Information Acquisition and Provision in School Choice*

Yan Chen Yinghua He

January 10, 2016

Abstract

When participating in school choice, students usually do not know perfectly their preferences over schools, while acquiring such information is costly. We study how two popular school choice mechanisms, the Immediate Acceptance and the Deferred Acceptance mechanisms, incentivize students' information acquisition. Under the Immediate Acceptance, students pay more to acquire information on their own preferences as well as that on others' preferences. We then show the potential welfare improvement when the education authority provides more information on school quality. Evidence from our lab experiment is consistent with the theoretical predictions. Furthermore, students often over-pay for information, especially among those who expect that others are paying more for information and among those who are more curious. Taken together, our results underscore the crucial role of information provision by education authorities.

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

JEL Classification Numbers: C78, C92, D82

*We thank Federico Echenique, Guillaume Fréchette, Astrid Hopfensitz, Tanjim Hossain, Fuhito Kojima, Erin Krupka, Steve Leider, David Miller, Yesim Orhun, Tanya Rosenblat, Andy Schotter, Séverine Toussaert, and seminar participants at Michigan, NYU, Tsinghua and the 2015 ESA World (Sydney, Australia) and North American (Dallas, TX) meetings for helpful discussions and comments, and Ming Jiang and Yingzhi Liang for excellent research assistance. Financial support from the National Science Foundation through grants no. SES-0720943 and 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. Chen: School of Information, University of Michigan, 105 South State Street, Ann Arbor, MI 48109-2112. Email: yanchen@umich.edu. He: Toulouse School of Economics, 21 allée de Brienne, 31000 Toulouse, France. Email: yinghua.he@gmail.com.

1 Introduction

"'*It was very hard, and very time-consuming,' New Orleans resident Carrie Fisher 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.*'" Writes Prothero (2015) when reporting the obstacles that parents confront as school choice expands. Upon participating in school choice, students may not know their preferences over the candidate schools, which makes it necessary for them to acquire more information.

Incomplete information on one's own preferences is not uncommon in real life, and future outcomes are particularly uncertain in educational settings (Dustan, de Janvry and Sadoulet 2015). Sometimes students face too many choices, e.g., when applying to colleges. They might find it difficult to navigate the vast quantity of information about college attributes. At other times, it is costly to acquire information even when there are not that many choices. For instance, to evaluate a candidate school, students may need to know the academic performance, teacher quality, school facilities, extra-curricular activities offered, and peer quality. As some of these aspects are less observable, one may have to pay substantial cost to obtain such information. As a result of informational frictions, a student may just choose the same or a similar school as her older sibling (Dustan 2015).

However, classical matching models typically assume that students know their own preferences, at least the ordinal ones. This assumption becomes unrealistic when the market is large or when the cost of acquiring information is high. Nonetheless, preferences are fundamental to evaluate mechanism performance, and more importantly, the information students need to play optimal strategies in games induced by matching mechanisms might not be the same.

Relaxing the full information assumption, we fill the gap in the literature by investigating how mechanisms incentivize student information acquisition in school choice and how information provision by educational authorities promote efficiency. We focus on the two widely used mechanisms, the Immediate Acceptance and the Gale-Shapley Deferred-Acceptance (hereafter shortened as DA) mechanisms. As it takes into account both the benefit and cost of information acquisition and information provision, this study thus provides a more comprehensive and realistic evaluation of the mechanism performance.

Specifically, in a school choice setting with unknown preferences and costly information acquisition, we show theoretically that, while both the strategy-proof DA and the non-strategy-proof Immediate Acceptance mechanisms incentivize students to acquire information on their own ordinal preferences which improves efficiency, non-strategy-proof mechanisms also induce students to acquire information on one's own cardinal preferences as well as that on others' preferences.

We then investigate the welfare effects of information provision by education authorities. In a special setting where every student has the same ordinal preferences, we show that the *ex ante* welfare is constant under the DA mechanism, while providing information on one's own cardinal preferences improves welfare under the Immediate Acceptance mechanism. However, providing information on others' preferences has ambiguous welfare effects.

As our theory predicts that incentives to acquire information depend on the strategic properties of the mechanism, naturally-occurring field data typically do not enable a systematic evaluation of the theory, especially at the welfare level, as there are often only one mechanism involved at a time (Hastings and Weinstein 2008). We resort to a laboratory experiment to quantify the net welfare effect of endogenous information acquisition. In our experiment, we use the Becker-DeGroot-Marschak mechanism to elicit participant willingness-to-pay (WTP) for information (Becker, DeGroot and Marschak 1964a).

We find that participants' WTP for their own and others' preferences under the non-strategy-proof Immediate Acceptance mechanism is significantly greater than that under the strategy-proof DA mechanism, which is consistent with theory. However, their WTP is systematically higher than the theoretical prediction. The excess WTP can be decomposed into understanding the game, conformity, curiosity, cognitive load and learning. Students who do not understand well the school choice game and those expecting that others are paying more for information tend to over-pay for information.

At the welfare level, we find that free information on own preferences improves the performance of both mechanisms, whereas free information on others' preferences does not improve the performance of either mechanism. Lastly, costly information acquisition on own or others' preferences improves efficiency, if costs are not taken into account. Our results suggest that more transparent information provision on school performance can significantly improves welfare.

The rest of this paper is organized as follows. Section 2 reviews the information acquisition

and school choice literature. Section 3 presents the theoretical results. Section 4 describes the experimental design. Section 5 summarizes the results of the experiments. Section 6 concludes.

2 Literature Review

Most of papers in the matching literature deal with the case in which students know their preferences (Gale and Shapley 1962, Roth and Sotomayor 1990, Abdulkadiroğlu and Sönmez 2003). One exception is Chade, Lewis and Smith (2014), where colleges only observe signals of students' ability. Other examples are Lee and Schwarz (2012) and Rastegari, Condon and Immorlica (2013) who study models where firms' preferences over workers are unknown or partially known and are revealed only through interviews.

To our knowledge, the only two theoretical papers in market design that explicitly studies endogenous information acquisition are Bade (2015) and Harless and Manjunath (2015). In the setting of house allocation, Bade finds that there is a unique *ex ante* Pareto optimal, strategy-proof and non-bossy allocation mechanism: serial dictatorship. Harless and Manjunath (2015), however, show that top trading cycles mechanism dominates serial dictatorship under progressive measures of social welfare, e.g., inequality-averse social welfare functions. In both papers, the authors focus on ordinal mechanisms, and thus information acquisition is essentially on one's ordinal preferences. As we show below, in any strategy-proof ordinal mechanism, students only have incentives to learn their ordinal preferences. However, information on cardinal preferences may be efficiency improving, especially when students have similar ordinal preferences (Abdulkadiroğlu, Che and Yasuda 2011).

The growing experimental school choice literature has not dealt with endogenous information acquisition. Most studies focus on strategy, stability and welfare comparisons among various mechanisms, given that students know their own preferences (Chen and Sönmez 2006, Featherstone and Niederle 2008, Calsamiglia, Haeringer and Klijn 2010, Klijn, Pais and Vorsatz 2012). In addition, Pais and Pintér (2008) and Pais, Pintér and Veszteg (2011) study matching mechanisms in different information settings and provide evidence on how more information may change individual strategies and overall efficiency of a mechanism, but the information structure is taken as exogenous in their setting. Therefore, their results are more on the robustness of the mechanisms

to information. More recently, several experimental studies of school choice focus on peer information sharing in networks (Ding and Schotter 2015b) and intergenerational advice (Ding and Schotter 2015a), whereas Guillen and Hing (2014) investigate top down advice.

Another feature of our study is the acquisition of information on others' preferences, which is in contrast with the usual studies focusing on acquiring information on one's own preferences. One exception is Kim (2008) who compares two information structures in a common value first-price auction with two bidders. In one of them, one of the bidders learns her opponent's signal as well.

Costly information acquisition is considered in many other fields as well, 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), law and economics (Lester, Persico and Visschers 2009). In particular, there is a large theoretical literature on information acquisition in mechanism design, especially in auction design, e.g., Persico (2000), Compte and Jehiel (2007), Crémer, Spiegel and Zheng (2009), Shi (2012), and papers surveyed in Bergemann and Valimaki (2006). Notably, Bergemann and Valimaki (2002) show that in every private value environment the Vickrey-Clark-Groves mechanism guarantees both ex-ante and ex-post efficiency.

While the literature on information acquisition is mostly theoretical, there are a few experimental investigations in various contexts. Gabaix, Laibson, Moloche and Weinberg (2006) use two experiments with costly information acquisition to test the directed cognition model against the fully rational model. Combining theory and experiment, Choi, Guerra and Kim (2015) compare the second-price (sealed-bid) auction with the English auction in a setting where bidders have independent values and are heterogeneously informed. Bhattacharya, Duffy and Kim (2015) study endogenous information acquisition in voting.

3 Theoretical Analysis

In this section, we outline a theoretical model of endogenous information acquisition for own and others' preferences under two commonly used school choice mechanisms, the Immediate and Deferred Acceptance mechanisms.

3.1 The Setup

There is a finite set of students, I, to be assigned to a finite set of schools, S, through a centralized school choice mechanism. S contains a "null school" or outside option s^0 that denotes the possibility of not being matched with any school in $S \setminus \{s^0\}$. For each $s \in S$, there is a finite supply of seat, $q_s \in \mathbb{N}$, and there are enough seats to accommodate all students, $\sum_{s \in S \setminus \{s^0\}} q_s = |I|$, while $q_s > 0$ for all s. By assumption, $q_{s^0} \geq |I|$. In the mechanism, schools do not have pre-determined priority ranking over students but rank students by a common and even lottery (single tie-breaking) whose realization is unknown by students when entering the mechanism.

For each student $i \in I$, her valuations of schools are i.i.d. draws from a distribution F and are denoted by a vector $V_i = [v_{i,s}]_{s \in S}$, where $v_{i,s} \in [0, 1]$ is i's von Neumann-Morgenstern utility associated with school s.

To simplify notations, student preferences are assumed to be strict: For any pair of distinct schools s and t, $v_{i,s} \neq v_{i,t}$ for all i. Furthermore, we define strict ordinal preferences P on S such that sP_it if and only if $v_{i,s} > v_{i,t}$. We augment the set of all possible strict ordinal preferences $\mathcal{P}$ with a "null preference" $P^\phi \equiv \emptyset$ denoting that one has no information on her ordinal preference, which leads us to the notation $\bar{\mathcal{P}} = \mathcal{P} \cup \emptyset$. The distribution of V conditional on P is denoted as $F(V|P)$, while the distribution of P implied by F is $G(P|F)$. We impose a full-support assumption on $G(P|F)$, i.e., $G(P|F) > 0, \forall P \in \mathcal{P}$, while, on the other hand, $G(P^\phi|F) = 0$. That is, given the distribution of cardinal preferences, every strict ordinal preference ranking is possible.

In the following, we first introduce two centralized mechanisms for student placement. Regarding students' information on their preferences, we consider various cases, while it is maintained that the distribution of preferences, $F(V)$ and thus $G(P|F)$, are always common knowledge. More importantly, deviating from the previous literature on school choice, 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.

3.2 School Choice Mechanisms

We focus on two school choice mechanisms that are popular both in the research literature and in practice: the Immediate-Acceptance mechanism (also known as the Boston mechanism) and the

Gale-Shapley Deferred-Acceptance mechanism.

The **Immediate-Acceptance** mechanism (IA) asks students to submit rank-ordered lists (ROL) of schools. Together with the pre-announced capacity of each school, the mechanism uses pre-defined rules to determine school priority ranking over students and has the following rounds:

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 kth choice in order of their priority at that school until either there is no seats left at that school or no such student left.

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

The **Gale-Shapley Deferred-Acceptance** mechanism (DA) can be either student-proposing or school-proposing. We focus on student-proposing DA in this study. The mechanism collects school capacities and students' submitted ROL of schools. With strict rankings of schools over students that are determined by pre-specified rules, the process also has several rounds:

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 who are held from Round ($k - 1$) and rejects the least ranked students in excess of its capacity. Those who are not rejected are temporarily held by the schools.

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

3.3 Acquiring Information on Own Preferences

We first investigate student incentives to acquire information on one's own value (type). The timing of the game and the corresponding information structure are described as follows and also in Figure

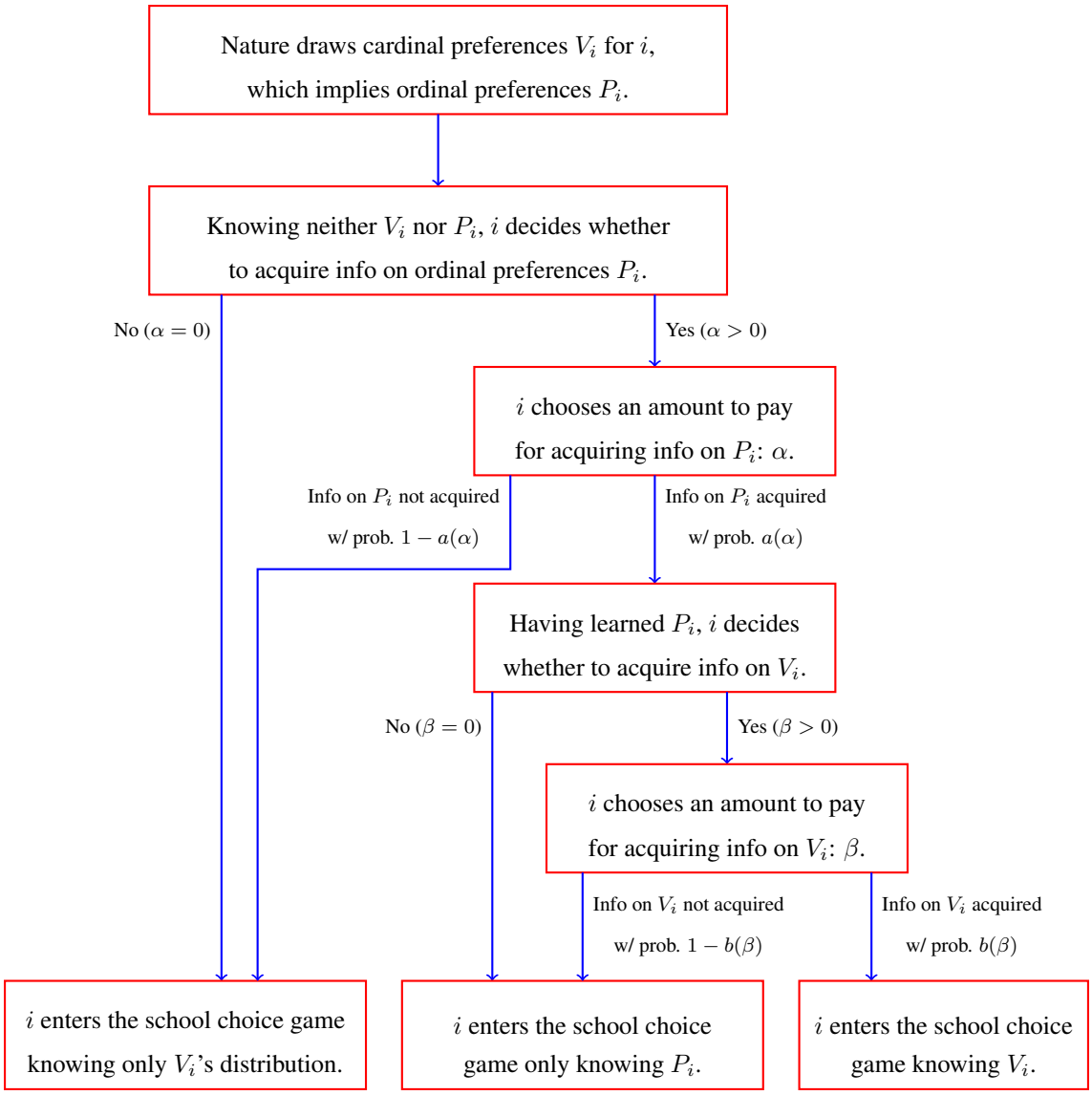

Figure 1: Acquiring Information on One's Own Preferences.

1:

(i) Nature draws individual valuation, V_i, from $F(V)$ for each i, but i knows the value distribution $F(V)$ only;

(ii) Each individual i decides whether or not to acquire a signal on her ordinal preferences; If yes, she decides how much she is willing to invest in information acquisition; Her willingness to pay to acquire her ordinal preferences is denoted by $\alpha \in [0, \bar{\alpha}]$.

(iii) If i learns her ordinal preferences, she then decides whether to acquire a signal on her cardinal

preferences; If yes, she decides how much she is willing to invest in information acquisition, denoted by $\beta \in [0, \bar{\beta}]$.

(iv) Regardless of the decision and outcomes of the information acquisition stage, every student plays the school choice game under a given school choice mechanism, the IA or DA mechanism.

3.3.1 Technology of Information Acquisition

Information acquisition in the model is covert. That is, i knows that everyone else is engaging in information acquisition, but does not know what information has been acquired by whom.

The information acquisition has two stages (see Figure 1): i first pays a cost α to acquire a signal on ordinal preference, $\omega_{1,i} \in \bar{\mathcal{P}}$. With probability $a(\alpha)$, she learns perfectly, $\omega_{1,i} = P_i$; with probability $1 - a(\alpha)$ nothing is learned, $\omega_{1,i} = P^{\phi}$. At 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 $\bar{\mathcal{V}} \equiv \mathcal{V} \cup \emptyset \equiv \bar{\mathcal{V}} \cup \emptyset$. With probability $b(\beta)$, one learns her cardinal preferences, $\omega_{2,i} = V_i$; with probability $1 - b(\beta)$, she does not, $\omega_{2,i} = V^{\phi}$, where $V^{\phi} \equiv \emptyset$ denotes no information on cardinal preferences.

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 $a'(0) = b'(0) = +\infty$.[1] 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 may limit our attention to $\alpha \in [0, \bar{\alpha}]$ and $\beta \in \left[0, \bar{\beta}\right]$, where $c(\bar{\alpha}, 0) = c(0, \bar{\beta}) = 1$. This is because $c(\alpha, \beta)$ can exceed the maximum possible payoff (=1), when α and β are too large.

We summarize the signals that i may observe after the two-stage information acquisition by

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

$\omega_i = (\omega_{1,i}, \omega_{2,i}) \in \bar{\mathcal{P}} \times \bar{\mathcal{V}}$. If i pays (α, β), the distribution of signals is $H(\omega_i|\alpha, \beta)$:

$$\begin{aligned}
H\left(\omega_i = \left(P^\phi, V^\phi\right)|\alpha, \beta\right) &= 1 - a(\alpha), \\
H\left(\omega_i = \left(P_i, V^\phi\right)|\alpha, \beta\right) &= a(\alpha)(1 - b(\beta)), \\
H(\omega_i = (P_i, V_i)|\alpha, \beta) &= a(\alpha) b(\beta), \\
H(\omega_i = (P, V)|\alpha, \beta) &= 0 \text{ if } (P, V) \notin \{(P^\phi, V^\phi), (P_i, V^\phi), (P_i, V_i)\}.
\end{aligned}$$

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).

3.3.2 Game of School Choice with Information Acquisition

After observing the signal ω_i, students enter the school choice game under DA or IA as pre-announced. 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.[2] The payoff if i submits L_i and others submit L_{-i} is:

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

where $a_s(L_i, L_{-i})$ is the probability that i is accepted by s given (L_i, L_{-i}). The mechanism determines $A(L_i, L_{-i})$. We further distinguish between two types of mechanisms: strategy-proof and non-strategy-proof.

