Information Acquisition in Matching Markets: The Role of Price Discovery*

Nicole S. Immorlica Jacob D. Leshno Irene Y. Lo
Brendan J. Lucier†

October 5, 2020

We explore the acquisition and flow of information in matching markets through a model of college admissions with endogenous costly information acquisition. We extend the notion of stability to this partial information setting, and introduce regret-free stability as a refinement that additionally requires optimal student information acquisition. We show regret-free stable outcomes exist, and finding them is equivalent to finding appropriately-defined market-clearing cutoffs.

To understand information flows, we recast matching mechanisms as price-discovery processes. No mechanism guarantees a regret-free stable outcome, because information deadlocks imply some students must acquire information suboptimally. Our analysis suggests approaches for facilitating efficient price discovery, leveraging historical information or market sub-samples to estimate cutoffs. We show that mechanisms that use such methods to advise applicants on their admission chances yield approximately regret-free stable outcomes. A survey of university admission systems highlights the practical importance of providing applicants with information about their admission chances.

*We thank Rediet Abebe, Eduardo Azevedo, Eric Budish, Caterina Calsamiglia, Ozan Candogan, Lorenzo Casaburi, Yeon-Koo Che, Li Chen, Laura Doval, Piotr Dworczak, Rustam Hakimov, Yinghua He, Emir Kamenica, Yash Kanoria, Youngwoo Koh, Fuhito Kojima, David Manlove, Daniel Münich, Muriel Niederle, Chris Neilson, Canice Prendergast, Al Roth, Sven Seuken, Ran Shorrer, Jakub Steiner, Andrea Stella, Vasilis Syrgkanis, Utku Unver, and Leeat Yariv, for helpful suggestions, feedback and assistance in surveying university admission systems. We are also grateful to seminar participants in the AEA, Chicago Booth micro workshop, Columbia, Conference on Economic Design, the Econometric Society World Congress, ICERM, INFORMS, NBER market design, MATCH-UP 2019, NY Computational Economics, Penn State, Princeton CS theory, Stony Brook game theory festival, Simons Institute, Stanford Market Design, and UPenn. Ethan Che, Anand Shah and Raymond Lee provided invaluable research support.

†Immorlica and Lucier: Microsoft Research. Leshno: University of Chicago Booth School of Business. Lo: Stanford University. This work is supported by the Robert H. Topel Faculty Research Fund at the University of Chicago Booth School of Business.

1 Introduction

In matching settings such as school choice and college admissions it is common for applicants to spend a significant amount of effort investigating potential placements before forming their preferences. This has motivated several informational interventions (Corcoran et al. 2018, Dynarski et al. 2020, Grenet et al. 2019), and raises the question of how the design of the marketplace facilitates the acquisition and flow of information in the market. However, common matching models do not incorporate these features.

We present a model of many-to-one matching that incorporates costly information acquisition. College priorities are common knowledge. Students have independent private values, but only know the distribution of these values. They can refine these distributions by acquiring signals at a cost. One example is the Pandora's Box model (Weitzman 1979), where each student has a prior over her values and can pay a college-specific cost to learn her value at a selected college. Since the set of colleges available to a student depends on the decisions of other students, the information acquisition decisions of different students are interlinked.

By incorporating costly information acquisition, our model captures important features of matching markets. A market outcome consists of an assignment as well as the information acquired. Student utilities depend on their values for their assigned partners as well as their information acquisition costs. Students can rationally choose to remain partially informed, or to delay decision-making even when they have multiple offers of admission.

We focus on stable outcomes where, ex-post, no agent can benefit by changing the outcome. Under full information, stability requires that after the outcome is revealed no agents can form a blocking pair. That is, there is no pair of agents who prefers each other to their assigned partners. In our setting, we define an outcome to be stable if, after observing the entire match and forming preferences based on the information they collected, no student can form a blocking pair with a college or would wish to collect more information.

Our main question is whether the market can reach a stable outcome while facilitating efficient information acquisition by all students. Not every stable outcome is informationally efficient: for example, if students collect all information and are matched as in a full-information stable matching, the outcome is stable, but students incur unnecessary information acquisition costs. To motivate our benchmark for efficient information acquisition, consider a student who delays acquiring information until after she sees the market outcome. Such a 'last to market' student can use all market information to guide her information acquisition. Because students have independent private values, the only market information that is relevant for this student's information acquisition is her budget set: the set of colleges at which she has sufficiently high priority to be admitted. Motivated by this, we define an outcome to be *regret-free stable* if every student has acquired information as if she knew her budget set in advance. We ask whether an appropriately designed market can facilitate efficient information acquisition in the sense that it achieves regret-free stable outcomes.

Regret-free stable outcomes have a natural formulation in terms of demand and cutoffs. Given an outcome, the corresponding cutoff of each college is equal to the lowest priority of a student assigned to that college. Cutoffs specify a budget set for each student, equal to the set of colleges where her priority for the college is above the college's cutoff. A student's demand, given cutoffs, is defined as the outcome of the process where the student first optimally acquires information given the budget set specified by the cutoffs, and then selects her most preferred college in her budget set under the resulting preferences. Our formulation abstracts away from the details of each student's information acquisition process by encoding it in her demand.

We show that cutoffs that clear the market under this demand are equivalent to regret-free stable outcomes. Cutoffs function like prices, summarizing all the market information a student needs in order to decide what information to acquire. The market-clearing condition ensures stability since the resulting aggregate demand is consistent with the cutoffs. Together, this implies that regret-free stable outcomes are fully determined by aggregate demand. The formulation of regret-free stable outcomes in terms of demand and cutoffs allows us to directly apply results from the theory of stable matching under full information, which significantly simplifies our analysis. We focus on economies where student demand satisfies the weak axiom of revealed preferences (WARP), such as the Pandora's Box model. Aggregate demand in such an economy is identical to the aggregate demand of a related full-information economy. This implies that a regret-free stable outcome always exists for such economies, and that the set of regret-free stable outcomes forms a non-empty lattice. There exists a student-optimal regret-free stable outcome which gives all students the highest ex-ante expected utility. Additionally, this perspective allows us to show that, generally, there is a unique regret-free stable outcome.

To answer which market mechanisms implement regret-free stable outcomes, we recast market mechanisms as price-discovery tools that both learn and communicate cutoffs. We introduce communication processes which formalize the price-discovery process under a market mechanism by specifying its initial information, its information flows and information acquisition processes, and the resulting outcome. The information provided by the mechanism to students guides their information acquisition, and the market will reach a regret-free stable outcome if the mechanism is able to guide students' information acquisition to learn market-clearing cutoffs without incurring unnecessary costs. For example, for certain economies, iterative implementations of college-proposing deferred acceptance can implement regret-free stable outcomes by sequentially approaching students whose budget sets are fully known to the mechanism and communicating budget sets to these students. The mechanism collects preference information from these students, allowing it to identify more students to approach. In contrast, the standard one-shot implementation of student-proposing deferred acceptance asks students to acquire information and report their preferences without guidance from the mechanism, which is likely to result in outcomes that are not regret-free stable.

Despite the existence of regret-free stable outcomes, our main theorem shows that essentially no mechanism is regret-free stable for general economies. While the mechanism knows that some cutoffs will implement a regret-free stable outcome, the difficulty lies in finding such cutoffs without making students incur additional costs. In other words, price discovery is costly. We show this by demonstrating that general economies can exhibit *information deadlocks* that prevent mechanisms from reaching regret-free stable outcomes. Information deadlocks arise when there is a cycle of students in which each student's budget set and information acquisition decisions depend on the demand of the others. In order to learn market-clearing cutoffs, some student in the cycle must acquire information before she has sufficient market information to do so optimally. We emphasize that this impossibility result holds despite the guaranteed existence of a regret-free stable outcome. In addition, it is not driven by student or college incentives, and continues to hold even if students are assumed to follow the instructions of the mechanisms without strategizing. Rather, the challenge stems from the fact that students need information to know which information they should gather.

Despite this impossibility result, regret-free stability is a desirable and useful benchmark. The cutoff structure of regret-free stable outcomes suggests a natural two-stage approach for obtaining approximately regret-free stable outcomes. In a first stage, the mechanism learns market-clearing cutoffs. In a second stage, it publishes the cutoffs, lets students acquire information using the implied budget sets, and then computes the assignment. In other words, the mechanism first engages in price discovery and then publishes prices that guide students' information acquisition and clear the market.

We provide a few examples of this approach that achieve approximately regret-free stable outcomes. One option is to leverage external information such as historical cutoffs. In many settings, cutoffs are stable over time, and price discovery can be performed using information across years rather than within a given year. Another option is to calculate cutoffs using a demand estimation approach. There may be students with "free information" whose initial information is sufficient for them to acquire information optimally. Such students will, in general, comprise a non-representative sub-sample, but under structural assumptions their demand can be used to estimate aggregate demand. While such approaches may only give a noisy estimate of the cutoffs, we show that committing to these cutoffs will give an approximately regret-free stable outcome in the sense that it is regret-free stable for perturbed capacities. Another alternative is to estimate demand from a random sub-sample of students. The sampled students will bear additional information acquisition costs for the benefit of others, resulting in an approximately regret-free stable outcome in the sense that almost all students have acquired information optimally.

To supplement our theoretical analysis, we surveyed college admission systems across the world. Our theoretical results highlight the importance of providing applicants with information about their admission chances. In our survey, we find that many admission systems make an effort to provide applicants with an estimate of their admission chances before they apply, and this information is more readily accessible in many cases than information about the matching algorithm itself. In particular, many systems provide score calculators and/or post historical admission cutoffs. Two prominent examples are Australia and Israel: both post admission cutoffs for students, estimated using historical data. Some universities in Australia commit to the posted cutoffs and absorb extra demand by perturbing capacities. Israel does not have a centralized admission clearinghouse, and relies on posted cutoffs and score calculators to coordinate admissions. Overall, our survey and theoretical results suggest market designers should pay careful attention to information flows in marketplaces, both in determining market-clearing cutoffs and in communicating that information to students.

Related Work

A large body of empirical work emphasizes the importance of information and its availability in determining educational choices. Hoxby & Avery (2012) find that the majority of low-income, high-achieving students do not apply to selective colleges, even though these colleges are likely to offer them higher quality at a lower cost, and argue this is driven by the lack of proper information. Evidence from multiple field experiments in many settings shows that access to information on educational options has significant impacts on student choices and educational outcomes (Hastings & Weinstein 2008, Hoxby & Turner 2015, Andrabi et al.

2017, Dynarski et al. 2020, Corcoran et al. 2018, Neilson et al. 2019).[1]

A growing literature highlights the implication of information for matching mechanisms. Kapor et al. (2020), Dur et al. (2015), Luflade (2017) and Narita (2018) provide empirical evidence showing that lack of information harms students in standard matching mechanism. Grenet et al. (2019) analyze how students respond to admission offers in a dynamic university admission process and argue it can be explained by students undergoing costly information acquisition. Chen & He (2017) and Chen & He (2020) study student incentives to acquire information about their own cardinal preferences and other's preferences under the DA and Boston mechanisms.

A rich strand of literature explores notions of stability with partially informed agents (Chakraborty et al. 2010, Liu et al. 2014, Liu 2020, Chen & Hu 2020, Bikhchandani 2017, Kloosterman & Troyan 2020, Ehlers & Massó 2015). A major challenge in such models is that stable matchings may not necessarily exist. We circumvent this challenge by assuming agents have independent private values, and in our model stable matchings exist under a natural condition.

Our paper contributes to the broader literature on information acquisition and information provision in mechanism design. Within the matching literature, Coles et al. (2010), Coles et al. (2013), and Lee & Niederle (2015) theoretically and experimentally evaluate signaling in the job market for new economists and dating markets. Bade (2015) shows that serial dictatorship is the unique mechanism that is Pareto-optimal, strategy-proof and non-bossy under endogenous information acquisition. Harless & Manjunath (2015) consider an allocation problem with common values and information acquisition. A number of papers, including Aziz et al. (2016) and Rastegari et al. (2013, 2014), develop algorithms to efficiently find a stable matching through interviews that reveal ordinal preference information. A related paper from the auctions literature is Kleinberg et al. (2016) which shows that descending price auctions create optimal incentives for value discovery.

A growing body of work explores informational efficiency in matching. Segal (2007) studies the communication complexity of social choice rules. Gonczarowski et al. (2015) consider the communication complexity of finding a stable matching and show that it requires $\Omega(n^2)$ boolean queries. Ashlagi et al. (2017) find that the communication complexity of finding a stable matching can be low under assumptions on the structure of the economy if the mechanism can use a Bayesian prior. In contrast to our model, these papers assume

[1] Abdulkadiroğlu et al. (2020) find that parents' preferences for schools are mainly driven by peer quality, and suggest this finding can be explained by parents' difficulty in collecting other measures of school effectiveness.

that agents know their preferences (for example, can report their first choice) and focus on the cost of communicating that information.

Organization of the paper

Section 2 provides our model of matching with costly information acquisition and defines regret-free stable outcomes. Section 3 provides theoretical results showing that finding regret-free stable outcomes is equivalent to finding market cutoffs, implying the existence and lattice structure of regret-free stable outcomes. Given the existence result, the question arises of whether a clearinghouse can efficiently discover market clearing cutoffs. Section 4 defines communication processes to precisely capture information flows in a clearinghouse, and shows that, except for in specific settings, a clearinghouse cannot guarantee a regret-free stable outcome. Given this impossibility, Section 5 provides a price-discovery approach to clearing-house design that can achieve approximately regret-free stable outcomes. Section 6 presents results from our survey of admission systems, discusses the similarity between the mechanisms suggested in Section 5 and the mechanisms used in practice, and presents concluding remarks. Omitted proofs and examples are in the appendixes. The survey of admission systems can be found in the online appendix.

2 A Model for Matching with Costly Information Acquisition

We present a model where colleges priorities are known and students learn their preferences through costly information acquisition. The set of colleges is denoted by $\mathcal{C} = \{1, \ldots, n\}$, and each college $i \in \mathcal{C}$ has capacity to admit $q_i > 0$ students. We use ϕ to denote being unmatched and write $\mathcal{C}_\phi = \mathcal{C} \cup \{\phi\}$.

Under full information, a student is given by $(r, v) \in \mathcal{R} \times \mathcal{V} = [0, 1]^{\mathcal{C}} \times \mathbb{R}^{\mathcal{C}}$, where r_i is the student's priority or rank at college $i \in \mathcal{C}$, and v_i is the student's value for attending college $i \in \mathcal{C}$. College i prefers student (r, v) over student (r', v') if and only if $r_i > r'_i$.

We set up a model where students are partially informed about their independent private values for colleges, and can adaptively acquire costly signals to refine their information. At any given moment, each student is associated with a tuple $\omega \in \Omega$ that encodes the student's beliefs, the information she can acquire or has acquired, and her realized values. We write

$$\omega = \left(r^{\omega}, F^{\omega}, \Pi^{\omega}, c^{\omega}, \chi^{\omega}, \{\pi\left(v^{\omega}\right)\}_{\pi\in\chi^{\omega}}; v^{\omega}\right) \in \Omega$$

where r^{ω} is the student's priorities, F^{ω} is her prior over her possible private values $\mathcal{V}$, Π^{ω} is a finite set of possible signals that can be acquired, $c^{\omega}\colon 2^{\Pi^{\omega}} \to \mathbb{R}_{\geq 0}$ is the cost to student ω of acquiring signals, and $v^{\omega} \in \mathcal{V}$ is the realization of the student's values. Each signal $\pi \in \Pi^{\omega}$ is a partition of $\mathcal{V}$ into F^{ω}-measurable sets,[2] and we denote its realization by $\pi\left(v^{\omega}\right) \subset \mathcal{V}$. $\chi^{\omega} \subset \Pi^{\omega}$ denotes the set of signals the student has acquired so far, and $\{\pi\left(v^{\omega}\right)\}_{\pi\in\chi^{\omega}}$ are the signal realizations observed by the student. We write Ω for the set of all such tuples ω.

Each tuple ω consists of three parts: the information initially available to the student, the student's realized values (which are initially unobserved), and the information acquired by the student through signals. It will be helpful to introduce notation for different subsets of this tuple. We say that student ω has a *state* $\theta = \theta(\omega)$ given by

$$\theta\left(\omega\right) = \left(r^{\omega}, F^{\omega}, \Pi^{\omega}, c^{\omega}, \chi^{\omega}, \{\pi\left(v^{\omega}\right)\}_{\pi\in\chi^{\omega}}\right) \in \Theta.$$

We refer to θ as a student's *information state* because θ encodes all information available to a student at a given moment. That is, the student knows $r^{\omega}, F^{\omega}, \Pi^{\omega}, c^{\omega}$ and observes the realizations of acquired signals $\{\pi\left(v^{\omega}\right)\}_{\pi\in\chi^{\omega}}$ but does not know her realized values v^{ω} beyond those signals. Denote the set of all information states by Θ. We will sometimes abuse notation and associate θ with the set of all ω such that $\theta(\omega) = \theta$. Such $\omega \in \theta$ can differ only on their realized values v^{ω}, so we will often write r^{θ} to represent the unique r^{ω} for all $\omega \in \theta$ and similarly for F^{θ}, Π^{θ}, etc.

The *type* of a student consists of all information initially observable to the student. We assume that initially students have not acquired any signals. Thus, the set of types is the set of initial information states $\Theta_0 \subset \Theta$, where $\theta_0 \in \Theta_0$ precisely if $\chi^{\theta_0} = \emptyset$. The *initial value realization* ω_0 of a student consists of all the information included in their type θ_0 as well as their realized values v^{ω}. We write $\Omega_0 \subset \Omega$ to denote the set of possible realizations.

Given a state $\theta \in \Theta$, let $F^{|\theta}$ denote the posterior distribution over $\mathcal{V}$ given the information available to θ. Overloading notation, we will consider $F^{|\theta}$ also as a posterior distribution over $\omega \in \theta$. Let $\hat{v}^{\theta} = \mathbb{E}_{v\sim F^{|\theta}}\left[v\right]$ denote the corresponding expected values. We will let $\hat{v}^{\omega} = \hat{v}^{\theta(\omega)}$ denote the perceived expected value of a student ω.

[2]Implicitly, we are restricting attention to signals that are deterministic given $v \in \mathcal{V}$; this is for clarity of notation.

Definition 1. *A **continuum economy with information acquisition** is specified by $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$, where $q = \{q_i\}_{i\in\mathcal{C}}$ is the vector of quotas at each college, and η is a measure over the set Ω_0 of initial value realizations.*

Note that η specifies the joint distribution over student types and value realizations; note also that students initially do not know these realized preferences.

We make the following assumptions. The distribution of value realizations is consistent with the student priors. That is, there exists a measure ν over Θ_0 such that for any sets $A \subset \Theta_0$ and $V \subset \mathcal{V}$, we have that

$$\eta\left(\{\omega = (\theta, v) \mid \theta \in A, v \in V\}\right) = \int_{\theta\in A} F^{\theta}\left(V\right) d\nu\left(\theta\right).$$

All students and colleges are acceptable. The rank r_i^{ω} is normalized to be the student's percentile, college priorities are strict, and there is an excess of students.[3] For ease of exposition, we also assume that student preferences are strict by imposing that for any $\theta \in \Theta$ we have that $F^{\theta}\left(\{v \mid v_i = v_j\}\right) = 0$ for all $i \neq j \in \mathcal{C}$ and by assuming that any remaining indifferences are resolved in favor of the college with lower index.[4] The posterior $F^{|\theta}$ and expected values $\hat{v}^{\theta}$ are well-defined for any θ.

An outcome specifies both an assignment of students to colleges as well as the information acquired by each student.

Definition 2. *An **outcome** (μ, χ) consists of an assignment μ and acquired information χ. An assignment μ is an η-measurable mapping $\mu : \Omega_0 \to \mathcal{C}_{\phi}$ specifying the assignment of $\omega_0 \in \Omega_0$. Acquired information χ is an η-measurable mapping specifying the information $\chi\left(\omega_0\right) \subseteq \Pi^{\omega_0}$ acquired by $\omega_0 \in \Omega_0$.*[5]

Overloading notation, for college $i \in \mathcal{C}$ let $\mu\left(i\right)$ denote the set $\mu^{-1}\left(i\right) \subseteq \Omega_0$ of initial value realizations of students assigned to college i. We denote the set of students with positive measure under an outcome (μ, χ) by $\Omega_{\chi} = \{\omega \in \Omega \mid \chi\left(\omega\right) = \chi^{\omega}\}$, and we abuse notation and write $\chi\left(\omega\right)$ and $\mu\left(\omega\right)$ for $\omega \in \Omega$ to mean $\chi\left(\omega_0\right)$ and $\mu\left(\omega_0\right)$, respectively (where ω_0 is the initial realization associated with student ω). Given (μ, χ), the utility of a student ω is $v^{\omega}_{\mu(\omega)} - c^{\omega}\left(\chi^{\omega}\right)$.

[3] That is, for any $i \in \mathcal{C}$ and $x \in [0,1]$, we have that $\eta\left(\{\omega \in \Omega_0 | \, r_i^{\omega} \leq x\}\right) = x$, as well as $\eta\left(\{\omega \in \Omega_0 | \, r_i^{\omega} = x\}\right) = 0$, and we also have that $\sum_{i\in\mathcal{C}} q_i < \eta\left(\Omega\right) = 1$.

