwnValue} and CI* $=$ *{Acquiring OtherValue}. Furthermore, the allocation is always (not always) efficient for either CI or {Acquiring OtherValue} under IA (DA).*

Hypothesis 8 (Efficiency under Information Acquisition: Mechanisms)**.** *With costly information acquisition and not taking into account costs, in terms of the ex ante subject welfare and the fraction of efficient allocations, IA* $>$ *DA under either {Acquiring OwnValue} or {Acquiring OtherValue}.*

Result 8 (Efficiency under Information Acquisition)**.** *When we do not take information acquisition cost into account, we obtain the following results: (i) Acquiring OwnValue improves both subject payoffs and allocation efficiency for both IA and DA, with a greater effect for IA; (ii) Acquiring OtherValue does not affect either subject payoffs or allocation efficiency for IA or DA; (iii) Outcomes of IA under OwnValue or OtherValue are closer to the efficient outcome relative to those of DA under OwnValue or OtherValue.*

Support: Table C.9 in Appendix C reports the means and standard deviations (in parentheses) of subject payoffs, the fraction of efficient allocations by information structure, the fraction of having successfully acquired information, the WTP for information, and the costs of information acquisition. Similar to our earlier analyses, for each treatment, we focus on the same subjects who play a pair of the school choice games in both the no-information and the costly-information scenarios in each round, where the order of the two scenarios is randomized in each round. This design feature enables us to perform both within- and between-treatment tests (parts (i) and (ii)). Part (iii) is obtained from simple calculations. With the acquisition of information about OwnValue, IA achieves 89% of maximum payoffs and efficient allocations among 83% of all games; as a comparison, DA achieves 80% of maximum payoffs and efficient allocations among 73% of all games. Similarly, with the acquisition of information about OtherValues, IA achieves 97% of maximum payoffs and 96% efficient allocations, whereas DA achieves only 87% and 82%, respectively. ∎

Table C.9 also presents the fraction of times each subject successfully acquires the information as well as her expressed WTP. These are positively correlated with each other due to our experimental design. In the IA OwnValue treatment, we find that 44% of subjects obtain the desired information in each round, which is exactly the ratio of the average WTP (6.56) to the upper

Table C.8: Effects of Information Provision on Payoffs and Allocation Efficiency

	Payoff		Fraction(Efficient Allocation)	
	Observed in Experiment (1)	Theoretical Prediction (2)	Observed in Experiment (3)	Theoretical Prediction (4)
A: IA OwnValue (# obs.:720)				
UI: UnInformed	42.51	43.33	0.69	0.68
	(51.12)		(0.49)	
CI: Cardinally Informed	50.67	52.93	0.94	1.00
	(52.52)		(0.29)	
Test: H_0: UI = CI; H_1: UI <CI				
p-value	0.01		0.01	
B: IA OtherValue (# obs.: 720)				
CI: Cardinally Informed	49.13	52.93	0.89	1.00
	(51.90)		(0.35)	
PI: Perfectly Informed	49.12	52.93	0.91	1.00
	(52.20)		(0.34)	
Test: H_0: CI = PI; H_1: CI $\neq$ PI				
p-value	0.92		0.35	
C: DA OwnValue (# obs.: 720)				
UI: UnInformed	42.96	43.33	0.71	0.68
	(48.93)		(0.48)	
CI: Cardinally Informed	47.22	48.67	0.84	0.87
	(54.92)		(0.43)	
Test: H_0: UI=CI; H_1: UI<CI				
p-value	0.04		0.04	
D: DA OtherValue (# obs.: 1080)				
CI: Cardinally Informed	45.90	48.67	0.81	0.87
	(49.96)		(0.39)	
PI: Perfectly Informed	46.09	48.67	0.81	0.87
	(49.53)		(0.40)	
Test: H_0: CI = PI; H_1: CI $\neq$ PI				
p-value	0.86		0.86	
E: Comparison between IA & DA				
Test: H_0: (IA UI) = (DA UI); H_1:(IA UI) $\neq$ (DA UI)				
p-value	1.00		1.00	
Test: H_0: (IA CI) = (DA CI); H_1:(IA CI)> (DA CI)				
p-value: OwnValue[a]	0.01		0.01	
p-value: OtherValue[a]	0.02		0.02	
Test: H_0: (IA PI) = (DA PI); H_1:(IA PI) > (DA PI)				
p-value	0.02		0.01	