Definition 1. *A mechanism is **strategy-proof** if*

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

[2] 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 including s^0, which differs from reality where the outside option is not included in ROL. However, we assume that i is always assigned to s^0 when she is rejected by all s s.t. sL_is^0. It is therefore equivalent to say that the effective ROL is truncated at s^0, which makes our restriction of $L_i \in \mathcal{P}$ innocuous.

i.e., reporting true ordinal preferences is a dominant strategy.

It is well-known that the student-proposing DA mechanism is strategy-proof (Dubins and Freedman 1981, Roth 1982), whereas the IA mechanism is not (Abdulkadiroğlu and Sönmez 2003).

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

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

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

That is, given her own signal ω, everyone best responds to others, taking into account that others have paid $\left(\alpha^*_{-i}, \beta^*_{-i}\right)$ to acquire information. We further define the value function given $\left(\omega, \alpha^*_{-i}, \beta^*_{-i}\right)$:

$$\begin{aligned} & \Pi\left(\omega, \alpha^*_{-i}, \beta^*_{-i}\right) \\ = \; & \max_{\sigma} \left\{ \int \int \int u(V, \sigma, \sigma^*(\omega_{-i}))\, dF(V|\omega)\, dF(V_{-i}|\omega_{-i})\, dH\left(\omega_{-i}|\alpha^*_{-i}, \beta^*_{-i}\right) \right\}. \end{aligned}$$

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

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

$\beta^*(P, \alpha^*)$ is therefore the optimal decision given that one has learned her ordinal preference (P) after paying α^* for acquiring P.

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

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

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

We are now ready to show our existence result (Lemma 1) and analyses of information acquisition behavior in equilibrium (Proposition 1).

Lemma 1. *Under DA or IA, a symmetric Bayesian Nash equilibrium exists.*

Proposition 1. *In a symmetric Bayesian Nash equilibrium* $(\alpha^*, \beta^*(P, \alpha^*), \sigma^*(\omega))$ *under DA or IA,*

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

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

(iii) under IA, there always exists a preference distribution F *such that* $\beta^*(P) > 0$ *for some* P*.*

Remark 1. *Similar to the results on DA, for all strategy-proof mechanisms that elicits ordinal preferences from students, students have no incentive to learn their own cardinal preferences.*[3]

3.4 Acquiring Information on Others' Preferences

This section studies students' incentive to acquire information on others' types/preferences. We now assume that everyone knows exactly her own cardinal preferences (V_i) but not others' preferences (V_{-i}), and that the distribution of V_i, $F(V_i)$, is still common knowledge and has the same properties as before.

The process and technology for information acquisition are depicted in Figure 2, which are similar to the acquisition of information on own preferences. 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 $1 - d(\delta)$, $\omega_{3,i} = V_{-i}^\phi \equiv \emptyset$, i.e., i learns nothing. The distribution of signals and the posterior distribution of preferences are:

$$K\left(\omega_{3,i} = V_{-i}^\phi | \delta\right) = 1 - d(\delta),$$
$$K\left(\omega_{3,i} = V_{-i} | \delta\right) = d(\delta),$$
$$K\left(\omega_{3,i} = V'_{-i} | \delta\right) = 0 \text{ if } V'_{-i} \notin \{V_{-i}, V_{-i}^0\};$$

$$F(V_{-i}|\omega_{3,i}) = \begin{cases} F(V_{-i}) & \text{if } \omega_{3,i} = V_{-i}^\phi; \\ 1_{V_{-i}} & \text{if } \omega_{3,i} = V_{-i}. \end{cases}$$

[3]This statement does not extend to mechanisms elicit cardinal preferences, e.g., Hylland and Zeckhauser (1979) and He, Miralles, Pycia and Yan (2015).

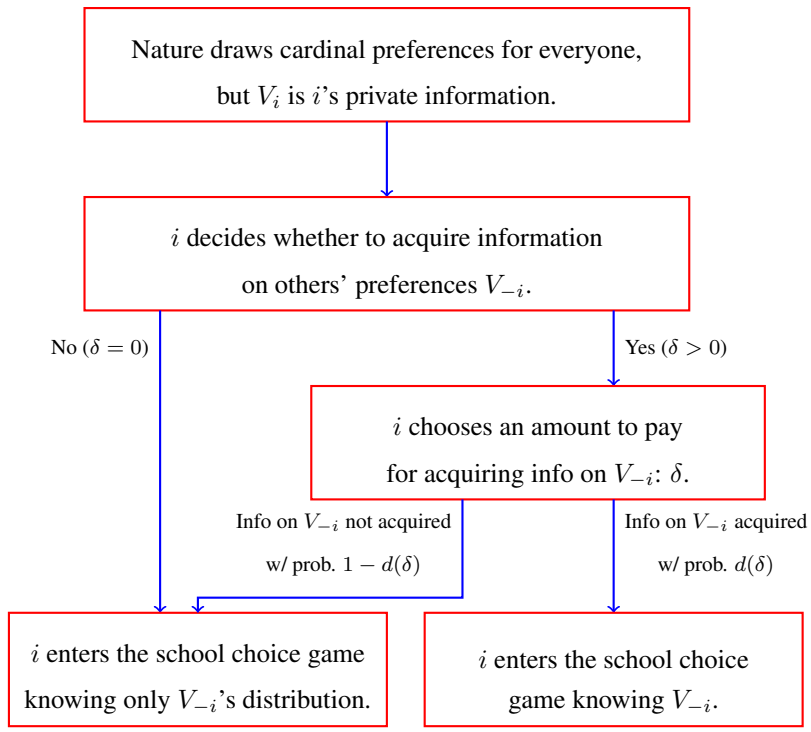

Figure 2: Acquiring Information on Others' Preferences.

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 may restrict our attention to $\delta \in \left[0, \bar{\delta}\right]$, where $e\left(\bar{\delta}\right) = 1$.

Information acquisition is again covert, i.e., everyone knows that others engage in acquiring information but does not know if they succeed or not. 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 her own signal $\omega_{3,i}$, everyone best responds to others, taking into account that they have 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 on others' preferences $\delta^*(V) : \bar{\mathcal{V}} \to [0, \bar{\delta}], \forall V$:

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

$\delta^*(V_i)$ is therefore the optimal information acquisition.

The existence of such an equilibrium can be proven by similar arguments in the proof of Lemma 1, and the properties of information acquisition in equilibrium is summarized as follows:

Proposition 2. *Suppose* $(\delta^*(V), \sigma^*(\omega_3, V))$ *is a symmetric Bayesian Nash equilibrium under a given mechanism.*

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

(ii) There always exists a preference distribution F *such that* $\delta^*(V) > 0$ *under IA for* V *in some positive-measure set.*

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

The above result thus provides another perspective on strategy-proofness as a desideratum in market design: Strategy-proofness makes the school choice game easier to play by reducing the incentive to acquire information on others' preferences to zero.

3.5 Welfare Effects of Information Provision

The above results thus show students always have incentives to acquire information on own preferences and sometimes even information on others' preferences. However, such incentives do not always lead to successful information acquisition which depends on how costly it is. Moreover, what education authority should care about is the efficiency of school-student matching instead of information acquisition itself.

Therefore, we now analyze the effects of information provision by the education authority. We assume that information provision decreases the cost of information acquisition to zero, while the lack of it increases such cost to infinity. For simplicity, we focus on a special setting where everyone has the same ordinal (but different cardinal) preferences, similar to Abdulkadiroğlu et al.

(2011) and Troyan (2012). Realizing its restrictiveness, we relax this assumption in the experiment (cf. section 4).

In the same model as above, we start with a prior F and thus $G(P|F)$ such that after a P is drawn, it becomes everyone's ordinal preferences. Besides, every school is acceptable: $v_{i,s} > 0$ for all i and s. We use F_{v_s} to denote the marginal distribution of the cardinal preference for school s.

The education authority decides 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 $\bar{\omega}_{2,i}$ are 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 $\bar{\omega}_{3,i} = \left\{V^{\phi}_{-i}, V_{-i}\right\}$, i.e., they are either perfectly informative or completely uninformative.

Under each of the following four information structures , we investigate the *ex ante* welfare in equilibrium:

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

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

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

(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, or PI, while under UI, nobody knows the realization of ordinal preference, but everyone knows that the ordinal preference will be the same across students.

These four information structures can be considered as the outcomes of various policy interventions. When the education authority makes it difficult for students and parents to acquire information on schools, we are likely to be in the UI scenario. When it makes some information easy to access, students may find it costless to learn their ordinal preferences, and thus we are likely in the OI scenario. If all information on own preferences is readily available, we are likely to be in the CI scenario.

We are also interested in the PI scenario, which relates to the gaming part of school choice under a non-strategy-proof mechanism. From Proposition 2, individual students have incentives

to acquire information on others' preferences under the Immediate Acceptance mechanism. The literature has shown that this additional strategic behavior may create additional inequalities in access to public education. More precisely, if one does not understand the game and does not invest enough to acquire information on others' preferences, she may have a disadvantage when playing the school choice game. As a policy intervention, the education authority can choose to make this information public or at least easier to obtain. For example, in Amsterdam, strategies of applicants are published, and students are allowed to submit and revise their strategies upon seeing others' strategies (De Haan, Gautier, Oosterbeek and Van der Klaauw 2015).

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

Proposition 3. *Under DA, ex ante welfare of every student under any of the four information structures (UI, OI, CI, and PI) is the same and equals to* $\sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}(v_{i,s})$ *in any equilibrium.*

It therefore implies that there is no gain in terms of *ex ante* student welfare by providing more information under DA.

Proposition 4. *Under IA, we obtain the following ex ante student welfare comparisons in terms of Pareto dominance:*

(i) under UI and OI, the ex ante welfare of every student is $\sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}(v_{i,s})$*;*

(ii) CI $>$ *OI* $=$ *UI in symmetric equilibrium;*

(iii) PI $>$ *OI* $=$ *UI in symmetric equilibrium;*

(iv) However, PI can either Pareto dominate or be Pareto dominated by CI.

Therefore, it is always beneficial to provide more information on one's own cardinal preferences, although the effect of providing information on others' preferences is ambiguous. We use two examples to prove part (iv) in the proposition: in Appendix A, Section A.5.4 shows that PI can dominate CI in symmetric equilibrium; and the example in section A.5.5 shows the opposite. The reason that PI may be dominated is that PI sometimes leads multiple students of high type at a school to play mixed strategies in Nash equilibrium, instead of always top-ranking that school. This can happen because multiple high-type students are competing for the same school seats knowing

the presence of other high-type students, which makes top-ranking that school sub-optimal. As a result, sometimes that school is assigned to students of low type, which leads to welfare loss. This may not happen under CI in symmetric Bayesian Nash equilibrium, as high-type students do not know precisely the number of other high-type students in the game.

3.6 Possible Extensions

There are several potential extensions to our model. For example, one may allow students to acquire information on one's own and others' preferences simultaneously. Similarly, one may consider general signals to be acquired instead of the "hard news" in our model. However, given the lack of strategy-proofness and the role of cardinal utility under the IA mechanism, our results should be robust to such generalizations.

For our analysis of information provision, we assume that the lack of information provision makes the information acquisition cost infinite. While the education authority can manage to increase the cost, it is impossible to increase it to infinity. Our results from the lab experiment will show this limitation is less of a concern, as they call for more information provision by education authority (see Section 5).

Another restriction is the limited student preference domain. That is, students have the same ordinal preferences. This is not uncommon when studying 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 more general preference domain will be fruitful and leave this to future study.

Thus far, our theoretical analyses rely on the common knowledge of rationality assumption, which is implausible in the lab or the field. Besides, our school choice game is augmented with an information-acquisition stage, which makes the game even more complex. One may explore a theoretical model with students of heterogenous sophistication levels as in Pathak and Sönmez (2008), but it is necessary to restrict ourselves to a finite number of types of student sophistication.

Such considerations thus make it fruitful to study school choice with information acquisition in the laboratory, where we can allow students to have any type of behavioral responses in an otherwise controlled environment.

4 Experimental Design

We design our experiment to compare student incentives to acquire information under the IA and the DA mechanisms, and the subsequent welfare implications.

When designing the experiment, we relax the common-ordinal-preference assumption. Students may have different ordinal preferences. On the other hand, to simplify the possibly over-complex game, we design the payoff/preference distribution so that the two-step acquisition of information on own preferences is reduced to one step. That is, upon learning one's own ordinal preferences, a student learns her cardinal preferences as well, because there is only one possible realization of cardinal preferences consistent with such ordinal preferences. Therefore, the follow-up acquisition of cardinal preferences is not needed, which simplifies the game considerably.

4.1 The Environment

Following our theoretical analyses, we consider a simple environment with three students, $i \in \{1, 2, 3\}$, and three schools, $s \in \{a, b, c\}$. Each school has one slot and rank students with lotteries. Student cardinal preferences are presented in Table 1.

Table 1: Payoff Table for the Game Played in the Experiment

Students	$s = a$	$s = b$	$s = c$
1	1	0.1 with prob. 4/5; 1.1 with prob. 1/5	0
2	1	0.1 with prob. 4/5; 1.1 with prob. 1/5	0
3	1	0.1 with prob. 4/5; 1.1 with prob. 1/5	0

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

In this environment, the uncertainty comes from the realization of the value of school b, which can be either better or worse than school a. We thus relax the common-ordinal-preference assumption. *Ex ante*, the expected payoff of being assigned to b is 0.3 ($= \frac{4*0.1}{5} + \frac{1*1.1}{5}$), which is less than 1/3 of the payoff from school a for any student. In terms of welfare, the inefficiency has only one source, 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-$(1, 1.1, 0)$.

Under the assumption that every student is a risk-neutral expected-utility maximizer, we solve the (symmetric) equilibrium of the school choice game under either IA or DA for any given infor-

mation structure with or without costly information acquisition. Detailed derivations and results are presented in Appendix B. We also derive the results under the assumption that students are risk averse (Appendix C). While risk-averse students are often willing to pay less for information on either one's own or others' preferences, the same directional comparison between the IA and DA mechanisms maintains. In particular, we quantify student incentives to acquire information on own preferences or others' preferences. When we measure the incentive to acquire information on own preferences (denoted as "OwnValue"), every student is endowed with the same prior that everyone knows only the preference distribution not the realizations. For each student, we then calculate the payoff difference between playing the game knowing and not knowing own preferences, taking into account that the other two students may or may not know their own preferences. Therefore, we reach the theoretical predictions on student willingness to pay (WTP, henceforth) for OwnValue. Similarly, to measure student WTP for information on others' preferences (denoted as "OtherValue"), preference realizations are private information, and for any given student, we derive the payoff difference between playing the game with and without the knowledge of others' preferences.

4.2 Treatments

Following our theoretical analyses, 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): The two mechanisms are the major contenders for the school choice reforms. While DA is dominant strategy incentive compatible, the IA mechanism is manipulable. However, welfare comparisons of these two mechanisms are ambiguous depending on the information condition as well as parameters of the environment.

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

(iii) Free vs. costly information acquisition (within-subject): While the free information condition allows us to evaluate information provision policies, the costly information acquisition

condition reflects the reality. As this is implemented within-subject, we also take into account the order effect: For half of the sessions, subjects first experience free-info periods and then costly-info periods (denoted as "Free-Costly"); and for the other half of the sessions, subjects experience costly-info periods first and then free-info periods (denoted as "Costly-Free").

For the free information part, participants are provided the information regarding their own value (or others' values) for free. In comparison, in the costly information acquisition part, we use the Becker-Degroot-Marshak (BDM) mechanism (Becker, DeGroot and Marschak 1964b) to elicit participant's WTP for the relevant piece of information, own value or others' values for school B. Specifically, each subject is asked to enter her WTP for her own value (or others' values) in the interval of $[0, 15]$. The server collects WTP from each participant and generates a random number between $[0, 15]$ for each participant independently. If her WTP is greater than the random number, she finds out the relevant information and pay an amount equal to the random number; otherwise, she does not find out the information and pays zero. To facilitate participant understanding of the BDM mechanism, we use numerical examples in the instructions as well as quiz questions at the end of the instructions to illustrate how it works. Our instructions for the BDM mechanism is adapted from 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). Compared to the quadratic scoring rule (QSR), the BSR is more robust in the sense that it is incentive compatible under different risk attitude and even when the decision maker is not an expected utility maximizer (Schotter and Trevino 2013). Specifically, each subject submits a guess for the average WTP of the other two participants. The server computes the MSE, i.e., the square root of the squared difference between the guess and the actual average. We then randomly draw a number, R, uniformly from $[0, U]$. If the MSE is less than or equal to R, the subject gets a fixed prize of 5 points. Otherwise, she gets zero. Based on our pilot sessions, we find that 90% of the mean squared errors (MSEs) are at or below 49. Therefore, we use 49 as the upper bound for BSR, i.e., $U = 49$. The random number, R, is drawn independently for each subject, and for each round.

4.3 Experimental Procedures

In each experimental session, each participant is randomly assigned an ID number and is seated in front of a terminal in the laboratory. The experimenter then reads the instructions for the first ten periods aloud. Subjects have the opportunity to ask questions, which are answered in public. Subjects are then given 10 minutes to read the instructions at their own pace and to finish the review questions. After everyone finishes the review questions, the experimenter distributes the answers and goes over the answers in public. Afterwards, participants go through 10 periods of a school choice experiment. After the first ten periods, the experimenter reads the instructions for the second ten periods aloud and answers questions. Then the participants complete the review questions for the second half of the instructions, and go through the second ten periods of the school choice experiment.

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

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

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

The server collects rankings, draws the school B value for each subject, generates the tie-breaker, and allocates the schools, but withholds the outcomes till the end of the round.

(iii) Each participant acquires her value for school B, either for free or by paying a cost:

(a) For the *free* information treatment, we then provide each subject her own value for school B for free.

(b) For the *costly* information treatment only, we then use the BDM mechanism to elicit each participant's willingness to pay for learning her own value. 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 value or not.

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

(iv) Afterwards, each subject receives the following feedback on her screen:

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

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

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

ii. if no, only the number of other subject(s) in her group who also observe their value(s).

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

The server again collects the rankings, draws a new set of values for every subject, generates a new tie-breaker for each participant, and allocates the schools.

(vi) Each participant receives the following feedback on her screen:

(a) For the allocation before Own-Value acquisition: her ranking, her value, the tie-breaker, her allocation and her payoff; and

(b) For the allocation after Own-Value 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 making any school ranking decisions. The information provided for free or acquired in a costly manner using the BDM is the values for school B of the other two participants in her group.

In each treatment, each session lasts 20 periods with no practice rounds. Each session uses either costly or free information for the first ten periods, and another for the next ten periods. The order is counterbalanced for each treatment. At the end of 20 periods, we implement the Holt and Laury lottery choice procedure to elicit subject risk attitude. After telling each subject their payoff from the risk elicitation task, we offer an opportunity for subjects to acquire information on the realization of the lottery, again using the BDM mechanism. Their WTP for this useless information is a measurement of their *curiosity*, which we will use in subsequent analysis.

At the end of the experiment, each participant fills out a demographics and strategy survey on the computer. Each participant is paid in private at the end of the experiment. The experiment is programmed in z-Tree (Fischbacher 2007).

