Appendix

A Proofs and additional theoretical analysis

In the subsequent analysis, we omit the subscript i when referring to any student.

A.1 Proof of Proposition 4

The proof of Proposition 4, as well as several subsequent results, will use the following lemma. Suppose that $\lambda^1(x)$ and $\lambda^2(x)$ are two probability mass functions (PMFs) of distributions over the same discrete domain Ψ, and that and $\Lambda^1(x)$ and $\Lambda^2(x)$ are their corresponding cumulative distribution functions (CDFs). Let $\eta(x)$ be the difference between these two PMFs, that is, $\eta(x) \equiv \lambda^1(x) - \lambda^2(x)$.

Lemma 1. *If there exists a threshold $\hat{x} \in \Psi$ such that $\eta(x) \leq 0$ for $x \leq \hat{x}$ and $\eta(x) > 0$ otherwise, then Λ^1 first-order stochastically dominates Λ^2, that is, $\Lambda^1(x) \leq \Lambda^2(x), \forall x$.*

Proof. Denote the smallest and largest values in Ψ as $\underline{x}$ and $\bar{x}$, respectively. Denote x^+ as the smallest element in Ψ that is greater than x (for $x < \bar{x}$). Given the definition of $\hat{x}$, we know that when $x \leq \hat{x}$,

$$\Lambda^1(x) - \Lambda^2(x) = \sum_{x'=\underline{x}}^{x} \eta(x') \leq 0.$$

When $x > \hat{x}$,

$$\begin{aligned}
\Lambda^1(x) - \Lambda^2(x) &= \sum_{x'=\underline{x}}^{x} \eta(x') \\
&= \sum_{x'=\underline{x}}^{\hat{x}} \eta(x') + \sum_{x'=\hat{x}^+}^{x} \eta(x') \\
&\leq \sum_{x'=\underline{x}}^{\hat{x}} \eta(x') + \sum_{x'=\hat{x}^+}^{x} \eta(x') + \sum_{x'=x^+}^{\bar{x}} \eta(x') \\
&= \sum_{x'=\underline{x}}^{\bar{x}} \eta(x') \\
&= 0.
\end{aligned}$$

The last step is due to the definition of $\eta(x)$. Therefore, we have $\Lambda^1(x) \leq \Lambda^2(x), \forall x$. The inequality holds strictly for some x as long as the two distributions are not identical. Hence, Λ^1 first-order stochastically dominates Λ^2. □

Now we prove Proposition 4.

Proof. We write the expected utility for those unsearched universities in $C \setminus C^S$ as

$$V(0) = \sum_{j=1}^{m} f^0(j,\alpha) u^j$$

and the updated expected utility for the university relatively ranked γth in C^S ($\gamma = 1, \ldots, \alpha+1$) as

$$V^\gamma(\alpha) = \sum_{j=1}^{m} f^\gamma(j,\alpha) u^j,$$

in which

$$f^0(j,\alpha) = \frac{1}{m}$$

and

$$f^\gamma(j,\alpha) = \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}}$$

are the PMFs of the distributions over the set of cardinal utilities $\{u^1, ..., u^m\}$; let $F^0(j,\alpha)$ and $F^\gamma(j,\alpha)$ be the corresponding CDFs.

(1) We first show that $V^1(\alpha) > V(0)$ for any $\alpha = 1, 2, \ldots, m-1$. Let $g^{1,0}(j,\alpha)$ be the difference between the two PMFs $f^1(j,\alpha)$ and $f^0(j,\alpha)$, that is,

$$g^{1,0}(j,\alpha) \equiv f^1(j,\alpha) - f^0(j,\alpha) = \frac{\begin{pmatrix} m-j \\ \alpha \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{1}{m}.$$

We can see from the above definition that (i) given α and m, $g^{1,0}(j,\alpha)$ is decreasing in j;[31] (ii) $g^{1,0}(m,\alpha) = -\frac{1}{m} < 0$; and (iii) $g^{1,0}(1,\alpha) = \frac{\alpha+1}{m} > 0$. Therefore, there exists an integer $\hat{j}$ such that $g^{1,0}(j,\alpha) \leq 0$ when $\hat{j} \leq j \leq m$, and $g^{1,0}(j,\alpha) > 0$ when $1 \leq j \leq \hat{j}' - 1$. Because $u^1 > u^2 > \ldots > u^m$, $u^{\hat{j}}$ is equivalent to the threshold $\hat{x}$ in Lemma 1. According to Lemma 1, $F^1(j,\alpha)$ first-order stochastically dominates $F^0(j,\alpha)$, that is, $V^1(\alpha) > V(0)$ for any $\alpha = 1, 2, \ldots, m-1$.

Next, we show $V^{\alpha+1}(\alpha) < V(0)$ for any $\alpha = 1, 2, \ldots, m-1$. Let $g^{0,\alpha+1}(j,\alpha)$ be the difference

[31] $f^1(j,\alpha)$ equals zero when $j > m-\alpha$ and is strictly decreasing in j when $j \leq m-\alpha$.

between the two PMFs $f^0(j,\alpha)$ and $f^{\alpha+1}(j,\alpha)$, that is,

$$g^{0,\alpha+1}(j,\alpha) \equiv f^0(j,\alpha) - f^{\alpha+1}(j,\alpha) = \frac{1}{m} - \frac{\begin{pmatrix} j-1 \\ \alpha \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}}.$$

We know from the above definition that (i) given α and m, $g^{0,\alpha+1}(j,\alpha)$ is decreasing in j;[32] (ii) $g^{0,\alpha+1}(m,\alpha) = -\frac{\alpha}{m} < 0$; and (iii) $g^{0,\alpha+1}(1,\alpha) = \frac{1}{m} > 0$. Therefore, there exists an integer $\hat{j}'$ such that $g^{0,\alpha+1}(j,\alpha) \leq 0$ when $\hat{j}' \leq j \leq m$, and $g^{0,\alpha+1}(j,\alpha) > 0$ when $1 \leq j \leq \hat{j}'-1$. Again, according to Lemma 1, $F^0(j,\alpha)$ first-order stochastically dominates $F^{\alpha+1}(j,\alpha)$, that is, $V^{\alpha+1}(\alpha) < V(0)$ for any $\alpha = 1,2,\ldots,m-1$.

(2) We first show that $V^\gamma(\alpha) > V^{\gamma+1}(\alpha)$ for any $\gamma = 1,2,\ldots,\alpha+1$ and $\alpha = 1,2,\ldots,m-1$. Let $g^{\gamma,\gamma+1}(j,\alpha)$ be the difference between the two PMFs $f^\gamma(j,\alpha)$ and $f^{\gamma+1}(j,\alpha)$, that is,

$$\begin{aligned} g^{\gamma,\gamma+1}(j,\alpha) &\equiv f^\gamma(j,\alpha) - f^{\gamma+1}(j,\alpha) \\ &= \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\begin{pmatrix} j-1 \\ \gamma \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} \end{aligned}$$

Because $f^\gamma(j,\alpha) = f^{\gamma+1}(j,\alpha) = 0$ when $j > m-\alpha+\gamma$ or $j < \gamma$, we re-define $f^\gamma(j,\alpha)$, $f^{\gamma+1}(j,\alpha)$, and $g^{\gamma,\gamma+1}(j,\alpha)$ to be the PMFs over the set $\{u^\gamma, ..., u^{m-\alpha+\gamma}\}$. For $\gamma < j < m-\alpha+\gamma$, $f^\gamma(j,\alpha) > 0$, $f^{\gamma+1}(j,\alpha) > 0$, and

$$g^{\gamma,\gamma+1}(j,\alpha) \propto (m+1)\gamma - (\alpha+1)j.$$

Because $g^{\gamma,\gamma+1}(\gamma,\alpha) = f^\gamma(\gamma,\alpha) - 0 > 0$ and $g^{\gamma,\gamma+1}(m-\alpha+\gamma,\alpha) = 0 - f^{\gamma+1}(m-\alpha+\gamma,\alpha) < 0$, we know $g^{\gamma,\gamma+1}(\gamma,\alpha) \leq 0$ if $\frac{(m+1)\gamma}{\alpha+1} \leq j \leq m-\alpha+\gamma$ and $g^{\gamma,\gamma+1}(\gamma,\alpha) > 0$ if $\gamma \leq j < \frac{(m+1)\gamma}{\alpha+1}$. According to Lemma 1, $F^\gamma(j,\alpha)$ first-order stochastically dominates $F^{\gamma+1}(j,\alpha)$, that is, $V^\gamma(\alpha) > V^{\gamma+1}(\alpha)$ for any $\alpha = 1,2,\ldots,m-1$.[33]

Now we have shown that given any $\alpha = 1,2,\ldots,m-1$, $V^1(\alpha) > V(0)$, $V^{\alpha+1}(\alpha) < V(0)$, and $V^\gamma(\alpha) > V^{\gamma+1}(\alpha), \forall\gamma = 1,2,\ldots,\alpha+1$. Therefore, by the mean value theorem, there exists a threshold $\hat{\gamma}(\alpha)$ at which (i) $V^\gamma(\alpha) > V(0)$ for all $\gamma \leq \hat{\gamma}(\alpha)$, and (ii) $V^\gamma(\alpha) \leq V(0)$ otherwise. □

[32] $f^{\alpha+1}(j,\alpha)$ equals zero when $j < \alpha+1$ and is strictly increasing in j when $j \geq \alpha+1$.