[4] To simplify notation for resolving such indifferences, if z_i and z_j are quantities related to colleges i and j respectively (e.g. v_i^{ω} and v_j^{ω} for some student ω), we abuse notation and let $z_i > z_j$ denote that either $z_i > z_j$ or $z_i = z_j$ and $i < j$.

[5] Implicitly, we assume that information acquisition and assignment are deterministic given a student's initial value realization.

We now consider the feasibility of an outcome. One natural condition is that the assignment at each college does not exceed the quotas, $\eta(\mu(i)) \leq q_i$. In addition, acquiring sufficient information may be necessary for assignment to a college. Let $\Psi(\theta) \subset \mathcal{C}_\phi$ denote the subset of colleges a student with state θ can be assigned to given inspections χ^θ, and let $\Psi(\omega) = \Psi(\theta(\omega))$. An outcome (μ, χ) is *feasible* if for each college $i \in \mathcal{C}$, we have that $\mu(i)$ is η-measurable, $\eta(\mu(i)) \leq q_i$, and for all ω we have $\mu(\omega) \in \Psi(\omega)$.

2.1 Stability with Information Acquisition

We extend the standard definition of stable matchings to economies with information acquisition. Intuitively, an outcome is stable if every student who observes the outcome (μ, χ) does not form a blocking pair with some college, and does not want to acquire more information. That is, the outcome is both allocatively stable and informationally stable.

There are multiple equivalent formulations of allocative stability. We will express stability conditions in terms of demand and budget sets. This formulation will be convenient as it allows us to encode the information acquisition process within our notion of demand. Given an assignment μ, student ω has sufficient priority to be admitted to the set of colleges $B^\omega(\mu)$, given by

$$B^\omega(\mu) = \left\{ i \in \mathcal{C}_\phi \ \middle|\ r_i^\omega \geq \inf\{r_i^{\omega'} \mid \omega' \in \mu(i)\} \text{ or } \eta(\mu(i)) < q_i \right\}.$$

We refer to $B^\omega(\mu)$ as the student's *budget set*. Note that $B^\omega(\mu)$ depends only on r^ω, and we can write $B^\omega(\mu) = B^{\theta(\omega)}(\mu) = B^{\omega_0}(\mu)$ where ω_0 is the initial value realization associated with student ω. Any college $i \in B^\omega(\mu)$ is willing to block with ω. Thus, student ω would like to form a blocking pair if there is a college $i \in B^\omega(\mu) \cap \Psi(\omega)$ (which is a feasible match, given inspections) such that $\hat{v}_i^\omega > \hat{v}_{\mu(\theta)}^\omega$. This mimics the stability constraint in the full information model.

Additional stability concerns arise from the possibility of further information acquisition. Having observed an outcome (μ, χ), a student ω may want to acquire additional information. In particular, after learning her budget set $B^\omega(\mu)$, the student may wish to acquire additional information to better inform her selection from that budget set. Because students have independent private values, learning the outcome provides no further information beyond $B^\omega(\mu)$. Therefore, we can capture the choice of subsequent information acquisition for a student who knows the outcome (μ, χ) by an inspection rule that specifies the information acquired by a student ω given her current information state $\theta(\omega)$ as well as the knowledge that $B^\omega(\mu) = B \subset \mathcal{C}$.

To formally define an optimal inspection rule χ^*, consider a sequential inspection rule $\varphi : \Theta \times 2^{\mathcal{C}} \to \Pi \cup \{\phi\}$ that specifies for each possible state θ and budget set $B \subset \mathcal{C}$ whether the student should acquire another signal $\pi \in \Pi^\theta \setminus \chi^\theta$ or stop (denoted by ϕ). We will sometimes refer to the acquisition of a signal as an *inspection.* Each inspection in the sequence of inspections can depend on value realizations, but only through the information revealed by prior inspections.

Given a sequential inspection rule φ, consider a student ω whose budget set is $B \subset \mathcal{C}$. Let $\theta^\varphi(\omega)$ denote the state reached by sequential applications of φ until it stops (which it must, because Π^θ is finite). Define $\chi^\varphi(\omega, B) \subset \Pi^\omega$ to be the set of all signals acquired by $\theta^\varphi(\omega)$. Define the demand of ω given B and φ to be $D^{\omega,\varphi}(B) = \arg\max \left\{ \hat{v}_i^{\theta^\varphi(\omega)} \mid i \in B \cap \Psi\left(\theta^\varphi(\omega)\right) \right\} \in B$, which is the most preferred college in B for a student in state $\theta^\varphi(\omega)$. Note that the effects of information acquisition are captured within the definition of $D^{\omega,\varphi}(\cdot)$, as we assume that student ω acquires information according to the information acquisition strategy φ and available information $\theta(\omega), B$. For each inspection type $\theta \in \Theta$, we also define the demand of θ under information acquisition strategy φ to be $D_i^{\theta,\varphi}(B) = F^{|\theta}\left(\{\omega \in \theta : D^{\omega,\varphi}(B) = i\}\right)$. Note that the college demanded by a student $\omega \in \Omega$ is a deterministic function of ω (since the inspection strategy and values are fixed), but the college demanded by a type $\theta \in \Theta$ is probabilistic.

The optimal information acquisition strategy is the result of the utility-maximizing sequential inspection rule, defined as follows. For a student ω with budget set B, her utility after applying the sequential information acquisition rule φ will be $v^\omega_{D^{\omega,\varphi}(B)} - c^\omega(\chi^\varphi(\omega, B))$. The expected utility of a student ω in state $\theta = \theta(\omega)$ with budget set B is the expectation of this quantity over realizations drawn from $F^{|\theta}$. We let φ^* denote the sequential rule that maximizes this expected utility over all choices of φ.[6] For notational convenience, we will write $\chi^*(\omega, B) = \chi^{\varphi^*}(\omega, B)$ for the outcome of the optimal information acquisition rule and $\theta^*(\omega, B)$ for the resulting state.

The demand of student ω given budget set B is defined to be the student's most preferred college from B given the optimally acquired information $\chi^*(\omega, B)$,

$$D^\omega(B) = \arg\max \left\{ \hat{v}_i^{\theta^*(\omega,B)} \mid i \in B \cap \Psi\left(\theta^*(\omega, B)\right) \right\}.$$

We write

$$D_i^\theta(B) = F^{|\theta}\left(\{\omega \in \theta : D^\omega(B) = i\}\right)$$

[6]Note that this maximum is obtained as there are only finitely many signals to acquire from each state.

for the stochastic demand of a student in state θ. The expected utility of a student in state θ and budget set B is given by

$$\mathbb{E}_{\omega\sim F|\theta}\left[v^{\omega}_{D^{\omega}(B)} - c^{\omega}\left(\chi^{*}\left(\omega, B\right)\right)\right].$$

We are now ready to define our notion of stability that extends the standard stability notion.

Definition 3. *An outcome (μ, χ) is **stable** if it satisfies:*

1. *Any student $\omega \in \Omega_{\chi}$ is assigned to the college in her budget set that is most preferred given current information:*

$$\mu\left(\omega\right) = \arg\max\{\hat{v}^{\theta(\omega)}_{i} \mid i \in B^{\omega}(\mu) \cap \Psi(\theta(\omega))\}.$$

2. *For any $\omega \in \Omega_{\chi}$, student ω would not like to acquire more information:*

$$\chi^{\omega} = \chi^{*}\left(\omega, B^{\omega}\left(\mu\right)\right).$$

Note that conditions 1 and 2 together imply that $\mu(\omega) = D^{\omega}\left(B^{\omega}\left(\mu\right)\right)$ for each student $\omega \in \Omega_{\chi}$. Condition 1 is equivalent to that standard stability notion that there are no student and school that can benefit by forming a blocking pair.

An immediate observation is that a stable outcome exists for any $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$. For example, if all students acquire all possible information (that is, for all students $\chi\left(\omega\right) = \Pi^{\omega}$), condition 2 is trivially satisfied. The resulting economy is equivalent to a full information economy, which has a stable assignment. Together, these give a stable outcome. However, such an outcome requires students to pay large information acquisition costs and may be wasteful.

We refine the set of stable outcomes to ask that students acquire the appropriate information given what they can learn from the market. To motivate the refinement, consider a student who waits to see the market outcome before acquiring any information. Such a student $\omega_0 \in \Omega_0$ will learn her budget set $B^{\omega_0}\left(\mu\right)$ before starting her information acquisition, and will optimally acquire the information $\chi^{*}\left(\omega_0, B^{\omega_0}\left(\mu\right)\right)$. In contrast, a student $\omega' \in \Omega_{\chi}$ with initial value realization ω'_0 who has acquired information before knowing her budget set $B^{\omega'}\left(\mu\right) = B^{\omega'_0}\left(\mu\right)$ may have acquired $\chi^{\omega'} \neq \chi^{*}\left(\omega', B^{\omega'}\left(\mu\right)\right)$. Such a student regrets not waiting to learn $B^{\omega'}\left(\mu\right)$ before inspecting. The following definition of regret-free stable outcomes requires that all students acquire information optimally, as if they were provided all the information that is eventually available in the market.

Definition 4. *An outcome (μ, χ) is **regret-free stable** if it is stable, and for every $\omega \in \Omega_\chi$ with corresponding initial value realization ω_0 we have that*

$$\chi^\omega = \chi^* \left(\omega_0, B^\omega(\mu)\right).$$

Stability ensures that the student has not under-inspected, and regret-free stability additionally ensures that the student has not over-inspected given knowledge of her budget set (see also Example 2). If $\chi^\omega \neq \chi^*(\omega_0, B^\omega(\mu))$ we say that the student acquired information suboptimally.

We make a few technical remarks about the definition of regret-free stable outcomes. First, while the definition of regret-free stability is stated in terms of each student's realized type ω, it only requires that students conduct the optimal inspections given their observable information θ and the budget set $B^\omega(\mu) = B^\theta(\mu)$. In particular, following χ^* can only be optimal in expectation.[7] Second, an outcome (μ, χ) can be verified to be regret-free stable based on the revealed information $\theta(\omega)$ for any $\omega \in \Omega_\chi$.

2.2 Tractable expressions via Pandora's Box

We utilize the Pandora's Box model of Weitzman (1979) to illustrate our definitions, and to give a tractable information acquisition framework with closed form solutions. In this specification, values of colleges are independently distributed according to known priors, and the student can adaptively inspect a college and learn its value. This model assumes that a student can be only be assigned to a college they have inspected.[8]

More formally, the **Pandora's Box model** (equivalently, **Pandora's Box economy**) is an economy $(\mathcal{C}, \Omega, \eta, q)$ with students $\omega \in \Omega$ whose prior F^ω is the product of marginal distributions $\{F_i^\omega\}_{i \in \mathcal{C}}$, and whose signals $\Pi^\omega = \{\pi_i\}_{i \in \mathcal{C}}$ specify the value v_i^ω of student ω at college i, i.e., $\pi_i(v^\omega) = \{v \mid v_i = v_i^\omega\}$. Furthermore, students can only be assigned to colleges

[7]A student that knows the outcome (μ, χ) still faces uncertainty about their values for uninspected colleges. Thus, the optimal inspection policy may have the student inspect a college only to discover that the college is undesirable.

[8]Doval (2018) argues that without the assumption that the agent inspects the selected item the problem is not generally tractable, but derives optimal policies under sufficient parametric conditions. Beyhaghi & Kleinberg (2019) provide approximately optimal policies under non-obligatory inspections. Matějka & McKay (2015) and Steiner et al. (2017) provide an alternative tractable formulation based on rational inattention models (Sims 2003).

they inspect, i.e., $\Psi(\theta) = \left\{i \mid \pi_i \in \chi^\theta\right\}$. With slight abuse of notation, we refer to a signal by the college it inspects (i.e., $\pi_i = i$), the outcome of signal i with the value of college i (i.e., $\pi_i(v) = v_i$), and the cost of signal i with $c_i \geq 0$ (i.e., $c(\pi_i) = c_i$). We likewise identify each acquired signal χ^ω with the corresponding set of inspected colleges.

In the Pandora's Box model, a student who can choose a college out of a set of colleges $B \subset \mathcal{C}$ aims to adaptively acquire information χ^ω to maximize

$$\max_i \{v_i^\omega \mid i \in B \cap \chi^\omega\} - \sum_{i \in \chi^\omega} c_i^\omega.$$

The student's optimal information acquisition policy is given by the following known result.

Lemma 1. *(Weitzman 1979) Consider a student in state θ who can choose a college from $B \subset \mathcal{C}$. For each college $i \in B$, define the index $\underline{v}_i^\theta$ to be the unique solution to*

$$\mathbb{E}_{v_i \sim F_i^{|\theta}}\left[\max\{0, v_i - \underline{v}_i^\theta\}\right] = c_i^\theta.$$

The student's optimal adaptive information acquisition is to sequentially inspect colleges in decreasing order of their indices $\underline{v}_i^\theta$, and stop if the maximal realized value $\max\left\{v_i^\omega \mid i \in \chi^\theta\right\}$ is higher than the index of any remaining uninspected college in B.

Lemma 1 fully characterizes the optimal inspection policy $\chi^*(\omega, B)$, and implies this corollary.

Corollary 1. *In the Pandora's Box model, an outcome (μ, χ) is stable if for each $\omega \in \Omega_\chi$ with state $\theta = \theta(\omega)$ and each $i \in B^\theta(\mu) \setminus \{\mu(\theta)\}$, we have that either:*

- *$i \in \chi^\theta$ and $v_{\mu(\theta)}^\theta > v_i$; or*
- *$i \notin \chi^\theta$ and $\mathbb{E}_{v_i \sim F_i^\theta}\left[\max\{v_{\mu(\theta)}^\theta, v_i\}\right] - c_i^\theta < v_{\mu(\theta)}^\theta$.*

Student ω can potentially block with any college i in her budget set. The matching is stable if each such college that student ω has inspected is less preferred than her assigned college $\mu(\theta)$, and each such college that student ω has not inspected is not worth inspecting.

In a matching market setting, a student's budget set may depend on the preferences of other students. We give an example showing that, in a Pandora's Box economy, students benefit by not inspecting before they know their budget set.

Example 1. *Suppose that* $\mathcal{C} = \{1, 2, 3\}$*, and consider a student with* $v_1 \sim F_1 = [8; 1/2]$*,* $v_2 \sim F_2 = [6; 1/2]$*, and* $v_3 \sim F_3 = [7; 1/3]$*, where* $[x; p]$ *denotes the probability distribution which assigns probability* p *to the value* x *and* $1 - p$ *to* 0*. Suppose the student's inspection costs are* $c_1 = c_2 = c_3 = 2$*. This implies* $\underline{v}_1 = 4$*,* $\underline{v}_2 = 2$*,* $\underline{v}_3 = 1$*.*

If $B = \{1, 2, 3\}$*, the optimal inspection strategy is to first inspect college 1, then inspect college 2 only if* $v_1 = 0$*, and then inspect college* 3 *only if* $v_2 = v_1 = 0$*. If instead* $B = \{2, 3\}$*, the optimal inspection strategy is to first inspect college* 2*, and then inspect college* 3 *only if* $v_2 = 0$*. In particular, if* $B = \{2, 3\}$ *the student will not inspect college* 1*, and if in addition* $v_2 = 6$*, the student will not inspect* 3*.*

Example 1 shows that it is valuable for a student to know the set of colleges B in her budget set. If the student does not know her budget set B, her inspection strategy may be sub-optimal in two ways. First, the student may inspect college 1 when it is not in her budget set, wasting the cost c_1. Second, the student may inspect college 2 (3) when she is able to attend college 1 (1 or 2). This is likely to waste the cost c_2 (c_3); it is optimal for the student to first inspect college 1 if college 1 is in her budget set, and so, with 50% chance, $v_1 = 8$ and the student's optimal inspection strategy is to only inspect college 1.[9] It follows that a student who is uncertain whether $B = \{1, 2, 3\}$ or $B = \{2, 3\}$ would prefer to wait to learn her exact budget set before inspecting.

We build on Example 1 to illustrate the difference between stability and regret-free stability.

Example 2. *Consider an economy* $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ *in which* $\mathcal{C} = \{1, 2, 3\}$*,* $q_1 = q_2 = q_3 = 1/6$*, all colleges have identical priority ranking over students, all students have priors and signals as described in Example 1, and any student* ω *prefers a college* i *to being unassigned only if* $v_i^\omega > 0$*. With slight abuse of notation, we let* r^ω *be the common priority for student* ω *at all colleges. In this economy, there is a stable outcome* (μ, χ) *in which students inspect all schools (i.e.,* $\chi(\omega) \equiv \mathcal{C}$*) and the matching is given by*

$$\begin{aligned} \mu(1) &= \{\omega \in \Omega_0 \mid r^\omega \geq 2/3, v_1^\omega = 8\}, \\ \mu(2) &= \{\omega \in \Omega_0 \mid r^\omega \geq 1/3, v_2^\omega = 6, v_1^\omega = v_3^\omega = 0\}, \\ \mu(3) &= \{\omega \in \Omega_0 \mid r^\omega \geq 1/3, v_1^\omega = 0, v_3^\omega = 7\}. \end{aligned}$$

The outcome (μ, χ) *is not regret-free stable: for example, students in* $\mu(2)$ *with* $r^\omega < 2/3$

[9] If $B = \{1, 2, 3\}$ then inspecting college 1 first yields expected utility 2.58, but first inspecting one of colleges 2 or 3 yields a lower expected utility of at most 2.08.

have wasted c_1 (as 1 is not in their budget set) and have also wasted c_3 (as given $v_2^\omega = 6$ the expected benefit of inspecting 3 is not worth the cost).

This economy has a unique regret-free stable outcome $(\mu^\dagger, \chi^\dagger) \neq (\mu, \chi)$. Here $\mu^\dagger$ is given by

$$\begin{aligned}
\mu^\dagger(1) &= \{\omega \in \Omega_0 \mid r^\omega \geq 2/3, v_1^\omega = 8\}, \\
\mu^\dagger(2) &= \{\omega \in \Omega_0 \mid r^\omega \geq 1/2, v_2^\omega = 6\} \setminus \mu^\dagger(1), \\
\mu^\dagger(3) &= \{\omega \in \Omega_0 \mid r^\omega \geq 1/6, v_3^\omega = 7\} \setminus \left(\mu^\dagger(1) \cup \mu^\dagger(2)\right).
\end{aligned}$$

The information acquired $\chi^\dagger$ is given by the budget sets determined by μ and the optimal inspection rule described in Example 1.

3 Cutoff Structure of Regret-Free Stable Outcomes

In this section, we provide several results about the structure of regret-free stable outcomes. We show, perhaps surprisingly, that regret-free stable outcomes always exist and form a non-empty lattice. We prove these results by giving a concise characterization of regret-free stable outcomes in terms of market-clearing cutoffs. The cutoffs provide a sufficient statistic for describing both components of a regret-free stable outcome, namely the matching and the optimally acquired information. In addition, cutoffs allow the interconnected information acquisition problems to be disaggregated across different students. These results allow us to shed light on the challenges in implementing such outcomes. In section 5, we leverage these results to construct mechanisms.

3.1 Equivalence of market-clearing outcomes and regret-free stable outcomes

Cutoffs $\boldsymbol{P} = \{P_i\}_{i\in\mathcal{C}} \in \mathbb{R}^{\mathcal{C}}$ are admission thresholds for each college, determining a budget set $B^\omega(\boldsymbol{P})$ for each student $\omega \in \Omega$ equal to the set of colleges where their priority is above the college's cutoff,

$$B^\omega(\boldsymbol{P}) = \{i \in \mathcal{C} \mid r_i^\omega \geq P_i\}.$$

The demand $D^\omega(\boldsymbol{P})$ of student ω given cutoffs $\boldsymbol{P}$ is the college selected by ω from budget set $B = B^\omega(\boldsymbol{P})$, where the student first optimally acquires the information $\chi^*(\omega, B^\omega(\boldsymbol{P}))$

and then selects her most preferred college given the revealed information,

$$D^{\omega}(\boldsymbol{P}) = D^{\omega}(B^{\omega}(\boldsymbol{P})).$$

Finally, aggregate demand for college i given cutoffs $\boldsymbol{P}$ in economy $\mathcal{E}$ is defined to be the measure of initial student realizations that demand college i,

$$D_i(\boldsymbol{P} \mid \eta) = \eta(\{\omega_0 \in \Omega_0 | \, D^{\omega_0}(\boldsymbol{P}) = i\}).$$

We write $D_i(\boldsymbol{P})$ when η is clear from context, and denote overall demand by $D(\boldsymbol{P}) = (D_i(\boldsymbol{P}))_{i \in \mathcal{C}}$.

Next, we define market-clearing cutoffs (as in Azevedo & Leshno (2016)) and show there is a one-to-one correspondence between market-clearing cutoffs and regret-free stable outcomes.