Table 2: Features of Experimental Sessions

Information to Be Acquired	Immediate Acceptance Mechanism		Deferred Acceptance	
OwnValue: Own Preferences	Free-Costly	3×12	Free-Costly	3×12
	Costly-Free	3×12	Costly-Free	3×12
OtherValue: Others' Preferences	Free-Costly	3×12	Free-Costly	3×12
	Costly-Free	3×12	Costly-Free	3×12

Notes: Each session has 10 periods with free information and another 10 periods with costly information. For any given treatment, sessions with free information periods first are denoted as "Free-Costly"; and the others with costly information first are denoted as "Costly-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. Each session consists of 12 subjects. The subjects are students from the University of Michigan. No one participates in more than one session. This gives us a total of 24 independent sessions and 288 participants. In addition, three sessions of DA-OtherValue (Free-Costly) treatment use a z-Tree program with a coding error in the second ten periods of the experiment, i.e., own value was not provided in the second ten periods, thus rendering the second half of the data useless. Nonetheless, we use the data from the first ten periods from these three sessions in our data analysis, since the instructions and program for the first half are both correct. If we include these three sessions, we have a total of 27 independent sessions with 324 participants. Our subjects are University of Michigan students, recruited using ORSEE (Greiner 2015)

Experimental instructions are included in the Appendix D, and the data are available from the authors upon request.

5 Experimental Results

Our theoretical analyses provide a set of benchmarks for our data analyses, where we focus on individual WTP and ROL, as well as efficiency.

We introduce several shorthand notations in presenting the results. First, let $x > y$, $x, y \in \{IA, DA\}$, denote that a measure under mechanism x is greater than the corresponding measure under mechanism y at the 5% significance level or less. Second, let $x \geq y$ denote that a mea-

sure under mechanism x is greater than the corresponding measure under mechanism y, but the comparison is not statistically significant at the 5% level.

5.1 Willingness to Pay for Information

Our theoretical analyses provide the WTP comparisons across mechanisms, summarized by the first two hypotheses. Hypothesis 1 is implied by Proposition 1, while Hypothesis 2 is implied by Proposition 2. For our experimental environment, we derive the WTP for risk-neutral and risk averse students in Appendices B and C, respectively, which shows the comparison between DA and IA.

Hypothesis 1 (WTP for OwnValue)**.** *Subject WTP to acquire information on own preferences under the IA mechanism is greater than that under the DA mechanism, while both being positive, i.e., IA $>$ DA $>$ 0.*

Result 1 (WTP for OwnValue)**.** *Subject WTP to acquire information on OwnValue under the IA mechanism is significantly greater than that under the DA mechanism, while both are positive.*

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

Furthermore, Figure 3 presents the time-series of the average WTP for OwnValue (upper panel) and OtherValue (lower panel) over time. Theoretical predictions for risk neutral students are presented as the red horizontal line(s). Note that, while the session-level average WTP for own value under the IA mechanism is mostly within the range predicted by theory, $[5.2, 8]$, that under under the DA mechanism is way above the risk neutral prediction of 0.67. We will explain this excess WTP in subsequent subsections.

We now investigate subject WTP to acquire information on others' preferences.

Hypothesis 2 (WTP for OtherValue)**.** *Subject WTP for OtherValue is zero under the DA mechanism regardless of risk attitude, whereas it is positive under the IA mechanism. Therefore, IA $>$ DA $=$ 0.*

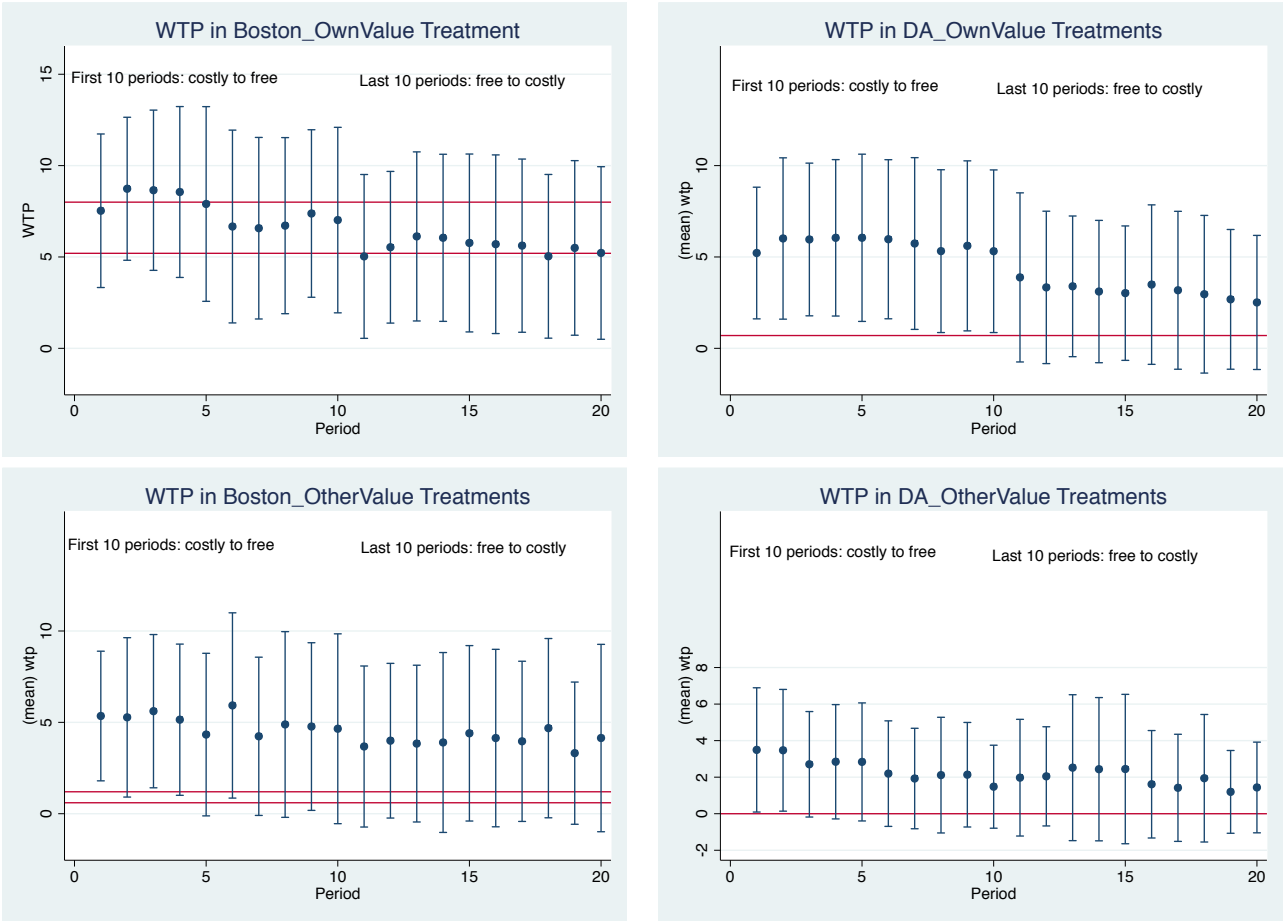

Figure 3: Average WTP for own value (upper panel) and others' values (lower panel) over time

Notes: Red horizontal lines denote theoretical predictions for risk neutral students. Error bars represent one standard deviation.

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

Treatment	All 6 Sessions	Free-Costly Sessions	Costly-Free Sessions	Theoretical Prediction[a]
IA OwnValue	6.56	5.56	7.57	[5.2, 8]
	(4.78)	(4.59)	(4.75)	
IA OtherValue	4.51	4.00	5.02	[0, 0.24]
	(4.55)	(4.55)	(4.49)	
DA OwnValue	4.44	3.16	5.72	0.67
	(4.38)	(4.05)	(4.33)	
DA OtherValue	2.21	1.90	2.52	0
	(3.16)	(3.25)	(3.04)	

Notes: The WTPs are measured in experiment points. There are 6 sessions under each treatment, and each session has 10 periods with free information and another 10 periods with costly information. For any given treatment, three sessions have free information periods first (denoted as Free-Costly); and the other three have costly information first (denoted as Costly-Free). Standard deviations are in parentheses and are calculated by treating each subject-period outcome as an observation. Therefore, for each treatment, there are in total 720 observations from 72 subjects, half of which are Costly-Free, while the other half are Free-Costly.

a. The predictions are for risk-neutral subjects, which are derived in details in Appendix B. Under the two treatments of IA, the predictions are intervals, because the WTPs depend on how many other students successfully acquire such information. As it is uncertain how many others can successfully acquire information *ex ante*, the prediction is thus an interval that takes into account all possibilities.

Result 2 (WTP for OtherValue)**.** *Subject WTP for others' values under the IA mechanism is significantly greater that that under the DA mechanism, but both are significantly different from zero: IA* $>$ *DA* > 0.

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

Furthermore, we find that, under either the IA or the DA mechanisms, subject WTP for OwnValue is significantly greater than that for OtherValue ($p = 0.01$, one-sided Wilcoxon rank-sum test), consistent with our theoretical predictions. While the WTP for OtherValue between the two mechanisms is in the direction predicted by theory, we again find that the WTP for OtherValue lie above the range predicted by theory, which is $[0, 0.24]$ under the IA and zero under the DA mechanism.

To explain the excess WTP for information, we first investigate determinants of WTP for information at the subject level (subsection 5.1.1) and then using panel regressions (subsection 5.1.2). Based on these investigations, we then decompose the excess WTP into various behavioral and

cognitive factors in subsection 5.2.

5.1.1 Determinants of WTP for Information: Subject-Average

To investigate the determinants of subject WTP, we analyze the subject-level average WTP in a Tobit model as follows:

$$\overline{WTP}_i^* = Controls_i + \varepsilon_i,$$
$$\overline{WTP}_i = \max\{0, min\{\overline{WTP}_i^*, 15\}\}.$$

The dependent variable is the subject-average WTP, $\overline{WTP}_i$, which is obtained by averaging over all WTPs of subject i. We use the Tobit model to take into account that $\overline{WTP}_i$ is censored between $[0, 15]$ among 18% of observations (43 out of 241 consistent subjects[4]) in our main sample.[5] For independent variables (or controls), we include the four treatment dummies (without constant term) and some demographic variables. Moreover, we consider effects of the following factors:

(i) Misunderstanding DA: This measure is negatively correlated with how well subjects understand the school choice game under DA. We take the periods where no information acquisition is involved and determine if a subject plays a dominated strategy in each period given her information. We then average over all such periods to obtain this variable. Unfortunately, there is no similar measures for IA, as there is no dominant strategy.

(ii) Costly-Free: When a session has a Costly-Free order, i.e., subjects play the game with costly information first, it leaves little room for subjects to learn and understand the game. Therefore, subjects may have sub-optimal WTP due to this greater cognitive load, which is reflected in Table 3.

(iii) Curiosity: We measure subject's curiosity by her WTP for the lottery realization in the Holt-Laury choice game. As such information is irrelevant to payoffs, this WTP thus provides a good measure of subject's "curiosity," or her general taste for information. One may ex-

[4]We define consistent subjects as those with only one switching point and who switch before the last decision in the Holt-Laury lottery choice task.

[5]As a robustness check, results from linear models, which are available in Table E15 in Appendix E, are both quantitatively and qualitatively similar.

pect that a more curious subject is willing to pay more for information on OwnValue or on OtherValue in the school choice game.

(iv) Risk aversion: Risk aversion is measured by the Holt-Laury lottery choice switching point, and more risk-averse subjects switch later. As we show in Appendix C, risk aversion decreases WTP for information. Indeed, 78% of our subjects are risk averse.[6]

Note that the first two factors provide measures of how well subjects understand the game, while the last two are behavioral factors. In Table 4, where we present the results, Column (1) includes the full sample, Columns (2)-(4) include only consistent subjects, and Column (4) further excludes observations with missing demographic information and includes additional controls.

While the treatment effects estimated from the Tobit model are largely consistent with Results 1 and 2, we observe the following additional findings. First, the more a subject misunderstands DA (i.e., playing dominated strategies more often under DA), the more she is willing to pay for information. Second, measured by her WTP for useless information (i.e. the lottery realization in the Holt and Laury lottery choice game), subject curiosity is positively correlated with WTP for information on school values. Third, the timing for introducing the costly information acquisition game matters. When subjects have to learn both the mechanism and information acquisition in the first ten periods, i.e., Costly-Free $= 1$, subject WTP is substantially higher. This is consistent with previous experimental findings that higher cognitive load can cause suboptimal play (Bednar, Chen, Liu and Page 2012). Lastly, consistent with the theoretical prediction, risk aversion decreases WTP, but the results are not robust in terms of statistical significance.

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

To explore within-subject across-period variations, we next take the subject-period observations and use panel data methods. Similar to the analysis of subject-average WTP (Table 4), we take

[6] Among 241 subjects who have one switching point and make rational choices in the Holt-Laury game, 7% (16/241) are risk loving; 16% (38/241) are risk-neutral; and 78% (187/241) are risk averse.

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

	(1) Full Sample	(2) Sub-sample 1	(3) Sub-sample 1	(4) Sub-sample 2
IA_OwnValue	6.45***	6.26***	5.22***	5.77***
	(0.56)	(0.57)	(1.10)	(1.74)
IA_OtherValue	4.32***	4.05***	3.46***	3.91**
	(0.62)	(0.72)	(1.21)	(1.91)
DA_OwnValue	4.13***	3.78***	2.94***	3.60**
	(0.71)	(0.82)	(1.07)	(1.71)
DA_OtherValue	1.47***	1.01**	0.91	1.98
	(0.45)	(0.47)	(1.13)	(1.79)
Misunderstanding DA[a]			6.85***	6.29***
			(2.02)	(2.21)
Curiosity			0.34***	0.33***
			(0.05)	(0.04)
Costly-Free			1.88***	1.87***
			(0.45)	(0.36)
Risk Aversion			-0.28**	-0.20
			(0.13)	(0.13)
Female				-0.87*
				(0.48)
Graduate Student				-0.72
				(0.51)
Black				-1.09*
				(0.59)
Asian				-1.75**
				(0.69)
Hispanic				-0.83
				(0.57)
N	288	241	241	233

Notes: Outcome variable is subject-level average WTP for information. There are 42 (out of 241, 17%) subjects with average WTP $= 0$ and one subject with WTP $= 15$. Columns (2)-(4) include only consistent subjects. Column (4) further excludes observations with missing age/gender/ehnicity information and includes other controls: age, ACT score, SAT score, dummy for ACT score missing, dummy for SAT score Missing, and dummy for degree missing.

a. Standard errors clustered at the session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

b. "Misunderstanding DA" is defined as the percentage of times when the subject played dominated strategies in the OwnValue or OtherValue treatment of DA in periods without information acquisition. Mean $= 0.09$, standard deviation $= 0.14$ among all subjects ($n = 144$) played the information acquisition game under DA. Only periods without information acquisition, i.e., with no information or free information provision, are considered. This variable equals to zero for the IA treatments, because dominant strategies are not well defined under the IA mechanism.

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 for periods (within each session). Given the non-linear nature of the Tobit model, we cannot consistently estimate α_i as subject fixed effects with a short panel (ten periods), which leads us to focus on a random effects Tobit model. For the above specification, we run them with all four treatments pooled as well as treatment-by-treatment. [7]

In terms of controls, we include $high\_B \times IA\_OtherValue_{i,t}$ which equals to one if in period t subject i has a high valuation of School B ($= 110$) under the treatment IA OtherValue. Otherwise, it is equal to 0. $high\_B \times DA\_OtherValue_{i,t}$ is similarly defined for subjects under the DA OtherValue treatment. For the other treatments, IA or DA OwnValue, it is impossible to define such a variable, because subjects do not know the valuation of School B in advance. Our theory predicts that the coefficient on $high\_B \times IA\_OtherValue_{i,t}$ should be positive, while the one on $high\_B \times DA\_OtherValue_{i,t}$ should be zero (see Appendix B.6). This is confirmed in our results.

One key dependent variable is $WTP\_guess_{i,t}$, i's guess of her opponents' average WTP in period t. Our theory predicts that i's own WTP should be negatively correlated with $WTP\_guess_{i,t}$, albeit weakly (see Appendices B.5 and B.6). [8]

Other controls include accumulated wealth at the beginning of the period, period (i.e., a linear time trend), period in Free-Costly sessions, and lagged average WTP of other players. Depending on specification, sometimes we also include lagged guess of others' WTP ($WTP\_guess_{i,t-1}$).

Results are shown in Table 5, where the three columns only differ in terms of inclusion or exclusion of $WTP\_guess_{i,t}$ and its lag.

There are a few notable findings in the table. First, the coefficient on $WTP\_guess_{i,t}$ is significantly positive, contrary to what the theory predicts. Second, subjects learn over time in the sense

[7] For robustness checks, we present the respective fixed effects (Table E16) and random effects panel regressions (Table E17) in Appendix E.

[8] One might be concerned with the endogeneity of $WTP\_guess_{i,t}$. That is, there might be some common shocks in period t which make i's $WTP_{i,t}$ and $WTP\_guess_{i,t}$ higher. We address this with an IV approach in Tables E16 and E17 in Appendix E, from which we interpret that the endogeneity issue is not a big concern.

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

	(1)	(2)	(3)
IA_OwnValue	2.37**	4.06***	2.28**
	(0.91)	(0.93)	(0.96)
IA_OtherValue	0.48	1.20	0.43
	(1.05)	(1.20)	(0.91)
DA_OwnValue	2.11*	2.31*	2.08**
	(1.14)	(1.28)	(1.03)
high_B $\times$ IA_OtherValue	3.38**	3.47**	3.38***
	(1.41)	(1.35)	(1.11)
high_B $\times$ DA_OtherValue	-0.49	-0.42	-0.50
	(1.03)	(0.83)	(1.05)
Accumulated wealth	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Successfully acquired info in $t-1$	0.30	0.13	0.27
	(0.30)	(0.28)	(0.27)
Period	-0.20**	-0.27**	-0.20**
	(0.09)	(0.11)	(0.08)
Period $\times$ Free-Costly	0.11	0.12	0.11
	(0.10)	(0.11)	(0.09)
Average WTP of others in $t-1$	-0.04	0.12***	-0.04
	(0.03)	(0.03)	(0.03)
$WTP\_guess_{i,t-1}$: Guess of others' WTP in $t-1$	-0.04	0.23***	
	(0.08)	(0.06)	
$WTP\_guess_{i,t}$: Guess of others' WTP in t	0.91***		0.90***
	(0.11)		(0.09)
Misunderstanding DA	7.06***	8.86***	7.04***
	(2.69)	(3.08)	(2.70)
Curiosity	0.38***	0.47***	0.38***
	(0.06)	(0.07)	(0.06)
Costly-Free	1.78*	2.93**	1.76*
	(1.07)	(1.27)	(1.05)
Risk Aversion	-0.29	-0.27	-0.29
	(0.20)	(0.22)	(0.19)
# of observations	2097	2097	2097
# of subjects	233	233	233

Notes: The regression sample is the same as that in Column (5) in Table 6. There are 233 subjects each of whom has 9 observations from 9 periods. Estimates are form 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, and dummy for degree missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

that they lower their excess WTP. Third, the coefficients on the four factors (Misunderstanding DA, Curiosity, Costly-Free, and Risk Aversion) are similar to those in Table 4, although the one on Risk Aversion is never significant.

5.2 Decomposition of the WTP

From our above analyses of the determinants of WTP, several results are contrary to the theoretical predictions. Indeed, given the complexity of the game, it is unlikely that subjects would behave exactly as the model predicts. This subsection then investigates how much of the observed WTP can be explained by the factors that we have found important in the above analyses. Specifically, we decompose WTP treatment-by-treatment by quantifying the effects of the following factors:

(i) Cognitive load: Playing Costly-info periods before Free-info periods.

(ii) Learning: How WTP changes over periods.

(iii) Misunderstanding DA: How WTP is correlated with playing dominated strategies under DA.