[33] Since $\frac{(m+1)\gamma}{\alpha+1}$ is not necessarily an integer, the threshold in Lemma 1 can be considered as $u^{\left\lceil\frac{(m+1)\gamma}{\alpha+1}\right\rceil}$, where $\left\lceil\frac{(m+1)\gamma}{\alpha+1}\right\rceil$ is the ceiling of $\frac{(m+1)\gamma}{\alpha+1}$, i.e., the smallest integer greater than $\frac{(m+1)\gamma}{\alpha+1}$.

A.2 Information acquisition under DirSD

In this section, we discuss the role of information and students' information acquisition strategy under DirSD in a one-tier market.

Proposition 5. *In a one-tier market under DirSD, the marginal benefit of additional information is non-negative and can be non-monotonic.*

The proof of Proposition 5 will use the following two lemmas.

Lemma 2. *For any* $\alpha = 2, \ldots, m-1$ *and* $\gamma = 1, 2, \ldots, \alpha+1$, $V^{\gamma}(\alpha) > V^{\gamma}(\alpha-1)$.

Proof. Suppose a student has completed $(\alpha - 1)$ steps of searching and is considering the benefit of step α, $\alpha = 2, \ldots, m-1$. When this additional search step is conducted, the change in the expected value is given by

$$\begin{aligned} V^{\gamma}(\alpha) - V^{\gamma}(\alpha-1) &= \sum_{j=1}^{m} f^{\gamma}(j,\alpha)\, u^{j} - \sum_{j=1}^{m} f^{\gamma}(j,\alpha-1)\, u^{j} \\ &= \sum_{j=\gamma}^{m-\alpha+\gamma-1} \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma+1}}{\binom{m}{\alpha+1}} u^{j} - \sum_{j=\gamma}^{m-\alpha+\gamma} \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma}}{\binom{m}{\alpha}} u^{j}. \end{aligned}$$

Let $h(j)$ be the difference between the two PMFs $f^{\gamma}(j,\alpha)$ and $f^{\gamma}(j,\alpha-1)$, that is,

$$\begin{aligned} h(j) &\equiv f^{\gamma}(j,\alpha) - f^{\gamma}(j,\alpha-1) \\ &= \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma+1}}{\binom{m}{\alpha+1}} - \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma}}{\binom{m}{\alpha}}. \end{aligned}$$

When $j = m-\alpha+\gamma$, $h(m-\alpha+\gamma) = 0 - \frac{\binom{m-\alpha+\gamma-1}{\gamma-1}}{\binom{m}{\alpha}} < 0$. When $j \leq m-\alpha+\gamma-1$,

$$h(j) \propto (m+1)\gamma - (\alpha+1)j.$$

We can see that $h(j) \leq 0$ if $\frac{(m+1)\gamma}{\alpha+1} \leq j \leq m-\alpha+\gamma$ and $h(j) > 0$ if $\gamma \leq j < \frac{(m+1)\gamma}{\alpha+1}$. According to Lemma 1, $F^{\gamma}(j,\alpha)$ first-order stochastically dominates distribution $F^{\gamma}(j,\alpha-1)$. Hence, $V^{\gamma}(\alpha) > V^{\gamma}(\alpha-1)$ for any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha+1$. □

Lemma 3. *For any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha$, $V^{\gamma}(\alpha - 1) > V^{\gamma+1}(\alpha)$.*

Proof. The proof of this lemma is similar to the proof of Lemma 3. Given $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha$,

$$\begin{aligned} V^{\gamma}(\alpha-1) - V^{\gamma+1}(\alpha) &= \sum_{j=1}^{m} f^{\gamma}(j, \alpha-1) u^{j} - \sum_{j=1}^{m} f^{\gamma+1}(j, \alpha) u^{j} \\ &= \sum_{j=\gamma}^{m-\alpha+\gamma} \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma}}{\binom{m}{\alpha}} u^{j} - \sum_{j=\gamma+1}^{m-\alpha+\gamma} \frac{\binom{j-1}{\gamma}\binom{m-j}{\alpha-\gamma}}{\binom{m}{\alpha+1}} u^{j}. \end{aligned}$$

Let $h'(j)$ be the difference between the two PMFs $f^{\gamma}(j, \alpha-1)$ and $f^{\gamma+1}(j, \alpha)$, that is,

$$\begin{aligned} h'(j) &\equiv f^{\gamma}(j, \alpha-1) - f^{\gamma+1}(j, \alpha) \\ &= \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma}}{\binom{m}{\alpha}} - \frac{\binom{j-1}{\gamma}\binom{m-j}{\alpha-\gamma}}{\binom{m}{\alpha+1}}. \end{aligned}$$

When $j = \gamma$, $h'(\gamma) = \frac{\binom{m-\gamma}{\alpha-\gamma}}{\binom{m}{\alpha}} - 0 > 0$. When $j \geq \gamma + 1$,

$$h'(j) \propto (m+1)\gamma - (\alpha+1) j.$$

Therefore, $h'(j) \leq 0$ if $\frac{(m+1)\gamma}{\alpha+1} \leq j \leq m - \alpha + \gamma$ and $h'(j) > 0$ if $\gamma \leq j < \frac{(m+1)\gamma}{\alpha+1}$. According to Lemma 1, $F^{\gamma}(j, \alpha-1)$ first-order stochastically dominates $F^{\gamma+1}(j, \alpha)$. Thus, $V^{\gamma}(\alpha-1) > V^{\gamma+1}(\alpha)$ for any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha$. □

Now we move to prove Proposition 5: under DirSD, the marginal benefit of additional information (1) is non-negative, and (2) can be non-monotonic.

Proof. (1) Suppose the student submits a list $\hat{\succ}_i$ under DirSD. Let Q_i^{θ} be the probability that she will be accepted by the θth ranked university in $\hat{\succ}_i$. Recall that in each step of DirSD, the student whose turn it is is assigned to the highest-ranked university in her submitted list from those that still have vacant seats. Thus, for any probability distribution over one's budget set, DirSD ensures that $Q^{\theta} \geq Q^{\theta'}$ if $\theta < \theta'$, that is, a student is more likely to be admitted by a university if it is ranked higher in her submitted list.

A student who stops searching at step α and chooses the optimal strategy of truth-telling under DirSD, according to propositions 4 and 1, would rank the unsearched universities below the $\hat{\gamma}(\alpha)$th-ranked searched university, but above the $(\hat{\gamma}(\alpha)+1)$th-ranked searched university, and would rank the searched universities according to the discovered relative preferences. Hence, her expected utility is given by

$$u^{DirSD}(\alpha) = \sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^{\theta} V^{\theta}(\alpha) + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^{\theta} V(0) + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha}^{m} Q^{\theta} V^{\theta-m+\alpha+1}(\alpha) - \alpha k,$$

in which αk is the total cost of information. For any $\alpha = 2, \ldots, m-1$, the benefit of conducting an additional search step under DirSD is given by $A(\alpha) - A(\alpha-1)$, where

$$A(\alpha) \equiv \sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^{\theta} V^{\theta}(\alpha) + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^{\theta} V(0) + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha}^{m} Q^{\theta} V^{\theta-m+\alpha+1}(\alpha),$$

and thus

$$A(\alpha-1) = \sum_{\theta=1}^{\hat{\gamma}(\alpha-1)} Q^{\theta} V^{\theta}(\alpha-1) + \sum_{\theta=\hat{\gamma}(\alpha-1)+1}^{\hat{\gamma}(\alpha-1)+m-\alpha} Q^{\theta} V(0) + \sum_{\theta=\hat{\gamma}(\alpha-1)+m-\alpha+1}^{m} Q^{\theta} V^{\theta-m+\alpha}(\alpha-1).$$

Recall that $\hat{\gamma}(\alpha)$ is the threshold at which $V^{\gamma}(\alpha) > V(0)$ for all $\gamma \leq \hat{\gamma}(\alpha)$ and $V^{\gamma}(\alpha) \leq V(0)$ otherwise. From Lemma 2, we know that $V^{\gamma}(\alpha) > V^{\gamma}(\alpha-1)$, $\forall \gamma$, therefore we have $\hat{\gamma}(\alpha-1) \leq \hat{\gamma}(\alpha)$.