Notes: This table reports the means and standard deviations (in parentheses) of payoffs and the fraction of efficient allocations by information structure. It also presents p-values for the Wilcoxon matched-pairs signed-ranks tests or the Wilcoxon rank-sum (or Mann-Whitney) tests. All tests are performed with the session averages of payoffs or efficiency. All data are weighted at the session level so that the probability of having a high valuation for school b equals $1/5$.

a. These two p-values are calculated with the samples of IA and DA OwnValue treatments and the one with OtherValues treatments, respectively.

bound of WTP (15). By contrast, we find that subjects acquire the information less often in the other treatments, ranging from 14% in the DA OtherValue treatment to 30% in the DA OwnValue treatment.

To evaluate the net effects of information acquisition, it is necessary to consider the costs. Section 5.3.2 provides more details on the welfare implications of these costs.

Table C.9: Effects of Information Acquisition on Payoffs and Allocation Efficiency

	Payoff	Fraction(Efficient Allocation)	Pr(Info Acquired)	WTP	Costs Paid[a]
		A: IA OwnValue (# obs.: 720)			
UI: UnInformed	42.50	0.69			
	(51.00)	(0.54)			
Acquiring OwnValue	47.05	0.83	0.44	6.56	2.25
	(52.77)	(0.47)	(0.50)	(4.78)	(3.56)
Test: H_0: UI = (Acquiring OwnValue); H_1: UI $<$ (Acquiring OwnValue)					
p-value	0.01	0.01			
		B: IA OtherValue (# obs.:720)			
CI: Cardinally Informed	49.98	0.92			
	(56.75)	(0.42)			
Acquiring OtherValue	51.36	0.96	0.28	4.49	1.29
	(54.07)	(0.35)	(0.45)	(4.56)	(2.66)
Test: H_0: CI = (Acquiring OtherValue); H_1: CI $\neq$ (Acquiring OtherValue)					
p-value	0.25	0.25			
		C: DA OwnValue (# obs.: 720)			
UI: UnInformed	42.73	0.70			
	(52.73)	(0.51)			
Acquiring OwnValue	43.80	0.73	0.30	4.44	1.35
	(48.73)	(0.49)	(0.46)	(4.38)	(2.88)
Test: H_0: UI = (Acquiring OwnValue); H_1: UI $<$ (Acquiring OwnValue)					
p-value	0.06	0.06			
		D: DA OtherValue (# obs.: 720)			
CI: Cardinally Informed	46.77	0.82			
	(50.46)	(0.43)			
Acquiring OtherValue	46.27	0.82	0.14	2.21	0.48
	(52.46)	(0.45)	(0.35)	(3.15)	(1.50)
Test: H_0: CI = (Acquiring OtherValue); H_1: CI $\neq$ (Acquiring OtherValue)					
p-value	0.92	0.92			
		E: Comparison between IA & DA			
Test: H_0: (IA Acquiring OwnValue) = (DA Acquiring OwnValue) H_1: (IA Acquiring OwnValue) $>$ (DA Acquiring OwnValue)					
p-value	0.00	0.00			
Test: H_0: (IA Acquiring OtherValue) = (DA Acquiring OtherValue) H_1: (IA Acquiring OtherValue) $>$ (DA Acquiring OtherValue)					
p-value	0.00	0.00			

Notes: This table presents p-values for the Wilcoxon matched-pairs signed-ranks tests and Wilcoxon rank-sum (or MannWhitney) tests. All data are weighted at the session level so that the probability of having high valuations of school b equals $1/5$. All tests are performed with the session averages of payoffs or efficiency.
a. "Costs paid" measures the actual costs subjects paid in the experiment.