(iv) Conformity: How own WTP is affected by the expectation of others' WTP.

(v) Curiosity: How WTP is correlated with WTP for useless information.

(vi) Risk aversion: How WTP is correlated with risk aversion.

Again, the first four factors provide some measure of how well subjects understand the game, which includes "Conformity." If one understands the game perfectly, her own WTP for information should be lower when she expects others paying more. The above analyses however show the contrary, which indicates that subjects do not understand the game very well.

To perform the decomposition, we first estimate the Tobit model as in equation system (1) for each treatment separately with sessions under that treatment only (Columns (1) - (4) in Table 6). It is more flexible to do so, as this allows each factor has a different effect in a given treatment. Indeed, coefficients of some of the key variables change significance level (e.g., "Period" and "Risk Aversion") or even switch signs across treatments (e.g., "Period $\times$ Free-Costly" and "Costly-Free"). As a comparison, results based on the pooled regression are in Column (5)), which is the same regression as in Column (3) of Tabel 5.

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)
high_B × IA_OtherValue		3.95***			3.38***
		(1.14)			(1.19)
high_B × DA_OtherValue				-0.51	-0.50
				(0.97)	(1.03)
Accumulated wealth	-0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Successfully acquired info in $t-1$	0.18	-0.36	0.45**	0.40	0.27
	(0.49)	(0.62)	(0.20)	(0.39)	(0.26)
Period	-0.20	-0.16	-0.26	-0.43	-0.20**
	(0.13)	(0.14)	(0.17)	(0.28)	(0.09)
Period × Free-Costly	0.22***	0.24	0.07	-0.16	0.11
	(0.07)	(0.20)	(0.10)	(0.20)	(0.09)
Average WTP of others in $t-1$	-0.05**	0.05	-0.11**	-0.01	-0.04
	(0.02)	(0.07)	(0.05)	(0.05)	(0.02)
Guess of others' WTP in t	0.88***	0.75***	0.95***	1.04***	0.90***
	(0.17)	(0.19)	(0.11)	(0.20)	(0.09)
Misunderstanding DA			6.55	8.02	7.04***
			(4.54)	(5.83)	(2.45)
Curiosity	0.31***	0.63***	0.26	0.46***	0.38***
	(0.12)	(0.18)	(0.21)	(0.09)	(0.07)
Costly-Free	-0.36	2.72	2.98	1.47	1.76
	(2.73)	(3.15)	(3.47)	(5.41)	(1.11)
Risk Aversion	-1.49***	-0.31	-0.41	0.37**	-0.29
	(0.41)	(0.39)	(0.37)	(0.18)	(0.20)
IA_OwnValue					2.28**
					(0.92)
IA_OtherValue					0.43
					(0.99)
DA_OwnValue					2.08**
					(1.00)
# of observations	549	495	558	495	2097
# of subjects	61	55	62	55	233

Notes: Estimates are form random effects panel Tobit models separately for each treatment and then for all treatments pooled. The regression sample in Column (5) is the same as the one in Table 5. 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, and dummy for degree missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Based on these estimated coefficients, Table 7 presents the decomposition of subject WTP for information.[9] In short, the six factors can explain the majority of the observed WTP; without these factors, WTP for information is close to the theoretical prediction, except in the IA OwnValue treatment, where subjects have insufficient WTP after controlling the factors. We now discuss the effects of each factor in details.

Table 7: Decomposition of Subject WTP for Information

	(1) **IA OwnValue**	(2) **IA OtherValue**	(3) **DA OwnValue**	(4) **DA OtherValue**
WTP: data	6.49	4.29	4.30	1.78
	(4.86)	(4.67)	(4.30)	(2.81)
Model prediction[a]	6.40	4.18	4.26	1.69
	(3.16)	(2.98)	(2.97)	(1.95)
(i) Cognitive load[b]	-0.70	0.36	0.99	0.45
	(0.74)	(0.49)	(1.02)	(0.66)
(ii) Learning over periods[b]	0.30	0.10	0.63	0.71
	(0.46)	(0.34)	(0.52)	(0.81)
(iii) Conformity[b]	4.10	2.10	2.78	1.30
	(1.99)	(1.62)	(2.13)	(1.62)
(iv) Misunderstanding DA[b]			0.41	0.25
			(0.70)	(0.62)
(v) Curiosity[b]	1.36	1.69	0.75	0.57
	(1.35)	(2.04)	(1.01)	(1.17)
(vi) Risk aversion[b]	-1.43	-0.24	-0.51	0.31
	(1.28)	(0.28)	(0.48)	(0.37)
Total[c]	3.60	3.17	3.73	1.68
	(2.49)	(2.79)	(2.92)	(1.95)
Explained by all other factors[d]	2.88	1.13	0.58	0.11
	(3.93)	(3.82)	(2.98)	(2.05)
Theoretical prediction[e]	[5.2,8]	[0, 0.24]	0.67	0
# of Observations	549	495	558	495
# of Subjects	61	55	62	55

Notes: Decompositions are based on 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 be in $[0, 15]$.
b. The WTP explained by the corresponding factor is the difference between the model prediction with the factor and that without the factor. The former is predicted with the current values of all variables; and the latter is calculated by setting the relevant variable value to zero (for factors "Cognitive load," "Conformity," "Misunderstanding DA," 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 period," we set "Period" to be the last period, i.e., "Period" $= 10$).
c. "Total" 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. "Explained by other factors" is the difference between the observed WTP and the total WTP explained by the six factors.
e. These are the theoretical predictions for risk neutral subjects.

[9] As a robustness check, decomposition based on pooled regressions (Column (5) in Table 6) are presented in Table E18 in Appendix E, which shows similar results.

(i) Cognitive Load

We focus on the cognitive load caused by Costly-Free (i.e., playing the costly information acquisition before the free information game). From our regressions, the Costly-Free order is associated with 1.76 points of extra WTP in every period on average among all treatments (Table 6, Column (5)), while the effect is not present in the IA OwnValue treatment (Table 6, Column (1)). Moreover, the order also affects learning over periods based on the coefficients on "Period" and "Period $\times$ Free-Costly": In sessions with Costly-Free, learning over periods decreases the WTP faster than those with Free-Costly. In the 10th period, WTP in Costly-Free sessions is reduced by 1.80 points relative to the 1st period, while the reduction is only 0.81 in sessions with Free-Costly.

To quantify the effect of cognitive load, we consider the counterfactual of replacing Costly-Free by Free-Costly. That is, we set all "Costly-Free" to zero and all "Period $\times$ Free-Costly" equal to "Period." The effect of cognitive load is then measured by the difference between two predictions: The model prediction based on current variable values and the prediction under the counterfactual. Both predictions are censored to guarantee that the predicted WTP is between 0 and 15. Table 7 reports average effect along with its standard deviation. The presence of cognitive load increase WTP by 0.36 to 0.99 points, except in the IA OwnValue treatment where the sign is opposite.

(ii) Learning over periods

Learning here refers to the fact that WTP changes over periods, and moreover the changes are in the direction of correcting the excess WTP. We consider the counterfactual of replacing behavior in periods 2-9 by that in period 10. Note that period 1 is not included in our regression, as we do not have the lagged values for any variable in that period. Again, the same censoring as above is adopted when predicting WTP. The estimated effect of learning, which is the difference between the prediction with the observed variable values and the prediction under the counterfactual, is from 0.10 to 0.70 points of WTP.

(iii) Conformity

Conformity here summarize that subjects positively respond to their beliefs about others'

behavior, WTP_guess. When expecting others to pay one extra point, each subject pays additional 0.90 points on average (Column (5) in Table 6).

Although the theory predicts a negative correlation between own WTP and WTP_guess, we consider the counterfactual where the correlation is zero. That is, in the counterfactual, there is no effect of WTP_guess on WTP. After calculations similar to above, it turns out that this can explain 1.30 to 4.10 points of the observed WTP, or 49% to 73%, which is by far the single most important factor.

(iv) Misunderstanding DA

Misunderstanding DA is the fraction of times that a subject plays dominated strategies in periods of the DA treatments without information acquisition. This is only defined for the two DA treatments, and Table 6 shows that it significantly increases WTP.

Similarly, to quantify its effect in the two DA treatments, we predict the WTP under the counterfactual without any misunderstanding of the mechanism (i.e., setting the variable to zero), and then calculate the difference between the model prediction with the observed variable values and the prediction under the counterfactual. Columns (3) and (4) in Table 7 show that the effect is 0.41 under DA OwnValue and 0.25 under DA OtherValue.

(v) Curiosity

The WTP to pay for useless information is considered to be a measure of curiosity. From the regression in Column (5) of Table 6, 1 point increase in WTP to pay for useless information is associated with 0.38 additional points of WTP in each period.

We consider the counterfactual where WTP for information in the school choice game is not associated with curiosity, and thus we set the coefficient on Curiosity to zero. After similar calculations, curiosity explains 0.57 to 1.69 points of the observed WTP.

(vi) Risk Aversion

Risk aversion is measured by subject's switching point in the Holt-Laury lottery choice game. Our theory predicts that risk-averse subjects pay less for information, which is consistent with our findings. On average, being more risk averse is correlated with lower WTP, although

insignificantly (Table 6, Column (5)). However, this correlation is heterogeneous across treatments, and it becomes insignificantly positive in the DA OtherValue treatment.

The counterfactual we consider is to have every subject risk neutral (i.e., switching at the 5th choice in the Holt-Laury game), which requires us to change about 78% of subjects from risk averse to risk neutral. The measured effect is that risk aversion decreases WTP by 0.51 to 1.43 points, except that it increases WTP in the DA OtherValue treatment.

Taking the above six factors together, the total explained WTP ranges from 1.68 to 3.73 points, which amounts to 55% to 94% of the observed WTP. The rest of the observed WTP that are explained by factors other than the six becomes very close to the theoretical prediction for the two DA treatments. However, that part is much below the theoretical prediction in the IA OwnValue treatment (2.88 versus [5.2, 8]), while being substantially above the theoretical prediction in the IA OtherValue (1.13 versus [0, 0.24]). These results on IA also imply the difficult of playing the game under IA.

5.3 Rank Ordered List

We now investigate the effects of information provision and the effects of information acquisition on individual strategies. In Appendix B, we derive the equilibrium strategies under various information structure some of which are augmented with information acquisition.

The first information structure is UI (UnInformed: no one is informed about her valuation of School B), under which we have the following hypothesis based on our theoretical results.

Hypothesis 3 (ROL: UI)**.** *A risk neutral player submits a ROL of ABC as a dominant strategy under either IA or DA.*

Result 3 (ROL: UI)**.** *When subjects play the game under UI, more subjects play BAC instead of ABC under IA than under DA. Under IA, ABC accounts for 72% of the ROLs, followed by BAC 25%; under DA, 90% play ABC, and 8% submit BAC. The rest plays some other strategies. A session-level Wilcoxon rank-sum (or Mann-Whitney) test rejects the hypothesis that the ABC or the BAC strategy is played equally often under IA and DA (both p-values* < 0.01*).*

Note that the strategy ABC is not a dominant strategy for subjects who are sufficiently risk-averse under IA, which implies that ABC may be less played by more risk-averse subjects. On

the contrary, after categorizing the subjects into two almost-equal-sized groups by risk aversion measured in the Holt-Laury lottery choice game, we find that ABC (BAC) are played by 71% (27%) of the less risk-averse subjects who switch choices before or at the 6th Holt-Laury lottery, while ABC (BAC) are played by 77% (21%) of the rest subjects who are more risk averse. This finding is consistent with Klijn et al. (2012) who also show that more risk-averse subjects are not more likely to play "safer" strategies under IA.

Recall that another information structure considered is CI (Cardinally Informed: everyone is informed about her own valuation of School B but not others' valuations).[10] Also note that under the treatment of OwnValue, one can acquire information on her own preferences by paying some costs, which results in a game with some informed players and some uninformed. The next hypothesis is about the informed players' strategies. When testing the next hypotheses, the reported p-value is from the session-level Wilcoxon rank-sum (or Mann-Whitney) test, unless noted otherwise.

Hypothesis 4 (ROL: CI and Acquiring OwnValue)**.** *When a subject knows her own preferences but does not know others' preferences, it is a BNE (dominant strategy) to submit a ROL truthfully under IA (DA), regardless of the number of opponents who know their own preferences.*

Result 4 (ROL: CI and Acquiring OwnValue)**.** *Under IA, when the valuation of School B is* 10*, informed subjects are truth-telling at a similar rate –* 87% *with free information,* 88% *with costly acquired information. When the valuation of School B is* 110*, there are more subjects playing BAC with acquired information (*90%*) than those with free information (*85%*). However, this difference is not significant (p-value* 0.52*).*

Under DA, when the valuation of School B is 10*, informed subjects are truth-telling at insignificantly different rates –* 95% *with free information, 91% with costly acquired information (p-value* $= 0.87$*). When the valuation of School B is* 110*, however, there are significantly more subjects playing BAC with acquired information (95%) than that with free information (*79%*) (p-value* $= 0.01$*).*

Lastly, we consider information structure PI (Perfectly Informed: valuations of School B are

[10] By design, in this experiment, CI is equivalent to OI (Ordinally Informed: everyone is informed of her ordinal preferences but not others).

common knowledge) as a result of information provision and also the OtherValue treatment. Our theoretical prediction regarding the ROL is summarized below.

Hypothesis 5 (ROL: PI and Acquiring OtherValue)**.** *When a subject knows both her own preferences and the preferences of her two opponents, it is a dominant strategy to rank the schools truthfully under DA; the optimal strategy under IA for low-B-valuation subjects report truthfully, while that for high-B-valuation subjects depends on the preference profile as well as the number of informed players.*

Result 5 (ROL: PI and Acquiring OtherValue)**.** *Under DA, when the valuation of School B is* 10*, informed subjects are truth-telling at insignificantly different rates –* 92% *with free information,* 84% *with costly acquired information (p-value* $= 0.29$*). When the valuation of School B is* 110*, there are fewer subjects playing BAC with acquired information (*75%*) than that with free information (*91%*). The difference is again insignificant (p-value =* 0.86*), partly because there are only 16 subjects successfully acquired information.*

Under IA, when the valuation of School B is 10*, informed subjects are truth-telling at a similar rate – 86% with free information,* 84% *with costly acquired information. When the valuation of School B is* 110*, there are insignificantly more subjects playing BAC with acquired information (*85%*) than that with free information (*81%*) (p-value* $= 0.75$*).*[11]

We consider our above results to be consistent with the theoretical predictions. Moreover, the only incidence where costly acquired information and freely provided information have significant effects is that acquired information on OwnValue makes subjects more likely to play dominant strategy.

5.4 Welfare Analysis

We now turn to welfare analysis by investigating the effects of information provision and the effects of information acquisition. Welfare is measured on two dimensions. The first is the payoffs that subjects receive in the experiment; the other is the efficiency of the allocation outcome. An efficient allocation is such that (i) the high-B subject, whose School-B valuation is 110, is matched with School B, if there is any such subject; and (ii) otherwise every allocation is efficient.

[11] One may be tempted to investigate subjects' strategies conditional the preference profile of all subjects. This however makes the samples very small, especially among those who successfully acquired information (61 in total).

5.4.1 Effects of Information Provision

We start with the theoretical results on information provision, where some information is freely provided while the rest is impossible to be acquired. In Appendix B, we derive the equilibrium strategies and outcomes under various information structure. Similar to the theoretical analyses, we first consider three information structures without information acquisition: (i) UI (UnInformed), (ii) CI (Cardinally Informed),[12] and (iii) PI (Perfectly Informed). The *ex ante* welfare for risk-neutral subjects and the allocation efficiency are summarized in Table 8, based on which we derive the following three hypotheses:

Table 8: Effects of Information Provision on Subject Welfare and Allocation Efficiency

	IA			DA		
	UI	CI	PI	UI	CI	PI
Subject *ex ante* payoff	43.33	52.93	52.93	43.33	48.67	48.67
Fraction of efficient allocations	68%	100%	100%	68%	87.2%	87.2%

Notes: The payoff and allocation efficiency are derived in symmetric equilibrium for risk neutral agents in Appendix B. UI = UnInformed (No one is informed about her valuation of School B); CI = Cardinally Informed (Everyone is informed about her own valuation of School B but not others' valuations); PI = Perfectly Informed (Valuations of School B are common knowledge).

Hypothesis 6 (Efficiency: IA)**.** *With free information provision, the ex ante subject welfare and the fraction of efficient allocations under IA should follow the order of UI $<$ CI $=$ PI. Moreover, the allocation is always efficient under either CI or PI.*

Hypothesis 7 (Efficiency: DA)**.** *With free information provision, the ex ante subject welfare and the fraction of efficient allocations under DA should follow the order of UI $<$ CI $=$ PI. However, the allocation is not always efficient under either CI or PI.*

Hypothesis 8 (Efficiency: Comparing IA with DA)**.** *With free information provision, in terms of the ex ante subject welfare and the fraction of efficient allocations, IA $=$ DA under UI, IA $>$ DA under either CI or PI.*

Result 6. *(i) No difference in subject payoffs or allocation efficiency between DA and IA under UI; (ii) Information provision transforming UI into CI improves subject payoffs and allocation efficiency of both DA and IA, but more so for IA; (iii) Information provision transforming CI to PI*

[12] By design, in this experiment, CI is equivalent to OI (Ordinally Informed: everyone is informed of her ordinal preferences but not others).

does not improve the performance of DA or IA; and (iv) Outcomes of IA under CI or PI are closer to the efficient outcome relative to those of DA under CI or PI.

These conclusions are drawn based on the statistics in Table 9. For each treatment, we focus on the same subjects who in each period play a pair of the school choice games under the no-info and the free-info scenarios, where the order of the two scenarios are randomized in each period. For each session, we also weight the data to account for small sample variation in the probabilities of having a high valuations of School B.

Given the design of the experiment, we perform both within-treatment as well as between-treatment tests. For instance, to test the effect of information provision under IA that changes information structure from UI to CI, we use a within-treatment test: the Wilcoxon matched-pairs signed-ranks test. This is feasible because the same group of subjects play the game under both UI and CI. An example for between-treatment tests is to test if IA and DA reach the same level of efficiency under CI, for which we use the Wilcoxon rank-sum (or MannWhitney) test for two independent samples, given that no individual experiences the two treatment.

In the result, parts (i)-(iii) are directly from the test results in the table. For Part (iv), outcomes of IA under CI or PI achieve 93-96% of maximum payoffs and result in efficient allocations among 89-94% of all games; as a comparison, outcomes of DA under CI or PI on average achieve only 87-89% of maximum payoffs and result in efficient allocations among 81-84% of all games.

5.4.2 Effects of Information Acquisition

We now turn to the effects of costly information acquisition. Similar to our theoretical model, the information acquisition technology results in a endogenous probability receiving the "hard news." Therefore, when entering the school choice game, there are probably some informed subjects as well as some uninformed subjects. We thus expect the outcomes to be between no information and free information provision, which leads us to three hypotheses similar to the above.