First, when $\hat{\gamma}(\alpha-1) = \hat{\gamma}(\alpha)$,

$$\begin{aligned} & A(\alpha) - A(\alpha-1) \\ = & \sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^{\theta} \Bigg[\underbrace{V^{\theta}(\alpha) - V^{\theta}(\alpha-1)}_{>0} \Bigg] + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^{\theta} \Bigg[\underbrace{V(0) - V(0)}_{=0} \Bigg] \\ & + Q^{\hat{\gamma}(\alpha)+m-\alpha} \Bigg[\underbrace{V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)}_{\leq 0} \Bigg] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha+1}^{m} Q^{\theta} \Bigg[\underbrace{V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)}_{<0} \Bigg]. \end{aligned}$$

In the above equation, $\left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right]$ is positive according to Lemma 2, the second term is zero, $\left[V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)\right]$ is non-positive according to the definition of $\hat{\gamma}(\alpha)$, and $[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)]$ is negative according to Lemma 3. Since for any α the total expected value of all universities is a constant equal to $\sum_{j=1}^{m} u^j$, we have

$$\begin{aligned}
&\sum_{\theta=1}^{\hat{\gamma}(\alpha)} \left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right] \\
&= -\left\{ \left[V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)\right] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha+1}^{m} \left[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)\right] \right\}.
\end{aligned}$$

Because Q^{θ} weakly increases as θ decreases, the positive term outweighs the negative, that is,

$$\begin{aligned}
&\sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^{\theta} \left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right] \\
&\geq -\left\{ Q^{\hat{\gamma}(\alpha)+m-\alpha} \left[V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)\right] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha+1}^{m} Q^{\theta} \left[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)\right] \right\}.
\end{aligned}$$

Therefore, we can conclude that $A(\alpha) \geq A(\alpha-1)$ for any $\alpha = 2, \ldots, m-1$.

Next, when $\hat{\gamma}(\alpha-1) < \hat{\gamma}(\alpha)$,

$$\begin{aligned}
&A(\alpha) - A(\alpha-1) \\
&= \sum_{\theta=1}^{\hat{\gamma}(\alpha-1)} Q^{\theta} \Bigg[\underbrace{V^{\theta}(\alpha) - V^{\theta}(\alpha-1)}_{>0}\Bigg] + \sum_{\theta=\hat{\gamma}(\alpha-1)+1}^{\hat{\gamma}(\alpha)} Q^{\theta} \Bigg[\underbrace{V^{\theta}(\alpha) - V(0)}_{>0}\Bigg] + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha-1)+m-\alpha} Q^{\theta} \Bigg[\underbrace{V(0) - V(0)}_{=0}\Bigg] \\
&\quad + \sum_{\theta=\hat{\gamma}(\alpha-1)+m-\alpha+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^{\theta} \Bigg[\underbrace{V(0) - V^{\theta-m+\alpha}(\alpha-1)}_{\geq 0}\Bigg] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha}^{m} Q^{\theta} \Bigg[\underbrace{V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)}_{<0}\Bigg].
\end{aligned}$$

In the above equation, $\left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right]$ is positive according to Lemma 2, $\left[V^{\theta}(\alpha) - V(0)\right]$ is positive according to the definition of $\hat{\gamma}(\alpha)$, the third term is zero, $\left[V(0) - V^{\theta-m+\alpha}(\alpha-1)\right]$ is non-negative according to the definition of $\hat{\gamma}(\alpha-1)$, and $\left[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)\right]$ is negative according to Lemma 3. Similar to the previous case, since the positive difference equals the absolute value of the negative difference but has more weight, we can again conclude that $A(\alpha) \geq A(\alpha-1)$ for any $\alpha = 2, \ldots, m-1$.

Lastly, when $\alpha = 1$, we know from Proposition 4 that $V^1(1) > V(0)$ and $V^2(1) < V(0)$. Thus,

$$\begin{aligned}
A(1) &= Q^1 V^1(1) + \sum_{\theta=2}^{m-1} Q^{\theta} V(0) + Q^m V^2(1) \\
&\geq Q^1 V(0) + \sum_{\theta=2}^{m-1} Q^{\theta} V(0) + Q^m V(0) \\
&= V(0) \equiv A(0)
\end{aligned}$$

Again, the inequality is due to the fact that $V^1(1) - V(0) = -[V^2(1) - V(0)]$ and $Q^1 \geq Q^m$.

Therefore, we can conclude that $A(\alpha) \geq A(\alpha')$ for any $\alpha > \alpha'$. That is, the marginal benefit of additional information is non-negative under DirSD.

(2) Under DirSD, the benefit of information is rescaled by the probabilities Q^θ's. Therefore, depending on the ex-ante probability distribution of a student's budget set, the marginal benefit of additional information is not necessarily decreasing.

With Assumption 1 and uniform within-tier priors, each student knows that the rank-order list submitted by any other student is equally likely to be any ranking in Ω. Thus, from the perspective of student i, she always has an equal chance at every university, and this chance is given by $\left\{P_i\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$. This makes the "name" of a university irrelevant to the student. Let $\{p_i(\beta)\}_{\beta=1,2,\ldots,m}$ be the probability distribution of the number of universities in student i's budget set, that is, $p_i(\beta) = \Pr[|B_i| = \beta]$, $\beta = 1, 2, \ldots, m$.[34] We thus have $P_i\left(\tilde{B}\right) = P_i\left(\tilde{B}'\right) = p_i(\beta) / \binom{m}{\beta}$ for all $\left|\tilde{B}\right| = \left|\tilde{B}'\right| = \beta$. That is, only the number of universities in a student's budget set but not its specific composition matters for her decisions under DirSD. For instance, consider a market with three universities $C = \{c_1, c_2, c_3\}$, each of which has two seats. The budget set of the student ranked third in the exam depends on the submitted rank-order lists of the two students ranked above her. If, for example, they both place university c_3 at the top of their lists, which occurs with probability $\frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$, then the budget set of the student ranked third contains only c_1 and c_2. The same probability $\frac{1}{9}$ should be assigned to all possible two-university compositions of her budget set: $\{c_1, c_2\}$, $\{c_1, c_3\}$, and $\{c_2, c_3\}$.

Suppose a student submits a list $\hat{\succ}$ under DirSD. Then, given $\{p(\beta)\}_{\beta=1,2,\ldots,m}$, the probability that she is accepted by the θth ranked university in $\hat{\succ}$ is given by

$$Q^\theta = \sum_{\beta=1}^{m-\theta+1} \frac{\binom{m-\theta}{\beta-1}}{\binom{m}{\beta}} p(\beta).$$

If a student is assigned to her θth choice, her budget set B has to include her θth choice and exclude the $(\theta - 1)$ universities listed above it. With probability $p(\beta)$, B includes β universities. One of them has to be her θth choice and the remaining $(\beta - 1)$ ones cannot be her top θ choices, which means $\binom{m-\theta}{\beta-1}$ out of the $\binom{m}{\beta}$ possible compositions can occur. Thus, Q^θ sums up, for all possible values of β, the probability that the student will be accepted by her θth choice. We can see that for any probability distribution over one's budget set, DirSD ensures that $Q^\theta \geq Q^{\theta'}$

[34] A student has at least one university in her budget set because we assume the total number of seats exceeds the total number of students. This assumption simplifies our analysis, but is not crucial.

if $\theta < \theta'$, that is, a student is more likely to be admitted by a university if it is ranked higher in her submitted list. This, again, proves the optimality of the truth-telling strategy stated in Proposition 1.

Now, consider our experimental market with one tier. There are six universities and each one has two seats. The cardinal utilities of every student are determined by the experimental payments $\{u^1, u^2, u^3, u^4, u^5, u^6\} = \{40, 34, 28, 22, 16, 10\}$.

For the student ranked first in the exam, $\{p(1), p(2), p(3), p(4), p(5), p(6)\} = \{0, 0, 0, 0, 0, 1\}$ and the marginal benefit of each additional step of searching is $A(1) - A(0) = 7$, $A(2) - A(1) = 3.5$, $A(3) - A(2) = 2.1$, $A(4) - A(3) = 1.4$, $A(5) - A(4) = 1$, which is decreasing.