Definition 5. *A vector of cutoffs $\boldsymbol{P}$ is* market-clearing *if it matches supply and demand for all colleges with non-zero cutoffs: $D_i(\boldsymbol{P}) \leq q_i$ for all i and $D_i(\boldsymbol{P}) = q_i$ if $P_i > 0$.*

Theorem 1. *An outcome (μ, χ) is regret-free stable if and only if there exist market-clearing cutoffs $\boldsymbol{P}$ such that for all $\omega \in \Omega_\chi$ with corresponding initial value realization ω_0, we have*

$$\mu(\omega) = D^{\omega_0}(\boldsymbol{P}) \quad \textit{and} \quad \chi(\omega) = \chi^*(\omega_0, B^{\omega}(\boldsymbol{P})).$$

Proof. It is immediate to verify that if P is market-clearing, then (μ, χ), for $\mu(\omega) = D^{\omega_0}(\boldsymbol{P})$ and $\chi^{\omega} = \chi^*(\omega_0, B^{\omega}(\boldsymbol{P}))$, is a regret-free stable outcome. For the opposite direction, given a regret-free stable outcome (μ, χ), define $P_i = \inf\left\{r_i^{\omega'} \mid \omega' \in \mu(i)\right\}$ for any college i such that $\eta(\mu(i)) = q_i$ and $P_i = 0$ for any college i such that $\eta(\mu(i)) < q_i$. Then we have that $B^{\omega}(\mu) = B^{\omega}(\boldsymbol{P})$ for all $\omega \in \Omega$. Regret-free stability of (μ, χ) implies that $\chi^{\omega} = \chi^*(\omega_0, B^{\omega}(\boldsymbol{P}))$ for all $\omega \in \Omega_\chi$, and stability thus implies that $\mu(\omega) = D^{\omega_0}(\boldsymbol{P})$ for all $\omega \in \Omega_\chi$. Therefore, $\boldsymbol{P}$ are market-clearing cutoffs. □

Theorem 1 shows an equivalence between market clearing cutoffs and regret-free stable outcomes. Thus, existence of a regret-free stable outcome is equivalent to the existence of cutoffs $\boldsymbol{P}$ that clear the demand $D(\cdot)$.

3.2 Existence and uniqueness of regret-free stable outcomes in the Pandora's Box model

We first demonstrate that market-clearing cutoffs exist in the Pandora's Box model, and so regret-free stable outcomes exist in the Pandora's Box model. We focus on the Pandora's Box model, as this tractable setting allows us to explicitly construct regret-free stable outcomes and provide intuition for differences from complete-information settings. Notably, unlike the standard proof of existence in complete-information settings, the construction in our proof does not provide an algorithm that implements a regret-free stable outcome.

To prove the existence of market-clearing cutoffs, we note that the demand $D^{\omega_0}(B)$ in the Pandora's Box model can be rationalized by a strict ordering over colleges which is independent of the student's budget set. Namely, given $\omega_0 \in \Omega_0$, we construct a full information preference ordering $\succ^{\omega_0}$ such that for any budget set B, the demand $D^{\omega_0}(B)$ is identical to the demand of a student with preferences $\succ^{\omega_0}$ in a corresponding full information economy (despite the dependence of χ^* on the budget set B).[10] It follows that all structural results about the set of stable outcomes in a full information economy can be directly carried over to the Pandora's Box model.

Proposition 1 (Reduction to demand from complete information). *Let $\omega_0 \in \Omega_0$ be an initial value realization in the Pandora's Box model. Let $\succ^{\omega_0}$ be an ordering of $\mathcal{C}$ defined by*

$$i \succ^{\omega_0} j \Leftrightarrow \min\left\{\underline{v}_i^{\omega_0}, v_i^{\omega_0}\right\} > \min\left\{\underline{v}_j^{\omega_0}, v_j^{\omega_0}\right\}.$$

Then, for all $B \subset \mathcal{C}$,

$$D^{\omega_0}(B) = \max_{\succ^{\omega_0}}(B).$$

The proof of Proposition 1 follows from similar arguments in Kleinberg et al. (2016), and is provided in the appendix.

Note that Proposition 1 implies that even though the set of colleges B in a student's budget set will affect her inspection decisions, her final demand is the same as if she chose according to the ordering $\succ^{\omega_0}$. In particular, to determine the demanded college $D^{\omega_0}(B)$, it suffices to consider the relationship between the n values $\min\{\underline{v}_i^{\omega_0}, v_i^{\omega_0}\}$ without considering the effects of the set B on how information is acquired. It is worth noting, however, that the preferences $\succ^{\omega_0}$ depend on the realized values v^{ω_0}. Hence, any student ω_0 who only has initial information $\theta = \theta(\omega_0)$ does not a priori know their corresponding preferences

[10] In a full information economy, the demand of a student $\omega = (r, v)$ with ordinal preferences $\succ^{\omega}$ (as induced by her values v) is $D^{\omega}(B) = \max_{\succ^{\omega}}(B)$.

$\succ^{\omega_0}$, and so a mechanism cannot rely on students reporting $\succ^{\omega_0}$ in order to implement a regret-free stable outcome.[11]

By Proposition 1, for any Pandora's Box economy $\mathcal{E}$, we can construct a full information economy $\tilde{\mathcal{E}}$ (as in Azevedo & Leshno (2016)) that has the same demand for any cutoffs.

Corollary 2. *Let $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ be a Pandora's Box economy. There exists a full information economy $\tilde{\mathcal{E}} = (\mathcal{C}, \tilde{\eta}, q)$ such that for any cutoffs $\boldsymbol{P}$, we have $D(\boldsymbol{P} \mid \eta) = D(\boldsymbol{P} \mid \tilde{\eta})$.*

Proof. Define the measure $\tilde{\eta}$ over $[0,1]^{\mathcal{C}} \times \mathcal{L}(\mathcal{C})$ by[12] $\tilde{\eta}(A) = \eta\left(\left\{\omega_0 \in \Omega_0 \mid \left(r^{\theta}, \succ^{\omega_0}\right) \in A\right\}\right)$. The result follows from the definition of demand and Proposition 1. $\square$

From Corollary 2, the demand $D(\cdot)$ given any Pandora's Box economy $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ is identical to the demand of a full information economy. Since the same market-clearing condition characterizes stable matching in full information economies, it follows that the regret-free stable outcomes of a Pandora's Box economy $\mathcal{E}$ have the same attractive structural properties as the stable outcomes of a full information economy.

Proposition 2. *For every Pandora's Box economy $\mathcal{E}$, there exists a regret-free stable outcome, and the set of regret-free stable outcomes is a non-empty lattice.*

Proof. The set of market cutoffs for the full information economy $\tilde{\mathcal{E}}$ constructed in Corollary 2 is a non-empty lattice (Blair 1988, Azevedo & Leshno 2016). Since demand under $\mathcal{E}$ and $\tilde{\mathcal{E}}$ is identical, the set of market clearing cutoffs for $\mathcal{E}$ is also a non-empty lattice. Therefore, regret-free stable outcomes exist and form a lattice defined as follows. Let $(\mu, \chi), (\mu', \chi')$ be regret-free stable outcomes and let $\boldsymbol{P}, \boldsymbol{P'}$ be the corresponding market clearing cutoffs. Define the order $\triangleright$ over outcomes by $(\mu, \chi) \triangleright (\mu', \chi')$ if and only if $P_i \geq P'_i$ for all i. $\square$

Remark 1. *The proof of Proposition 2 does not require that $\mathcal{E}$ is a continuum economy. If E is a discrete economy (formally defined in Section 5), we can construct a full information discrete economy $\tilde{E}$ that has identical demand. Since market clearing cutoffs for $\tilde{E}$ form a non-empty lattice, the set of regret-free stable outcomes of E forms a non-empty lattice.*

[11] Furthermore, the demand of a student ω_0 (equivalently, the corresponding preferences $\succ^{\omega_0}$) may not correspond to the preference ordering of a student who acquires all signals. For example, suppose there are two colleges at which student ω_0 has Pandora's Box indices $\underline{v}_1 = 4$ and $\underline{v}_2 = 3$ and realized values $v_1 = 4$ and $v_2 = 5$. Then the student will inspect the first college and stop, demanding the first college. Also, as required by Proposition 1, the preference ordering $\succ^{\omega_0}$ ranks college one first and then college two second. However, the preference ordering of a student who has acquired all signals is to rank college 2 first and then college 1 second.

[12] We use $\mathcal{L}(\mathcal{C})$ to denote all strict orderings over $\mathcal{C}$.

One consequence is that there is a unique regret-free stable matching that is ex-ante optimal for all students.

Proposition 3. *For every Pandora's Box economy $\mathcal{E}$, there exists a unique student-optimal regret-free stable outcome $\left(\mu^{\dagger}, \chi^{\dagger}\right)$ that achieves the highest ex-ante expected utility for each student type out of all regret-free stable outcomes. That is, for any $\theta_0 \in \Theta_0$*

$$\mathbb{E}_{\omega \sim F^{\theta_0}}\left[v^{\omega}_{\mu^{\dagger}(\omega)} - c^{\omega}\left(\chi^{\dagger}\left(\omega\right)\right)\right] \geq \mathbb{E}_{\omega \sim F^{\theta_0}}\left[v^{\omega}_{\mu(\omega)} - c^{\omega}\left(\chi\left(\omega\right)\right)\right]$$

for any regret-free stable outcomes (μ, χ).

The proof of Proposition 3 can be found in the appendix.

3.3 Existence and uniqueness of regret-free stable outcomes under WARP

In this section, we show that our results on the existence and structure of market-clearing cutoffs extend beyond the Pandora's Box model. Specifically, regret-free stable outcomes exist and form a non-empty lattice in any economy where student demand satisfies the weak axiom of revealed preferences (WARP). This is somewhat surprising, given that a regret-free stable outcome has the strong requirement that each student has acquired information optimally, as if she were last to market.

Definition 6. *Demand $D^{\omega}(\cdot)$ for a student ω satisfies the **weak axiom of revealed preferences (WARP)** if $D^{\omega}(B) \neq i \Rightarrow D^{\omega}(B') \neq i$ for all budget sets $B \subseteq B'$ and any college $i \in B$. We say that $\mathcal{E}$ is an economy where demand satisfies WARP if $D^{\omega_0}(\cdot)$ satisfies WARP for all $\omega_0 \in \Omega_0$.*

In other words, a student's demand satisfies WARP if, whenever colleges are added to the student's budget set, the student either demands the same college or one of the added colleges. An immediate corollary of Proposition 1 is that demand always satisfies WARP in the Pandora's Box model. Beyond the Pandora's Box model, the condition that demand satisfies WARP is not without loss of generality. Indeed, WARP may be violated when there are informational complementarities between colleges. For example, a student may have a signal (visiting a city) that is informative about both colleges i_1, i_2 (both in the same city), but the signal is costly and only worth acquiring if $\{i_1, i_2\} \subset B$. For such a student ω, it may be that $D^{\omega}(\{i_1, j\}) = j$ but $D^{\omega}(\{i_1, i_2, j\}) = i_1$.

Theorem 2. *For any economy $\mathcal{E}$ where demand satisfies WARP, the set of regret-free stable outcomes forms a non-empty lattice.*

Theorem 2 can be proved analogously to the proof of Proposition 2 by constructing a full-information economy where each student has the same demand as in $\mathcal{E}$; the existence of such an economy is guaranteed by WARP. Alternatively, Theorem 2 can be proved by replicating the proof of Azevedo & Leshno (2016), which uses the weaker condition that aggregate student demand satisfies weak gross substitutes. As it turns out, all the results in this subsection would continue to hold if we relaxed the assumption that individual demands satisfy WARP and required only that aggregate student demand satisfies weak gross substitutes. Moreover, while this gross substitutes property is necessary for our theoretical result, we conjecture that existence of a regret-free stable outcome is likely to hold more generally in practice, similarly to the case of matching with couples (Ashlagi et al. 2014, Kojima et al. 2013).

The equivalence between full-information economies and economies where demand satisfies WARP also allows us to carry over sufficient conditions for the uniqueness of regret-free stable outcomes. Intuitively, the condition is that the implied aggregate demand given the distribution of students η is sufficiently smooth, as captured in the following definition.

Definition 7. *(Azevedo & Leshno 2016) A measure η is* regular *if the image under $D(\cdot \mid \eta)$ of the closure of the set $\left\{\boldsymbol{P} \in (0,1)^{\mathcal{C}} \mid D(\cdot|\eta) \text{ is not continuously differentiable at } \boldsymbol{P}\right\}$ has Lebesgue measure 0.*

For example, any measure that has a piecewise continuous density satisfies regularity.

Corollary 3. *Suppose η is a regular measure. Then for almost every q with $\sum_i q_i < 1$, if demand in the economy $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ satisfies WARP, then $\mathcal{E}$ has a unique regret-free stable outcome.*

Corollary 3 is a direct analog of Theorem 1 in Azevedo & Leshno (2016), and we omit the proof.[13]

4 Communication Processes & Information Deadlocks

In this section, we show it is impossible for a clearinghouse to guarantee a regret-free stable outcome. To formalize this impossibility, we define communication processes, which capture

[13] We note that while the model in (Azevedo & Leshno 2016) is not precisely the same as the one considered here, their argument is more general than stated and indeed their proof follows without change when applied to our model.

how information is collected from and provided to students by a mechanism. The language of communication processes also allows us to compare the information provided to students under different mechanisms.

4.1 Communication Processes

We first define communication processes. Our notation is meant to explicitly capture what the mechanism knows; in particular, we will not necessarily assume the mechanism has access to a prior. Formally, a *communication process* is given by $\mathcal{P} = (\theta_R, \sigma; \iota, a, m)$. The reporting function θ_R and information acquisition strategy σ describe students' behavior. The initial information ι, allocation function a, and message function m describe the mechanism. Possible student reports are given by the set Θ_R, and possible mechanism messages are given by the set $\mathcal{M}$. In each period t, the students acquire signals as dictated by their information acquisition strategy and then send a report to the mechanism; the mechanism in turn observes these reports and either chooses an allocation or sends a message back to the students.

We first describe student behavior. The reporting function $\theta_R : \Theta \to \Theta_R$ specifies the report $\theta_R(\theta)$ a student sends to the mechanism when her inspection state is θ. Throughout, we assume that $\theta_R(\theta)$ correctly reports the publicly available priorities r^θ, and further incorporates all the information the mechanism has about student θ.[14] We sometimes consider $\theta_R(\theta)$ as an equivalence class of Θ. The information acquisition strategy $\sigma : \Omega \times \mathcal{M} \to \cup_{\omega \in \Omega} 2^{\Pi_\omega}$ specifies how students acquire information, where $\sigma(\omega, m) \subset \Pi^\omega$ is the set of all signals acquired by a student ω who receives message m. We assume any information the student has about the economy $\mathcal{E}$ is provided through the communication process. To capture this, we let $m_0 \in \mathcal{M}$ denote the empty message, and let $\sigma(\omega_0, m_0)$ denote the information acquired by a student $\omega_0 \in \Omega_0$ in the first period before the student receives any messages from the mechanism.

We now describe the mechanism. The initial information available to the mechanism is given by ι. We say that the mechanism has no initial information if $\iota = (q)$. The mechanism proceeds in discrete time periods indexed by $t \geq 1$. At each time t, the mechanism learns the distribution η_R^t over student reports Θ_R; we denote this distribution by $\eta_R = \eta_R^t$ when t is clear from context.[15] The allocation function a takes the information η_R, ι, t and either

[14] That is, we implicitly assume that $\theta_R(\theta)$ also incorporates any relevant information that was disclosed in previous reports. This is for notational convenience; we could alternatively record the history in the mechanism's state space.

[15] With slight abuse of notation, the mechanism also learns the marginal density of η_R^t conditional on any

decides to continue the process, which we denote by $a(\eta_R, \iota, t)$ = "continue," or decides to terminate and compute an allocation $a(\eta_R, \iota, t) = \mu$. To capture that the mechanism can only distinguish between students in states θ, θ' if $\theta_R(\theta) \neq \theta_R(\theta')$, we require that the outputted assignment is given by $\mu : \Theta_R \to \mathcal{C}_\phi$.[16] If the mechanism decides to continue, the next round starts with a message from the mechanism to students. The message function $m(\eta_R, \iota, t+1) \in \mathcal{M}^{\Theta_R}$ sends to each student who in round t reported some $\theta_R \in \Theta_R$ a message in $\mathcal{M}$; i.e. we allow the mechanism to send individual messages to students based on their reports. We use $m(\theta_R; \eta_R, \iota, t+1) \in \mathcal{M}$ to denote the message received in round $t+1$ by a student whose report in round t is θ_R.

We say that an economy $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ is *compatible* with a communication process $\mathcal{P} = (\theta_R, \sigma; \iota, a, m)$ if the reporting function $\theta_R : \Theta \to \Theta_R$ is defined for the type space Θ that corresponds to Ω. We restrict attention to communication processes that terminate with an allocation for any compatible economy, i.e., where student behavior and mechanism computation and communication is such that for any economy $\mathcal{E}$ there exists a period $t = t(\mathcal{P}, \mathcal{E})$ at which the mechanism decides to terminate.[17] We let $\mathcal{P}(\mathcal{E}) = (\mu, \chi)$ denote the allocation μ specified by the terminal message and the information χ that has been acquired by students by the terminal time t, and we call $\mathcal{P}(\mathcal{E})$ the *outcome* of $\mathcal{P}$ on $\mathcal{E}$.

Definition 8. *Let $\mathcal{P} = (\theta_R, \sigma; \iota, a, m)$ be a communication process. We say that $\mathcal{P}$ is **regret-free-stable** if for any compatible economy $\mathcal{E}$ we have that $\mathcal{P}(\mathcal{E}) = (\mu, \chi)$ is a regret-free stable outcome for $\mathcal{E}$.*

4.2 Examples of common communication processes and their properties

Our definition of a communication process is general enough to capture many common mechanisms. For example, the one-shot college-proposing deferred acceptance mechanism (Gale & Shapley 1962) asks students (and colleges) to submit ordinal preference lists and then runs an algorithm to determine the allocation. In canonical descriptions, the mechanism does not provide any information to students to aid them in forming preference lists. Translating this to our language, the corresponding communication process sends the empty message

possible rank r.

[16]That is, the assignment of a student θ whose most recent report is $\theta_R(\theta)$ is $\mu(\theta_R(\theta))$.

[17]One technical issue is that in continuum economies, the standard DA algorithm may not terminate in finite time, but rather converges to a stable matching. To avoid this complication, we allow for the communication process to terminate at transfinite time.

m_0, students inspect colleges according to some information acquisition process (defining σ) and send a message containing their ordinal preference lists and college priorities (defining θ_R). The mechanism runs the DA algorithm on reported preferences and priorities (captured by η_R), and outputs the resulting assignment ($a(\eta_R, q, 1) = \mu$).

The lack of information provision by one-shot DA can lead to very inefficient information acquisition. Since students acquire information before learning their budget sets, it is straightforward to see that under any student information acquisition strategy σ, the resulting communication process can either lead to a non-stable outcome (e.g., if students only inspect one college), or to an outcome where students regret their inspection decisions (e.g., if students inspect all signals). In Appendix A, we provide examples of economies in which any one-shot communication process leads to regret for an arbitrarily large fraction of the students.

Some clearinghouses have recognized the importance of providing information to students, and so have implemented versions of DA that incorporate information provision. For example, the SISU mechanism used in Brazil (Bo & Hakimov 2019) and the assignment mechanism used in Inner Mongolia (Chen & Pereyra 2015, Gong & Liang 2016) aim to provide applicants with better information by running several simulation rounds in which students participate in non-binding DA before running a final and binding round of DA. The SISU mechanism gives a communication process with T periods. In each period: (1) students acquire signals and (2) report a preference list; and (3) the mechanism runs DA on the submitted preferences and sends each student their assignment under the computed student-optimal stable matching. The periods $1, \ldots, T-1$ serve as practice rounds, and the assignment is entirely determined by the students' final reports in the T-th period. Translating this mechanism to our language requires us to specify students' information acquisition strategies σ and report functions θ_R. Two immediate concerns arise. First, reports sent by students in periods $1, \ldots, T-1$ are "cheap talk," and thus messages sent by the mechanism in these periods may not be informative.[18] Second, if messages are informative, students may want to observe these messages before acquiring information, and so may want to delay their information acquisition to later periods. In other words, messages may not be informative, and if they are informative then this may incentivize students to delay acquiring the

[18] For example, some students may change their reported preferences from period to period as they acquire new information, which could lead other students to make inaccurate inferences about their budget sets. Bo & Hakimov (2019) analyze whether students may want to misreport their preferences in cheap-talk rounds in a full information environment. Theorem 3 will imply an impossibility result regardless of whether students are truthful or not.

information required to make them informative.

An iterative implementation of the college-proposing deferred acceptance mechanism (ICPDA) circumvents these issues by taking advantage of the fact that colleges know their ranking of students. We prove that this mechanism outputs regret-free stable outcomes for a class of economies. For expository convenience, we focus attention on the Pandora's Box model. We define the ICPDA process $\mathcal{P}^{ICPDA} = (\theta_R, \sigma; \iota, a, m)$ using our language of communication processes. The reports $\theta_R \in \Theta_R$ label each college as "tentatively accepted" or "rejected."[19] The messages $m \in \mathcal{M}$ correspond to subsets of colleges. The information acquisition strategy $\sigma(\omega, m)$ acquires signals from all colleges $i \in m$. The mechanism has no initial information (i.e., $\iota = (q)$).

ICPDA proceeds over a number of periods, where colleges propose to students and students acquire more information, accept at most one proposal, and reject the rest. Students inspect nothing and report nothing in period $t = 1$.[20] In every period $t \geq 1$, colleges propose to the top ranked students who have not yet rejected them; i.e., college i proposes to all students ω for whom there are at most q_i students with higher rank than ω at i who have not yet rejected i. This is implemented by sending the message $m_t = m\left(\theta_R\left(\omega\right); \eta_R, \iota, t\right) \subset \mathcal{C}$ consisting of the set of colleges that propose to student ω in period t. In periods $t > 1$, students follow the all-inspecting strategy $\sigma\left(\omega, m\right) = \{i \in \cup_{t' \leq t} m_{t'}\}$, where they collect signals from all colleges that have proposed to them. They then report a message which tentatively accepts the single college they most prefer out of $\cup_{t' \leq t} m_{t'}$ and rejects all other colleges in $\cup_{t' \leq t} m_{t'}$. The mechanism stops and outputs the current assignment as specified by the tentative acceptances when there are no further proposals.