Hypothesis 9 (Efficiency: IA)**.** *With costly information acquisition and not taking into account its costs, the ex ante subject welfare and the fraction of efficient allocations under IA should follow the order of UI $<$ (Acquiring OwnValue) and CI $=$ (Acquiring OtherValue). Moreover, the allocation is always efficient under either CI or (Acquiring OtherValue).*

Table 9: Effects of Information Provision on Payoffs and Allocation Efficiency

	Payoff	Fraction(Efficient Allocation)
IA OwnValue (# observations = 720)		
UI	42.51	0.69
	(51.12)	(0.49)
CI	50.67	0.94
	(52.52)	(0.29)
	H_0: UI = CI; H_1: UI <CI	
p-value	0.01	0.01
IA OtherValue (# observations = 720)		
CI	49.13	0.89
	(51.90)	(0.35)
PI	49.12	0.91
	(52.20)	(0.34)
	H_0: CI = PI; H_1: CI $\neq$ PI	
p-value	0.92	0.35
DA OwnValue (# observations = 720)		
UI	42.96	0.71
	(48.93)	(0.48)
CI	47.22	0.84
	(54.92)	(0.43)
	H_0: UI=CI; H_1: UI<CI	
p-value	0.04	0.04
DA OtherValue (# observations = 1080)		
CI	45.90	0.81
	(49.96)	(0.39)
PI	46.09	0.81
	(49.53)	(0.40)
	H_0: CI = PI; H_1: CI $\neq$ PI	
p-value	0.86	0.86
Comparison between IA & DA		
	H_0: (IA UI) = (DA UI); H_1:(IA UI) $\neq$ (DA UI)	
p-value	1.00	1.00
	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
	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 fraction of efficient allocation by information structure. It also presents p-values for the Wilcoxon matched-pairs signed-ranks tests or for 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 high valuations of School B equals to $1/5$.

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

Hypothesis 10 (Efficiency: DA)**.** *With costly information acquisition and not taking into account its costs, the ex ante subject welfare and the fraction of efficient allocations under DA should follow the order of UI < (Acquiring OwnValue) and CI = (Acquiring OtherValue). However, the allocation is not always efficient under either CI or (Acquiring OtherValue).*

Hypothesis 11 (Efficiency: Comparing IA with DA)**.** *With costly information acquisition and not taking into account its 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 7. *When we do not taking into account the information acquisition cost, the results are: (i) Acquiring OwnValue improves subject payoffs and allocation efficiency for both IA and DA, but more so for IA; (ii) Acquiring OtherValue does not affect improves subject payoffs and allocation efficiency of IA or DA; (iii) and (iv) Outcomes of IA under OwnValue or OtherValue are closer to the efficient outcome relative to those of DA under OwnValue or OtherValue.*

These conclusions are drawn based on the statistics in Table 10. Similar to before, for each treatment, we focus on the same subjects who in each period play a pair of the school choice games in both the no-information and the costly-information scenarios, where the order of the two scenarios are randomized in each period. This design of the experiment allows us to perform both within-treatment as well as between-treatment tests.

Parts (i) and (ii) are the test results in the table, while Part (iii) is concluded by some simple calculations. With acquisition of information on OwnValue, IA on average achieves 89% of maximum payoffs and efficient allocations among 83% of all games; as a comparison, while DA leads to 80% of maximum payoffs and efficient allocations among 73% of all games. Similarly, with acquisition of information on OtherValues, IA obtains 97% of maximum payoffs and 96% efficient allocations, whereas the numbers are only 87% and 82% under DA.

Such welfare effects are certainly because some subjects manage to acquire the information. Table 10 also presents the fraction of times each subject successfully acquiring the information as well as their expressed WTP, which are positively correlated with each other due to our experimental design. In the IA OwnValue treatment, on average 44% of subjects obtain the information in each period, which is exactly the ratio of the average WTP (6.56) to the upper bound of WTP (15). Subjects acquire the information less often in other treatments, ranging from 14% in the DA

Table 10: Effects of Information Acquisition on Payoffs and Allocation Efficiency

	Payoff	Fraction(Efficient Allocation)	Info Acquired	WTP	Costs Paid
IA OwnValue (# observations = 720)					
UI	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)
	H_0: UI = (Acquiring OwnValue); H_1: UI $<$ (Acquiring OwnValue)				
p-value	0.01	0.01			
IA OtherValue (# observations = 720)					
CI	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)
	H_0: CI = PI; H_1: CI $\neq$ PI				
p-value	0.25	0.25			
DA OwnValue (# observations = 720)					
UI	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)
	H_0: UI = (Acquiring OwnValue); H_1: UI $<$ (Acquiring OwnValue)				
p-value	0.06	0.06			
DA OtherValue (# observations = 720)					
CI	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)
	H_0: CI = (Acquiring OtherValue); H_1: CI $\neq$ (Acquiring OtherValue)				
p-value	0.92	0.92			
Comparison between IA & DA					
	H_0: (IA Acquiring OwnValue) = (DA Acquiring OwnValue) H_1: (IA Acquiring OwnValue) $>$ (DA Acquiring OwnValue)				
p-value	0.00	0.00			
	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 reports the means and standard deviations (in parentheses) of payoffs, fraction of efficient allocation by information structure, fraction of having successfully acquired info, WTP for information, and costs of information acquisition. It also presents p-values for the Wilcoxon matched-pairs signed-ranks tests or for the 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 to $1/5$. All tests are performed with the session averages of payoffs or efficiency.

OtherValue treatment to 30% in the DA OwnValue treatment.

To take into account the cost for information acquisition, we have to make some assumptions on the acquisition technology. The one used in the experiment can be considered as providing a lower bound of such costs, because one only pays a half of her WTP in expectation conditional on information being acquired. Table 10 shows that the actual costs paid by subjects in each treatment, which is on average 22-34% of the WTP.[13] Under this technology, costly information acquisition is welfare-improving in the IA OwnValue treatment; it increases the average payoff for each subject in each period by 2.3 points. However, in the IA OtherValue treatment, it is essentially welfare-neutral: costly information acquisition brings merely 0.09 points in profits. Moreover, in both treatments of DA, costly information acquisition is welfare-decreasing.

An alternative technology of information acquisition, which can be considered as an upward correction of the costs to the previous assumption, is to assume that everyone has to pay her WTP if successfully acquiring information, while the probability of acquiring information is $WTP/15$.[14] This alternative can be implemented with the same design as in the experiment, as it keeps the incentive properties unchanged. It is equivalent to assume that the random lottery, which is drawn between $[0, 15]$ to determine if one receives the information or not, is always turns out to be WTP, and thus the subject always pays WTP when successfully acquiring the information. This leads to the same WTP as well as the same probability of acquiring information, which allows us to use the experimental data to calculate welfare.[15] Under this assumption, information acquisition is welfare-decreasing except in the IA OwnValue treatment, based on the results in Table 10. The net loss for each subject in each period ranges from 1.23 in the IA OtherValue treatment to 1.57 in the DA OwnValue, while the net gain in the IA OwnValue is only 0.22.

5.4.3 Total Welfare Effects of Information Provision

Based on the above results, we are now ready to calculate the welfare effects of policies of information provision. More specifically, we focus on the following three regimes, whose welfare is

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

[14] The expected cost is then $w^2/15$, where w is the WTP.

[15] Under this alternative technology, the mean cost paid is 4.33 (s.d. 5.50) in the IA OwnValue, 2.61 (s.d. 4.69) in the IA OtherValue, 2.64 (s.d. 4.61) in the DA OwnValue, and 0.95 (s.d. 2.64) in the DA OtherValue treatment, Note that the costs in the two OtherValue treatments are weighted to control for the probability of having high B value ($= 110$).

measured relative to the UI baseline where no one knows her own preferences:

(i) Laissez-Faire Policy: Education authority does not provide any information, but let students conduct information acquisition as they wish with either the lower-cost or the higher-cost technologies.

(ii) Free OwnValue and Free OtherValue: Education authority makes every piece of information available.

(iii) Free OwnValue and Costly OtherValue: Education authority makes every piece of information relevant to OwnValue available but does not provide information on OtherValues. This policy corresponds to policies employed by many school districts where information about school characteristics is readily available, but not the information on others' strategies. Or students have to rely on historical strategies to infer others' strategies this year.

Based on the estimated effects of information acquisition and provision, we can calculate the welfare, which measured by student average payoff in each period, of each policy regime. Results are summarized in Table 11. Taking the free-OwnValue-free-OtherValue policy as an example, its welfare effect under IA is the sum of the welfare gain of providing OwnValue (8.16) and that of providing OtherValue (-0.01). Here we do not taking into account the cost of information provision for education authority.

To analyze the welfare of 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 of the cost assumption ($4.55-2.25$ or $4.55-4.33$), and then only those who have successfully acquired OwnValue can engage in acquiring OtherValue, which implies only 44% of subjects. We further assume that this leads to 44% of the net payoff gain of acquiring OtherValue ($44\% \times (1.38 - 1.29)$ or $44\% \times (1.38 - 2.61)$).[16] Similarly, for the free-OwnValue-costly-OtherValue policy, we take into account the effect of providing OwnValue and that of letting them acquire OtherValue (given that they know OwnValue already).

The results clearly show the advantages of making both OwnValue and OtherValue freely available, which leads to additional 8.15 points for every subject in each period under IA, or 4.68 points

[16]This is a simplification assumption, as we ignore the fact that the game with this two-stage information acquisition will have both informed and uninformed players. However, given the effects of acquiring OtherValue, we do not expect our conclusion on the policy comparison to change under more plausible assumptions.

Table 11: Estimated Effects and Counterfactual Analyses

Panel A: Estimated Effects

	Info Provision	Information Acquisition			
	Payoff gain[a]	Payoff gain[a]	Lower cost[b]	Higher cost[b]	P(Info Acquired)
IA OwnValue	8.16	4.55	2.25	4.33	0.44
IA OtherValue	-0.01	1.38	1.29	2.61	0.28
DA OwnValue	4.26	1.07	1.35	2.64	0.30
DA OtherValue	0.42	-0.50	0.48	0.95	0.14

Panel B: Estimated Welfare Effects (relative to UI)

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

Notes: Panel A shows estimates from Table 10 and Footnote 15; Panel B presents welfare effects of information acquisition and information provision relative to UI (i.e., nobody knows her own preferences).
a. Payoff gain measures the difference in average subject payoff in each period with free information provision or costly information acquisition and that without it under each mechanism, without taking into account the costs.
b. "Lower costs" and "higher costs" are two technologies for information acquisition. The former is exactly the one used in the experiment, where subject in expectation pays a half of her WTP when successfully acquiring the information; in the latter, subject always pays her WTP if successfully acquiring information.

under DA. If the costs of information acquisition is low, the laissez-faire policy increases average payoffs under IA but not under DA; this benefit goes away for IA as well if the cost is high. the free-OwnValue-costly-OtherValue policy is always welfare-improving. When the cost is low, this policy increases the average payoffs by 0.11 points more than that of the free-OwnValue-free-OtherValue policy. However, this small advantage disappears when the cost is high. Under DA, the welfare under the free-OwnValue-costly-OtherValue policy is always lower than that under the free-OwnValue-free-OtherValue policy, regardless of the costs.

6 Conclusions

This paper theoretically and experimentally studies endogenous information acquisition and the effects of information provision. In a school choice setting, we provide evidence that mechanisms provide heterogeneous degrees of incentives for students to acquire information.

We distinguish two types of information acquisitions. One is to learn one's own preferences over schools, and the other is to learn others' preferences. Our findings provide new insights for designing better school choice programs.

Acquiring information on own preferences is necessary in school choice, given the complex nature of education production and the usual lack of information on schools. We show that better information on own preferences in general improves student welfare, which is in line with the recent calls for better information provision on school quality. However, our experimental results also show that students tend to over-invest in information acquisition for own preferences. This thus unveils another dimension of the benefits of information provision. Namely, free or less costly information reduces the wasteful information acquisition by students.

Acquiring information on others' preferences is related to the gaming aspect of school choice, because one may enjoy a strategic advantage after learning others' preference. Our results show that this incentive is determined by a mechanism's strategy-proofness, or the lack of it. Theoretically, only a non-strategy-proof mechanism incentivizes students to acquire information on others' preferences. Moreover, if students have learned their preferences, learning others preferences affects more the re-distribution of welfare instead of promoting better matches.

More importantly, we find that a strategy-proof mechanism is not enough to prevent wasteful information acquisition on others' preferences. In our experiment, students still over-invest in acquiring information on others preferences even when facing a strategy-proof mechanism that makes such information useless. And very often, students often over-pay for information, especially among those who do not understand well the school choice game and those expecting that others are paying more for information. Such wasteful investment can be only avoided or reduced if the information on others' preferences or strategies is made freely available. This result thus calls for information provision that is beyond what has been considered in practice.

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," *American Economic Review*, 2011, *101* (1), 399–410.

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.

__ , __ , and __ , "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.

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.

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 CharlesZ 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," 2015. New York University Manuscript.

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

Dustan, Andrew, "Peer Networks and School Choice under Incomplete Information," *Unpublished manuscript*, 2015.

__ **, 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. 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.

Harless, Patrick and Vikram Manjunath, "Incentives to learn as a criterion for object allocation rules," 2015.

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," 2012.

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

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

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*, 2009.

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.

Pathak, Parag A. and Tayfun Sönmez, "Leveling the playing field: sincere and sophisticated

players in the Boston mechanism," *American Economic Review*, September 2008, *98*, 1636–1652(17).

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.

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.

A Appendix: Omitted 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 thus omit the formal proof.[17]

Lemma 2. *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 $s \in S \backslash \{s^0\}$ and t are swapped such that tL_is, $sL_i't$, $sL_i's^0$, and $\#\{s'' \in S | s'L_is''\} = \#\{s'' \in S | s'L_i's''\}$ for all $s' \in S \backslash \{s, t\}$, then:

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

and the inequality is strict when $L_j = L_i$, $\forall j \neq i$ and there exits $s^* \in S \backslash \{s^0, s, t\}$ such that $s^*L_is^0$.

(ii) Guaranteed share in first choice: If school $s \in S \backslash \{s^0\}$ is top ranked in L_i by i, $a_s(L_i, L_{-i}) \geq q_s / |I|$, for all L_{-i}.

(iii) Guaranteed assignment: $\sum_{s \in \{s \in S | sL_is^0\}} a_s(L_i, L_{-i}) + a_{s^0}(L_i, L_{-i}) = 1$ for all L_{-i}.

A.1 Proof of Lemma 1.

The 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 one's own investments in information acquisition, although it depends on the signal that one has received (ω).

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

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

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

[17] Similar results on IA and their proofs are available in He (2012).

order conditions are satisfied by the assumptions on the functions $a()$, $b()$, and $c()$):

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

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

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.2 Proof of Proposition 1.

A.2.1 Proof of $\alpha^* > 0$

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

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

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 only 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 $c'(0, 0) < a'(0) = \infty$.

Suppose that in a given symmetric equilibrium σ^*, the finiteness of the strategy space implies that in the equilibrium pure strategies $\left(L^{(1)}, ..., L^{(N)}\right)$ are played with positive probabilities $\left(p^{(1)}, ..., p^{(N)}\right)$ ($N \in \mathbb{N}$). s^* is ranked the last in $L^{(1)}$ among non-outside-option schools, i.e., $\#\left\{s \in S | s^* L^{(1)} s,\, s \neq s^0\right\} = 0$ and there is at least another school s_1 s.t. $s_1 L^{(1)} s^0$ and $s_1 L^{(1)} s^*$. Given the full support of ordinal preferences and Lemma 2, such an $L^{(1)}$ must exist (otherwise, it cannot be an equilibrium). Similarly, there exists an ordinal preference P^* such that $s^* P^* s^0 P^* s$ for all $s \neq s^*, s^0$.

Since $\Pi\left(\left(P^*, V^\phi\right), 0, \mathbf{0}\right) = \Pi\left(\left(P, V^\phi\right), 0, \mathbf{0}\right)$, it 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 P^*.

By monotonicity of the mechanism (Lemma 2), $a_{s^*}\left(P^*, L_{-i}\right) \geq a_{s^*}\left(P, L_{-i}\right)$ for all L_{-i}, where P does not top rank s^*. Moreover, $a_{s^*}\left(P^*, L_{-i}\right) > a_{s^*}\left(P, L_{-i}\right)$ when everyone else submits $L^{(1)}$ in L_{-i}.

σ^* leads to a probability distribution over a finite number of possible profiles of other's strategies (L_{-i}). With a positive probability, everyone else plays $L^{(1)}$. In this event, therefore, by submitting P^*, i strictly increases the probability of being accepted by s^*, the only acceptable school, comparing 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 P^*. 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.2.2 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. It implies that:

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

or,

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

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.2.3 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. Suppose that F implies a distribution of ordinal preferences $G(P|F)$ such that for s_1 and s_2:

$$G(P|F)=\begin{cases}(1-\varepsilon) & \text{if } P=\bar{P},\ \text{s.t. } s_1\bar{P}s_2\bar{P}s^0\bar{P}s_3...\bar{P}s_{|S|};\\ \frac{\varepsilon}{|\mathcal{P}|-1} & \text{if } P\neq\bar{P}.\end{cases}$$

The distribution of cardinal preferences is:

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

(ε,η,ξ) are all small positive numbers in $(0,1)$. Otherwise, there is no restriction on $F(V|P)$ for $P\neq\bar{P}$ nor on $v_s,\forall s\in S\backslash\{s_1,s_2,s^0\}$.

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

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

And (η,ξ) are small enough such that $E\left(v_{i,s_2}|\bar{P}\right)<q_{s_1}/|I|$, and therefore obtaining s_2 for sure is less preferable than obtaining $q_{s_1}/|I|$ of s_1. In equilibrium with such 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 some students deviate from such equilibrium strategies. Suppose that i had learned her ordinal preferences and $P_i = \bar{P}$. If furthermore she succeeds in acquiring information on V_i, there is a positive probability that $(v_{s_1}, v_{s_2}, v_{s^0}) = (1, 1 - \xi, \xi/2)$. In this case, if she plays L_i s.t., $s_2 L_i s_1 L_i s^0 L_i s_3 ... L_i s_{|S|}$ (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:

$$(1 - \varepsilon)^{(|I|-1)} \left[\frac{q_{s_1}}{|I|} + \left(1 - \frac{q_{s_1}}{|I|}\right)\frac{\xi}{2}\right] + \left(1 - (1 - \varepsilon)^{(|I|-1)}\right).$$

This upper bound is obtained under the assumption that one is always assigned to s_1 when 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:

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

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.3 Proof of Proposition 2.

For the first part, by the definition of strategy-proofness, information on 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:

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

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

Suppose that for student i, $V_i = V^{(3)}$. If $\delta^*(V) = 0$ for all V, the best response for i in equilibrium is P_i or L_i s.t. $s_2 L_i s_1 L_i s^0 L_i s$ for all other s. Without knowing V_{-i}, i has to play a pure strategy, P_i or L_i, or a mixed strategy $\Delta(\{P_i, L_i\})$ in equilibrium. $\delta^*(V_i) = 0$ implies that either the two pure strategies are always payoff-equivalent or one always dominates the other, given any realization of V_{-i}.

Given $F(V)$, there is a positive probability, $\left(\frac{1}{2} - \varepsilon\right)^{|I|-1}$, that every other student has $V^{(1)}$ and submit her true preferences. In this case, the payoff for i submitting P_i is less than $q_{s_1}/|I| + \xi$, while the one when submitting L_i is $(1-\eta)$.

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

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}), thus:

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

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 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.4 Proof of Proposition 3.

Under UI, the only information i has is that her preferences follow the distribution $F(V)$. Denote W_i^E as the expected (possibly weak) ordinal preferences of i such that $sW_i^E t$ 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_i^E, $\left(P_i^{E,1}, ..., P_i^{E,M}\right) \in \mathcal{P}$ are all the strict ordinal preferences that can be generated by randomly breaking ties in W_i^E if there is any. Therefore, $M \geq 1$.