However, for the student ranked 10th in the exam, $\{p(1), p(2), p(3), p(4), p(5), p(6)\} \approx \{0, 0.57, 0.43, 0, 0, 0\}$ and the marginal benefit of each additional step of searching is approximately $A(1) - A(0) \approx 2.84$, $A(2) - A(1) \approx 1.42$, $A(3) - A(2) \approx 1.87$, $A(4) - A(3) \approx 1.25$, $A(5) - A(4) \approx 1.14$, which is clearly non-monotonic.

This implies that under DirSD, the optimal information acquisition strategy is not necessarily unique in the general setting. However, we ensure the uniqueness for every student in each treatment of our experimental design. □

A.3 Information acquisition under SeqSD

In this section, we discuss the role of information and students' information acquisition strategy under SeqSD in a one-tier market.

Proposition 6. *In a one-tier market under SeqSD,*

(1) the marginal benefit of an additional step of searching among available universities is non-negative and decreasing;

(2) the optimal stopping point α^{SeqSD} in a student's search process is characterized as (i) $\alpha^{SeqSD} = 0$ if $V^1(1) - V(0) < k$; (ii) $\alpha^{SeqSD} = 1$ if $V^1(1) - V(0) > k$ and $V^1(2) - V^1(1) \leq k$; and (iii) α^{SeqSD} solves $\left[V^1\left(\alpha^{SeqSD}\right) - V^1\left(\alpha^{SeqSD} - 1\right)\right] > k$ and $\left[V^1\left(\alpha^{SeqSD} + 1\right) - V^1\left(\alpha^{SeqSD}\right)\right] \leq k$ otherwise.

Proof. (1) First, we show that the marginal benefit of an additional step of searching is non-negative.

Under SeqSD, each student, when being considered, is asked to select from all universities that still have vacant seats, that is, from all universities in her budget set B. Obviously, a student would not search outside her budget set because the information about unavailable universities cannot affect her selection. When searching within B, a student who stops at step α and chooses the optimal strategy of truth-telling under SeqSD, according to Proposition 4, would choose the university with the highest expected utility. Hence, her expected utility at this point is given by $V^1(\alpha) - \alpha k$, in which αk is the total cost of information.

Suppose that when a student is considered by SeqSD, there is only one university left available, that is, her budget set includes only one university ($|B| = 1$). Thus, she obviously has no incentive to invest in any information and the marginal benefit of additional information is constantly zero.

Suppose a student is asked by SeqSD to choose from multiple universities ($|B| > 1$). According to Proposition 1, it is an optimal strategy for her to choose the university with the highest expected utility. Then, the marginal benefit of conducting the first step of searching is $V^1(1) - V(0)$ and the marginal benefit of conducting an additional subsequent search step is given by $V^1(\alpha) - V^1(\alpha - 1)$, $\alpha = \{2, \ldots, |B| - 1\}$. According to Lemma 2, we know that $V^1(\alpha) > V^1(\alpha - 1)$ for any $\alpha = 2, \ldots, m-1$, and we have already shown that $V^1(1) > V(0)$ in Proposition 4. Therefore, $V^1(\alpha) > V^1(\alpha')$ for any $\alpha > \alpha'$ when $|B| > 1$.

Combining the cases of $|B| = 1$ and $|B| > 1$, we can conclude the marginal benefit of additional information is non-negative under SeqSD.

Next, we consider the change in marginal benefit during a student's search process under SeqSD.

We only consider a student with $|B| > 2$ because one with $|B| \leq 2$ would not conduct multiple search steps. The difference in marginal benefits between an increase from $(\alpha - 1)$ to α and an increase from α to $(\alpha + 1)$, $\alpha = 2, \ldots, |B| - 2$ is given by

$$\begin{aligned}
&\left[V^1(\alpha) - V^1(\alpha - 1)\right] - \left[V^1(\alpha + 1) - V^1(\alpha)\right] \\
=&2V^1(\alpha) - V^1(\alpha + 1) - V^1(\alpha - 1) \\
=&2\sum_{j=1}^{m-\alpha} f^1(j, \alpha) u^j - \sum_{j=1}^{m-\alpha-1} f^1(j, \alpha + 1) u^j - \sum_{j=1}^{m-\alpha+1} f^1(j, \alpha - 1) u^j.
\end{aligned}$$

Define $\chi(j)$ as the corresponding difference in PMFs:

$$\begin{aligned}
\chi(j) &\equiv 2f^1(j, \alpha) - f^1(j, \alpha + 1) - f^1(j, \alpha - 1). \\
&= 2\frac{\begin{pmatrix} m-j \\ \alpha \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\begin{pmatrix} m-j \\ \alpha+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+2 \end{pmatrix}} - \frac{\begin{pmatrix} m-j \\ \alpha-1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}}.
\end{aligned}$$

We can calculate that $\chi(m - \alpha + 1) = -\frac{1}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} < 0$, $\chi(m - \alpha) = 2\frac{1}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\alpha}{\begin{pmatrix} m \\ \alpha \end{pmatrix}}$, and thus

$$\begin{aligned}\chi(m-\alpha+1)+\chi(m-\alpha) &= \frac{2}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\alpha+1}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} \\ &\propto \alpha+2-m \\ &\leq 0.\end{aligned}$$

Therefore, the difference in CDFs is non-positive when $j \geq m-\alpha$. For $1 \leq j \leq m-\alpha-1$,

$$\chi(j) \propto (j-1)\left[2(m+1)-(\alpha+2)j\right].$$

We can see that $\chi(j) < 0$ if $\frac{2(m+1)}{\alpha+2} < j \leq m-\alpha-1$ and $\chi(j) \geq 0$ if $1 \leq j \leq \frac{2(m+1)}{\alpha+2}$. According to Lemma 1, the difference in CDFs are non-positive at any j, which indicates first-order stochastic dominance. Hence, we conclude that $\left[V^1(\alpha)-V^1(\alpha-1)\right] > \left[V^1(\alpha+1)-V^1(\alpha)\right]$ for any $\alpha = 2,\ldots,|B|-2$.

When $\alpha = 1$, the difference in marginal benefits between an increase from $(\alpha-1)$ to α and an increase from α to $(\alpha+1)$ is given by

$$\begin{aligned}&\left[V^1(1)-V(0)\right]-\left[V^1(2)-V^1(1)\right] \\ =&2\sum_{j=1}^{m-1} f^1(j,1)\,u^j - \sum_{j=1}^{m-2} f^1(j,2)\,u^j - \frac{1}{m}\sum_{j=1}^{m} u^j\end{aligned}$$

Define $\chi^1(j)$ as the corresponding difference in PMFs:

$$\begin{aligned}\chi^1(j) &\equiv 2f^1(j,1)-f^1(j,2)-\frac{1}{m} \\ &= 2\frac{m-j}{\begin{pmatrix} m \\ 2 \end{pmatrix}} - \frac{\begin{pmatrix} m-j \\ 2 \end{pmatrix}}{\begin{pmatrix} m \\ 3 \end{pmatrix}} - \frac{1}{m}\end{aligned}$$

When $j \leq m-2$,

$$\chi^1(j) = \frac{(j-1)(2(m+1)-3j)}{m(m-1)(m-2)}.$$

Therefore, $\chi^1(j) < 0$ when $\frac{2(m+1)}{3} < j \leq m-2$, and $\chi^1(j) \geq 0$ when $1 \leq j \leq \frac{2(m+1)}{3}$. We can also calculate that $\chi^1(m) = -\frac{1}{m} < 0$ and $\chi^1(m-1) = \frac{4}{m(m-1)} - \frac{1}{m}$. When $m > 4$, $\chi^1(m-1) \leq 0$ and thus $\chi^1(j) \leq 0$ for $\frac{2(m+1)}{3} < j \leq m$ and $\chi^1(j) \geq 0$ for $1 \leq j \leq \frac{2(m+1)}{3}$. According to Lemma 1, the difference in CDFs is non-positive at any j, which indicates first-order stochastic dominance and thus $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$. When $m = 3$, we know $V(0) = \frac{1}{3}(u^1 + u^2 + u^3)$, $V^1(1) = \frac{2}{3}u^1 + \frac{1}{3}u^2$, $V^1(2) = u^1$, and thus $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$ as $u^1 > u^2 > u^3$. When $m = 4$, we know $V(0) = \frac{1}{4}(u^1 + u^2 + u^3 + u^4)$, $V^1(1) = \frac{1}{2}u^1 + \frac{1}{3}u^2 + \frac{1}{6}u^3$, $V^1(2) = \frac{3}{4}u^1 + \frac{1}{4}u^2$, and thus $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$ as $u^1 > u^2 > u^3 > u^4$. Therefore, $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$ holds for any $m > 2$.