The following proposition shows that the information provided by the iterative process can help facilitate more efficient information acquisition, albeit in a restrictive setting where students wish to inspect any college that is interested in admitting them.[21]

Proposition 4. *Let $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ be a Pandora's Box economy. Suppose that students wish to inspect any college in their budget set. That is, for any $\omega \in \Omega$ and budget set B, we have that $\chi^*\left(\omega, B\right) = B$. Then $\mathcal{P}^{ICPDA}\left(\mathcal{E}\right)$ is a regret-free stable outcome.*

The proof uses the fact that when $\chi^*\left(\omega, B\right) = B$, in a regret-free stable outcome, stu-

[19] The reports also contain the college priorities of a student; we suppress that in this notation.

[20] That is, $\sigma\left(\omega, m_0\right) = \phi$ and $\theta_R\left(\theta\right) = \phi$ for any $\theta \in \Theta_0$.

[21] The informational benefits of ICPDA have been identified previously in the literature, e.g. Rastegari et al. (2014) show that ICPDA minimizes interview costs, and that having the informed side move first improves informational efficiency. Similarly, Ashlagi et al. (2017) and Kanoria & Saban (2020) also find that it is more efficient to have the informed party propose.

dents collect signals exactly from all colleges in their budget set. Under $ICPDA$, a student receives offers exactly from all the colleges in her budget set. In particular, students are provided enough information to ensure that every inspection is part of the optimal information acquisition, and student reports allow the mechanism to make sufficient progress.

When colleges have identical rankings over students, there is a different version of ICPDA, often called iterative serial dictatorship (ISD), that results in regret-free stable outcomes. The ISD communication process identifies students whose budget sets are fully determined, and asks only those students to complete their information acquisitions and report their demand. Other students do not acquire any information. It does so by tracking the remaining capacity q_i^t for each college i and period t. At period t, the budget set for the top $\min_{i:q_i^t>0} q_i^t$ students who have not yet acquired information is fully determined and equal to $\{i : q_i^t > 0\}$. The process sends a message $m = \{i : q_i^t > 0\}$ to these top remaining students, and messages all other remaining students they should wait. Students optimally acquire the information $\chi^*(\omega, B)$ immediately after learning their budget set $B = m$, and the process updates remaining capacities according to their reported demands. Intuitively, this results in a regret-free stable outcome because students only acquire information once their entire budget set is revealed to them.

While both our initial descriptions of ICPDA and ISD can be thought of as iterative college-proposing DA, they take very different approaches to guiding students' information acquisition. This can be seen in their different messaging and information acquisition functions: ICPDA sends a message to a student about a given college as soon as it is in their budget set; ISD sends a message to a student only once their budget set is fully determined. The two mechanisms exploit different special features of the economy. The success of ICPDA (Proposition 4) relies on the fact that students find it optimal to inspect a college given knowledge that it is in their budget set (which is not the case for the student in Example 1). The success of ISD relies on the fact that it can always fully determine budget sets for some remaining students. It is clear that neither process will be regret-free stable without these special assumptions.[22] In the next section, we explore whether a communication process can reach regret-free stable outcomes in general settings.

[22] Roth & Xing (1997) provide a detailed description of the market for clinical psychologists that used a version of ICPDA, and find there is not enough time for the market to clear. This is consistent with students holding to offers and delaying their decisions while they wait for more information from the market.

4.3 Information Deadlocks and an Impossibility Result

Given that regret-free stable outcomes exist, a natural question is whether it is always possible to find them without causing regret. We focus on a class of communication processes that do not use initial information beyond the college capacities (i.e., $\iota = (q)$). This class includes many standard implementations of common mechanisms (including ICPDA, DA, SISU, etc.). We find that no process from this class can guarantee a regret-free stable outcome on general markets. In other words, price-discovery is costly: these processes cannot find the market clearing cutoffs without imposing some regret on some students.

The impossibility stems from the existence of information deadlocks, where the communication process cannot uncover sufficient information to safely collect more information. In particular, the impossibility result holds even if the mechanism could dictate how students should inspect (i.e., the impossibility result holds even ignoring any incentive constraints).

Example 1 provides some intuition for why information deadlocks can arise. The student in Example 1 has partial knowledge of her budget, but this knowledge is insufficient to determine which signal to acquire (in particular, the information provided by ICPDA is insufficient). A student's budget set may depend on other students' signal realizations. An information deadlock arises when all students are simultaneously in this situation: each student requires more information about their budget set in order to proceed, but this depends on the signal realizations of other students who are likewise waiting for more information before obtaining additional signals.

We now formalize the notion that a student may wish to delay information acquisition until she has obtained more information about her budget set.

Definition 9. *A student of type $\theta_0 \in \Theta_0$ is* stagnant *given budget sets B, B' if $\chi^*(\omega_0, B) \neq \phi$ and $\chi^*(\omega_0, B') \neq \phi$ for all $\omega_0 \in \theta_0$, but*

$$\cap_{\omega_0 \in \theta_0} \left(\chi^*(\omega_0, B) \cap \chi^*(\omega_0, B')\right) = \phi \, .$$

A student is stagnant given B, B' if she finds it optimal to acquire more signals if her budget set is either B or B', but *any* signal she acquires would lead to suboptimal information acquisition under one of B or B'. In other words, if a student is uncertain whether her budget set is B or B', no signal she acquires guarantees optimal information acquisition. Note that if a student is stagnant given B and B' then it must be that $B \neq B'$. Indeed, if $B = B'$ then $\cap_{\omega_0 \in \theta_0} \left(\chi^*(\omega_0, B)\right) \neq \phi$ is not empty since it contains $\varphi^*(\theta_0, B)$, which is the first signal acquired by a student of type θ_0 given budget set B.

There will always exist stagnant students for any sufficiently rich student state space. Formally, the following lemma shows that for any two distinct budget sets B and B', the Pandora's Box model contains student types that are stagnant given B, B'.

Lemma 2. *For any two distinct budget sets $B \neq B' \subseteq \mathcal{C}$, the Pandora's Box model includes a student type θ_0 that is stagnant given B, B'.*

To prevent the process from "abusing" the continuum model, we require that the reporting function satisfies a mild assumption. In a discrete economy, aggregate demand is uncertain even if the mechanism collects all of the initial information available to students. In a continuum economy, there is no aggregate uncertainty, as the demand of a mass of students with known types can be perfectly predicted. The following assumption ensures the process faces uncertainty about aggregate demand even after receiving initial reports from all students.

Definition 10. *A reporting function $\theta_R : \Theta \to \Theta_R$* maintains aggregate uncertainty *if for every $\theta_0 \in \Theta_0$ and budget sets B, B' such that θ_0 is stagnant on B, B', and every $\varepsilon > 0$ and $i \in B$, there exists $\theta_0' \in \Theta_0$ such that $\theta_R(\theta_0') = \theta_R(\theta_0)$, θ_0' is stagnant on B, B', and $D_i^{\theta_0'}(B) > 1 - \varepsilon$.*

Formally, the definition asks that given the information reported by a mass of students, together with the information that these students are stagnant given B, B', one still cannot rule out that almost the entirety of this mass of students will demand any college from B. This assumption will be violated if students fully report their types (which requires students to report their prior distributions). We show that it is satisfied for a natural reporting function for the Pandora's Box model, in which students report both their strategies for acquiring information as well as all signals they have acquired.

Lemma 3. *Consider the Pandora's Box model and the reporting function θ_R, where students report all their acquired signals and their indices. That is,*

$$\theta_R(\theta) = \left(r^\theta, \Pi^\theta, c^\theta, \chi^\theta, \{\pi(v)\}_{\pi \in \chi^\theta}; \underline{v}^\theta\right)$$

where $\underline{v}^\theta$ is the vector of the student's inspection indices (as defined in Lemma 1). Then θ_R maintains aggregate uncertainty.

We are now ready to state our main impossibility result: if the set of possible student types includes all Pandora's Box types, then no communication process can guarantee a regret-free stable outcome.

Theorem 3. *Let $\mathcal{P} = (\theta_R, \sigma; \iota, a, m)$ be a communication process defined over Θ. Suppose that Θ includes all the Pandora's Box types, θ_R maintains aggregate uncertainty, and the process has no prior information (i.e., $\iota = q$). Then $\mathcal{P}$ is not regret-free stable. Moreover, there is a constant $\beta > 0$ (independent of $\mathcal{P}$) such that there exists an economy where at least a β fraction of the students acquire information suboptimally under $\mathcal{P}$.*

Theorem 3 together with Theorem 2 shows that the difficulty lies in *discovering* regret-free stable outcomes without incurring additional costs, rather than guaranteeing their existence.[23] Previous impossibility results for matching under incomplete information found that a stable matching may not exist. In contrast, in our model when demand satisfies WARP, a regret-free stable outcome is guaranteed to exist. The challenge stems from the fact that even given the knowledge that a regret-free stable outcome exists, the communication process still needs to collect information from students to find such an outcome. Theorem 3 shows that collecting the required information to find a regret-free stable outcome necessitates making a positive fraction of all students acquire information suboptimally and incur regret.

The following example provides some intuition for the proof of Theorem 3. It describes a Pandora's Box economy in which each student's budget set can be one of two possibilities. The student is stagnant given the two possible budget sets, and her actual budget set depends on other students' signal realizations. These dependencies form a cycle that constitutes an information deadlock.

Example 3. *We construct a collection of Pandora's Box economies with three groups of students X, Y, Z each of mass $1/3$, and three colleges $\{1, 2, 3\}$ each with capacity $2/3$.*[24] *Students in X and Y are top-ranked at college 3; students in Y and Z are top-ranked at college 1; and students in Z and X are top-ranked at college 2. As students in X are top-ranked at colleges 2 and 3, and bottom-ranked at college 1, they either have the budget set $B = \{1, 2, 3\}$ or $B' = \{2, 3\}$, depending on the demand of students in Y and Z. We can construct Pandora types for students in X (similarly to Example 1) such that all students in X are stagnant given B, B' (and symmetrically for Y, Z).*

Consider a communication process that aims to learn the actual budget sets by having the students acquire information. Students in X need to know the decisions of students in $Y \cup Z$ to know their budget sets, and likewise for students in Y, Z. Thus, any students who

[23] A related result by Ashlagi & Gonczarowski (2018) shows that while the student-proposing DA is strategy-proof, no stable matching is obviously strategy-proof (Li 2017).

[24] Strictly speaking, our model requires that total capacity not exceed the mass of students; this can be satisfied in any economy by adding dummy students with low priority at every college.

are the first to inspect may acquire information suboptimally and the outcome of the process may not be regret-free stable.

This example can be formalized to show that any communication process which maintains aggregate uncertainty must ask some stagnant student to acquire information. The formal construction creates a collection of economies that cannot be distinguished based on reports before students have acquired information. For each possible choice of a first student to inspect and each of her possible inspections, there exists an economy in the collection where this inspection leads to suboptimal information acquisition.

Theorem 3 applies to common implementations of assignment mechanisms, including ICPDA, Probabilistic Serial, and the one-shot Boston mechanism, and implies that they are not regret-free stable. That is, under commonly-used mechanisms, students can benefit from delaying their information acquisition until after the market resolves. Practical implementations of these mechanisms therefore impose deadlines, "exploding offers," or other activity rules to force students to collect information early and make decisions before the rest of the market resolves. Theorem 3 implies that these approaches must necessarily impose additional price-discovery costs on students. But even though some price-discovery costs are unavoidable, it is not clear that the costs imposed by these commonly-used mechanisms are anywhere near the minimal possible costs. In the next section, we employ an alternative design approach to better inform students and reduce price-discovery costs.

5 Implementing Approximately Regret-Free Stable Outcomes

Our results highlight the role of matching markets in facilitating *price discovery*. Section 3 showed that regret-free stable outcomes exist, and are equivalent to market-clearing cutoffs. Section 4 highlights that the challenge in implementing a regret-free stable matching is in efficiently discovering such market-clearing cutoffs, without incurring excessive price-discovery costs. This difficulty suggests that one should instead aim to implement approximately regret-free stable outcomes, and in this section, we explore that goal.

The connection to market-clearing cutoffs suggests a two-stage approach to designing matching mechanisms. In a first stage, the mechanism engages in price discovery to learn market-clearing cutoffs. In a second stage, the mechanism publishes the learned cutoffs, thereby determining the allocation and guiding student information acquisition. While such a two-stage approach differs from the mechanisms typically suggested in the matching literature, such mechanisms are common practice in combinatorial auctions (Ausubel & Baranov 2014, Levin & Skrzypacz 2016) and our survey in Section 6 finds such approaches are common in university admission.[25] Since Theorem 3 shows that no mechanism can guarantee a regret-free stable outcome without prior information, any such mechanism will either have to exploit some prior information, or only obtain approximate regret-free stability.

We address challenges with this approach by discussing several implementations of two-stage mechanisms. One challenge is that learned cutoffs may be noisy, and will not exactly clear the market. Section 5.1 models the approximation error when cutoffs are estimated from external historical data, and shows that a second stage with flexible capacities can address approximation errors. Another challenge is that, absent external information sources, the first stage needs to learn the cutoffs at minimal cost to students. Section 5.2 considers two possible first-stage methods of learning cutoffs by estimating demand from a sample of students.

5.1 Price discovery from historical information: Handling estimation errors

Many admission systems operate year after year, and many real life mechanisms exploit the information revealed by admission in previous years (see Section 6). We model the annual variations in the economy from year to year by modeling each year as a finite sample of students from the same continuous population distribution. This allows us to assess the resulting estimation error for a mechanism that uses historical data to learn approximate market-clearing cutoffs, and suggest how the mechanism can address this challenge.

Formally, a *finite economy*[26] is given by $E = (\mathcal{C}, \Omega, S, q)$, where $S \subset \Omega_0$ is a finite set of students and $q \in \mathbb{N}^{\mathcal{C}}$ is the number of seats at each college. We interpret such an economy as being equivalent to a continuum economy $\left(\mathcal{C}, \Omega, \eta^S, q\right)$ where η^S has $|S|$ equally sized atoms on S, with three changes. First, the distribution of realized values does not need to be consistent with students' priors. Second, the total mass of students is scaled to be $|S|$,

[25] Several papers argue that students have or should have knowledge of the cutoffs. Artemov et al. (2017) argue that many such misreports can be explained by advanced knowledge of admission chances, and give empirical methods that account for this. Fack et al. (2019) argue that applicants with knowledge of admission chances may deviate from reporting their full rank-ordered list.

[26] We note that the definitions of stability and regret-free stability ignore the effect of a student on her budget set, and there are examples of finite economies where our definition of $B^{\omega}(\mu)$ does not properly capture the set of colleges a student ω can be admitted to (see Appendix E). However, previous theoretical results (Azevedo & Leshno 2016, Menzel 2015) show that such issues arise only in knife-edge cases.

rather than 1. Third, indifference curves contain at most one student.[27]

Given population $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$, we let $E^k = \left(\mathcal{C}, \Omega, S^k, q^k\right)$ denote a finite economy of k randomly sampled students, where S^k is a set of k students drawn independently from η and the scaled capacities are $q^k = \lfloor qk \rfloor$.

Consider a mechanism that operates over many years, and each year faces a sampled economy E^k from the same population $\mathcal{E}$. Over the years, the mechanism can learn the population $\mathcal{E}$, while each year E^k is unknown until students report their preferences. We thus consider the following communication process which has oracle access to the market clearing cutoffs of the population $\mathcal{E}$.

Definition 11. *The Historical Cutoff Process $\mathcal{P}^H = (\theta_R, \sigma; \iota^H, a, m)$ is defined for k sampled economies E^k. The process knows the market-clearing cutoffs $\boldsymbol{P}^*$ of the population $\mathcal{E}$ through the oracle $\iota^H = (q, \boldsymbol{P}^*)$. The process posts $\boldsymbol{P}^*$ and assigns each student $\omega \in \Omega$ to $\mu(\omega) = D^\omega(\boldsymbol{P}^*)$.*

While historical data will only provide an approximation to market-clearing cutoffs, the following proposition shows that the noisy estimates are sufficient if it is possible to perturb capacities,[28] allowing $\mathcal{P}^H$ to produce an approximately regret-free stable outcome.

Proposition 5. *Given any population $\mathcal{E}$ and $\varepsilon > 0$ there exists K such that for any $k > K$ the outcome of the process $\mathcal{P}^H$ on economy E^k sampled from $\mathcal{E}$ is regret-free stable for perturbed college capacities $\hat{q}^k$ where $\mathbb{P}\left(\|\hat{q}^k - q^k\|_1 > k \cdot \varepsilon\right) < \varepsilon$.*

The Historical Cutoff Process engages in price discovery over many years, rather than fully discovering market-clearing cutoffs in a single year. If the mechanism does not have access to the population cutoffs, similar arguments can be used to show that posting the market-clearing cutoffs from a previous year (i.e. another randomly sampled economy $\tilde{E}^k$ from the same underlying distribution) will result in an outcome that is regret-free stable with respect to slightly more perturbed capacities (see Appendix D for further details.). Also note that this approach does not require any additional knowledge beyond historical market-clearing cutoffs, and its second stage can be a very simple assignment mechanism.

[27]That is, we generally have that $\eta^S\left(\{\omega = (\theta, v) \,|\, \theta \in A, v \in V\}\right) \neq \int_{\theta \in A} F^\theta(V)\, d\nu(\theta)$, for any x we ask that $\eta^S\left(\{\omega \in \Omega_0 |\, r_i^\omega = x\}\right) \in \{0, 1\}$, and the total mass of students is $\eta^S(\Omega_0) = k$. In addition, feasible outcomes must be integral (i.e., assign the same information and college to each atom of students).

[28]Our survey of admission systems finds that many universities do use flexible capacities (see Section 6). However, Che & Koh (2016) argue that such over-capacity risk is costly for Korean universities, and comment that such risk could be eliminated by a sequential centralized mechanism.

5.2 Price discovery by sampling students

When historical information is not available, the mechanism must engage students in price discovery. Fortunately, the cutoff structure pinpoints the information that the mechanism needs to obtain. In particular, it is sufficient for the mechanism to learn aggregate demand, which can be readily learned from a subsample of students. Thus, a natural approach is for the mechanism to survey a subset of students in a first stage, obtain an estimate of demand, and calculate the cutoffs. In the second stage, the mechanism posts the cutoffs, and deals with approximation errors with perturbed capacities, as in Section 5.1.

We consider two versions of this natural sampling approach. The first, random sampling, can be applied to general economies, but requires surveyed students to inefficiently acquire information. The second targets specific students to minimize price discovery costs, but requires structural assumptions.

5.2.1 Random sub-sampling

A natural approach to price discovery is to estimate demand by surveying a random sample of students.[29] For any cutoffs $\boldsymbol{P}$, we can identify the demand $D(\boldsymbol{P})$ from samples of student demand $D^{\omega_0}(\boldsymbol{P})$, $\omega_0 \sim \eta$. However, note that a sampled student ω_0 needs to acquire the signals $\chi^*(\omega_0, B^{\omega_0}(\boldsymbol{P}))$ and report what would be her most preferred college if she only acquired these signals.

To illustrate why a naive approach that asks students to report their preferences (given their current information) will fail to correctly identify demand, consider sampling a population of students θ_0 with preferences as defined in Example 1. Suppose that we are interested in estimating $D^{\theta_0}(\boldsymbol{P})$ such that the student's budget set is $B^{\theta_0}(\boldsymbol{P}) = \{2, 3\}$. If the student is in state θ_0 with $\chi^{\theta_0} = \phi$, (i.e., she did not collect any signals) and determines her preferences by comparing expected values given current information, she will report the preferences $2 \succ 3$ (as her perceived expected values are $\hat{v}_2^{\theta_0} = 3 > \hat{v}_3^{\theta_0} = 7/3$). If the student is asked to collect all signals, she will report that her top choice from $\{2, 3\}$ is $2, 3, \phi$ w.p. $1/3, 1/3, 1/3$, respectively. Both fail to give the demand of the student, because if the student is asked to acquire information optimally given the budget set $B = \{2, 3\}$, she will report that her top choice from $\{2, 3\}$ is $2, 3, \phi$ w.p. $1/2, 1/6, 1/3$ respectively.

The mechanism can identify the entire demand function $D(\cdot)$ by randomly sampling

[29] A line of literature within optimal auction design leverages samples from valuation profiles, for example using samples to set optimal reserves (Cole & Roughgarden 2014, Elkind 2007, Medina & Mohri 2014, Morgenstern & Roughgarden 2015).

students from Ω_0 according to the measure η, and asking each sampled student to report their demand $D^{\omega_0}(B)$ for each possible budget set $B \subset \mathcal{C}$. This requires each sampled student ω_0 to acquire the set of signals $\cup_B \chi^*(\omega_0, B)$ and incur the costs associated with doing so.