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

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

Since DA with single tie breaking is essentially the 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:

$$\sum_{s \in S} a_s\left(P_i^{E,m}, L_{-i}^*\right) \int v_{i,s} dF_{v_s}(v_{i,s}) = \sum_{s \in S \setminus \{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}(v_{i,s}), \forall m. \tag{3}$$

Note that for any L_{-i}^*, $\sum_{s \in S} a_s\left(P_i^{E,m}, L_{-i}\right) \int v_{i,s} dF_{v_s}(v_{i,s})$ 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:

$$\sum_{s \in S \setminus \{s^0\}} q_s \int v_{i,s} dF_{v_s}(v_{i,s}) \tag{4}$$

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

$$\sum_{s\in S} a_s\left(P_i^{E,m},\hat{L}^*_{-i}\right)\int v_{i,s}dF_{v_s}\left(v_{i,s}\right) > \sum_{s\in S\backslash\{s^0\}}\frac{q_s}{|I|}\int v_{i,s}dF_{v_s}\left(v_{i,s}\right),\forall m. \tag{5}$$

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

$$\sum_{s\in S} a_s\left(P_j^{E,m},\hat{L}^*_{-j}\right)\int v_{j,s}dF_{v_s}\left(v_{j,s}\right) < \sum_{s\in S\backslash\{s^0\}}\frac{q_s}{|I|}\int v_{j,s}dF_{v_s}\left(v_{j,s}\right), \tag{6}$$

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 the true ordinal preferences, and thus the expected payoff (ex ante) is:

$$\begin{aligned}
&\int\int\sum_{s\in S} a_s\left(P,L_{-i}\left(P\right)\right)v_{i,s}dF\left(V|P\right)dG\left(P|F\right)\\
&=\int\int\sum_{s\in S\backslash\{s^0\}}\frac{q_s}{|I|}v_{i,s}dF_{v_s}\left(v_{i,s}|P\right)dG\left(P|F\right)\\
&=\sum_{s\in S\backslash\{s^0\}}\frac{q_s}{|I|}\int v_{i,s}dF_{v_s}\left(v_{i,s}\right),
\end{aligned}$$

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

A.5 Proof of Proposition 4.

A.5.1 Welfare under UI and OI

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

Under UI, the game can be transformed into one similar to PI but 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 thus must 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:

$$\sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}\right).$$

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:

$$\sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}|P\right),$$

which leads to an *ex ante* payoff:

$$\int \sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}|P\right) dG\left(P|F\right) = \sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}\right).$$

A.5.2 Proof of CI > UI = OI under IA

We then show CI > OI = UI.

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

$$\int\int\left(\int A\left(\sigma^{BN}\left(V_i\right), \sigma^{BN}\left(V_{-i}\right)\right) dF\left(V_{-i}|P\right)\cdot V_i\right) dF\left(V_i|P\right) dG\left(P|F\right)$$
$$\geq \sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} \int v_{i,s} dF_{v_s}\left(v_{i,s}\right).$$

The following uses the same idea as in the proof of Proposition 2 in (Troyan 2012). Note that $\int a_s\left(\sigma^{BN}\left(V_i\right),\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)$ 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

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

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

$$|I|\int\int\int a_s\left(\sigma^{BN}\left(V_i\right),\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)dF\left(V_i|P\right)dG\left(P|F\right)=q_s,\forall s\in S\backslash\left\{s^0\right\}, \tag{7}$$

as in equilibrium all seats at all $s\in S\backslash\left\{s^0\right\}$ must be assigned.

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

$$\begin{aligned}
&\int\left(\int A\left(\sigma_i,\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)\cdot V_i\right)dF\left(V_i|P\right)\\
&=\int\left(\int\left(\int\int A\left(\sigma^{BN}\left(V_i\right),\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_i|P\right)dG\left(P|F\right)\right)dF\left(V_{-i}|P\right)\cdot V_i\right)dF\left(V_i|P\right)\\
&=\int\left(\sum\nolimits_{s\in S}\left(\int\int\int a_s\left(\sigma^{BN}\left(V_i\right),\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_i|P\right)dG\left(P|F\right)dF\left(V_{-i}|P\right)\right)v_{i,s}\right)dF\left(V_i|P\right)\\
&=\int\left(\sum_{s\in S\backslash\{s^0\}}\frac{q_s}{|I|}v_{i,s}\right)dF\left(V_i|P\right).
\end{aligned}$$

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 for *ex ante* welfare:

$$\begin{aligned}
&\int\int\left(\int A\left(\sigma^{BN}\left(V_i\right),\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)\cdot V_i\right)dF\left(V_i|P\right)dG\left(P|F\right)\\
&\geq\int\int\left(\int A\left(\sigma_i,\sigma^{BN}\left(V_{-i}\right)\right)dF\left(V_{-i}|P\right)\cdot V_i\right)dF\left(V_i|P\right)dG\left(P|F\right)\\
&=\sum_{s\in S\backslash\{s^0\}}\frac{q_s}{|I|}\int v_{i,s}dF_{v_s}\left(v_{i,s}\right),
\end{aligned}$$

which proves CI $>$ OI = UI in terms of Pareto dominance of ex ant student welfare.

A.5.3 Proof of PI $>$ 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, we must have PI Pareto dominates OI and UI.

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

$$\int\int\int \left(A\left(\sigma^{NE}(V_i, V_{-i}), \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF(V_{-i}|P)\, dF(V_i|P)\, dG(P|F)$$
$$\geq \sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|}\int v_{i,s} dF_{v_s}(v_{i,s}).$$

Note that $a_s\left(\sigma^{NE}(V_i, V_{-i}), \left[\sigma^{NE}(V_j, V_{-j})\right]_{j\in I\setminus\{i\}}\right)$ 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

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

which must be the same across students by symmetry. Therefore, we must have, $\forall s \in S\setminus\{s^0\}$:

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

as in equilibrium all seats at all $s \in S\setminus\{s^0\}$ must be assigned.

Suppose i plays an alternative strategy σ_i such that

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

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:

$$
\begin{aligned}
&A\left(\sigma_i, \left[\sigma^{NE}\left(V_j, V_{-j}\right)\right]_{j\in I\setminus\{i\}}\right)\cdot V_i \\
&= \left(\int\int\int A\left(\sigma^{NE}\left(V_i, V_{-i}\right), \left[\sigma^{NE}\left(V_j, V_{-j}\right)\right]_{j\in I\setminus\{i\}}\right) dF\left(V_{-i}|P\right) dF\left(V_i|P\right) dG\left(P|F\right)\right)\cdot V_i \\
&= \sum\nolimits_{s\in S}\left(\int\int\int a_s\left(\sigma^{NE}\left(V_i, V_{-i}\right), \left[\sigma^{NE}\left(V_j, V_{-j}\right)\right]_{j\in I\setminus\{i\}}\right) dF\left(V_i|P\right) dG\left(P|F\right) dF\left(V_{-i}|P\right)\right) v_{i,s} \\
&= \sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} v_{i,s}.
\end{aligned}
$$

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

$$
\begin{aligned}
&\int\int\left(A\left(\sigma_i, \left[\sigma^{NE}\left(V_j, V_{-j}\right)\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF\left(V_{-i}|P\right) dF\left(V_i|P\right) \\
&= \int\left(\sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|} v_{i,s}\right) dF\left(V_i|P\right).
\end{aligned}
$$

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

$$
\begin{aligned}
&\int\int\int\left(A\left(\sigma^{NE}\left(V_i, V_{-i}\right), \left[\sigma^{NE}\left(V_j, V_{-j}\right)\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF\left(V_{-i}|P\right) dF\left(V_i|P\right) dG\left(P|F\right) \\
&\geq \int\int\int\left(A\left(\sigma_i, \left[\sigma^{NE}\left(V_j, V_{-j}\right)\right]_{j\in I\setminus\{i\}}\right)\cdot V_i\right) dF\left(V_{-i}|P\right) dF\left(V_i|P\right) dG\left(P|F\right) \\
&= \sum_{s\in S\setminus\{s^0\}} \frac{q_s}{|I|}\int v_{i,s} dF_{v_s}\left(v_{i,s}\right),
\end{aligned}
$$

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.5.4 shows that PI can dominate CI in symmetric equilibrium; and the example in section A.5.5 shows the opposite.

A.5.4 Example: PI dominates CI in symmetric equilibrium under IA

There are 3 schools (a, b, c) and 3 students whose cardinal preferences are i.i.d. draws from the following distribution:

$$\Pr\left((v_a, v_b, v_c) = (1, 0.1, 0)\right) = 1/2$$
$$\Pr\left((v_a, v_b, v_c) = (1, 0.5, 0)\right) = 1/2$$

Each school has one seat. For any realization of preference profile, we can find a symmetric Nash equilibrium as in Table A1.

Table A1: Symmetric Nash Equilibrium for Each Realization of the Game under PI

Realization of Preferences	Probability Realized	Strategy given realized type (1, 0.1, 0)	Strategy given realized type (1, 0.5, 0)	Payoff given realized type (1, 0.1, 0)	Payoff given realized type (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 that:

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

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

There are 3 schools (a,b,c) and 3 students whose cardinal preferences are i.i.d. draws from the following distribution:

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

Each school has one seat. For any realization of preference profile, we can find a symmetric Nash equilibrium as in Table A2. The *ex ante* welfare under PI with the above symmetric equilibrium profile is:

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

Under CI, i.e., when one'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:

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

Table A2: 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

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.

B Analyses of the Game in the Experiment under Risk Neutrality

Given the payoff table introduced in Section 4, this appendix derives in details the equilibrium strategies and payoffs under the assumption that every student is risk neutral. We also vary information structure and derive the incentive to acquire information. The results on risk-averse students are presented in Appendix C. Throughout, students do not know the realization of tie breakers when playing the game.

B.1 Information Structure

We consider the following 5 scenarios where information structure differs:

(1) Complete information on preferences: Everyone knows her own and others' realized preferences;

(2) Incomplete information on preferences: Everyone knows her own realized preferences but only the distribution of others';

(3) Unknown preferences: Everyone only knows the distribution of her own preferences and of others';

(4) Unknown preferences (Scenario (3)) with acquisition of information on one's own preferences;

(5) Incomplete information (Scenario (2)) with acquisition of information on others' preferences.

The literature on school choice, or on matching in general, focuses on the first two scenarios – complete or incomplete information. By introducing scenarios (3)-(5), we extend the literature by endogenizing the acquisition of information on one's own or on others' preferences.

Figure B1 shows the relationship among the five scenarios.

B.2 Scenario (1): Complete Information on Preferences

The Immediate Acceptance Mechanism Given any realization of the preferences, we have the following symmetric equilibrium strategies and payoffs under the Immediate Acceptance mechanism (Table B3).

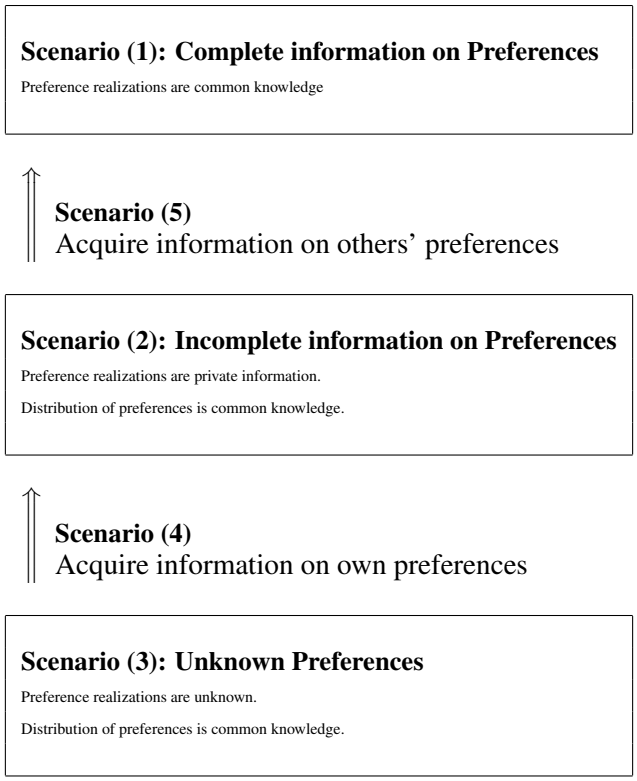

Figure B1: Scenarios Considered and the Corresponding Information Structure

Table B3: Symmetric Equilibrium under IA given Each Realization of Preference Profiles

Realization of Preference	Probability Realized	Strategy given realized type (1, 0.1, 0)	Strategy given realized type (1, 1.1, 0)	Payoff given realized type (1, 0.1, 0)	Payoff given realized type (1, 1.1, 0)
(1, 0.1, 0) (1, 0.1, 0) (1, 0.1, 0)	64/125	(a, b, c)	-	11/30	-
(1, 1.1, 0) (1, 0.1, 0) (1, 0.1, 0)	48/125	(a, b, c)	(b, a, c)	1/2	11/10
(1, 1.1, 0) (1, 1.1, 0) (1, 0.1, 0)	12/125	(a, b, c)	(b, a, c)	1	11/20
(1, 1.1, 0) (1, 1.1, 0) (1, 1.1, 0)	1/125	-	(a, b, c) w/ prob. 3/7[a] (b, a, c) w/ prob. 4/7[a]	-	7/10

a. We may allow one student to play (a,b,c) and the other two to play (b,a,c), which is a pure-strategy Nash equilibrium. As long as everyone has the same probability to play (a,b,c), the expected payoff of everyone is also 7/10.

Ex ante, before the realization of the preferences, given that they know they will play the game with complete information under IA, the expected payoff of each student is:

$$\frac{4}{5}\left(\frac{11}{30}\frac{16}{25}+\frac{1}{2}\frac{8}{25}+\frac{1}{25}\right)+\frac{1}{5}\left(\frac{11}{10}\frac{16}{25}+\frac{11}{20}\frac{8}{25}+\left(\frac{7}{10}\right)\frac{1}{25}\right)=\frac{4}{5}\frac{326}{750}+\frac{1}{5}\frac{681}{750}=\frac{397}{750}.$$

The DA Mechanism Before looking at equilibrium, we use the following table to clarify the assignment probabilities given students' actions (Table B4). Note that we always use DA with single tie-breaking.

Table B4: Assignment Probability under DA given Each Strategy Profile

Submitted List	Probability of Being Assigned to Each School if Playing (a, b, c)			Playing (b, a, c)		
	a	b	c	a	b	c
(a, b, c) (a, b, c) (a, b, c)	1/3	1/3	1/3	-	-	-
(b, a, c) (a, b, c) (a, b, c)	1/2	1/6	1/3	0	2/3	1/3
(b, a, c) (b, a, c) (a, b, c)	2/3	0	1/3	1/6	1/2	1/3
(b, a, c) (b, a, c) (b, a, c)	-	-	-	1/3	1/3	1/3

Given any realization of the preferences, we have the following equilibrium strategies and payoffs under DA (Table B5).

Ex ante, before the realization of the preferences, given that they know they will play the game with complete information under DA, the expected payoff to each student is:

$$\frac{4}{5}\left(\frac{11}{30}\frac{16}{25}+\frac{31}{60}\frac{8}{25}+\frac{2}{3}\frac{1}{25}\right)+\frac{1}{5}\left(\frac{22}{30}*\frac{16}{25}+\frac{43}{60}\frac{8}{25}+\left(\frac{21}{30}\right)\frac{1}{25}\right)=\frac{365}{750}.$$

B.3 Scenario (2): Incomplete Information on Preferences

The Immediate Acceptance Mechanism When one's own preferences are private information and the distribution of preferences is common knowledge, there is a unique symmetric equilibrium

Table B5: Equilibrium under DA given each Realization of Preference Profiles

Realization of Preference	Probability Realized	Strategy given realized type (1, 0.1, 0)	Strategy given realized type (1, 1.1, 0)	Payoff given realized type (1, 0.1, 0)	Payoff given realized type (1, 1.1, 0)
(1, 0.1, 0) (1, 0.1, 0) (1, 0.1, 0)	64/125	(a, b, c)	-	11/30	-
(1, 1.1, 0) (1, 0.1, 0) (1, 0.1, 0)	48/125	(a, b, c)	(b, a, c)	$\frac{1}{2}+\frac{1}{60}$ $= 31/60$	$\frac{2}{3}\frac{11}{10}$ $= 22/30$
(1, 1.1, 0) (1, 1.1, 0) (1, 0.1, 0)	12/125	(a, b, c)	(b, a, c)	2/3	$\frac{11}{20}+\frac{1}{6}$ $= 43/60$
(1, 1.1, 0) (1, 1.1, 0) (1, 1.1, 0)	1/125	-	(b, a, c)	-	21/30

under IA:

$$\sigma_{BM}^{(2)}\left((1,1.1,0)\right)=(b,a,c)\,;\sigma_{BM}^{(2)}\left((1,0.1,0)\right)=(a,b,c)\,.$$

For any given student, there are three possibilities of opponents' types:

Types	Probability	Others' Strategy Profile
(1, 0.1, 0) (1, 0.1, 0)	16/25	(a, b, c) (a, b, c)
(1, 1.1, 0) (1, 0.1, 0)	8/25	(b, a, c) (a, b, c)
(1, 1.1, 0) (1, 1.1, 0)	1/25	(b, a, c) (b, a, 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:

$$\frac{16}{25}\left(\frac{1}{3}\left(1+\frac{1}{10}+0\right)\right)+\frac{8}{25}\frac{1}{2}+\frac{1}{25}=\frac{326}{750}.$$

If a type-$(1,0.1,0)$ student deviates to (b,a,c), she obtains:

$$\frac{16}{25}\left(\frac{1}{10}\right)+\frac{8}{25}\left(\frac{1}{2}\left(\frac{1}{10}+0\right)\right)+\frac{1}{25}\left(\frac{11}{30}\right)=\frac{71}{750}.$$

For a type-$(1, 1.1, 0)$ student, given others follow $\sigma_{BM}^{(2)}$, playing (b, a, c) results in a payoff of:

$$\frac{16}{25}\frac{11}{10} + \frac{8}{25}\left(\frac{1}{2}\left(\frac{11}{10} + 0\right)\right) + \frac{1}{25}\left(\frac{1}{3}\left(\frac{11}{10} + 1 + 0\right)\right) = \frac{681}{750}.$$

If a type-$(1, 1.1, 0)$ student deviates to (a, b, c), she obtains:

$$\frac{16}{25}\left(\frac{1}{3}\left(\frac{11}{10} + 1 + 0\right)\right) + \frac{8}{25}\left(\frac{1}{2}(1 + 0)\right) + \frac{1}{25}(1) = \frac{486}{750}.$$

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).