To sum up, we can conclude that the marginal benefit of information acquisition within one's budget set decreases under SeqSD.

(2) Since the marginal benefit of an additional search step within B decreases and the marginal cost is constantly k, it is optimal for a student to adopt another search step as long as the marginal benefit exceeds the marginal cost, and stop searching otherwise. Specifically, the optimal stopping point α^{SeqSD} in the search process is characterized as

(i) $\alpha^{SeqSD} = 0$ if $V^1(1) - V(0) < k$;

(ii) $\alpha^{SeqSD} = 1$ if $V^1(1) - V(0) > k$ and $V^1(2) - V^1(1) \leq k$; and

(iii) α^{SeqSD} solves $\left[V^1\left(\alpha^{SeqSD}\right) - V^1\left(\alpha^{SeqSD} - 1\right)\right] > k$ and $\left[V^1\left(\alpha^{SeqSD} + 1\right) - V^1\left(\alpha^{SeqSD}\right)\right] \leq k$ otherwise.

Due to the discreteness of the problem, under some parameters a student may be indifferent between two optimal stopping points if the marginal benefit of the last search step equals k; here, we assume the student chooses the smaller one. Otherwise the optimal stopping point is unique. □

A.4 Proof of Theorem 1

Proof. This proof does not rely on a particular search technology. We define student i's search decision as a choice s_i from the set S_i and denote its cost as $d_i(s_i)$. For the search technology specified in Section 2.3, student i's search decision s_i represents her stopping point $\alpha_i \in \{0, 1, \dots, m-1\}$ and its cost is $d_i(s_i) = \alpha_i k_i$.

According to Proposition 1, a student who adopts the optimal strategy of truth-telling ranks universities according to the expected utilities from high to low under DirSD and chooses the university with the highest expected utility under SeqSD. In both cases, the student bases her submission strategy on her updated beliefs about her preferences after search and is accepted by the university with the highest expected utility in her budget set. The expected utility of this university, denoted as $EU\left(s, \tilde{B}\right)$, is thus determined by the student's search decision and her budget set. With Assumption 1 and uniform within-tier priors, the ex-ante probability distribution of a student's budget set $\left\{P\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$ does not depend on the search strategies of others and is the same under DirSD and SeqSD.

Under DirSD, all students simultaneously submit their rank-order lists. A student takes her search decision s based on the ex-ante probability distribution of her budget set $\left\{P\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$ and needs to pay the information cost $d(s)$. Thus, the optimization problem under DirSD is given by

$$U^{DirSD} = \max_{s\in S}\left[\left(\sum_{\tilde{B}\subseteq C} P\left(\tilde{B}\right) EU\left(s,\tilde{B}\right)\right) - d(s)\right].$$

Under SeqSD, a student selects the preferred university after the higher-ranked students have made their choices. She therefore observes the realization of her budget set before she makes her search decision. Thus, the optimization problem under SeqSD is given by

$$U^{SeqSD} = \sum_{\tilde{B}\subseteq C} P\left(\tilde{B}\right)\max_{s\in S}\left[EU\left(s,\tilde{B}\right) - d(s)\right].$$

Therefore, we have

$$\begin{aligned} U^{DirSD} &= \max_{s\in S}\left[\left(\sum_{\tilde{B}\subseteq C} P\left(\tilde{B}\right) EU\left(s,\tilde{B}\right)\right) - d(s)\right] \\ &= \max_{s\in S}\left[\sum_{\tilde{B}\subseteq C} P\left(\tilde{B}\right)\left(EU\left(s,\tilde{B}\right) - d(s)\right)\right] \\ &\leq \sum_{\tilde{B}\subseteq C} P\left(\tilde{B}\right)\max_{s\in S}\left[EU\left(s,\tilde{B}\right) - d(s)\right] \\ &= U^{SeqSD}. \end{aligned}$$

Hence, we conclude that a student with any probability distribution for her budget set is weakly better off under SeqSD than under DirSD. □

A.5 Alternative search strategy: Information acquisition under SeqSD

The optimal search strategy under SeqSD can be generally characterized as a stopping rule in the proposition below.

Proposition 7. *In a one-tier market under SeqSD, student i's optimal search strategy is as follows.*

(1) Suppose $k_i \geq \frac{b-a}{8}$. The student does not acquire any information and randomly selects a school from her budget set B_i.

(2) Suppose $k_i < \frac{b-a}{8}$ and there is more than one unsearched university remaining in B_i. The student stops only if $\bar{u}_i^S \geq b-\sqrt{2(b-a)k_i}$, and selects university $\bar{c}_i^S$. Otherwise she keeps searching by randomly choosing an unsearched university to investigate.

(3) Suppose $k_i < \frac{b-a}{8}$ and only one unsearched university remains in B_i. If $\bar{u}_i^S \geq b - \sqrt{2(b-a)k_i}$, the student stops and selects $\bar{c}_i^S$. If $a + \sqrt{2(b-a)k_i} < \bar{u}_i^S < b - \sqrt{2(b-a)k_i}$, the student investigates that remaining university and selects the university with the highest utility

among all universities in B_i*. If* $\bar{u}_i^S \leq a + \sqrt{2(b-a)k_i}$*, the student selects that remaining university without investigating it.*

Proof. We omit the subscript i when referring to any student. A student observes her budget set B before conducting a search under SeqSD. For exposition purposes, let us assume $|B| \geq 3$.

Let $U(\alpha)$ be the utility of the student after she has conducted α steps of the search. If the student does not acquire any information, she randomly selects a university and receives the expected utility of

$$E[U(0)] = \frac{a+b}{2}.$$

With a cost of k, she can conduct the first search step by randomly choosing a university in B to investigate. Suppose that she investigates $c_1 \in B$ and discovers u^{c_1}. Then she would select c_1 if $u^{c_1} > \frac{a+b}{2}$ and select an unsearched university otherwise. Thus, the expected utility after the first search step is given by

$$\begin{aligned}
E[U(1)] &= E\left[\max\left\{u^{c_1}, \frac{a+b}{2}\right\} - k\right] \\
&= \Pr\left[u^{c_1} > \frac{a+b}{2}\right] E\left[u^{c_1} \mid u^{c_1} > \frac{a+b}{2}\right] + \Pr\left[u^{c_1} \leq \frac{a+b}{2}\right] \frac{a+b}{2} - k \\
&= \frac{1}{2} \times \frac{\frac{a+b}{2} + b}{2} + \frac{1}{2} \times \frac{a+b}{2} - k \\
&= \frac{3a+5b}{8} - k.
\end{aligned}$$

Therefore, the student will conduct the first search step if $E[U(1)] > \frac{a+b}{2}$, which solves

$$k < \frac{b-a}{8}.$$

Suppose $k < \frac{b-a}{8}$. With another cost of k, she can conduct the second step of her search. Suppose that she investigates $c_2 \in B$ and discovers u^{c_2}. Then, the expected utility after the second search step is given by

$$\begin{aligned}E\left[U\left(2\right)\right]=&E\left[\max\left\{u^{c_2},\max\left\{u^{c_1},\frac{a+b}{2}\right\}\right\}-2k\right]\\=&\Pr\left[u^{c_2}>\max\left\{u^{c_1},\frac{a+b}{2}\right\}\right]E\left[u^{c_2}\mid u^{c_2}>\max\left\{u^{c_1},\frac{a+b}{2}\right\}\right]\\&+\Pr\left[u^{c_2}\leq\max\left\{u^{c_1},\frac{a+b}{2}\right\}\right]\max\left\{u^{c_1},\frac{a+b}{2}\right\}-2k\\=&\frac{b-\max\left\{u^{c_1},\frac{a+b}{2}\right\}}{b-a}\times\frac{\max\left\{u^{c_1},\frac{a+b}{2}\right\}+b}{2}\\&+\frac{\max\left\{u^{c_1},\frac{a+b}{2}\right\}-a}{b-a}\times\max\left\{u^{c_1},\frac{a+b}{2}\right\}-2k\\=&\frac{b^2-2a\max\left\{u^{c_1},\frac{a+b}{2}\right\}+\left(\max\left\{u^{c_1},\frac{a+b}{2}\right\}\right)^2}{2\left(b-a\right)}-2k.\end{aligned}$$

Therefore, the student will conduct the first search step if $E\left[U\left(2\right)\right]>U\left(1\right)$, which solves

$$\max\left\{u^{c_1},\frac{a+b}{2}\right\}<b-\sqrt{2\left(b-a\right)k}.$$

We already know that $k<\frac{b-a}{8}$ and therefore $b-\sqrt{2\left(b-a\right)k}>\frac{a+b}{2}$. Hence, the student will conduct the second search step if $u^{c_1}<b-\sqrt{2\left(b-a\right)k}$.