Definition 12. *The* Random Sampling *communication process* $\mathcal{P}^{RS,\ell}$ *has two stages. In the first stage, the mechanism estimates demand by randomly sampling a set* S *of* $\ell = |S|$ *students and asking them to report their demand from each possible budget set. In the second stage, the mechanism publishes a market-clearing cutoff* $\hat{\boldsymbol{P}}$ *for the estimated demand, students* $\omega_0 \in \Omega_0 \setminus S$ *optimally acquire information given* $\hat{\boldsymbol{P}}$ *and are assigned to* $\mu(\omega_0) = D^{\omega_0}(\hat{\boldsymbol{P}})$.

We show this random sampling mechanism yields an outcome that is approximately regret-free in the sense that only sampled students S have not acquired information optimally. As in Section 5.1, we can absorb the noise due to sampling error by perturbing college capacities. We leave further study of the estimation process for future work.

Corollary 4. *For any economy* $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ *and* $\varepsilon > 0$ *there exists* K *such that for any* $\ell > K$ *the communication process* $\mathcal{P}^{RS,\ell}$ *on economy* $\mathcal{E}$ *produces an outcome that is stable with respect to capacities* $\hat{q}$*, where* $\mathbb{P}(\|\hat{q} - q\|_1 > \varepsilon) < \varepsilon$*. Moreover, all students in* $\Omega_0 \setminus S$ *have acquired information optimally and so have no regret.*

The random sampling communication process does not guarantee a regret-free stable outcome (as implied by Theorem 3) because it makes the sampled students bear the cost of price discovery.[30] The students sampled in the first stage provide a public good. By performing additional information acquisition, they facilitate price discovery, allowing the mechanism to learn the market-clearing cutoffs and reducing the information acquisition costs for all other students. However, sampled students would prefer others to bear this cost, giving rise to a situation akin to the Grossman & Stiglitz (1980) paradox, where no agent wants to be surveyed and bear the cost of price discovery.

5.2.2 Targeted sampling

Instead of randomly sampling students, the mechanism may try to collect preference information from students who can acquire information optimally without any input from the mechanism. A student θ with $r_i^\theta \geq 1 - q_i$ is guaranteed to have college i in her budget set,

[30]It also requires a perturbation to capacities that can be made arbitrary small with high probability by increasing the sample size.

regardless of the preferences of other students. Similarly, when all students find all colleges acceptable, student θ with $r_i^\theta < 1 - \sum_j q_j$ is guaranteed to not have college i in her budget set. Thus, a student θ such that $r_i^\theta \notin [1-\sum_j q_j, 1-q_i]$ for all $\in \mathcal{C}$ knows her budget set. Such students can be identified by their priorities. By asking these students to acquire information optimally in the first stage, the mechanism can access their demand as "free information."

Because such students are a selected sample, an additional assumption is necessary to allow the mechanism to identify demand from sampled students (as it is possible that demand of these students systematically differs from the demand of other student groups). As an illustration, we present a natural structural assumption that allows the mechanism to estimate demand and discover market-clearing cutoffs from free information.

Definition 13. *A Pandora's Box Multinomial Logit (MNL) economy $\mathcal{E}(\Gamma) = (\mathcal{C}, \Omega, \eta, q)$ is a Pandora's Box economy parametrized by $\Gamma = (\delta_1, \dots, \delta_n, c_1, \dots, c_n)$. For any student $\theta \in \Theta$, the cost of each signal is $c_i^\theta = c_i$ and the prior distribution F^θ over values v_i^ω is given by $v_i^\omega = \delta_i + \varepsilon$ (where δ_i is common to all students) and $\varepsilon \sim EV[0,1]$ is an i.i.d. extreme value draw (McFadden 1973). All students prefer all colleges over being unassigned.*

Note that the index $\underline{v}_i$ is a strictly monotonic function of c_i (Lemma 1), and therefore we can equivalently parameterize the economy by $\Gamma = (\delta_1, \dots, \delta_n, \underline{v}_1, \dots, \underline{v}_n)$. We say that the economy is *ordered* if colleges are labeled in order as their indexes, i.e., $\underline{v}_1 \geq \underline{v}_2 \geq \cdots \geq \underline{v}_n$.

Observing demand $D^{\theta_0}(\mathcal{C})$ of students with budget set $B = \mathcal{C}$ is insufficient for identifying all $2n$ parameters $\Gamma = (\delta_1, \dots, \delta_n, c_1, \dots, c_n)$. For instance, demand $D_i^{\theta_0}(\mathcal{C})$ may be low because the cost of inspection c_i is high, or because the expected value δ_i is low. We show that it is possible to identify all $2n$ parameters by observing demand given just two budget sets.

Proposition 6. *Let $\mathcal{E}(\Gamma)$ be an ordered Pandora's Box MNL economy. Then Γ is identified from $D^{\theta_0}(\mathcal{C})$ and $D^{\theta_0}(\mathcal{C} \setminus \{n\})$.*

The proof of Proposition 6 is constructive, and relies on closed-form expressions for demand in a Pandora's Box MNL economy, which may be of independent interest (see Appendix F).

We say that an economy has free information if both sets of students $\{\theta_0 \in \Theta_0 \mid r_i^{\theta_0} > 1 - q_i \; \forall i\}$ and $\{\theta_0 \in \Theta_0 \mid r_i^{\theta_0} > 1 - q_i \; \forall i \neq n, r_n^{\theta_0} < 1 - \sum_j q_j\}$ have positive mass. We abstract away from estimation error, and we assume that the mechanism can learn $D^{\theta_0}(\mathcal{C})$ and $D^{\theta_0}(\mathcal{C} \setminus \{n\})$ by asking students in the sets above to optimally acquire information given their known budget sets and report their resulting demand. The following communication process leverages the identification result of Proposition 6 to discover the market-clearing cutoffs from free information

Definition 14. *The* MNL targeted sampling *communication process $\mathcal{P}^{MNL-TS}$ is defined for ordered MNL economies with free information. In the first stage, the process identifies a set of students S with free information; it asks these students to acquire information according to their optimal information acquisition strategy and report back their demanded college. From Proposition 6, learning the demand of these students allows the process to identify Γ, and thus learn the demand distribution, and its market-clearing cutoff $\boldsymbol{P}$. In the second stage, the mechanism publishes $\boldsymbol{P}$, students $\omega_0 \in \Omega_0 \setminus S$ optimally acquire information, and all students $\omega_0 \in \Omega_0$ are assigned to $\mu(\omega_0) = D^{\omega_0}(\boldsymbol{P})$.*

Students in S have acquired information optimally, and, as in Section 5.2.1, students in $\Omega_0 \setminus S$ have acquired information optimally. As this process results in a precise determination of the market-clearing cutoffs, we conclude that the resulting mechanism is indeed regret-free stable.

Corollary 5. *The communication process $\mathcal{P}^{MNL-TS}$ produces a regret-free stable outcome for any ordered MNL economy with free information.*

As in Section 5.1, flexibility in the capacities can absorb errors in the estimation of Γ. We leave a more thorough econometric analysis of the estimation error in $\hat{\Gamma}$ in finite economies (and the corresponding required flexibility in capacities) to future work.

6 Survey of Admission Systems

Our theoretical results suggest admission systems ought to provide students with information to guide their information acquisition processes. To understand the prevalence of such guidance in practice, we surveyed college admission systems in OECD countries with a population of at least 8M, as well as the two most populous nations in each continent.[31] Table 1 provides a summary of the survey, which can be found in the online appendix. We found that, in accordance with our theory, many university admissions systems provide applicants access to information about their admission chances. In countries where such information is not provided by the universities or a centralized clearinghouse, third parties, such as coaching

[31] Neilson (2020) and Grenet et al. (2019) provide complementary surveys of admission systems. The website www.matching-in-practice.eu provided us with a lot of valuable information about European matching systems.

institutes, provide students with predictions of their admission chances. This information is provided in systems using a range of mechanisms, from centralized DA mechanisms, to decentralized admission systems. Our findings highlight the necessity of providing information on admission chances to students in a well-functioning admission system.

In practice, two challenges arise when systems try to provide this information to applicants. The first challenge is that students need to know how each program will evaluate their application. In several countries, notably the USA, applications include essays and other materials that are evaluated subjectively. This can make it difficult to provide students with personalized predictions about their admission chances. A similar issue arises when entrance is based on exams with undisclosed results, such as national exams with grades that arrive after applications are submitted (e.g. UK), or when university-specific exams force students to choose between programs early in the application process (e.g. Japan). In contrast, in most countries, universities rank students using scores based on objective common criteria, such as performance in national exams taken by all college-going students.[32] While admission criteria may differ from program to program (for example, different programs may use different weighted averages of exam grades), universities commonly enable students to calculate their scores (e.g. by disclosing the weights used, or providing online score calculators). In some places, third parties, such as coaching institutes, supplement the information provided by universities.

The second challenge is that students need to know if their score is sufficiently high for admission to a program in the current year. In our model, this is equivalent to knowing the cutoffs for each college. For programs that admit all qualified applicants (e.g. most programs in Austria, Belgium, and the Netherlands), the cutoff is known to applicants (a qualification requirement which is specified in advance). For programs with limited capacity, admission decisions are based on the applying student's score in comparison to other applicants, so the cutoff depends on the preferences and relative performance of other applicants. Uncertainty about other applicants therefore implies that cutoffs are uncertain. Many systems (e.g. Australia, Chile, Israel, Poland, Turkey) make historical cutoff information publicly available, providing applicants with an estimate of the cutoff they will face. In some countries (e.g. Brazil, Spain), cutoff information is aggregated and distributed by third parties. We were able to collect a time series of cutoffs over the years in several countries; we found that cutoffs

[32] We note that even in the USA, the SAT score is a major admission criterion available to students prior to the application process, and, while applicants face considerable uncertainty about their admission chances, they also have a great deal of personalized information provided by high school guidance counselors regarding their chances at different colleges.

Table 1: Summary of University Admission Processes

	What information do students have about their admission chances?				**Admission process**		
Question	Are students evaluated based on objective common criteria (e.g., exams taken by all college-going students)?	Can students calculate their university- and program-specific scores?	Do universities provide historical cutoffs or estimates?	Is information about cutoffs or admission chances available from other sources?	What is the admission process?	Since when has the admission process been in place?	Are applications centrally collected?
Australia*	Yes - ATAR	Yes	Yes	N/A	Centralized one-shot	10+ years	Yes
Austria	Yes - high school diploma	Yes	Yes - admit all	N/A	Decentralized	10+ years	No
Belgium	Yes - high school diploma	Yes	Yes - admit all	N/A	Decentralized	10+ years	No
Brazil	Yes - ENEM	Yes - weights published	Yes - some	Yes - full	Centralized dynamic	2010	Yes
Canada*	High school grades; some programs have subjective criteria	No	Coarse indicators	Students intuit their chances	Decentralized	10+ years	
Chile	Yes - PSU	Yes - calculators	Yes	N/A	Centralized one-shot	10+ years	Yes
China*	Yes - Gāokǎo	Yes - weights published	Yes	N/A	Centralized one-shot	Algorithm changed in 2012	Yes
Czech Republic	National matriculation exam (Maturita); many universities have university-specific exams	University-specific exams	Yes	N/A	Decentralized	10+ years	Yes
Ethiopia	Yes - EUEE	Yes - weights published	Yes	N/A	Centralized	10+ years	Yes
France	Baccalaureate; as well as subjective criteria	No	Yes - some	Students intuit their chances	Centralized dynamic	2018	Yes
Germany*	Yes - Abitur; high school grades	Yes - weights published	Yes - most	Third parties supplement weights	Centralized dynamic	Algorithm changed in 2012	Yes
Hungary	Yes - Érettségi and high school criteria	Yes - calculators	N/A	Yes - by the central clearinghouse	Centralized one-shot	10+ years	Yes
India (engineering)	Yes within the IIT system (JEE-Mains + JEE-Advanced); NIT and BIT use university-specific criteria	Yes - self-score by third party coaching institutes	Yes	Coaching institutes provide guidance	Separate centralized systems	10+ years, with some changes	Each admissions system separately
Israel	Yes - Bagrut, PET	Yes - calculators	Yes	N/A	Decentralized	10+ years	No
Italy	High school diploma; specific majors require a test	Yes	Yes - admit all	N/A	Decentralized	10+ years	No
Japan	Center exam; many universities have university-specific exams	Yes - self-score by third party coaching institutes	Coarse indicators	Yes - by third party coaching institutes	Decentralized	10+ years	No
Mexico	School-specific exams; subjective criteria	University-specific exams	Coarse indicators	Yes - third parties	Decentralized	10+ years	No
Netherlands	Yes - high school diploma	Yes	Yes - admit all	N/A	Decentralized	10+ years	No
Nigeria	UTME and SSC, high school grades; competitive universities require university-specific exams	Yes - weights published	Yes	Coaching institutes provide guidance	Centralized dynamic	10+ years	Yes
Poland	Yes - Matura	Yes - weights published	Yes	N/A	Decentralized	10+ years	No
Portugal	Yes - Diploma de Ensino Secundário, Concurso	Yes	N/A	Yes - central clearinghouse provides application simulator	Centralized one-shot	10+ years	Yes
South Korea	CSAT; high school grades; subjective criteria	No	No	Third parties provide coarse indicators	Decentralized	10+ years	No
Spain	Yes - Bachillerato, PAU	Yes - weights published	No	Yes - third parties	Centralized for public, decentralized for private	10+ years	
Sweden	SweSAT, high school grades, university-specific criteria	Yes - self-score with answer keys, weights published	Yes	N/A	Centralized dynamic	10+ years	Yes
Switzerland	Yes - high school diploma	Yes	Yes - admit all	N/A	Decentralized	10+ years	No
Turkey	Yes - OSYS	Yes - weights published	Yes	N/A	Centralized one-shot	10+ years	Yes
UK	A-Levels or IB; subjective criteria	No	Coarse indicators	Students intuit their chances	Decentralized	10+ years	Yes
US	SATs; subjective criteria	No	Coarse indicators	Yes - third parties help identify safety and reach schools	Decentralized	10+ years	Yes

Summary of admission processes, focusing on the main features of the general university admission process for domestic applicants (excluding exceptions such as medical schools). See the online appendix for details. Countries marked with an asterisk have some variation across provinces, states or systems. The following terms were used: "Yes - weights published" – students are evaluated using a weighted average of their exam scores and these weights are publicly available to students; "Yes - calculators" – universities provide online websites where students can enter their exam scores and determine whether they meet the program's criteria; "Yes - admit all" – majority of programs admit all applying students; "Yes - full" – information that is not available from universities is made available by third party providers.

were generally stable from year to year, and historical cutoffs provide a relatively accurate prediction of admission chances. Such systems are very much in line with our suggested Historical Cutoff Process (Definition 11).

Two of the admission systems we surveyed are notable in how their implementation aligns with our theory. The Australian Universities Admissions Center (UAC) provides a centralized admission system for universities in each state in Australia. The UAC informs students of their exact percentile rank at each program.[33] Some programs, including most programs at the University of Sydney, additionally publish admission cutoffs and commit to admitting any applicant whose score is above the cutoff.[34] These cutoffs are published well in advance of student applications and so, when students choose programs to apply to, they know where they will be accepted (see Figure 1). The university website does not provide students with information about the assignment mechanism.

Under the commitment to accept all applying students that pass the cutoff, programs have limited control over the sizes of their incoming classes, and need the flexibility to adjust capacities. We see the willingness of the University of Sydney to bear the associated costs as further evidence of the importance of providing information about admission chances to students. In accordance with our theory, a sample of the time series of annual cutoffs suggests that the amount of flexibility required is not large, as cutoffs and enrollments are stable over time. This system is therefore almost an exact parallel to our suggested Historical Cutoff Process (Definition 11).

The decentralized Israeli university admission process provides another illustrative example. Each university assigns a score to each student that is a function of their test scores in standardized national exams, and universities provide formulas and online calculators that allow students to determine their program-specific scores and admission chances.[35] Students submit a separate application directly to each university. The admission process progresses dynamically, with each university sending admission offers to admitted students and updating information about admission chances online. The resulting assignment respects program capacity constraints. Despite being decentralized, this admission system has been operating successfully for many years without unravelling.[36] This is in contrast with the unraveling of

[33] https://www.uac.edu.au/future-applicants/atar/how-to-get-your-atar

[34] https://www.sydney.edu.au/study/how-to-apply/undergraduate/guaranteed-atar.html, https://www.sydney.edu.au/content/dam/corporate/documents/study/how-to-apply/domestic-admission-criteria.pdf

[35] See for example: https://in.bgu.ac.il/welcome/Pages/Rishum/what_are_my_chances.aspx, http://bagrut-calculator.huji.ac.il/, https://go.tau.ac.il/calc3 (retrieved April 2020).

[36] In private conversations, Israeli university admission officers disclosed that the admission system re-

similar decentralized matching markets that do not provide similar information to applicants.

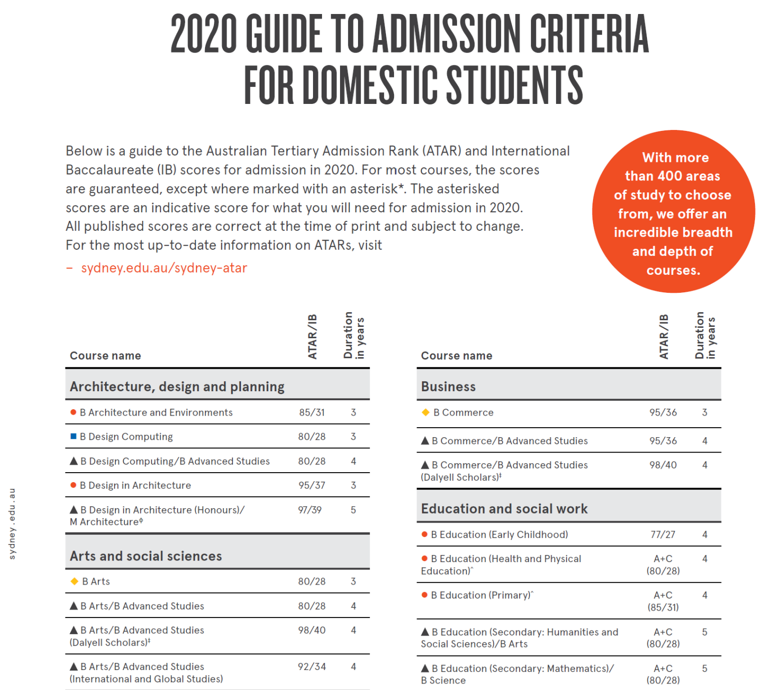
Overall, we find that most admission systems admit all qualified applicants or provide cutoffs that allow students to estimate their admission chances. Systems that do not do so generally have subjective evaluation, which makes it infeasible to disclose numerical cutoffs, or evaluate students based on exam scores that are unknown to students when they apply. We find anecdotal evidence that such admission systems are less well-functioning. For example, in the U.K. students do not know their exam scores when they apply, and the admission system includes second and third rounds for unmatched applicants. Another example is the U.S. admission system, in which evaluation criteria are subjective (e.g. essays), students can obtain only rough information about their admission chances, and early admission is prevalent.[37] The redesign of the French university centralized admission system provides another example in which lack of information about admission chances necessitated a redesign of a centralized system (see the online appendix to this paper).

7 Conclusion

A great deal of theoretical and empirical evidence suggests that stability is a crucial determinant of the success of a matching market (e.g. Roth 1984, 1991, Roth & Xing 1994). Motivated by this, a large body of literature has offered clearinghouses algorithmic solutions that ensure stability. These algorithms typically assume that students know and are able to report their full preferences, but do not account for the costs of acquiring this information. Our theory and survey show that, in many applications, providing students with information about their admission chances can help address both stability and information acquisition. For example, when the system has been running for many years, providing students with historical cutoff information can lead to a stable outcome while helping students acquire information efficiently. Moreover, our survey of college admissions systems indicates practitioners make substantial efforts to provide students with cutoffs and other information about their admission chances. Taken together, we hope that our work encourages further research by market designers on solutions that combine algorithmic design with information provision.

mained decentralized despite attempts by the government to centralize the process. This is because of the success of the current admission systems eliminated the need for a nationally centralized process.