Before the realization of their own preferences while knowing that they will play the game under DA with incomplete information, the ex ante payoff to every student is:

$$\frac{326}{750}\frac{4}{5} + \frac{681}{750}\frac{1}{5} = \frac{397}{750}.$$

Remark B1. *Note that the two scenarios, (1) and (2), result in the same ex ante payoffs under IA.*

Remark B2. *In neither scenarios, a type-$(1, 0.1, 0)$ student is ever matched with school b as long as there is at least one type-$(1, 1.1, 0)$ student.*

The DA Mechanism When one's own preferences are private information and the distribution of preferences is common knowledge, there is a unique equilibrium under DA:

$$\sigma_{DA}^{(2)}((1, 1.1, 0)) = (b, a, c)\,; \sigma_{DA}^{(2)}((1, 0.1, 0)) = (a, b, c)\,.$$

For any given student, there are three possibilities of opponents' types: For a type-$(1, 0.1, 0)$

	Types	Probability	Others' Strategy Profile
1	(1, 0.1, 0) (1, 0.1, 0)	16/25	(a, b, c) (a, b, c)
2	(1, 1.1, 0) (1, 0.1, 0)	8/25	(b, a, c) (a, b, c)
3	(1, 1.1, 0) (1, 1.1, 0)	1/25	(b, a, c) (b, a, c)

student, it is a dominant strategy to play (a, b, c). Conditional on her type, her equilibrium payoff is:

$$\frac{16}{25}\left(\frac{1}{3}\left(1 + \frac{1}{10} + 0\right)\right) + \frac{8}{25}\left(\frac{1}{2} + \frac{1}{60} + 0\right) + \frac{1}{25}\left(\frac{2}{3} + 0\right) = \frac{320}{750}.$$

If a type-$(1, 0.1, 0)$ student deviates to (b, a, c), she obtains:

$$\frac{16}{25}\left(\frac{2}{30}+0\right)+\frac{8}{25}\left(\frac{1}{20}+\frac{1}{6}+0\right)+\frac{1}{25}\left(\frac{1}{3}\left(1+\frac{1}{10}+0\right)\right)=\frac{95}{750}.$$

For a type-$(1, 1.1, 0)$ student, given others follow $\sigma_{DA}^{(2)}$, playing (b, a, c) results in a payoff of:

$$\frac{16}{25}\left(\frac{2}{3}\frac{11}{10}\right)+\frac{8}{25}\left(\frac{1}{2}\frac{11}{10}+\frac{1}{6}\right)+\frac{1}{25}\left(\frac{1}{3}\left(\frac{11}{10}+1+0\right)\right)=\frac{545}{750}.$$

If a type-$(1, 1.1, 0)$ student deviates to (a, b, c), she obtains:

$$\frac{16}{25}\left(\frac{1}{3}\left(\frac{11}{10}+1+0\right)\right)+\frac{8}{25}\left(\frac{1}{2}+\frac{11}{60}\right)+\frac{1}{25}\left(\frac{2}{3}\right)=\frac{520}{750}.$$

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, before knowing their own true preferences, is:

$$\frac{320}{750}\frac{4}{5}+\frac{545}{750}\frac{1}{5}=\frac{365}{750}.$$

Remark B3. *Note that the two scenarios, (1) and (2), result in the same ex ante payoffs under DA.*

Remark B4. *In both scenarios, there is a positive probability that a type-$(1, 0.1, 0)$ student is matched with school b when there is at least one type-$(1, 1.1, 0)$ student.*

B.4 Scenario (3): Unknown Preferences

The Immediate Acceptance Mechanism Under IA, the unique symmetric equilibrium is that everyone plays $\sigma_{BM}^{(3)} = (a, b, c)$. The expected payoff of this strategy is:

$$\frac{1}{3}\left(1+\left(\frac{1}{5}\frac{11}{10}+\frac{4}{5}\frac{1}{10}\right)+0\right)=\frac{13}{30}=\frac{325}{750}.$$

If an student deviates to (b, a, c), her payoff is:

$$\left(\frac{1}{5}\frac{11}{10}+\frac{4}{5}\frac{1}{10}\right)=0.3=\frac{225}{750}.$$

Remark B5. *In Scenario (2), the ex ante payoff is* $\frac{397}{750}$ *which is higher than that of Scenario (3),* $\frac{225}{750}$.

Remark B6. *Comparing Scenarios (1), (2), and (3), we can improve the social welfare by making it easier for students to learn their preferences and then transforming (3) into (2) or (1) under the Immediate Acceptance.*

The DA Mechanism The unique symmetric equilibrium under DA is that everyone plays $\sigma_{DA}^{(3)} = (a, b, c)$.

The expected payoff of this strategy is:

$$\frac{1}{3}\left(1 + \left(\frac{1}{5}\frac{11}{10} + \frac{4}{5}\frac{1}{10}\right) + 0\right) = \frac{13}{30} = \frac{325}{750}.$$

If an student deviates to (b, a, c), her payoff is:

$$\frac{1}{5}\left(\frac{2}{3}\frac{11}{10}\right) + \frac{4}{5}\left(\frac{2}{3}\frac{1}{10}\right) = 0.2 = \frac{150}{750}.$$

Remark B7. *In Scenario (2), the ex ante payoff is* $\frac{365}{750}$ *which is higher than that of Scenario (3),* $\frac{325}{750}$.

Remark B8. *Comparing Scenarios (1), (2), and (3), we can improve the social welfare by making it easier for students to learn their preferences and then transforming (3) into (2) or (1) under DA.*

Remark B9. *The benefit of providing free information on own preferences is higher under the Immediate Acceptance.*

Remark B10. *In Scenarios (3), the Immediate Acceptance mechanism achieves the same outcome as DA.*

In the following, we discuss students' incentives to acquire information on one's own preferences.

B.5 Scenario (4): (3) + acquisition of information on one's own preferences

The Immediate Acceptance Mechanism Now suppose that students only know the distribution of their own and others' preferences. We consider their incentives to acquire information on their

own preferences.

After acquiring the information, both informed and uninformed students know how many others are informed. However, informed students know their own preferences, while uninformed students only know the distribution of own preference.

Willingness to pay for information on own preferences can be defined in the following three cases:

$$w_0^{own}: \text{ when no other informed students;}$$

$$w_1^{own}: \text{ when there is another informed student;}$$

$$w_2^{own}: \text{ when there are two other informed students.}$$

The following table summarizes the equilibrium strategies and *ex ante* payoffs for informed and uninformed players (Table B6).

Table B6: Willingness to Pay for Information on Own Payoffs under IA

# of Players		Strategy:	Strategy: Informed		Ex Ante Payoff		Willingness to
Informed	Uninformed	Uninformed	(1, 0.1, 0)	(1, 1.1, 0)	Informed	Uninformed	pay for info
0	3	(a, b, c)	-	-	-	$\frac{325}{750}$	$\frac{60}{750}$
1	2	(a, b, c)	(a, b, c)	(b, a, c)	$\frac{385}{750}$	$\frac{335}{750}$	$\frac{49.5}{750}$
2	1	(a, b, c)	(a, b, c)	(b, a, c)	$\frac{384.5}{750}$	$\frac{358}{750}$	$\frac{39}{750}$
3	0	-	(a, b, c)	(b, a, c)	$\frac{397}{750}$	-	

Overt and covert information acquisition: In the current setting, we focus on overt information acquisition. Namely, all students, informed and uninformed, know how many students in total are informed. Note that, for uninformed students, knowing or not knowing how many students are informed does not change their strategy. If information acquisition is covert, an informed student should have other students behave as if she is uninformed. This is difficult to achieve in a lab experiment, as everyone knows that everyone is offered a chance to learn their own preferences. Our overt-information-acquisition approach provides a lower bound on information acquisition regarding one's own preferences. That is, one always has a greater incentive to acquire information covertly and choose to make it public only if she finds it profitable. Besides, the information

acquisition is purely about one's own preferences, while all other information is costless.

When no other students are informed and an student acquires this information, the unique equilibrium in the school choice game is:

$$(One)\ Informed : \sigma\left((1, 1.1, 0)\right) = (b, a, c) \text{ and } \sigma\left((1, 0.1, 0)\right) = (a, b, c)\,;$$
$$(Two)\ Uninformed : (a, b, c)\,,$$

The informed student obtains an expected payoff:

$$\frac{1}{5}\frac{11}{10} + \frac{4}{5}\left(\frac{1}{3}\left(\frac{1}{10} + 1 + 0\right)\right) = \frac{385}{750}.$$

When she chooses not to acquire information, the game is returned to Scenario (3) and her expected payoff is $\frac{325}{750}$. Therefore, given there is no other informed student, her willingness to pay for the information is:

$$w_0^{own} = \frac{385}{750} - \frac{325}{750} = \frac{60}{750}.$$

If there is one informed student already, an additional student acquires this information, and the game has two informed players and one uninformed. The unique equilibrium in this case is:

$$(Two)\ Informed : \sigma\left((1, 1.1, 0)\right) = (b, a, c) \text{ and } \sigma\left((1, 0.1, 0)\right) = (a, b, c)\,;$$
$$(One)\ Uninformed : (a, b, c)\,.$$

Informed students obtain an ex ante payoff:

$$\frac{1}{5}\left(\frac{1}{5}\frac{1}{2}\frac{11}{10} + \frac{4}{5}\frac{11}{10}\right) + \frac{4}{5}\left(\frac{1}{5}\frac{1}{2} + \frac{4}{5}\frac{11}{30}\right) = \frac{384.5}{750}.$$

If the student chooses not to acquire information, she plays against one informed and one uninformed players. The equilibrium is discussed above, and her payoff as an uninformed player is:

$$\frac{1}{5}\left(\frac{1}{5}\frac{1}{2} + \frac{4}{5}\frac{21}{30}\right) + \frac{4}{5}\left(\frac{1}{5}\frac{1}{2} + \frac{4}{5}\frac{11}{30}\right) = \frac{335}{750}$$

This implies that the willingness to pay for information in this case is:

$$w_1^{own} = \frac{384.5}{750} - \frac{335}{750} = \frac{49.5}{750}.$$

When the other two students are informed, if the third student also decides to acquire this information, the game turns into one with three informed players as in Scenario (2). We know that her expected payoff is $\frac{397}{750}$. If she decides not to do so, she remains uninformed and plays against two informed players. The equilibrium is discussed above and her expected payoff is:

$$\begin{aligned}
&\frac{1}{5}\left(\frac{16}{25}\left(\frac{1}{3}\left(\frac{11}{10}+1+0\right)\right)+\frac{8}{25}\left(\frac{1}{2}(1+0)\right)+\frac{1}{25}(1)\right)\\
&+\frac{4}{5}\left(\frac{16}{25}\left(\frac{1}{3}\left(\frac{1}{10}+1+0\right)\right)+\frac{8}{25}\left(\frac{1}{2}(1+0)\right)+\frac{1}{25}(1)\right)\\
&=\frac{358}{750}
\end{aligned}$$

Therefore, the willingness to pay is:

$$w_2^{own} = \frac{397}{750} - \frac{358}{750} = \frac{39}{750}.$$

Remark B11. *The willingness to pay depends on the number of informed students. When the cost is lower than w_2^{own}, all students choose to be informed.*

Remark B12. *When more students are informed, the incentive to acquire information is lower.*

Remark B13. *Information acquisition has externalities. Namely, when more students are informed, the payoffs to uninformed students are higher.*

Remark B14. *If we only elicit one amount of willingness to pay, an student reports a number in $\left[\frac{39}{750}, \frac{60}{750}\right]$, because she forms a probability distribution over the three possible realizations – playing against another 0-2 informed students.*

The DA Mechanism Now we consider DA. Students only know the distribution of their own and others' preferences. The following table, Table B7, summarizes the equilibrium strategies and ex ante payoffs for informed and uninformed players under DA.

Table B7: Willingness to Pay for Information on Own Payoffs under DA

# of Players		Strategy:	Strategy: Informed		Ex Ante Payoff		Willingness to
Informed	Uninformed	Uninformed	(1, 0.1, 0)	(1, 1.1, 0)	Informed	Uninformed	pay for info
0	3	(a, b, c)	-	-	-	$\frac{325}{750}$	$\frac{5}{750}$
1	2	(a, b, c)	(a, b, c)	(b, a, c)	$\frac{330}{750}$	$\frac{342.5}{750}$	$\frac{5}{750}$
2	1	(a, b, c)	(a, b, c)	(b, a, c)	$\frac{347.5}{750}$	$\frac{360}{750}$	$\frac{5}{750}$
3	0	-	(a, b, c)	(b, a, c)	$\frac{365}{750}$	-	

When no other students are informed and an student acquires this information, the unique equilibrium in the school choice game is:

$$(One)\ Informed : \sigma\left((1, 1.1, 0)\right) = (b, a, c) \text{ and } \sigma\left((1, 0.1, 0)\right) = (a, b, c)\,;$$
$$(Two)\ Uninformed : (a, b, c)\,,$$

The informed student obtains an expected payoff:

$$\frac{1}{5}\left(\frac{11}{10}\frac{2}{3}\right) + \frac{4}{5}\left(\frac{1}{3}\left(\frac{1}{10} + 1 + 0\right)\right) = \frac{330}{750}.$$

If she chooses not to acquire information, the game is returned to Scenario (3) and her expected payoff is $\frac{325}{750}$. Therefore, given there is no other informed student, her willingness to pay for the information is:

$$w_0^{own} = \frac{330}{750} - \frac{325}{750} = \frac{5}{750}.$$

If there is one informed student already, an additional student acquires this information, and the game has two informed players and one uninformed. The unique equilibrium in this case is:

$$(Two)\ Informed : \sigma\left((1, 1.1, 0)\right) = (b, a, c) \text{ and } \sigma\left((1, 0.1, 0)\right) = (a, b, c)\,;$$
$$(One)\ Uninformed : (a, b, c)\,.$$

Informed students obtain an ex ante payoff:

$$\frac{1}{5}\left(\frac{1}{5}\left(\frac{1}{2}\frac{11}{10}+\frac{1}{6}\right)+\frac{4}{5}\left(\frac{11}{10}\frac{2}{3}\right)\right)+\frac{4}{5}\left(\frac{1}{5}\left(\frac{1}{2}+\frac{1}{60}\right)+\frac{4}{5}\left(\frac{11}{30}\right)\right)=\frac{347.5}{750}.$$

If the student chooses not to acquire information, she plays against one informed and one uninformed players. The equilibrium is discussed above, and her payoff as an uninformed player is:

$$\frac{1}{5}\left(\frac{1}{5}\left(\frac{1}{2}+\frac{11}{60}\right)+\frac{4}{5}\frac{21}{30}\right)+\frac{4}{5}\left(\frac{1}{5}\left(\frac{1}{2}+\frac{1}{60}\right)+\frac{4}{5}\frac{11}{30}\right)=\frac{342.5}{750}$$

This implies that the willingness to pay for information in this case is:

$$w_1^{own}=\frac{347.5}{750}-\frac{342.5}{750}=\frac{5}{750}.$$

When the other two students are informed, if the third student also decides to acquire this information, the game turns into one with three informed players as in Scenario (2). We know that her expected payoff is $\frac{365}{750}$. If she decides not to do so, she remains uninformed and plays against two informed players. The equilibrium is discussed above and her expected payoff is:

$$\begin{aligned}&\frac{1}{5}\left(\frac{16}{25}\left(\frac{1}{3}\left(\frac{11}{10}+1+0\right)\right)+\frac{8}{25}\left(\frac{1}{2}+\frac{11}{60}\right)+\frac{1}{25}\left(\frac{2}{3}\right)\right)\\&+\frac{4}{5}\left(\frac{16}{25}\left(\frac{1}{3}\left(\frac{1}{10}+1+0\right)\right)+\frac{8}{25}\left(\frac{1}{2}+\frac{1}{60}\right)+\frac{1}{25}\left(\frac{2}{3}\right)\right)\\&=\frac{360}{750}\end{aligned}$$

Therefore, the willingness to pay is:

$$w_2^{own}=\frac{365}{750}-\frac{360}{750}=\frac{5}{750}.$$

Remark B15. *The willingness to pay is independent of the number of informed students.*

Remark B16. *Information acquisition has very large externalities.*

Remark B17. *If we only elicit one amount of willingness to pay, an student reports* $\frac{5}{750}$.

B.6 Scenario (5): (2) + acquisition of information on others' preferences

The Immediate Acceptance Mechanism Now suppose everyone knows her own preferences but not others', while the distribution of preferences is common knowledge. With some abuse of terminology, an student is informed if she knows the realization of others' preferences and whether each student is informed or uninformed. An uninformed student knows her own preferences, but neither others' preference realizations nor how many being informed is revealed to uninformed students.

Here, two pieces of information, i.e., other students' preferences and whether they are informed or not, are always acquired together, never separately. As we hypothesize that researching others' preferences is wasteful given independent preferences, we thus study cases where the incentives for wasteful information acquisition is high.

Note that a type-$(1, 0.1, 0)$ student has no incentive to acquire information. Therefore, the discussion of information acquisition is conditional on one's own type being $(1, 1.1, 0)$.

Willingness to pay for information on others' preferences can be similarly defined in the following three cases:

$$w_0^{other} : \text{ when no other informed students;}$$
$$w_1^{other} : \text{ when there is another informed student;}$$
$$w_2^{other} : \text{ when there are two other informed students.}$$

Table B8 summarizes the equilibrium strategies and ex ante payoffs for informed and uninformed players under the Immediate Acceptance mechanism.

Table B8: Willingness to Pay for Information on Others' Payoffs under IA

# of Players		Ex Ante Payoff		Exp. Payoff to Type-(1,1.1,0)		WTP for info
Informed	Uninformed	Informed	Uninformed	Informed	Uninformed	given type-(1,1.1,0)
0	3	-	$\frac{397}{750}$	-	$\frac{681}{750}$	$\frac{9}{750}$
1	2	$\frac{398.8}{750}$	$\frac{396.1}{750}$	$\frac{690}{750}$	$\frac{676.5}{750}$	$\frac{0.6428}{750}$
2	1	$\frac{396.22857}{750}$	$\frac{398.54}{750}$	$\frac{677.14286}{750}$	$\frac{688.71}{750}$	0
3	0	$\frac{397}{750}$	-	$\frac{681}{750}$	-	

When there is no other students informed, the third student can stay uninformed and obtain $\frac{397}{750}$ ex ante, or $\frac{681}{750}$ conditional on being type $(1, 1.1, 0)$, as in Scenario (2). If she acquires information on others and becomes informed, the school choice game has the following equilibrium:

$$(Two)\ Uninformed : \sigma\left((1, 1.1, 0)\right) = (b, a, c)\,; \sigma\left((1, 0.1, 0)\right) = (a, b, c)\,;$$

and the informed player's strategies are summarized in Table B9:

Table B9: Equilibrium Strategies of the Player Informed of Others' Payoffs under IA

Others' Preferences	Ex Ante Probability	Strategy: Informed Player (1, 0.1, 0)	(1, 1.1, 0)	Ex Post Payoff: Informed Player (1, 0.1, 0)	(1, 1.1, 0)
(1, 0.1, 0) (1, 0.1, 0)	16/25	(a, b, c)	(b, a, c)	11/30	11/10
(1, 1.1, 0) (1, 0.1, 0)	8/25	(a, b, c)	(b, a, c)	1/2	11/20
(1, 1.1, 0) (1, 1.1, 0)	1/25	(a, b, c)	(a, b, c)	1	1

The ex ante payoff to the informed player is:

$$\begin{aligned}
&\frac{4}{5}\left(\frac{11}{30}\frac{16}{25} + \frac{1}{2}\frac{8}{25} + \frac{1}{25}\right) + \frac{1}{5}\left(\frac{11}{10}\frac{16}{25} + \frac{11}{20}\frac{8}{25} + \frac{1}{25}\right) \\
&= \frac{4}{5}\frac{326}{750} + \frac{1}{5}\frac{690}{750} \\
&= \frac{398.8}{750}.
\end{aligned}$$

Therefore, conditional on being type $(1, 1.1, 0)$, the willingness to pay is:

$$w_0^{other} = \frac{690}{750} - \frac{681}{750} = \frac{9}{750}$$

The ex ante payoff to uninformed players, given that there is one informed student, is:

$$\begin{aligned}
&\frac{4}{5}\left(\frac{11}{30}\frac{16}{25} + \frac{1}{2}\frac{8}{25} + \frac{1}{25}\right) + \frac{1}{5}\left(\frac{11}{10}\frac{16}{25} + \frac{11}{20}\frac{8}{25} + \frac{11}{20}\frac{1}{25}\right) \\
&= \frac{4}{5}\frac{326}{750} + \frac{1}{5}\frac{676.5}{750} \\
&= \frac{396.1}{750}
\end{aligned}$$

They have no incentives to deviate, and they are worse off than in Scenario (2).