We can generalize the above argument as follows. Suppose the student has searched a set of universities $C^S\subset B$ and discovered the highest utility $\bar{u}^S$. Suppose further that there is more than one unsearched university remaining in B. Then the student will conduct an additional search step if

$$\max\left\{\bar{u}^S,\frac{a+b}{2}\right\}<b-\sqrt{2\left(b-a\right)k}.$$

That is, in this case the student stops only if $\bar{u}^S<b-\sqrt{2\left(b-a\right)k}$.

Suppose now that the student has searched a set of universities $C^S\subset B$ and only one unsearched university $c_{|B|}$ remains in B. The expected utility after the last search step is given by

$$\begin{aligned}E\left[U\left(|B|\right)\right]=&E\left[\max\left\{u^{c_{|B|}},\bar{u}^S\right\}-|B|\,k\right]\\=&\Pr\left[u^{c_{|B|}}>\bar{u}^S\right]E\left[u^{c_{|B|}}\mid u^{c_{|B|}}>\bar{u}^S\right]+\Pr\left[u^{c_{|B|}}\leq\bar{u}^S\right]\bar{u}^S-|B|\,k\\=&\frac{b-\bar{u}^S}{b-a}\times\frac{\bar{u}^S+b}{2}+\frac{\bar{u}^S-a}{b-a}\times\bar{u}^S-|B|\,k\\=&\frac{b^2-2a\bar{u}^S+\left(\bar{u}^S\right)^2}{2\left(b-a\right)}-|B|\,k.\end{aligned}$$

Therefore, the student will conduct the last search step if $E[U(|B|)] > U(|B|-1)$. Because

$$
\begin{aligned}
& E[A(|B|)] - A(|B|-1) \\
=& \left(\frac{b^2 - 2a\bar{u}^S + \left(\bar{u}^S\right)^2}{2(b-a)} - |B|k \right) - \left(\max\left\{ \bar{u}^S, \frac{a+b}{2} \right\} - (|B|-1)k \right) \\
=& \frac{b^2 - 2a\bar{u}^S + \left(\bar{u}^S\right)^2}{2(b-a)} - \max\left\{ \bar{u}^S, \frac{a+b}{2} \right\} - k \\
=& \begin{cases} \frac{\left(b-\bar{u}^S\right)^2}{2(b-a)} - k & \text{if } \bar{u}^S > \frac{a+b}{2} \\ \frac{\left(a-\bar{u}^S\right)^2}{2(b-a)} - k & \text{if } \bar{u}^S \leq \frac{a+b}{2} \end{cases},
\end{aligned}
$$

we know that (i) if $a+\sqrt{2(b-a)k} < \bar{u}^S < b-\sqrt{2(b-a)k}$, the student investigates $c_{|B|}$ and selects the university with the highest utility among all universities in B; (ii) if $\bar{u}^S \geq b - \sqrt{2(b-a)k}$, the student stops before the last step and selects $\bar{c}^S$ (the university with the highest utility among the searched universities in C^S); and (iii) if $\bar{u}^S \leq a + \sqrt{2(b-a)k}$, the student selects $c_{|B|}$ without investigating it. □

A.6 Alternative search strategy: Information acquisition under DirSD

As explained in the main text, instead of generally characterizing the optimal search strategy under DirSD, we numerically calculate it for every student in our experimental setup and illustrate the calculation process below. In our experimental market with one tier, there are 12 students and six universities $\{c_1, c_2, c_3, c_4, c_5, c_6\}$; each university has two seats.

Under DirSD, student i chooses her search strategy based on k_i and the ex-ante probability distribution of her budget set $\left\{P_i\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$. This can be used to calculate $\{Q_i^\theta\}_{\theta=1,2,\ldots,m}$, in which Q_i^θ is the probability that i is accepted by the θth ranked university in her submitted rank-order list. The first simple case is that for a student in ranks 1 and 2, we know that $\{Q^1, Q^2, Q^3, Q^4, Q^5, Q^6\} = \{1, 0, 0, 0, 0, 0\}$. They know for sure that their budget set consists of all universities. Thus, only their first choice matters and they search just like a student in SeqSD with $B = \{c_1, c_2, c_3, c_4, c_5, c_6\}$. The second simple case is for the student in rank 12, $\{Q^1, Q^2, Q^3, Q^4, Q^5, Q^6\} = \left\{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right\}$. That is, her budget set only contains one school and she is equally likely to be accepted by any school on her list. Therefore, she does not have an incentive to acquire any information.

Now consider a student in ranks 3 to 11. If the student does not acquire any information, she ranks the six universities randomly and has an expected utility of

$$E[U(0)] = \frac{a+b}{2}.$$

With a cost of k, she can conduct the first step of her search by randomly choosing a university to investigate. Different from SeqSD, under DirSD the student does not observe her budget set and can therefore investigate any university in C. Suppose that she investigates c_1 and discovers u^{c_1}. If $u^{c_1} > \frac{a+b}{2}$, she would rank c_1 above all the unsearched universities. If $u^{c_1} < \frac{a+b}{2}$, she would rank c_1 below all the unsearched universities. If $u^{c_1} = \frac{a+b}{2}$, she woud randomly rank the six universities. Thus, the expected utility after the first search step is given by

$$\begin{aligned} E\left[U\left(1\right)\right] =& \Pr\left[u^{c_1} > \frac{a+b}{2}\right] E\left[Q^1 u^{c_1} + \sum_{\theta=2}^{6} Q^\theta \times \frac{a+b}{2} \mid u^{c_1} > \frac{a+b}{2}\right] \\ &+ \Pr\left[u^{c_1} \leq \frac{a+b}{2}\right] E\left[\sum_{\theta=1}^{5} Q^\theta \times \frac{a+b}{2} + Q^6 u^{c_1} \mid u^{c_1} \leq \frac{a+b}{2}\right] - k \\ =& \frac{1}{2}\left(Q^1 \frac{\frac{a+b}{2}+b}{2} + \sum_{\theta=2}^{6} Q^\theta \times \frac{a+b}{2}\right) + \frac{1}{2}\left(\sum_{\theta=1}^{5} Q^\theta \times \frac{a+b}{2} + Q^6 \frac{a+\frac{a+b}{2}}{2}\right) - k \end{aligned}$$

Given the values of k and Q^θ's, the student conducts the first search step if $E\left[U\left(1\right)\right] - U\left(0\right) > 0$.

Suppose that $E\left[U\left(1\right)\right] - U\left(0\right) > 0$ holds for this student and she indeed conducts the first step of her search.

- If the student investigates c_1 and discovers $u^{c_1} > \frac{a+b}{2}$. With an additional cost of k, she can conduct the second step of search. Suppose that she investigates c_2 and discovers u^{c_2}. If $u^{c_2} \in \left[a, \frac{a+b}{2}\right]$, she would rank c_1 the first and c_2 the last, with the unseached universities in between. If $u^{c_2} \in \left(\frac{a+b}{2}, u^{c_1}\right]$, she would rank c_1 the first and c_2 the second, followed by the unseached universities. If $u^{c_2} \in \left(u^{c_1}, b\right]$, she would rank c_2 the first and c_1 the second, followed by the unseached universities. Thus, the expected utility after the second search step is given by

$$\begin{aligned} E\left[U\left(2\right)\right] =& \frac{1}{2}\left(Q^1 u^{c_1} + \sum_{\theta=2}^{5} Q^\theta \times \frac{a+b}{2} + Q^6 \frac{a+\frac{a+b}{2}}{2}\right) \\ &+ \left(\frac{u^{c_1} - \frac{a+b}{2}}{b-a}\right)\left(Q^1 u^{c_1} + Q^2 \frac{\frac{a+b}{2} + u^{c_1}}{2} + \sum_{\theta=3}^{6} Q^\theta \times \frac{a+b}{2}\right) \\ &+ \left(\frac{b - u^{c_1}}{b-a}\right)\left(Q^1 \frac{u^{c_1}+b}{2} + Q^2 u^{c_1} + \sum_{\theta=3}^{6} Q^\theta \times \frac{a+b}{2}\right) - 2k \end{aligned}$$