[37] See for example http://marketdesigner.blogspot.com/2019/01/college-admissions-early-decision-stats.html



2020 GUIDE TO ADMISSION CRITERIA FOR DOMESTIC STUDENTS

Below is a guide to the Australian Tertiary Admission Rank (ATAR) and International Baccalaureate (IB) scores for admission in 2020. For most courses, the scores are guaranteed, except where marked with an asterisk*. The asterisked scores are an indicative score for what you will need for admission in 2020. All published scores are correct at the time of print and subject to change. For the most up-to-date information on ATARs, visit
– sydney.edu.au/sydney-atar

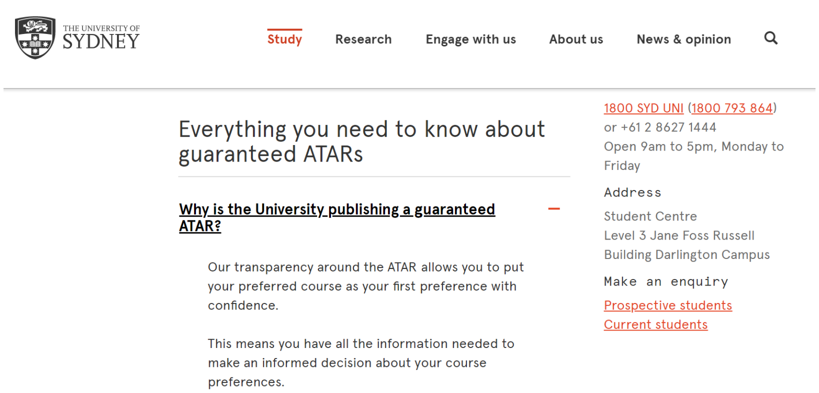
With more than 400 areas of study to choose from, we offer an incredible breadth and depth of courses.

sydney.edu.au

Course name	ATAR/IB	Duration in years
Architecture, design and planning		
● B Architecture and Environments	85/31	3
■ B Design Computing	80/28	3
▲ B Design Computing/B Advanced Studies	80/28	4
● B Design in Architecture	95/37	3
▲ B Design in Architecture (Honours)/ M Architecture[Φ]	97/39	5
Arts and social sciences		
◆ B Arts	80/28	3
▲ B Arts/B Advanced Studies	80/28	4
▲ B Arts/B Advanced Studies (Dalyell Scholars)[‡]	98/40	4
▲ B Arts/B Advanced Studies (International and Global Studies)	92/34	4

Course name	ATAR/IB	Duration in years
Business		
◆ B Commerce	95/36	3
▲ B Commerce/B Advanced Studies	95/36	4
▲ B Commerce/B Advanced Studies (Dalyell Scholars)[‡]	98/40	4
Education and social work		
● B Education (Early Childhood)	77/27	4
● B Education (Health and Physical Education)[^]	A+C (80/28)	4
● B Education (Primary)[^]	A+C (85/31)	4
▲ B Education (Secondary: Humanities and Social Sciences)/B Arts	A+C (80/28)	5
▲ B Education (Secondary: Mathematics)/ B Science	A+C (80/28)	5

(a) Screenshot from University of Sydney Domestic Admission Criteria 2020, available online at https://www.sydney.edu.au/content/dam/corporate/documents/study/how-to-apply/domestic-admission-criteria.pdf



THE UNIVERSITY OF SYDNEY

Study Research Engage with us About us News & opinion

Everything you need to know about guaranteed ATARs

Why is the University publishing a guaranteed ATAR? –

Our transparency around the ATAR allows you to put your preferred course as your first preference with confidence.

This means you have all the information needed to make an informed decision about your course preferences.

1800 SYD UNI (1800 793 864)
or +61 2 8627 1444
Open 9am to 5pm, Monday to Friday

Address

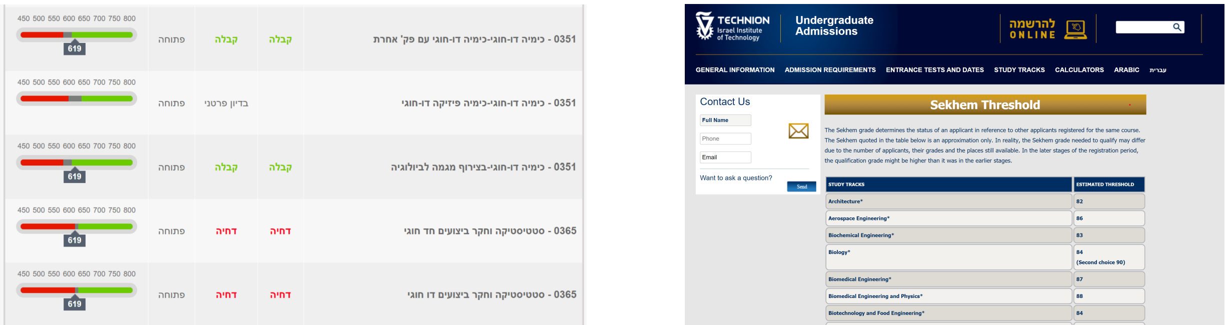
Student Centre
Level 3 Jane Foss Russell Building Darlington Campus

Make an enquiry

Prospective students
Current students

(b) Screenshot from University of Sydney's website providing information to prospective students, available online at https://www.sydney.edu.au/study/how-to-apply/undergraduate/guaranteed-atar.html

Figure 1: University of Sydney, Australia



(a) Screenshot from Tel Aviv University's admission website showing the student score (619) and the accept/reject thresholds for different programs. "Acceptance" is in green, "rejection" is in red. In the second row the student is neither accepted or rejected. The website is accessible online at https://go.tau.ac.il/b.a/calc.

(b) Screenshot from the Technion's admission website showing estimated admission thresholds. These estimates are available to students year round. The website is accessible online at https://admissions.technion.ac.il/en/english/general-admission-requirements/.

Figure 2: Tel Aviv University and Technion, Israel

References

Abdulkadiroğlu, A., Che, Y.-K. & Yasuda, Y. (2015), 'Expanding choice in school choice', *American Economic Journal: Microeconomics* **7**(1), 1–42.

Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J. & Walters, C. R. (2020), 'Do parents value school effectiveness?', *American Economic Review* **110**(5), 1502–39.

Andrabi, T., Das, J. & Khwaja, A. I. (2017), 'Report cards: The impact of providing school and child test scores on educational markets', *American Economic Review* **107**(6), 1535–1563.

Artemov, G., Che, Y.-K. & He, Y. (2017), Strategic 'mistakes': Implications for market design research, Technical report, mimeo.

Ashlagi, I., Braverman, M. & Hassidim, A. (2014), 'Stability in large matching markets with complementarities', *Operations Research* **62**, 713–732.

Ashlagi, I., Braverman, M., Kanoria, Y. & Shi, P. (2017), Communication requirements and informative signaling in matching markets, *in* 'EC '17', p. 263.

Ashlagi, I. & Gonczarowski, Y. A. (2018), 'Stable matching mechanisms are not obviously strategy-proof', *Journal of Economic Theory* **177**, 405–425.

Ausubel, L. M. & Baranov, O. V. (2014), 'Market design and the evolution of the combinatorial clock auction', *American Economic Review* **104**(5), 446–51.

Azevedo, E. M. & Leshno, J. D. (2016), 'A supply and demand framework for two-sided matching markets', *Journal of Political Economy* **124**(5), 1235–1268.

Aziz, H., Biró, P., Gaspers, S., de Haan, R., Mattei, N. & Rastegari, B. (2016), Stable matching with uncertain linear preferences, *in* 'International Symposium on Algorithmic Game Theory', Springer, pp. 195–206.

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

Beyhaghi, H. & Kleinberg, R. (2019), Pandora's problem with nonobligatory inspection, *in* 'Proceedings of the 2019 ACM Conference on Economics and Computation', pp. 131–132.

Bikhchandani, S. (2017), 'Stability with one-sided incomplete information', *Journal of Economic Theory* **168**, 372–399.

Blair, C. (1988), 'The lattice structure of the set of stable matchings with multiple partners', *Mathematics of Operations Research* **13**(4), 619–628.

Bo, I. & Hakimov, R. (2019), 'The iterative deferred acceptance mechanism', *Available at SSRN 2881880* .

Chakraborty, A., Citanna, A. & Ostrovsky, M. (2010), 'Two-sided matching with interdependent values', *Journal of Economic Theory* **145**(1), 85–105.

Che, Y.-K. & Koh, Y. (2016), 'Decentralized college admissions', *Journal of Political Economy* **124**(5), 1295–1338.

Chen, H. & He, Y. (2017), Information acquisition and provision in school choice: an experimental study, Technical report, Working paper.

Chen, H. & He, Y. (2020), Information acquisition and provision in school choice: A theoretical investigation, Technical report, Working paper.

Chen, L. & Pereyra, J. (2015), Time-constrained school choice, Technical report, ECARES Working Papers, 2016. 1, 1.1, 8.

Chen, Y.-C. & Hu, G. (2020), 'Learning by matching', *Theoretical Economics* **15**(1), 29–56.

Cole, R. & Roughgarden, T. (2014), The sample complexity of revenue maximization, *in* 'Proceedings of the forty-sixth annual ACM symposium on Theory of computing', pp. 243–252.

Coles, P., Cawley, J., Levine, P. B., Niederle, M., Roth, A. E. & Siegfried, J. J. (2010), 'The job market for new economists: A market design perspective', *Journal of Economic Perspectives* **24**(4), 187–206.

Coles, P., Kushnir, A. & Niederle, M. (2013), 'Preference signaling in matching markets', *American Economic Journal: Microeconomics* **5**(2), 99–134. NBER Working Paper No. 16185.

Corcoran, S. P., Jennings, J. L., Cohodes, S. R. & Sattin-Bajaj, C. (2018), Leveling the playing field for high school choice: Results from a field experiment of informational interventions, Technical report, National Bureau of Economic Research.

Doval, L. (2018), 'Whether or not to open pandora's box', *Journal of Economic Theory* **175**, 127–158.

Dur, U., Hammond, R. G. & Morrill, T. (2015), 'Identifying the harm of manipulable school-choice mechanisms', *American Economic Journal: Economic Policy* .

Dynarski, S., Libassi, C., Michelmore, K. & Owen, S. (2020), Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students, Technical report, National Bureau of Economic Research.

Ehlers, L. & Massó, J. (2015), 'Matching markets under (in) complete information', *Journal of Economic Theory* **157**, 295–314.

Elkind, E. (2007), Designing and learning optimal finite support auctions, *in* 'Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms', Society for Industrial and Applied Mathematics, pp. 736–745.

Erdil, A. & Ergin, H. (2008), 'What's the matter with tie-breaking? Improving efficiency in school choice', *American Economic Review* **98**(3), 669–689.

Fack, G., Grenet, J. & He, Y. (2019), 'Beyond truth-telling: Preference estimation with centralized school choice and college admissions', *American Economic Review* **109**(4), 1486–1529.

Gale, D. & Shapley, L. (1962), 'College admissions and the stability of marriage', *American Mathematical Monthly* **69**(1), 9–15.

Gonczarowski, Y. A., Nisan, N., Ostrovsky, R. & Rosenbaum, W. (2015), A stable marriage requires communication, *in* 'Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms', Society for Industrial and Applied Mathematics, pp. 1003–1017.

Gong, B. & Liang, Y. (2016), A dynamic college admission mechanism in Inner Mongolia: Theory and experiment, Technical report, Working paper, 2016. 2, 10.

Grenet, J., He, Y. & Kübler, D. (2019), Decentralizing centralized matching markets: Implications from early offers in university admissions, Technical report, WZB Discussion Paper.

Grossman, S. J. & Stiglitz, J. E. (1980), 'On the impossibility of informationally efficient markets', *American Economic Review* **70**(3), 393–408.

Harless, P. & Manjunath, V. (2015), 'The importance of learning in market design'.

Hastings, J. S. & Weinstein, J. M. (2008), 'Information, school choice, and academic achievement: Evidence from two experiments', *Quarterly Journal of Economics* **123**(4), 1373–1414.

Hoxby, C. M. & Avery, C. (2012), The missing" one-offs": The hidden supply of high-achieving, low income students, Technical report, National Bureau of Economic Research.

Hoxby, C. M. & Turner, S. (2015), 'What high-achieving low-income students know about college', *American Economic Review* **105**(5), 514–517.

Kanoria, Y. & Saban, D. (2020), 'Facilitating the search for partners on matching platforms', *Management Science (to appear)* .

Kapor, A. J., Neilson, C. A. & Zimmerman, S. D. (2020), 'Heterogeneous beliefs and school choice mechanisms', *American Economic Review* **110**(5), 1274–1315.

Kleinberg, R., Waggoner, B. & Weyl, E. G. (2016), 'Descending price optimally coordinates search', *arXiv preprint arXiv:1603.07682* .

Kloosterman, A. & Troyan, P. (2020), 'School choice with asymmetric information: Priority design and the curse of acceptance', *Theoretical Economics* **15**(3), 1095–1133.

Kojima, F., Pathak, P. A. & Roth, A. E. (2013), 'Matching with couples: Stability and incentives in large markets', *Quarterly Journal of Economics* **128**(4), 1585–1632.

Lee, S. & Niederle, M. (2015), 'Propose with a rose? signaling in internet dating markets', *Experimental Economics* **18**(4), 731–755.

Levin, J. & Skrzypacz, A. (2016), 'Properties of the combinatorial clock auction', *American Economic Review* **106**(9), 2528–51.

Li, S. (2017), 'Obviously strategy-proof mechanisms', *American Economic Review* **107**(11), 3257–87.

Liu, Q. (2020), 'Stability and bayesian consistency in two-sided markets', *American Economic Review* **110**(8), 2625–66.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20181186*

Liu, Q., Mailath, G. J., Postlewaite, A. & Samuelson, L. (2014), 'Stable matching with incomplete information', *Econometrica* **82**(2), 541–587.

Luflade, M. (2017), 'The value of information in centralized school choice systems'.

Matějka, F. & McKay, A. (2015), 'Rational inattention to discrete choices: A new foundation for the multinomial logit model', *American Economic Review* **105**(1), 272–98.

McFadden, D. L. (1973), 'Conditional logit analysis of qualitative choice behavior', *Frontiers in Econometrics* pp. 105–142.

Medina, A. M. & Mohri, M. (2014), Learning theory and algorithms for revenue optimization in second price auctions with reserve, *in* 'Proceedings of the 31st International Conference on Machine Learning (ICML-14)', pp. 262–270.

Menzel, K. (2015), 'Large matching markets as two-sided demand systems', *Econometrica* **83**(3), 897–941.

Morgenstern, J. H. & Roughgarden, T. (2015), On the pseudo-dimension of nearly optimal auctions, *in* 'Advances in Neural Information Processing Systems', pp. 136–144.

Narita, Y. (2018), 'Match or mismatch? learning and inertia in school choice', *Learning and Inertia in School Choice (June 18, 2018)* .

Neilson, C. A. (2020), 'The rise of centralized choice and assignment mechanisms in education markets around the world', *Working paper* .

Neilson, C., Allende, C. & Gallego, F. (2019), 'Approximating the equilibrium effects of informed school choice', *Working paper* .

Rastegari, B., Condon, A., Immorlica, N., Irving, R. & Leyton-Brown, K. (2014), Reasoning about optimal stable matchings under partial information, *in* 'Proceedings of the Fifteenth ACM Conference on Economics and Computation', ACM, pp. 431–448.

Rastegari, B., Condon, A., Immorlica, N. & Leyton-Brown, K. (2013), Two-sided matching with partial information, *in* 'Proceedings of the Fourteenth ACM Conference on Electronic Commerce', ACM, pp. 733–750.

Roth, A. E. (1984), 'The evolution of the labor market for medical interns and residents: a case study in game theory', *Journal of Political Economy* **92**(6), 991–1016.

Roth, A. E. (1991), 'A natural experiment in the organization of entry-level labor markets: regional markets for new physicians and surgeons in the united kingdom', *American Economic Review* pp. 415–440.

Roth, A. E. & Xing, X. (1994), 'Jumping the gun: Imperfections and institutions related to the timing of market transactions', *The American Economic Review* pp. 992–1044.

Roth, A. E. & Xing, X. (1997), 'Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists', *Journal of Political Economy* **105**(2), 284–329.

Segal, I. (2007), 'The communication requirements of social choice rules and supporting budget sets', *Journal of Economic Theory* **136**(1), 341–378.

Sims, C. A. (2003), 'Implications of rational inattention', *Journal of Monetary Economics* **50**(3), 665–690.

Steiner, J., Stewart, C. & Matějka, F. (2017), 'Rational inattention dynamics: Inertia and delay in decision-making', *Econometrica* **85**(2), 521–553.

Weitzman, M. L. (1979), 'Optimal search for the best alternative', *Econometrica* pp. 641–654.

A Informational Inefficiency of One-Shot Processes

We now demonstrate that standard communication processes can fail rather spectacularly in learning market-clearing cutoffs and alleviating the costs associated with information acquisition. Intuitively, in communication processes that maintain aggregate uncertainty, students need to know other students' choices in order to determine their optimal inspection strategy. Thus, in general, the student who performs the 'first' inspection will incur additional inspection costs.

Standard deferred acceptance processes, implemented as one-shot processes where students submit their full preference lists, perform especially poorly. This is because students are given almost no information about their choices before deciding on their inspection strategy. While in some settings regret can be eliminated by allowing for multi-round processes, we prove the stronger result that for general economies even multiple-round processes must either force some students to acquire information suboptimally, or create an information deadlock, where every student waits for others to acquire information first.

To demonstrate the issues in computing regret-free stable outcomes, consider an economy in which each student is willing to inspect any college as long as it is in their budget set. We may view such an economy as a setting where the costs affect *which* colleges students are willing to inspect, but not the *order* in which they are willing to inspect them. Examples of such economies include those in the Pandora's box model where $\mathbb{E}[v_i^\theta] = \infty$ for all $\theta \in \Theta$ that have not yet inspected i.

If each student is uncertain about aggregate demand, then the standard implementation of deferred acceptance as a one-shot process will not be regret-free even in this special case. This is because students' budget sets will depend on the preferences of other students, and so students who have low priority at the colleges they prefer are likely to acquire information suboptimally. We illustrate this in the following example.

Example 4. *Consider a continuum economy $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ with n colleges $\mathcal{C} = \{1, \ldots, n\}$ each with capacity $q_i = 1/n$. Suppose that college priorities are perfectly aligned, i.e. $r_i^\omega = r_j^\omega$ for all $\omega \in \Omega$, $i, j \in \mathcal{C}$.*

Let $\Omega^* \subset \Omega$ be the subset of students of measure $\eta(\Omega^*) = 1/n$ with highest rank. Let $\bar{\Omega} = \Omega \backslash \Omega^*$ be the remaining students.

The students in $\bar{\Omega}$ have identical priors and costs, which are also identical across colleges. For each $\omega \in \bar{\Omega}$ and $i \in \mathcal{C}$, the prior is F_i^ω is given by $F_i^\omega(x) = 0$ for all $x \in [0, 1)$, $F_i^\omega(x) = \frac{1}{4}$ for all $x \in [1, 2)$, $F_i^\omega(x) = 1 - \frac{1}{2^k}$ for all $k \geq 1$ and $x \in \left[2^k, 2^{k+1}\right)$ and the costs are $c_i^\omega = 1$.

Note that, under these priors and costs, the optimal inspection strategy for each student in $\bar{\Omega}$ given a budget set B is to inspect all colleges in B in an arbitrary order.

The students in Ω^* have identical and deterministic preferences, aligned on a particular college i^*. They each have value 1 for college i^* and value 0 for each college $j \neq i^*$ with probability 1, and cost 0 for inspecting any college. We think of i^* as being unknown to the communication process and to the students in $\bar{\Omega}$. One can formalize this by thinking of $\mathcal{E}$ as one of a family of n economies, one for each choice of i^*, and $\mathcal{E}$ is selected uniformly at random from this family.

In any one-shot process, a student ω will have no regret only if she chooses to examine precisely the set of all colleges whose capacities are not filled by higher-ranked students. This is because a student is willing to incur the cost to examine any college if and only if it is in her budget set. In particular, each student $\omega \in \bar{\Omega}$ has as their budget set some non-empty subset of $\mathcal{C}\backslash\{i^*\}$. Note that this might not be exactly $\mathcal{C}\backslash\{i^*\}$, depending on the inspection strategy used by students with higher rank than ω.

We therefore claim that, regardless of the choices of the other students, each $\omega \in \bar{\Omega}$ has probability at most $1/n$ of selecting their budget set precisely (over the uniform choice of i^* and the realization of other agents' values). To see this, note that the budget set of a student $\omega \in \bar{\omega}$ is precisely the subset of $\mathcal{C}$ whose capacities are not filled by students of higher rank. But, by symmetry with respect to colleges, from the perspective of ω this will be a uniformly random subset of $\mathcal{C}$ of a (possibly random) size $k \in \{1, 2, \ldots, n-1\}$. The number of such subsets is minimized when $k = 1$ or $k = n - 1$, in which case the number of possible budget sets that the student must select from is n.

We conclude that a mass of students of measure at least $\eta(\bar{\Omega}) \cdot (1 - 1/n) = (1 - 1/n)^2$ will regret their inspections, in expectation over the choice of i^* and any randomness in the communication process and the inspection strategies of the students. There must therefore exist some choice of i^* for which this measure of students experiences regret. The example can also be modified so that this fraction of students experiences unbounded regret.[38]

This example demonstrates that in settings with aggregate uncertainty one-shot processes cannot hope to find regret-free stable outcomes, even in settings where students are willing to incur the costs of searching any number of colleges. This is due to their inability to coordinate the search of worse-ranked students with the preferences of the better-ranked students.

[38]For each bound K the example can be modified so that each student who inspects suboptimally incurs unnecessary information costs at least K times their utility.

B Proofs from Section 3

B.1 Proof of Proposition 1

The proof follows a similar proof in Kleinberg et al. (2016). To simplify notation we write ω instead of ω_0. Define $i \succ^\omega j$ if and only if $\min\left\{\underline{v}_i^\omega, v_i^\omega\right\} > \min\left\{\underline{v}_j^\omega, v_j^\omega\right\}$.[39]

We verify that $D^\omega(B) = \max_{\succ^\omega}(B)$ for all $B \subseteq \mathcal{C}$. Let $i, j \in B$ be such that $\min\left\{\underline{v}_i^\omega, v_i^\omega\right\}$ is greater than $\min\left\{\underline{v}_j^\omega, v_j^\omega\right\}$. Suppose for the sake of contradiction that student ω demands college j, i.e. $D^\omega(B) = j$. Following the optimal inspection policy in Lemma 1, college j must be inspected and v_j^ω must be the maximal realized value and so $v_j^\omega \geq v_k^\omega$ for all inspected k.