When there is one other student informed, the third student can stay uninformed and obtain $\frac{396.1}{750}$ ex ante, or $\frac{676.5}{750}$ when being type $(1, 1.1, 0)$ as above. If she acquires information on others and becomes informed, the school choice game has the following equilibrium in pure strategies:

$$(One)\ Uninformed : \sigma\left((1, 1.1, 0)\right) = (b, a, c)\,; \sigma\left((1, 0.1, 0)\right) = (a, b, c)\,;$$

and the informed player's strategy is in the following table (Table B10):

Table B10: Equilibrium Strategies with One Informed and Two Uninformed Players under IA

Others' Preferences		Ex Ante	Strategy: Informed Player		Ex Post Payoff: Informed Player	
Uninformed	Informed	Probability	(1, 0.1, 0)	(1, 1.1, 0)	(1, 0.1, 0)	(1, 1.1, 0)
(1, 0.1, 0)	(1, 0.1, 0)	16/25	(a, b, c)	(b, a, c)	11/30	11/10
(1, 1.1, 0)	(1, 0.1, 0)	4/25	(a, b, c)	(b, a, c)	1/2	11/20
(1, 0.1, 0)	(1, 1.1, 0)	4/25	(a, b, c)	(b, a, c)	1/2	11/20
(1, 1.1, 0)	(1, 1.1, 0)	1/25	(a, b, c)	(a, b, c) w/ prob. 6/7[a] (b, a, c) w/ prob. 1/7[a]	1	4/7

a. We may allow one informed student to play (a,b,c) and the other informed to play (b,a,c), which is a pure-strategy Nash equilibrium. When either of the two informed students has the same probability to play (a,b,c), the expected payoff of everyone is $31/40$ ($> 4/7$). This leads to a type-(1,1.1,0) student willing to pay 6.75/750 to become informed, given that there is only one more informed student. Moreover, this makes the third uninformed student willing to pay $4.5/750$ to be informed. In any case, the interval prediction of WTP for information on others' preferences, which is $[0, 9/750]$ for a type-(1,1.1,0) student, includes all these values.

The ex ante payoff to an informed player is:

$$\begin{aligned}
&\frac{4}{5}\left(\frac{11}{30}\frac{16}{25} + \frac{1}{2}\frac{8}{25} + \frac{1}{25}\right) + \frac{1}{5}\left(\frac{11}{10}\frac{16}{25} + \frac{11}{20}\frac{8}{25} + \frac{4}{7}\frac{1}{25}\right) \\
&= \frac{4}{5}\frac{326}{750} + \frac{1}{5}\frac{158}{175} \\
&= \frac{396.22857}{750}.
\end{aligned}$$

Therefore, conditional on being type $(1, 1.1, 0)$, the willingness to pay given there is another informed agent is:

$$w_1^{other} = \frac{158}{175} - \frac{676.5}{750} = \frac{0.6428}{750}.$$

When there are two other agents are informed, if the third chooses to be informed, we are back

to Scenario (1). Conditional on being type $(1, 1.1, 0)$, her payoff is $\frac{681}{750}$ if being informed. When two other agents are informed, the third agent, if being uninformed, has a payoff of:

$$\begin{aligned}&\frac{4}{5}\left(\frac{11}{30}\frac{16}{25}+\frac{1}{2}\frac{8}{25}+\frac{1}{25}\right)+\frac{1}{5}\left(\frac{11}{10}\frac{16}{25}+\frac{11}{20}\frac{8}{25}+\frac{46.9}{49}\frac{1}{25}\right)\\&=\frac{4}{5}\frac{326}{750}+\frac{1}{5}\frac{688.71}{750}=\frac{398.54}{750}.\end{aligned}$$

Therefore,

$$w_2^{other}=\frac{681}{750}-\frac{688.71}{750}<0.$$

That is, when the other two students are informed, the third student does not have incentive to acquire information.

Remark B18. *When only one amount of willingness to pay is elicited, a type-*$(1, 1.1, 0)$ *student reports a number in* $\left[0, \frac{9}{750}\right]$. *Averaging over all student ex ante, the WTP for information on others's preferences is in* $[0, \frac{1.8}{750}]$.

The DA Mechanism Since reporting truthfully is a dominant strategy, there is no incentive to know others' preferences.

C Analyses of the Game in the Experiment with Risk-Averse Students

This appendix compares risk-neutral and risk-averse students in terms of their willingness to pay for information.

Risk-neutral students have the same cardinal preferences as before (Table 1), and risk-averse students have their von Neumann–Morgenstern utilities associated with each schools as in Table C11.

Table C11: Preference/Payoff Table for Risk-Averse Students

Students	$s = a$	$s = b$	$s = c$
1	1	$\sqrt{0.1}$w/ prob. 4/5; $\sqrt{1.1}$w/ prob. 1/5	0
2	1	$\sqrt{0.1}$w/ prob. 4/5; $\sqrt{1.1}$w/ prob. 1/5	0
3	1	$\sqrt{0.1}$w/ prob. 4/5; $\sqrt{1.1}$w/ prob. 1/5	0

Note that $\sqrt{0.1} \approx 0.316$, and $\sqrt{1.1} \approx 1.049$. In the following, we evaluate the ex ante welfare/payoff, i.e., before the realization of the utility associated with school b. Note that ex ante, the expected payoff of being assigned to b is 0.463 ($\approx \frac{4*\sqrt{0.1}}{5} + \frac{1*\sqrt{1.1}}{5}$) and is better than 1/3 of a for any student.[18]

Conclusion C1. *WTP for own values is smaller for risk-averse students; WTP for others' values is similar when measured as the percentage of expected utilities, but it is much lower when measured in dollars.*

C.1 Information on Own Values

Willingness to pay can be measured in dollars. However, one dollar does not mean the same in the two cases. Therefore, it is also measured as a percentage of the expected utility under complete information and then of the one under no information.

In the above table, the complete information expected utility with risk averse under IA is 0.558, while the one with no info is 0.488. The corresponding two expected values for the risk neutral students are $\frac{397}{750} = 0.529$ and $\frac{325}{790} = 0.411$, respectively.

[18] If $u(x) = \frac{x^{(1-r)}}{1-r}$, the expected utility from being matched with b is increasing in r which is also the coefficient of relative risk aversion.

Table C12: WTP for Info on Own Values: Risk-Averse and Risk-Neutral Students under IA

# of Other	in dollars		Pctg. of Complete Info EU		Pctg. of no Info EU	
Informed Players	Averse	Neutral	Averse	Neutral	Averse	Neutral
0	0.077	0.080	13%	15%	15%	18%
1	0.062	0.066	11%	12%	12%	15%
2	0.049	0.052	8%	10%	9%	12%

Notes: WTP in dollars with risk aversion is calculated as follows: we first obtain the certainty equivalence in dollars of the two expected utilities and then take the difference.

Table C13: Willingness to Pay for Info on Own Values: Risk-Averse and Risk-Neutral Students under the DA

# of Other	in EU		Pctg. of Complete Info EU		Pctg. of no Info EU	
Informed Players	Averse	Neutral	Averse	Neutral	Averse	Neutral
0	0.003	0.007	0.57%	1.37%	0.61%	1.54%
1	0.004	0.007	0.57%	1.37%	0.61%	1.54%
2	0.004	0.007	0.57%	1.37%	0.61%	1.54%

Notes: WTP in dollars with risk aversion is calculated as follows: we first obtain the certainty equivalence in dollars of the two expected utilities and then take the difference.

C.2 Information on Others' Values

Note that the willingness to pay for information given one's type being $(1, 0.1, 0)$ is always zero. Therefore, the table below is conditional on the student being type $(1, 1.1, 0)$.

Table C14: WTP for Info on Others' Values: Risk-Averse and Risk-Neutral Students under IA

# of Other	in EU		Pctg. of Complete Info EU		Pctg. of no Info EU	
Informed Players	Averse	Neutral	Averse	Neutral	Averse	Neutral
0	0.023	0.012	2%	2%	3%	3%
1	< 0	0.001	-	0%	-	0%
2	< 0	< 0	-	-	-	-

Notes: WTP in dollars with risk aversion is calculated as follows: we first obtain the certainty equivalence in dollars of the two expected utilities and then take the difference.

D Experimental Instructions: DA, Own Value

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	$\longrightarrow$	$\square$	
3	$\longrightarrow$	B	$\longrightarrow$	$\square$	
	$\longrightarrow$	C	$\longrightarrow$	$\square$	

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?

E Additional Analyses of Experimental Data

In this appendix, we present robustness checks on our earlier analyses in Section 5.1 on willingness to pay for information.

E.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 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, and the endogeneity issue is not a concern.

Corresponding to Table 4 in Section 5.1, we present Table E15 where 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 investigate determinants of WTP in random and fixed effects panel regressions. In all specifications, our outcome variable is the subject-period WTP. Our specification is as follows:

$$WTP_{i,t} = \alpha_i + \beta_1 highB\_IA_{i,t} + \beta_2 highB\_DA_{i,t} + \beta_3 WTP\_guess_{i,t} + ... + Controls_{i,t} + \varepsilon_{i,t},$$

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

The endogeneity of $WTP\_guess_{i,t}$ is plausible if there are some common shocks in period t makes everyone's $WTP_{i,t}$ and $WTP\_guess_{i,t}$ higher. We address this with an IV approach where the lagged $WTP\_guess_{i,t-1}$ is the instrumental variable. Clearly, $WTP\_guess_{i,t-1}$ is correlated with $WTP\_guess_{i,t}$, as there might some persistence in one's guess of others WTP. Moreover, conditional on what others have done in the previous period, $WTP\_average_{i,t-1}$, what i guessed in $t-1$ should not affect her decision in period t directly.

Table E15: Determinants of Subject-Average WTP: Linear Regression

	(1) Full Sample	(2) Sub-sample 1	(3) Sub-sample 1	(4) Sub-sample 2
IA_OwnValue	6.56***	6.41***	5.70***	6.11***
	(0.55)	(0.56)	(0.90)	(1.58)
IA_OtherValue	4.51***	4.31***	4.00***	4.34**
	(0.48)	(0.54)	(0.91)	(1.72)
DA_OwnValue	4.44***	4.16***	3.66***	4.18**
	(0.63)	(0.70)	(0.87)	(1.60)
DA_OtherValue	2.21***	1.92***	2.02**	2.80*
	(0.30)	(0.27)	(0.91)	(1.61)
Misunderstanding DA[a]			5.44***	4.87**
			(1.85)	(2.16)
Curiosity			0.29***	0.28***
			(0.05)	(0.04)
Order: Costly-Free			1.61***	1.62***
			(0.36)	(0.32)
Risk Aversion			-0.27**	-0.20*
			(0.11)	(0.11)
Female				-0.73
				(0.44)
Graduate Student				-0.71
				(0.45)
Black				-0.81
				(0.56)
Asian				-1.28*
				(0.63)
Hispanic				-0.67
				(0.48)
N	288	241	241	233
R^2	0.65	0.63	0.73	0.75

Notes:
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 making irrational choices. Column (4) further excludes observations with missing age/gender/ehnicity information and includes other controls: age, ACT score, SAT score, dummy for ACT score missing, dummy for SAT score Missing, and dummy for degree missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
a. "Misunderstanding DA" is defined as the percentage of times when the subject played dominated strategies (i.e., non-truth-telling) in the OwnValue or OtherValue treatment of DA in periods without information acquisition. Mean $= 0.09$, standard deviation $= 0.14$ among all subjects ($n = 144$) played the information acquisition game under DA. Only periods without information acquisition, i.e., with no information or free information provision, are considered. This variable equals to zero for both treatment of IA, because dominant strategies are not defined under IA.

Fixed-effect results are shown in Table E16. 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 $WTP\_guess_{i,t-1}$. Column (5) shows the first-stage result.

Table E16: Determinants of WTP: Fixed Effects Model and IV Regression Results

	(1) FE	(2) FE	(3) FE	(4) IV	(5) 1st Stage
high_B × IA_OtherValue	2.39**	2.36**	2.39**	2.38***	-0.06
	(0.94)	(0.98)	(0.94)	(0.88)	(0.25)
high_B × DA_OtherValue	-0.00	-0.03	-0.01	-0.01	-0.04
	(0.44)	(0.37)	(0.44)	(0.38)	(0.23)
Accumulated wealth	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Successfully acquired info in $t-1$	0.01	-0.11	-0.03	-0.03	-0.19*
	(0.26)	(0.25)	(0.24)	(0.22)	(0.11)
Period	-0.14**	-0.19**	-0.14**	-0.15**	-0.08
	(0.06)	(0.08)	(0.06)	(0.06)	(0.06)
Period × Free-Costly	0.09	0.08	0.09	0.08	-0.01
	(0.06)	(0.06)	(0.06)	(0.05)	(0.03)
Average WTP of others in $t-1$	-0.03	0.08***	-0.02	0.00	0.17***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Guess of others' WTP in $t-1$	-0.05	0.12**			0.26***
	(0.06)	(0.05)			(0.03)
Guess of others' WTP in t	0.63***		0.61***	0.44**	
	(0.08)		(0.08)	(0.18)	
N	2097	2097	2097	2097	2097
R^2	0.19	0.07	0.19	0.18	0.20

Notes: Outcome variable is WTP for information of each subject in each period. Regressions exclude participants with multiple switching points in the Holt-Laury lottery game or making irrational choices as well as subjects with missing age/gender/ehnicity information. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
"Guess of others' WTP in $t-1$" is used as IV for "Guess of others' WTP in t" (1st-stage results in the last column, i.e. outcome = "Guess of others' WTP in t").

Comparing column (1) with column (2), the WTP_guess explains 12% of the variations in WTP – excluding WTP_guess decreases the R squared from 0.19 to 0.07. Besides, when $WTP\_guess_{i,t}$ is included $WTP\_guess_{i,t-1}$ has an insignificant coefficient both statistically and economically.

We then consider $WTP\_guess_{i,t-1}$ as an IV for use $WTP\_guess_{i,t}$. Column (5) presents the first-stage result which shows that $WTP\_guess_{i,t-1}$ 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 increased. We then 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 $WTP\_guess_{i,t}$ is treated as exogenous. It turns out that the test statistic is 0.50 (p-value 0.48), which leads us to conclude that $WTP\_guess_{i,t}$ is exogenous.

In summary, the results in Table E16 are similar to those in Tables 5 and 6 from a random effect Tobit model. Moreover, the IV results in Column (4) are not that different from other results in Table E16.

When we repeat the same analyses with random effect panel regressions, we obtain similar results as well (Table E17).

E.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 the pooled regression (Table E18). Although results change to some extent, we still find "conformity" explains the most part of the excess WTP.

Table E17: Determinants of WTP: Random Effects Model and IV Regression Results

	(1) RE	(2) RE	(3) RE	(4) IV	(5) 1st Stage
IA_OwnValue	1.28***	2.15***	1.19***	1.42**	0.51***
	(0.42)	(0.42)	(0.43)	(0.58)	(0.19)
IA_OtherValue	0.18	0.58	0.14	0.25	0.19
	(0.51)	(0.49)	(0.49)	(0.68)	(0.13)
DA_OwnValue	1.13***	1.13**	1.10***	1.13**	-0.18
	(0.35)	(0.44)	(0.35)	(0.51)	(0.15)
high_B $\times$ IA_OtherValue	2.48***	2.46**	2.48***	2.48***	-0.04
	(0.92)	(0.96)	(0.93)	(0.93)	(0.21)
high_B $\times$ DA_OtherValue	0.10	0.07	0 0.10	0.10	-0.10
	(0.41)	(0.34)	(0.41)	(0.43)	(0.18)
Accumulated wealth	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Successfully acquired info in $t-1$	0.67***	0.65***	0.64***	0.66**	-0.10
	(0.23)	(0.22)	(0.21)	(0.27)	(0.08)
Period	-0.05	-0.07	-0.05	-0.05	-0.00
	(0.06)	(0.06)	(0.06)	(0.06)	(0.02)
Period $\times$ Free-Costly	0.08	0.08	0.08	0.08	0.03
	(0.06)	(0.06)	(0.06)	(0.05)	(0.02)
Average WTP of others in $t-1$	-0.06***	0.06***	-0.06***	-0.04**	0.19***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Guess of others' WTP in $t-1$	-0.04	0.26***			0.64***
	(0.06)	(0.05)			(0.04)
Guess of others' WTP in t	0.70***		0.67***	0.59***	
	(0.07)		(0.06)	(0.12)	
Misunderstanding DA	2.38	3.28*	2.35	2.53	0.84
	(1.70)	(1.94)	(1.69)	(1.99)	(0.62)
Curiosity	0.21***	0.26***	0.21***	0.22***	0.05***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.01)
Costly-Free	0.52	1.08*	0.50	0.61	0.51**
	(0.55)	(0.61)	(0.54)	(0.63)	(0.21)
Risk Aversion	-0.21**	-0.19**	-0.21**	-0.20*	0.01
	(0.10)	(0.10)	(0.10)	(0.11)	(0.03)
# of Observations	2097	2097	2097	2097	2097
# of Subjects	231	231	231	231	231

Notes: The regression sample is the same as that in Column (5) in Table 6. Each of the 231 subjects has 9 observations from 9 periods. Estimates are form 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, and dummy for degree missing. Standard errors clustered at session level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
"Guess of others' WTP in $t-1$" is used as IV for "Guess of others' WTP in t" (1st-stage results in the last column, i.e. outcome = "Guess of others' WTP in t").

Table E18: Decomposition of Subject WTP for Information Based on the Pooled Regression

	IA OwnValue	**IA OtherValue**	**DA OwnValue**	**DA OtherValue**
WTP: data	6.49	4.29	4.30	1.78
	(4.86)	(4.67)	(4.30)	(2.81)
Model prediction[a]	6.34	4.15	4.33	1.75
	(2.71)	(2.93)	(2.65)	(1.71)
Risk aversion	-0.29	-0.25	-0.32	-0.25
	(0.25)	(0.30)	(0.29)	(0.24)
Cognitive load	0.47	0.36	0.42	0.22
	(0.47)	(0.43)	(0.44)	(0.29)
Learning over periods	0.46	0.36	0.40	0.23
	(0.38)	(0.35)	(0.36)	(0.25)
Curiosity	1.70	1.21	1.01	0.40
	(1.67)	(1.45)	(1.37)	(0.87)
Conformity	4.27	2.70	2.49	1.07
	(1.94)	(2.09)	(1.85)	(1.18)
Misunderstanding DA			0.41	0.20
			(0.69)	(0.45)
Total[b]	5.45	3.49	3.30	1.34
	(2.68)	(2.76)	(2.56)	(1.63)
Explained by other factors[c]	1.03	0.81	0.81	0.37
	(4.00)	(3.88)	(3.17)	(2.21)
Theoretical prediction[d]	[5.2, 8]	[0, 0.24]	0.67	0
# of Observations	549	495	558	495
# of Subjects	61	55	62	55

Notes: Decompositions are based on a random effects panel Tobit model that pools observations from all four treatment (Columns (5) 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 truncated to be in $[0, 15]$.

b. "Total" 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 truncation at 0 and 15.

c. "Explained by other factors" is the difference between the observed WTP and the total WTP explained by the six factors.

d. These are the theoretical predictions for risk neutral subjects.