- If the student investigates c_1 and discovers $u^{c_1} \leq \frac{a+b}{2}$. With an additional cost of k, she can conduct the second step of her search. Suppose that she investigates c_2 and discovers u^{c_2}. If $u^{c_2} \in \left[a, u^{c_1}\right]$, she would first rank the unsearched universities, and then rank c_1 the fifth and c_2 the last. If $u^{c_2} \in \left(u^{c_1}, \frac{a+b}{2}\right]$, she would first rank the unseached universities, and then rank c_2 the fifth and c_1 the last. If $u^{c_2} \in \left(\frac{a+b}{2}, b\right]$, she would rank c_2 the first and c_1 the last, with the unseached universities in between. Thus, the expected utility after the second step of the search is given by

$$E\left[U\left(2\right)\right]=\left(\frac{u^{c_1}-a}{b-a}\right)\left(\sum_{\theta=1}^{4}Q^{\theta}\times\frac{a+b}{2}+Q^{5}u^{c_1}+Q^{6}\frac{a+u^{c_1}}{2}\right)$$
$$+\left(\frac{\frac{a+b}{2}-u^{c_1}}{b-a}\right)\left(\sum_{\theta=1}^{4}Q^{\theta}\times\frac{a+b}{2}+Q^{5}\frac{u^{c_1}+\frac{a+b}{2}}{2}+Q^{6}u^{c_1}\right)$$
$$+\frac{1}{2}\left(Q^{1}\frac{\frac{a+b}{2}+b}{2}+\sum_{\theta=2}^{5}Q^{\theta}\times\frac{a+b}{2}+Q^{6}u^{c_1}\right)-2k$$

Given the values of k and Q^{θ}'s, the student conducts the second search step if $E\left[U\left(2\right)\right]-U\left(1\right)>0$.

Suppose that $E\left[U\left(2\right)\right]-U\left(1\right)>0$ holds for this student and she indeed investigates c_2. To determine whether the student would continue searching, we need to discuss different contigencies in terms of how u^{c_2} compares to u^{c_1} and $\frac{a+b}{2}$. When calculating $E\left[U\left(3\right)\right]$ in each of these contigencies, we need to consider different cases in terms of how the realization of u^{c_3} compares to u^{c_2} , u^{c_1}, and $\frac{a+b}{2}$.

As we can see from the above analysis, information acquisition under DirSD is heavily history-dependent. That is, whether a student should stop searching under DirSD may depend on the utilities of *all* the universities she has searched (instead of only the highest dicovered utility $\bar{u}_i^S$ under SeqSD). As a result, the number of contingencies grows rapidly as the student conducts each additional search step and it is much more challenging to derive the optimal search strategy under DirSD than under SeqSD.

B Additional experimental results

B.1 Details on individual search strategies

We consider each environment separately, since optimal search strategies depend on search costs and whether the preferences are tiered.[35] Figure 7 presents the average cost of information acquisition by treatments when the cost is low at \$0.5. The left panel of Figure 7 presents optimal and actual search strategies for the two-tier environment. In DirSD under low information costs, the optimal search strategy for all subjects (except for subjects ranked sixth and 12th) is to invest \$1 to obtain full certainty about their own preferences in the respective tier. Note that subjects with score ranks 1-5 should only consider the universities in tier A while subjects with score ranks 7-11 should only consider universities in tier B. On average, we observe that subjects search too little, except for rank 7 subjects who, on average, over-search by investing in information about universities in both tiers. The excessive search by rank 7 subjects may be driven by optimism

[35]As explained in Section 3.5, we do not derive point predictions for optimal search strategies in the Cutoff treatment.

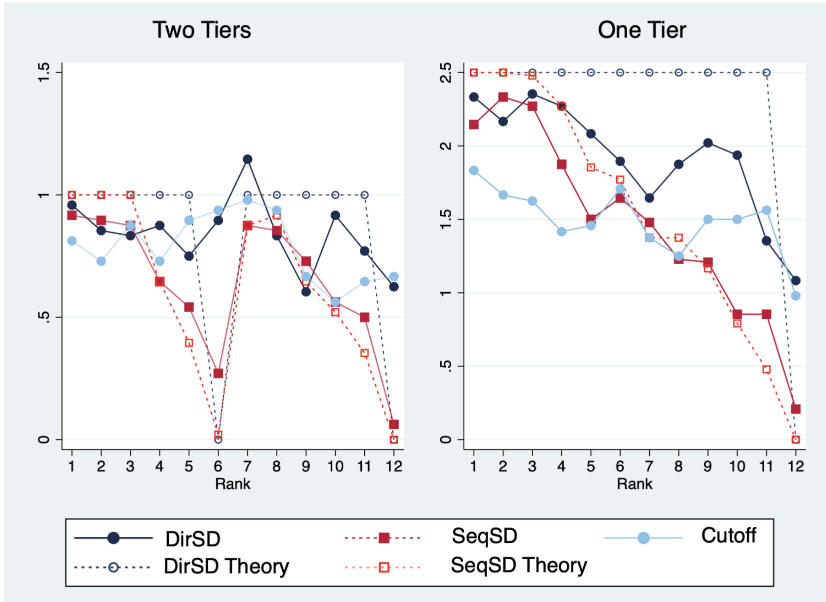

Figure 7: Average costs of information acquisition with low costs by treatments

that some of the subjects ranked 1 to 6 will be assigned to a tier B university, due to suboptimal preference submission.[36] In Cutoff, the behavior is similar to DirSD (p-value for the test of difference is 0.32). Thus, the cutoff provision does not have a significant effect on search strategies in the two-tier markets with low costs. On the one hand, the cutoffs are informative due to the full uncertainty resolution in the equilibrium of DirSD. On the other hand, the benefit of relying on cutoff information is relatively small, as the total cost of optimal information acquisition is just $1. Thus, subjects might not risk saving $1 by relying on cutoff information. As for subjects with ranks 6 and 12, they should not invest in information at all, as they both get the only free seat of the corresponding tier in equilibrium. However, we observe a high degree of over-search by these subjects.[37] As for SeqSD, the actual search behavior of subjects is, on average, remarkably in line with the theoretical predictions. The actual search costs are significantly lower than in DirSD and Cutoff (the p-value for the test of difference is $<$0.01 for both comparisons). Thus, the optimal search strategy in SeqSD is more straightforward for subjects than in DirSD. This is not surprising, as the optimal strategy consists of full investment in resolving uncertainty about one's available universities, and the only deviation could be under-search or searching prior to the start of the allocation procedure,–that is, before one learns which universities are available to them.

The right panel of Figure 7 presents the predictions and actual search strategies for the one-tier environment with low costs. In DirSD, the optimal search strategy for all subjects (except rank 12 subjects) is to invest $2.50 to obtain full certainty about their preferences. We observe that rank 1

[36]In total, a rank 7 participant had the potential choice between universities in the top tier due to suboptimal strategies of higher-ranked participants in only 1 out of 48 rounds of DirSD.

[37]Note that in our experimental setup all students had to submit the full rank-order list of universities in DirSD and Cutoff, or had to choose one university in SeqSD, thus making it impossible to remain unassigned. Therefore, not searching is an optimal strategy for rank 12 students.