If $v_j^\omega < \min\left\{\underline{v}_i^\omega, v_i^\omega\right\}$ then college i must be inspected since $\underline{v}_i^\omega > v_j^\omega$ which is the maximal realized value, and so college i must be demanded over college j since $v_i^\omega > v_j^\omega$, which is a contradiction.

If $\underline{v}_j^\omega < \min\left\{\underline{v}_i^\omega, v_i^\omega\right\}$ then college i must be inspected before college j, since college j has lower index, and so college i is inspected, as college j is inspected. But as $v_i^\omega > \underline{v}_j^\omega$ the student stops inspecting before college j, which contradicts that j is inspected.

B.2 Proof of Proposition 2

We provide an alternative proof of Proposition 2 by showing that any economy $\mathcal{E}$ in the Pandora's box domain satisfies WARP. We do this by showing that each individual student demand satisfies WARP.

Lemma 4. *Let ω be a realized student with Pandora demand. Then $D^\omega(\cdot)$ satisfies WARP, i.e. if $B \subseteq B'$ and $i \in B \setminus \{D^\omega(B)\}$ then $D^\omega(B') \neq i$.*

Proof. The proof follows from the characterization that $D^\omega(B) = \operatorname{argmin}_{j\in B} \min\{\underline{v}_j^\omega, v_j\}$ (see e.g. Proposition 1). Suppose for the sake of contradiction that $D^\omega(B') = i$. Let $j := D^\omega(B')$, then $\min\{\underline{v}_j^\omega, v_j\} > \min\{\underline{v}_i^\omega, v_i\}$ for $j \in B'$ which contradicts that $\operatorname{argmin}_{k\in B'} \min\{\underline{v}_k^\omega, v_k\} = D^\omega(B') = i$. $\square$

Proof of Proposition 2. The result follows from Theorem 2 and Lemma 4. $\square$

[39] If $\min\left\{\underline{v}_i^\omega, v_i^\omega\right\} = \min\left\{\underline{v}_j^\omega, v_j^\omega\right\}$ set $i \succ^\omega j$ if $i < j$.

B.3 Proof of Proposition 3

The following lemma is used to prove Proposition 3. It also allows us to provide a utility-based characterization of the partial order in the regret-free stable outcome lattice for Pandora's box economies.

Lemma 5 (Kleinberg et al. (2016))**.** *Let $\mathcal{E}$ be a Pandora's box economy, and let (μ, χ) be a regret-free stable outcome. Then expected utility of $\theta \in \Theta_0$ under (μ, χ) is*

$$\mathbb{E}_{\omega\sim F^\theta}\left[v^\omega_{\mu(\omega)} - c^\omega\left(\chi\left(\omega\right)\right)\right] = \mathbb{E}_{\omega\sim F^\theta}\left[\min\left\{v^\omega_{\mu(\omega)}, \underline{v}^\omega_{\mu(\omega)}\right\}\right].$$

Proof of Lemma 5. We reproduce the proof here using our notation.

$$\begin{aligned}
\mathbb{E}_{\omega\sim F^\theta}\left[v^\omega_{\mu(\omega)} - c^\omega\left(\chi\left(\omega\right)\right)\right] &= \mathbb{E}_{\omega\sim F^\theta}\left[v^\omega_{\mu(\omega)} - \sum_{i\in\chi(\omega)} c^\omega_i\right] \text{ (cost function for Pandora's Box)} \\
&= \mathbb{E}_{\omega\sim F^\theta}\left[\max_{i\in\chi(\omega)} v^\omega_i - \sum_{i\in\chi(\omega)} c^\omega_i\right] \text{ (since } (\mu,\chi) \text{ is regret-free stable)} \\
&= \mathbb{E}_{\omega\sim F^\theta}\left[\max_{i\in\chi(\omega)} v^\omega_i - \sum_{i\in\chi(\omega)} \mathbb{E}_{v_i\sim F^\theta_i}\left[\left(v_i - \underline{v}^\theta_i\right)^+\right]\right] \text{ (by definition of } \underline{v}^\theta_i\text{)} \\
&= \mathbb{E}_{\omega\sim F^\theta}\left[\max_{i\in\chi(\omega)} v^\omega_i - \sum_{i\in\chi(\omega)} \left(v^\omega_i - \underline{v}^\theta_i\right)^+\right] \text{ (} i\in\chi(\omega) \text{ independent of } v^\omega_i\text{)}
\end{aligned}$$

To further simplify the last expression, we show that $\mathbb{E}\left[\left(v^\omega_i - \underline{v}^\theta_i\right)^+ \mid i\in\chi\left(\omega\right), i\neq\mu\left(\omega\right)\right]$ is equal to 0. To show the contrary, consider ω, i such that $i\neq\mu\left(\omega\right)$, $v^\omega_i > \underline{v}^\omega_i$, and $i\in\chi\left(\omega\right)$. Under the optimal adaptive inspection, student ω will not inspect any other colleges after inspecting i. Since $i\in\chi\left(\omega\right)$, college i is the last college inspected by student ω. This implies for any college $i\neq j\in\chi\left(\omega\right)$ was inspected before i, and thus $v^\omega_j \leq \underline{v}^\omega_i < v^\omega_i$. Thus, $v^\omega_i = \max_{j\in\chi(\omega)} v^\omega_j$, providing a contradiction to the stability of (μ, χ). Thus we can simplify

$$\begin{aligned}
\mathbb{E}_{\omega\sim F^\theta}\left[v^\omega_{\mu(\omega)} - c^\omega\left(\chi\left(\omega\right)\right)\right] &= \mathbb{E}_{\omega\sim F^\theta}\left[\max_{i\in\chi(\omega)} v^\omega_i - \sum_{i\in\chi(\omega)} \left(v^\omega_i - \underline{v}^\theta_i\right)^+\right] \\
&= \mathbb{E}_{\omega\sim F^\theta}\left[v^\omega_{\mu(\omega)} - \left(v^\omega_{\mu(\omega)} - \underline{v}^\theta_{\mu(\omega)}\right)^+\right] - \mathbb{E}_{\omega\sim F^\theta}\left[\sum_{i\in\chi(\omega)\setminus\{\mu(\omega)\}} \left(v^\omega_i - \underline{v}^\theta_i\right)^+\right] \\
&= \mathbb{E}_{\omega\sim F^\theta}\left[v^\omega_{\mu(\omega)} - \left(v^\omega_{\mu(\omega)} - \underline{v}^\theta_{\mu(\omega)}\right)^+\right] \\
&= \mathbb{E}_{\omega\sim F^\theta}\left[\min\left\{v^\omega_{\mu(\omega)}, \underline{v}^\omega_{\mu(\omega)}\right\}\right]
\end{aligned}$$

The result follows. □

Proof of Proposition 3. It suffices to show that if two regret-free stable outcomes (μ, χ) and (μ', χ') satisfy $(\mu, \chi) \rhd (\mu', \chi')$ then

$$\mathbb{E}_{\omega \sim F^\theta}\left[v^\omega_{\mu(\omega)} - c^\omega(\chi(\omega))\right] \geq \mathbb{E}_{\omega \sim F^\theta}\left[v^\omega_{\mu'(\omega)} - c^\omega(\chi'(\omega))\right].$$

By Theorem 2 there exist market-clearing cutoffs $\boldsymbol{P} \leq \boldsymbol{P}'$ which correspond to (μ, χ) and (μ', χ') respectively, by Theorem 1 it holds that $\mu(\omega) = D^\omega(B^\omega(\boldsymbol{P}))$, and finally by Lemma 5 it holds that $\mathbb{E}_{\omega \sim F^\theta}\left[v^\omega_{\mu(\omega)} - c^\omega\left(\chi\left(\omega\right)\right)\right] = \mathbb{E}_{\omega \sim F^\theta}\left[\min\left\{v^\omega_{\mu(\omega)}, \underline{v}^\omega_{\mu(\omega)}\right\}\right]$. Hence

$$\begin{aligned}
\mathbb{E}_{\omega \sim F^\theta}\left[v^\omega_{\mu(\omega)} - c^\omega\left(\chi\left(\omega\right)\right)\right] &= \mathbb{E}_{\omega \sim F^\theta}\left[\min\left\{v^\omega_{\mu(\omega)}, \underline{v}^\omega_{\mu(\omega)}\right\}\right] \\
&= \mathbb{E}_{\omega \sim F^\theta}\left[\min\left\{v^\omega_{D^\omega(B^\omega(\boldsymbol{P}))}, \underline{v}^\omega_{D^\omega(B^\omega(\boldsymbol{P}))}\right\}\right] \\
&\geq \mathbb{E}_{\omega \sim F^\theta}\left[\min\left\{v^\omega_{D^\omega(B^\omega(\boldsymbol{P}'))}, \underline{v}^\omega_{D^\omega(B^\omega(\boldsymbol{P}'))}\right\}\right] \\
&= \mathbb{E}_{\omega \sim F^\theta}\left[v^\omega_{\mu'(\omega)} - c^\omega\left(\chi'\left(\omega\right)\right)\right].
\end{aligned}$$

where the inequality follows from the fact that $B^\omega(\boldsymbol{P}) \geq B^\omega(\boldsymbol{P}')$ (since $\boldsymbol{P} \leq \boldsymbol{P}'$), and that $D^\omega(B) = \text{argmax}_{i \in B}(\min\{\underline{v}^\omega_i, v^\omega_i\})$ (by Proposition 1).

□

B.4 Proof of Theorem 2

By Theorem 1, it is sufficient to show that market-clearing cutoffs form a non-empty lattice. This can be proved either by showing that demand under $\mathcal{E}$ is identical to demand under some full-information economy $\tilde{\mathcal{E}}$, or by arguing directly based on the properties of aggregate demand. While the former proof is much more succinct, the latter proof may allow the interested reader to understand which results from the full-information setting can similarly be directly translated into the incomplete information setting, and so we present both proofs.

First, we show that if demand $D^\omega(\cdot)$ for a student $\omega \in \Omega$ satisfies WARP, then it is identical to demand for a student in a full-information economy, i.e. we show there exists an ordering $\succ^\omega$ such that $D^\omega(B) = \max_{\succ^\omega}(B)$ for all budget sets $B \subseteq \mathcal{C}$.

Define $\succ = \succ^\omega$ by $i \succ j \Leftrightarrow D^\omega\left(\{i, j\}\right) = i$. Since $D^\omega(\cdot)$ satisfies WARP $\succ$ is transitive: in particular if $D^\omega(\{i, j\}) = i$ and $D^\omega(\{j, k\}) = j$ then WARP implies that $D^\omega(\{i, j, k\}) \neq j, k$ and so $D^\omega(\{i, j, k\}) = i$ and so again by WARP $D^\omega(\{i, k\}) = i$. Hence for all $B \subseteq \mathcal{C}$, if we let $i = \max_{\succ}(B)$ then by definition of $\succ$ $D^{\omega}(\{i, j\}) = i$ for all $j \in B \setminus \{i\}$, and so by WARP $D^{\omega}(B) = i$.

The proof of the lattice structure of the set of market-clearing cutoffs also follows almost exactly from the proof of the analogous result in complete information settings in Azevedo & Leshno (2016) (see also the proof in Abdulkadiroğlu et al. (2015)). For completeness we replicate the proof here.

We remark that the proof relies on the fact that when individual student demand satisfies WARP, aggregate demand $D(\cdot) : [0, 1]^{\mathcal{C}} \to [0, 1]^{\mathcal{C}}$ satisfies the following *weak gross substitutes* condition: if $D_i(\boldsymbol{P})$ is decreasing in P_i and increasing in P_j for all $j \neq i$.

We now begin the replicated proof. Given P_{-i} define the interval

$$I_i(\boldsymbol{P}_{-i}) = \{p \in [0; 1] : D_i(p; \boldsymbol{P}_{-i}) \leq q_i, \text{ with equality if } p > 0\}.$$

That is, $I_i(P_{-i})$ is the set of cutoffs for college i that clear the market for i given the cutoffs of other colleges. Define the map $T(\boldsymbol{P})$ as

$$T_i(\boldsymbol{P}) = \text{argmin }_{p \in I_i(\boldsymbol{P}_{-i})} |p - P_i| .$$

That is, the map T has college i adjust its cutoff as little as possible to clear the market for i, taking the cutoffs of other colleges as given.

We show that the map T monotone non-decreasing (in the standard partial order of $[0, 1]^{\mathcal{C}}$), and the set of fixed points of T coincides with the set of market clearing cutoffs.

We first show that T is well defined. Note that, because $D_i(1, \boldsymbol{P}_{-i}) = 0$ and D_i is continuous, then either there exists $p \in [0, 1]$ such that $D_i(p; \boldsymbol{P}_{-i}) = q_i$ or $0 \in I_i(\boldsymbol{P}_{-i})$. In either case, we have that $I_i(\boldsymbol{P}_{-i})$ is nonempty, and also compact (by monotonicity and continuity of D_i).

We now show that T is monotone. To see this, consider $\boldsymbol{P} \leq \boldsymbol{P}', t_i = T_i(\boldsymbol{P})$, and $t'_i = T_i(\boldsymbol{P}')$. To reach a contradiction, assume that $t'_i < t_i$. In particular $t_i > 0$. We have that

$$q_i = D_i(t_i, \boldsymbol{P}_{-i}) \leq D_i(t'_i, \boldsymbol{P}_{-i}) \leq D_i(t'_i, \boldsymbol{P}'_{-i}) \leq q_i,$$

where the second inequality holds since aggregate student demand satisfies weak gross substitutes. Likewise

$$q_i = D_i(t_i, \boldsymbol{P}_{-i}) \leq D_i(t_i, \boldsymbol{P}'_{-i}) \leq D_i(t'_i, \boldsymbol{P}'_{-i}) \leq q_i,$$

from which it follows that $D_i(t_i', \boldsymbol{P}_{-i}) = D_i(t_i, \boldsymbol{P'}_{-i}) = q_i$. Hence

$$[t_i', t_i] \subseteq I_i(\boldsymbol{P}_{-i}) \cap I_i(\boldsymbol{P'}_{-i})\,.$$

The fact that the closest point to $\boldsymbol{P}_i$ in $I_i(\boldsymbol{P}_{-i})$ is t_i implies that $\boldsymbol{P}_i \geq t_i$. Therefore $\boldsymbol{P}_i' \geq t_i$, and so $|t_i - \boldsymbol{P}_i'| < |t_i' - \boldsymbol{P}_i'|$ which contradicts $t_i' = T_i(\boldsymbol{P}')$. This contradiction establishes that T is monotone.

Since T is a monotone operator, by Tarski's Theorem the set of fixed points of T is a lattice under the standard partial order of $[0,1]^{\mathcal{C}}$. It is easy to verify that the set of fixed points of T coincide with the market clearing cutoffs, and the ordering over regret-free stable outcomes follows from Theorem 1.

C Proofs from Section 4

C.1 Proof of Proposition 4

The proof uses cutoffs P_c^t to describe the tth period of the process. We note that in every period t the set of students who have been proposed to by a college c is given by for $\{\omega : r_c^\omega \geq P_c^t\}$ for some cutoff P_c^t. For the first period, we have that $P_c^1 = 1 - q_c$.

We show that students receive offers exactly from all colleges in their budget set by analyzing the cutoffs P_c^t and showing that they are monotonically decreasing in t. Suppose $\boldsymbol{P}^\tau$ is monotonically decreasing in τ for all $\tau \leq t$. Under the specified σ students collect signals from a college i as soon as they are proposed to by the college (i.e. in the first period τ where $i \in m_\tau$). Hence by the end of period t student ω has acquired the optimal information $\chi^*(\omega, B(\boldsymbol{P}^t))$, and therefore she rejects all colleges in $B(\boldsymbol{P}^t)$ except for $D^\omega\left(\boldsymbol{P}^t\right)$. WARP guarantees that the student has not rejected their favorite college in $B(\boldsymbol{P}^t)$ in some previous round $\tau < t$, and so the set of students who have ever rejected a college c is the same as the set of students who reject c given cutoffs $\boldsymbol{P}^t$.

Hence in period $t+1$ college c proposes to the top q_c students who have not yet rejected the college if and only if they set $P_c^{t+1} = P_c^t - q_c + D_c\left(\boldsymbol{P}^t\right)$ and make new proposals to all students in $\{\omega : P_c^{t+1} \leq r_c^\omega < P_c^t\}$. Since $D_c\left(P_c^1\right) \leq q_c$, it follows that $P_c^t \geq P_c^{t+1}$. By induction $\boldsymbol{P}^t$ is decreasing in t, and $D_c\left(\boldsymbol{P}^t\right) \leq q_c$ for all t.

Thus ICPDA terminates at some (possibly transfinite) round t^*, and $D_c\left(\boldsymbol{P}^{t^*}\right) = q_c$ for all c such that $P_c^{t^*} > 0$ and so $\boldsymbol{P}^* = \boldsymbol{P}^{t^*}$ are market-clearing cutoffs. The outcome (μ, χ) of the process $\mathcal{P}^{ICPDA}$ is given by $\mu(\omega) = D^{\omega_0}(B^\omega(\boldsymbol{P}^*))$, $\chi^\omega = B^\omega$ and is regret-free stable.

C.2 Proof of Lemma 2

Let $i \in B \setminus B'$, and consider an initial inspection type θ_0 such that college i has the largest index in set B, i.e., $\underline{v}_i^{\theta_0} > \underline{v}_j^{\theta_0}$ for all $j \in B \setminus \{i\}$. Let θ_0' be the type obtained from θ_0 if we were to add a constant α to the value $v_i^{\theta_0}$, i.e. $v_i^{\theta_0'} = v_i^{\theta_0} + \alpha$, where α is chosen so that $\mathbb{P}\left(v_i^{\theta_0} + \alpha > \max_{j \in B \setminus \{i\}} \underline{v}_j^{\theta_0}\right) > 0$. Then for a student with type θ_0', given budget set B with positive probability the student will inspect only i and then stop and demand i. Hence $\cap_{\omega_0 \in \theta_0'} \chi^*(\omega_0, B) = i$, and $i \notin B'$ so $i \notin \cap_{\omega_0 \in \theta_0} \chi^*(\omega_0, B')$, from which it follows that $\cap_{\omega_0 \in \theta_0} \left(\chi^*(\omega_0, B) \cap \chi^*(\omega_0, B')\right) = \emptyset$. Therefore θ_0' is stagnant given B, B'.

C.3 Proof of Lemma 3

Let $\theta \in \Theta_0$ be an initial student state that is stagnant given budget sets $B, B' \subseteq \mathcal{C}$. For notational convenience, denote that student's inspection costs c^θ by c and inspection indices $\underline{v}^\theta$ by $\underline{v}$. Fix $i \in B$ and $\varepsilon > 0$.

We construct a state $\theta' \in \Theta_0$ such that $\theta_R(\theta') = \theta_R(\theta)$, θ' is stagnant given B, B', and $D_i^{\theta'}(B) > 1 - \varepsilon$ (that is, a student of type θ' demands i from budget set B with probability of at least a $1 - \varepsilon$).

Since both θ and θ' are initial inspection states in the Pandora's box model, we have that $\chi^\theta = \chi^{\theta'} = \phi$ and $\Pi^\theta = \Pi^{\theta'} = \mathcal{C}$. Set the priorities and costs of θ' to be $r^{\theta'} = r^\theta$ and $c^{\theta'} = c^\theta$.

We set the priors of θ' such that the inspection indices of θ' are $\underline{v}^{\theta'} = \underline{v}$ and θ' demands i with probability of at least a $1 - \varepsilon$. To define the priors, fix $m < \min_j \underline{v}_j$ and define $\rho_j = c_j/(M - \underline{v}_j)$, where M is chosen to be sufficiently large so that $\rho_j \leq \varepsilon/n$ for all j. Let the prior distribution $F_i^{\theta'}$ of v_i be defined by $\mathbb{P}\left(v_i = m\right) = 1 - \rho_i$ and $\mathbb{P}\left(v_i = M\right) = \rho_i$. For each $j \neq i$, let the prior distribution $F_j^{\theta'}$ of v_j be defined by $\mathbb{P}\left(v_j = m/2\right) = 1 - \rho_j$ and $\mathbb{P}\left(v_j = M\right) = \rho_j$. For this choice of priors we have that $\underline{v}^{\theta'} = \underline{v}$ since, for any college j (including $j = i$) we have that $\mathbb{E}[(v_j - \underline{v}_j)^+] = \rho_j(M - \underline{v}_j) = c_j$. Moreover θ' demands college i from B with probability of at least $\mathbb{P}\left(v_j \leq m \,:\, \forall j \neq i\right) \geq (1 - \frac{\varepsilon}{n})^n \geq 1 - \varepsilon$.