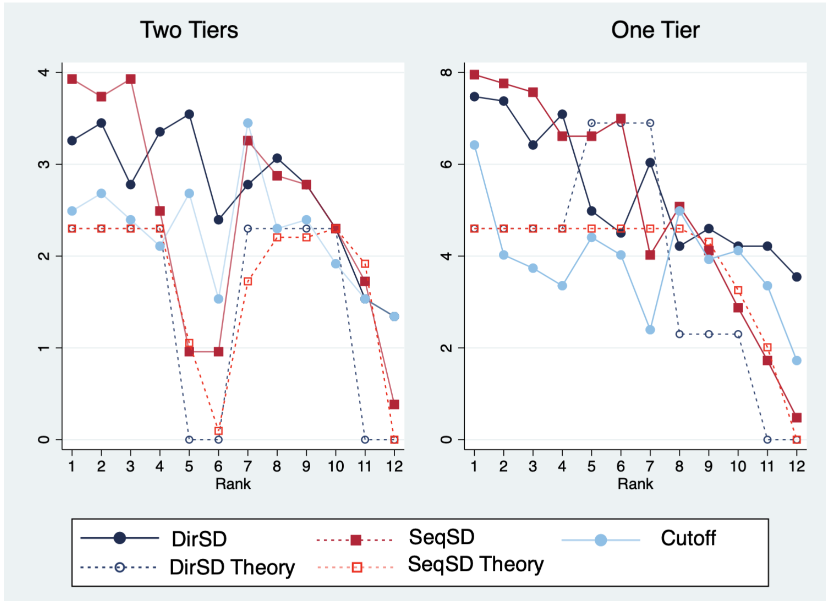

Figure 8: Average costs of information acquisition with high costs by treatments

to 11 subjects search too little, which is even more pronounced for subjects with ranks 6 to 11 than for subjects with ranks 1 to 5. Note that the relative benefit of search decreases with the rank, thus it can be partially driven by the risk aversion of subjects. Another possibility is that subjects perceive the preferences as correlated, and thus overestimate the chances that the most-preferred universities will be assigned to the higher-ranked subjects. In Cutoff, the actual search costs are significantly lower than in DirSD for ranks 1 to 10 (p-value for the test of difference is <0.01 for these ranks, and for all ranks). Again, cutoffs are informative due to full uncertainty resolution in the equilibrium of DirSD. In the two-tier environment, the potential benefit of relying on cutoffs for subjects is only \$1. In the one-tier environment, the optimal information cost is \$2.50 and the potential benefit of cutoffs from the perspective of saving search costs is higher. As for rank 12 subjects, they should not invest in information at all, yet, on average, they invest \$1.08 in DirSD and \$0.98 in Cutoff. This violates the optimal strategy of not searching. As for SeqSD, the actual search behavior of subjects is remarkably in line with the average theoretical predictions. The most substantial deviation is the under-search of the subjects ranked 1 to 5. Again, the optimal strategy in SeqSD consists of obtaining full certainty about the ranking of all available universities, and the only deviation could be under-search or search before the allocation procedure started. When the optimal strategy requires an investment of \$2.50, and thus five search steps, subjects often stop after four steps of the search, thus underestimating the probability of the last university being preferred to the other five universities. This under-search in SeqSD is similar to the under-search of rank 1 to 3 subjects in DirSD. Overall, the actual search costs in SeqSD are significantly lower than in DirSD, but not significantly different from Cutoff (p-value for the test of difference is <0.01 and equal to 0.79 respectively).

Figure 8 presents the average cost of information acquisition by treatments when the cost is high at \$2.30. The left panel presents predictions and actual search strategies for the two-tier environment. First, in DirSD the optimal search strategy for rank 1 to 4 and 7 to 10 subjects is to invest \$2.30 in resolving uncertainty about the relative ranking of any two universities in the respective tier. Thus, in the high-cost treatments subjects never obtain full certainty about the university rankings. Students ranked 5, 6, 11, and 12 should not invest in search at all. Unlike in treatments with low costs, we observe significant over-search in DirSD for all ranks. This finding is in line with previous experimental findings on information acquisition (see Chen and He, 2021, for school choice, Bhattacharya et al., 2017, for voting, and Gretschko and Rajko, 2015, for auctions). In Cutoff, the actual search costs are lower than in DirSD (p-value$<$0.01) with the highest difference for the lower-ranked students. Unlike the two-tier low-cost environment when the potential benefit of relying on cutoffs saves subjects only up to \$1, in the two-tier high-cost environment the optimal information cost is \$2.30. Thus, the potential benefit of cutoffs for saving information costs is much higher. Subjects rely on the cutoffs following the higher potential saving of information costs. In SeqSD with high costs, unlike in SeqSD with low costs where the actual search behavior of subjects is mostly in line with theoretical predictions, we observe a high degree of over-search for students ranked 1 to 3 and 7 to 9. The over-search for ranks 1 to 3 is even higher than in DirSD. As for ranks 5, 6, 11, and 12, the behavior is more in line with the theory than in the other treatments. Overall, in the two-tier high-cost environment, there is no significant difference in the average actual search costs between SeqSD and DirSD (p=0.12), and between SeqSD and Cutoff (p=0.16).

Finally, the right panel of Figure 8 presents predicted and actual search strategies for the one-tier high-cost environment. In DirSD, the optimal search strategy for rank 1 to 4 subjects is to invest \$4.60 in resolving the uncertainty about the ranking of any three universities. Similar to the two-tier environment with high costs, students ranked 1 to 4 over-search relative to the optimal strategy. Ranks 5 to 7 have an optimal strategy of investing \$6.90 to resolve uncertainty about the ranking of four out of six universities. Note that this is the only case where the lower-ranked subjects search more in theory than the higher-ranked subjects. This pattern, however, finds no support in the data, as the students ranked 5 to 7 search less than students ranked 1 to 4. As for ranks 8 to 12, on average, they all invest around \$4 in information acquisition, despite an optimum of \$2.30 for ranks 9 and 10 and an optimum of \$0 for ranks 11 and 12. In Cutoff, the actual search costs are lower than in DirSD ($p<0.01$), with larger differences for the higher-ranked students. Just as in the two-tier environment with high costs, subjects rely on the cutoffs leading to lower information costs than in DirSD. Yet again, they ignore the fact that in the high-cost environments, the cutoffs are less informative about the preferences of the previous cohort than in the low-cost environments, as many submissions of the previous cohort are made without resolving preference uncertainty. Note, however, that in both high-cost environments, in DirSD subjects over-invest in information relative to the optimal strategy. Thus, the cutoffs are more informative than in equilibrium. As for SeqSD, we observe a high degree of over-search for students ranked 1 to 6. As for ranks 7 to 12, the behavior is more in line with the theory than in the other treatments. Overall, in the one-tier high-cost environment, the average actual search costs in SeqSD are significantly higher than in Cutoff (p<0.01), and not significantly different from DirSD (p=0.52).

B.2 Normalized efficiency by environments

e

	Treatment			p-value for test of equality		
	DirSD (1)	SeqSD (2)	Cutoff (3)	DirSD=SeqSD (4)	DirSD=Cutoff (5)	SeqSD=Cutoff (6)
Two tiers & low cost	74.0%	82.7%	75.6%	0.01	0.63	0.01
Two tiers & high cost	81.5%	88.1%	87.0%	0.00	0.01	0.51
One tier & low cost	83.7%	91.7%	82.3%	0.01	0.74	0.00
One tier & high cost	64.7%	76.2%	72.7%	0.00	0.00	0.08
All	75.9%	84.7%	79.4%	0.00	0.08	0.00

Notes: For the tests in columns 4-6, we use the p-values for the coefficient of the treatment dummy in the OLS regression of efficiency on this dummy with standard errors clustered at the level of matching groups and with a sample restricted to the treatments that are of interest for the test.

Table 6: Average normalized efficiency by treatments and environments

B.3 Order effects

	Total payoff	Total payoff	Number of searches	Number of searches	Optimal strategy	Optimal strategy
Order	-0.00	-0.00	0.06	0.06	0.00	0.00
	(0.41)	(0.29)	(0.10)	(0.07)	(0.06)	(0.02)
SeqSD		2.05		-0.38		0.26
		(0.36)		(0.07)		(0.02)
Cutoff		0.90		-0.47		-0.14
		(0.43)		(0.10)		(0.02)
Tiers		-3.17		-1.19		0.09
		(0.49)		(0.12)		(0.03)
Cost of search		-2.34		-0.44		-0.03
		(0.12)		(0.03)		(0.01)
Period		0.07		0.04		0.01
		(0.10)		(0.03)		(0.01)
Observations	3384	3384	3384	3384	3384	3384
R	0.00	0.09	0.00	0.22	0.00	0.16

Note: Results of OLS regressions with clustering of standard errors on the level of matching groups. Order is a dummy variable equal to 0 when Low cost preceded High cost, and equal to 1 when High cost preceded Low cost. SeqSD is a dummy for treatment SeqSD, Cutoff is a dummy for treatment Cutoff. Tier is equal to 1 in One-tier environments and equal to 2 in Two-tier environments.

Table 7: Order effects

B.4 Alternative search technology: Information acquisition under Cutoff

	Dummy for search(2)	Dummy for search(2)
	(1)	(2)
Cost of search	-0.121	-0.29
	(0.05)	(0.05)
Higher cutoff, dummy	-0.07	
	(0.02)	
Higher cutoff, difference		-0.006
		(0.000)
Lower cutoff, difference		-0.003
		(0.000)
Observations	2304	2304

Note: Marginal effects of probit regressions regarding information acquisition about a university in Cutoff. "Higher cutoff, dummy" is a dummy that is equal to one if the cutoff score of the university minus the student's score is greater than zero. "Higher cutoff, difference" is equal to the cutoff score of the university minus the student's score if the difference is positive and zero otherwise. "Lower cutoff, difference" is equal to the student's score minus the cutoff score of the university if the difference is positive and zero otherwise. Standard errors are clustered at the level of matching groups.

Table 8: Probability of information acquisition about a university depending on the cutoff