Finally, we verify that θ' is stagnant given B, B'. For the sake of contradiction, suppose that student θ' always inspects college $j \in \mathcal{C}$ given either B or B', i.e., college j satisfies $j \in \cap_{\omega \in \theta'} \left(\chi^*\left(\omega, B\right) \cap \chi^*\left(\omega, B'\right)\right)$. By the characterization of the Pandora's box inspection policy in Lemma 1, it must be that for any $k \in B \cup B'$, either student θ' inspects college j before k (i.e., $\underline{v}_j \geq \underline{v}_k$), or she inspects college k first but will always subsequently choose to inspect j regardless of the realization of $v_k^{\theta'}$ (i.e., $\underline{v}_j \geq v_k^{\theta'}$). We therefore have that

$$\mathbb{P}\left(\underline{v}_j \geq \min\{\underline{v}_k, v_k^{\theta'}\} \ : \ \forall k \in B \cup B'\right) = 1. \tag{1}$$

Since θ is stagnant given B, B', it follows that

$$\text{argmax}_{k\in B}\underline{v}_k \cap \text{argmax}_{k\in B'}\underline{v}_k = \emptyset \ .$$

So in particular there exists $\ell \in B\cup B'$ such that $\underline{v}_\ell > \underline{v}_j$. Then (1) implies that $\mathbb{P}\left(v_\ell^{\theta'} < \underline{v}_\ell\right) = 1$. But by the definition of $F_\ell^{\theta'}$ we have that $\mathbb{P}\left(v_\ell^{\theta'} < \underline{v}_\ell\right) = 1 - \rho_\ell$, which implies that $\rho_\ell = \frac{c_\ell}{M - \underline{v}_\ell} = 0$, and therefore $c_\ell = 0$. This in turn implies that $\mathbb{P}\left(v_\ell^{\theta'} = m\right) = 1$ and $\underline{v}_\ell \leq m$, which contradicts the fact that $\mathbb{P}\left(v_\ell^{\theta'} < \underline{v}_\ell\right) = 1$.

C.4 Proof of Theorem 3

We provide initial information about an economy, and construct a set of 6 events that are indistinguishable when the reporting function maintains aggregate uncertainty.

Let colleges $\mathcal{C} = \{1, 2, 3\}$ have capacity $q_1 = q_2 = q_3 = 2\alpha$ for some $\alpha < \frac{1}{6}$. The student types are given by $\Omega = X \cup Y \cup Z \cup D \cup D'$. Students in X are stagnant given $\{2, 3\}, \{1, 2, 3\}$, students in Y are stagnant given $\{1, 3\}, \{1, 2, 3\}$, students in Z are stagnant given $\{1, 2\}, \{1, 2, 3\}$, and all students have Pandora demand, and so must inspect a college to attend it and do not wish to inspect colleges outside of their budget set (such students X, Y, Z exist by Lemma 2). D and D' are dummy students included to prevent trivially stable outcomes.

Priorities satisfy

$$\begin{aligned}
\text{priority at 1} &: r_1^y > r_1^z > r_1^d > r_1^x > r_1^{d'} \\
\text{priority at 2} &: r_2^z > r_2^x > r_1^d > r_2^y > r_2^{d'} \\
\text{priority at 3} &: r_3^x > r_3^y > r_1^d > r_3^z > r_3^{d'},
\end{aligned}$$

for all $(x, y, z, d, d') \in X \times Y \times Z \times D \times D'$, with single tie-breaking among students in the same group, i.e. $\forall x, x' \in X \ r_1^x > r_1^{x'} \Leftrightarrow r_2^x > r_2^{x'} \Leftrightarrow r_3^x > r_3^{x'}$, and similarly for $y, y' \in Y$ and $z, z' \in Z$.

Let η be a distribution of students that satisfies $\eta(X) = \eta(Y) = \eta(Z) = \alpha$, and $\eta(D) = 2\alpha$ and $\eta(D') = 1 - 5\alpha$. Note that this distribution implies that students in X

don't know if their budget set is $B = \{1, 2, 3\}$ or $B = \{2, 3\}$, as this depends on demand of students in $Y \cup Z \cup D$ (and symmetrically for Y and Z); and students in D don't know anything about their budget set so far, as there is a mass of 2α students in $X \cup Y \cup Z$ with higher priority at each college.

Given η and the communication process, let $\Delta = \frac{\alpha}{24}$, let X' denote the 2Δ students in X who inspect first, and equivalently define Y', Z'. Let D' be the 15Δ students in D who inspect first, and let X'', Y'', Z'' and D'' denote $X \setminus X', Y \setminus Y', Z \setminus Z'$ and $D \setminus D'$ respectively.

Consider the first point in time τ when either all students in X' have inspected, all students in Y' have inspected, all students in Z' have inspected, or all students in D' have inspected. If such a time does not exist, then some students in $X' \cup Y' \cup Z' \cup D'$ have not inspected and so the outcome is not regret-free stable.

Since the reporting function maintains aggregate uncertainty, it is unable to distinguish at time τ (i.e. after only observing the demand of a subset of students in $X' \cup Y' \cup Z' \cup D'$) between the following six events:

1. (a) Almost all students in $Y'' \cup Z'' \cup D''$ demand 1, i.e.

$$\eta\left(\omega \in Y'' \cup Z'' \cup D'' \mid D^{\omega}(\mathcal{C}) = 1\right) \geq (1-\varepsilon)\eta\left(Y'' \cup Z'' \cup D''\right);$$

   (b) Almost all students in $X'' \cup Z'' \cup D''$ demand 2;

   (c) Almost all students in $X'' \cup Y'' \cup D''$ demand 3;

2. (a) Almost all students in $X'' \cup Z''$ demand 2, i.e.

$$\eta\left(\omega \in X'' \cup Z'' \mid D^{\omega}(\mathcal{C}) = 2\right) \geq (1-\varepsilon)\eta\left(X'' \cup Z''\right).$$

   Almost all students in Y'' demand 3, i.e.

$$\eta\left(\omega \in Y'' \mid D^{\omega}(\mathcal{C}) = 3\right) \geq (1-\varepsilon)\eta\left(Y''\right).$$

   Students in D'' demand 3 if 3 is in their budget set, i.e.

$$\eta\left(\omega \in D'' \mid D^{\omega}(\mathcal{C}) = 3\right) \geq (1-\varepsilon)\eta\left(D''\right).$$

   (b) Almost all students in $X'' \cup Y''$ demand 3, almost all students in $Z'' \cup D''$ demand 1;

   (c) Almost all students in $Y'' \cup Z''$ demand 1, almost all students in $X'' \cup D''$ demand 2.

We show that any communication process with a reporting function that maintains aggregate uncertainty results in an outcome that is not regret-free stable.

1. **Case 1:** Δ students in X' inspect a college that is not in $\cap_{\omega_0 \in X} \chi^*(\omega_0, \{2,3\})$.
   Consider event 1a. If any student in X' has budget set $B = \{1,2,3\}$, then since students at Y, Z and D have higher priority at 1, all these students have 1 in their budget set. Since in event 1a it holds that a $(1-\varepsilon)$ fraction of students in $Y'' \cup Z'' \cup D''$ demand 1 from the full budget set $\mathcal{C}$, and since student demand satisfies the weak axiom of revealed preferences, it follows that at least $(1-\varepsilon)\eta\,(Y'' \cup Z'' \cup D'')$ students are assigned to college 1. But

   $$(1-\varepsilon)\eta\,(Y'' \cup Z'' \cup D'') = (1-\varepsilon)\,(4\alpha - \eta\,(Y' \cup Z' \cup D')) = (1-\varepsilon)\,(4\alpha - 19\Delta) > 2\alpha = q_1$$

   for sufficiently small ε, since $\Delta < \frac{2}{19}\alpha$. As this many students cannot be assigned to college 1, this shows that all students in X' have budget set $B = \{2,3\}$. It follows that a positive fraction of students in X' regret their inspection.
   Symmetrically, if Δ students in Y' inspect a college not definite for $B = \{1,3\}$, or if Δ students in Z' inspect a college not definite for $B = \{1,2\}$, then a positive fraction of students in Y' and Z' respectively regret their inspection under events 1b and 1c respectively.

2. **Case 2:** A set $\hat{X}$ of Δ students in X' inspect a college not in $\cap_{\omega_0 \in X} \chi^*(\omega_0, \{2,3\})$.
   Consider event 2a. If any student in $\hat{X}$ has budget set $B = \{2,3\}$, then the only students possibly with higher priority at 1 are Y, Z, D, X'' and the Δ students in $X' \setminus \hat{X}$. It follows that at least $q_1 - \Delta$ students in $Y \cup Z \cup D \cup X''$ demand college 1. We show that this cannot be the case.
   There are 4Δ students in $Y' \cup Z'$. Since all students in $X'' \cup Z''$ have 2 in their budget set, at least $(1-\varepsilon)\eta\,(X'' \cup Z'')$ students in $X'' \cup Z''$ demand 2, so at most $\varepsilon\eta\,(X'' \cup Z'')$ students in $X'' \cup Z''$ demand college 1. Since all students in Y'' have 3 in their budget set, at least $(1-\varepsilon)\eta\,(Y'')$ students in Y'' demand 3, so at most $\varepsilon\eta\,(Y'')$ students in Y'' demand college 1. Hence at most

   $$4\Delta + \varepsilon\eta\,(X'' \cup Z'') + \varepsilon\eta\,(Y'')$$

   students in $X'' \cup Y \cup Z$ demand college 1.
   We now consider the demand from students in D for college 1. We first show that

just under $\alpha = \frac{q_1}{2}$ students in D'' have 3 in their budget set. The only students higher ranked at college 3 than D'' are students in $X \cup Y \cup D'$. Of these, at least $(1-\varepsilon)\eta(X'')$ students in X'' demand 2, so at most

$$\eta(X') + \varepsilon\eta(X'') + \eta(Y) + \eta(D') = 2\Delta + \varepsilon\eta(X'') + \alpha + 15\Delta = \alpha + 17\Delta + \varepsilon\eta(X'')$$

students higher ranked than D'' demand 3, so at least $q_1 - (\alpha + 17\Delta + \varepsilon\eta(X'')) = \alpha - 17\Delta - \varepsilon\eta(X'')$ students in D'' have 3 in their budget set. Since at least $(1-\varepsilon)$ of these students demand college 3, at most $\eta(D) - (1-\varepsilon)(\alpha - 17\Delta - \varepsilon\eta(X'')) = \alpha + 17\Delta + \varepsilon((1-\varepsilon)\eta(X'') + \alpha - 17\Delta)$ students in D demand 1. Hence in total at most

$$\alpha + 21\Delta + \varepsilon\left(\eta(X'' \cup Z'') + \eta(Y'') + (1-\varepsilon)\eta(X'') + \alpha - 17\Delta\right)$$

students in $Y \cup Z \cup D \cup X''$ demand 1. Since $\Delta < \frac{\alpha}{22}$, this is less than $q_1 - \Delta = 2\alpha - \Delta$ for sufficiently small ε. Hence we have shown that none of the Δ students in $\hat{X}$ has budget set $B = \{2,3\}$, and so all of the students in $\hat{X}$ have budget set $B = \{1,2,3\}$ and a positive fraction of these students regret their inspection.
Symmetrically, if Δ students in Y' or Δ students in Z' inspect a college not definite for $B = \{1,2,3\}$, then a positive fraction of students in Y' and Z' respectively regret their inspection under events 2b and 2c respectively.

3. **Case 3:** 5Δ students in D' inspect a college i. Call this set $\hat{D}$, and assume without loss of generality that $i = 1$.
Consider event 1a. We show that almost Δ students in $\hat{D}$ do not have 1 in their budget set, and so regret inspecting 1. All students in $Y'' \cup Z''$ have higher priority at 1 than students in $\hat{D}$, and at least $(1-\varepsilon)\eta(Y'' \cup Z'')$ of these students are assigned to college 1. Hence at most

$$q_1 - (1-\varepsilon)\eta(Y'' \cup Z'') = 4\Delta + \varepsilon\eta(Y'' \cup Z'')$$

students in $\hat{D}$ have 1 in their budget set, so at least

$$\eta\left(\hat{D}\right) - 4\Delta + \varepsilon\eta(Y'' \cup Z'') = \Delta - \varepsilon\eta(Y'' \cup Z'')$$

students in $\hat{D}$ do not have 1 in their budget set.
Symmetrically, if 5Δ students in D' inspect college 2 or college 3 then a positive

proportion of students regret their inspection under events 1b and 1c respectively.

As one of cases 1, 2 or 3 must hold, it follows that a positive proportion of students regret their inspection.

D Additional results about the distribution of estimated market-clearing cutoffs

Proposition 7 (Distribution of approximately feasible capacities)**.** *Let $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ be a continuum economy and let $\boldsymbol{P}^* > 0$ be a corresponding market-clearing cutoff. Let $E^k = \left(\mathcal{C}, \Omega, S^k, q^k\right)$ be a finite economy of k randomly sampled students from $\mathcal{E}$. Let $\hat{q}^k = D\left(\boldsymbol{P}^* | \eta^k\right)$ be the capacities under which $\boldsymbol{P}^*$ is market-clearing for $\mathcal{E}^k$. Then*

$$\frac{1}{\sqrt{k}} \cdot \left(\hat{q}^k - q^k\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma\right),$$

where $\mathcal{N}\left(0, \Sigma\right)$ is the n-dimensional normal distribution with mean 0 and covariance matrix Σ given by $\Sigma_{ij} = -q_i q_j$ for $i \neq j$, $\Sigma_{ii} = q_i \left(1 - q_i\right)$.

Proof. The result follows from the central limit theorem by observing that $D_i\left(\boldsymbol{P}^* | \eta^k\right) = \sum_{\ell=1}^{k} X_i^\ell$, where $X_i^\ell = \mathbf{1}\left\{D^\theta\left(\boldsymbol{P}^*\right) = i\right\}$ for θ independently drawn according to η is a binary random variable with mean q_i. □

Proposition 7 implies that when demand is sampled from a known distribution, posting market-clearing cutoffs will result in an outcome that is regret-free stable with respect to perturbed capacities $\hat{q}^k$, where the required percent adjustment relative to the true capacities q^k is decreasing with the size of the market k.

Similar arguments can be used to show that posting the market-clearing cutoffs from a previous year (i.e. another randomly sampled economy $\tilde{E}^k$ from the same underlying distribution) will result in an outcome that is regret-free stable with respect to slightly more perturbed capacities. Specifically, the distribution of perturbed capacities will be $\frac{1}{\sqrt{k}}(\hat{q}^k - q) \xrightarrow{d} \mathcal{N}(0, 2\Sigma)$, where Σ is defined as in Proposition 7 .

E Budget Sets in Finite Economies

The following example shows that for finite economies our definition of a student ω's budget set $B^\omega\left(\mu\right)$ is not necessarily equivalent to the set of colleges the student ω can be admitted to.

Consider the following modification of an example due to Erdil & Ergin (2008). $E = (\mathcal{C} = \{1,2,3\}, S = \{x,y,z\}, q = \mathbf{1})$, where college priorities satisfy $r_1^y > r_1^z > r_1^x$ and $r_2^x > r_2^y > r_2^z$, student x prefers $1 \succ 2 \succ 3$ and has cost 0 at each college, student y prefers $2 \succ 1 \succ 3$ and has cost 0 at each college, and student z has information acquisition problem in the Pandora's box model (F^z, c^z, r^z) such that $\underline{v}_1^z > \underline{v}_2^z > \underline{v}_3^z$. There is a unique matching that corresponds to any stable outcome of $E \setminus z$, given by $\mu(x,y) = (1,2)$, so z's budget set arriving 'last to market' is $\{1,3\}$. If, given a budget set of $\{1,3\}$, z demands 3, then she is assigned to 3 and her budget set remains $\{1,3\}$. However, if z demands 1, this changes the unique stable matching in any stable outcome to be $\mu(x,y,z) = (2,1,3)$ and z now has budget set $\{3\}$.

While the two definitions of budget sets do not always coincide, previous theoretical results (Azevedo & Leshno 2016, Menzel 2015) show the situation illustrated by this example is unlikely to arise in randomly generated economies.

F Proofs from Section 5

F.1 Proof of Proposition 5

The proof relies on the following result, which is a restatement of Proposition 3 from Azevedo & Leshno (2016) and follows from the Vapnik-Chervonenkis Inequality.

Proposition 8 (Azevedo & Leshno (2016))**.** *Let $\mathcal{E} = (\mathcal{C}, \Omega, \eta, q)$ be an economy with market-clearing cutoffs $\boldsymbol{P}^*$. For any $\varepsilon > 0$, there exist constants $\alpha, \beta > 0$ such that for all k, if $\boldsymbol{P}^k$ are the market-clearing cutoffs of a finite economy E^k with k students randomly sampled from $\mathcal{E}$, then $|\boldsymbol{P}^* - \boldsymbol{P}^k|_1 > \varepsilon$ with probability at least $1 - \alpha k^{|\mathcal{C}|} \cdot e^{-\beta k}$.*

Let $\alpha, \beta > 0$ be constants given as in Proposition 8, and let K be such that $\alpha \cdot k^{|\mathcal{C}|} \cdot e^{-\beta k} < \varepsilon$ for all $k > K$. Then

$$\mathbb{P}\left(\|\hat{q}^k - q^k\|_1 > k \cdot \varepsilon\right) \leq \mathbb{P}\left(\|\boldsymbol{P}^k - \boldsymbol{P}^*\|_1 > \varepsilon\right) < \varepsilon$$

which completes the proof of Proposition 5.

F.2 Proof of Proposition 6

The proof of Proposition 6 is constructive, and relies on the following closed-form expression for demand in a Pandora's box MNL economy, which may be of independent interest.

Lemma 6. *Let $\mathcal{E}$ be a Pandora's box MNL economy, let $B \subseteq \mathcal{C}$ be an arbitrary budget set, and index the colleges so that $B = \{1, \ldots, m\}$ and $\underline{v}_1 \geq \underline{v}_2 \geq \cdots \geq \underline{v}_m$.*

Then for any type θ_0 demand is given by

$$D_i^{\theta_0}(B) = \left(1 - G\left(\underline{v}_i - \delta_i\right)\right) \prod_{j<i} G\left(\underline{v}_i - \delta_j\right) + \sum_{\ell=i}^{m} \frac{e^{\delta_i}}{\sum_{j \leq \ell} e^{\delta_j}} \left(\prod_{j \leq \ell} G\left(\underline{v}_\ell - \delta_j\right) - \prod_{j \leq \ell} G\left(\underline{v}_{\ell+1} - \delta_j\right) \right),$$

where $G(x) = e^{-e^{-x}}$ is the cdf of the extreme value distribution $EV[0, 1]$, and $\underline{v}_{m+1} = -\infty$.

The proof of Proposition 6 proceeds in three steps. In the first step the choice probabilities from $\mathcal{C} \backslash \{n\}$ and $\mathcal{C}$ are used to identify δ_i for all $i < n$. In the second step, given $\{\delta_1, \ldots, \delta_{n-1}\}$, the choice probabilities from $\mathcal{C} \setminus \{n\}$ are used to identify $\underline{v}_i$ for all $i < n$. In the third step, given $\{\delta_1, \ldots, \delta_{n-1}\}$ and $\{\underline{v}_1, \ldots, \underline{v}_{n-1}\}$, the demand for college n from $\mathcal{C}$ and some budget set $n \ni B \neq \{n\}, \mathcal{C}$ are used to identify δ_n and $\underline{v}_n$.

We now proceed with the proof. Without loss of generality, we can pick an arbitrary college $i \in \mathcal{C}$ (without loss, we choose $i = 1$) and normalize the parameters as follows: $\delta_i := \delta_i - \delta_1$ and $\underline{v}_i := \underline{v}_i - \delta_1$. The following procedure describes how we find the common value and index parameters for $i > 1$. There are only $2n-2$ parameters instead of $2n$ because the index parameter for the highest index college does not affect the choice probabilities and one of the common value terms is used as the normalization factor.

Let $u_i := e^{-e^{-\underline{v}_i}}$, $\alpha_i := e^{\delta_i}$ and $\gamma_i := \sum_{j \leq i} \alpha_j$. Note that since u_i and α_i are strictly monotonic transformations of the index $\underline{v}_i$ and common value δ_i parameters respectively, it is sufficient to identify these transformed parameters. Throughout, we will use the result in Lemma 6 that for a budget set $B = \{1, 2, \ldots, m\}$

$$D_i^{\theta_0}(B) = \left(1 - u_i^{\alpha_i}\right) u_i^{\gamma_{i-1}} + \sum_{\ell=i}^{m} \frac{\alpha_i}{\gamma_\ell} \left(u_\ell^{\gamma_\ell} - u_{\ell+1}^{\gamma_\ell}\right).$$

Step 1: We can pin down the α_i terms for $i < n$ by using the difference in choice probabilities $D_i^{\theta_0}\left(\mathcal{C} \backslash \{n\}\right) - D_i^{\theta_0}\left(\mathcal{C}\right)$. In particular, the difference in choice probabilities for college i is a constant multiple of α_i:

$$D_i^{\theta_0}(\mathcal{C}\backslash\{n\}) - D_i^{\theta_0}(\mathcal{C}) = \left(\frac{u_{n-1}^{\gamma_{n-1}}}{\gamma_{n-1}} - \frac{u_n^{\gamma_n}}{\gamma_n}\right) \cdot \alpha_i.$$

This implies that for all $i < n$

$$\alpha_i = \frac{D_i^{\theta_0}(\mathcal{C}\backslash\{n\}) - D_i^{\theta_0}(\mathcal{C})}{D_1^{\theta_0}(\mathcal{C}\backslash\{n\}) - D_1^{\theta_0}(\mathcal{C})}\alpha_1 = \frac{D_i^{\theta_0}(\mathcal{C}\backslash\{n\}) - D_i^{\theta_0}(\mathcal{C})}{D_1^{\theta_0}(\mathcal{C}\backslash\{n\}) - D_1^{\theta_0}(\mathcal{C})},$$

where the second equality holds since δ_1 is normalized to be 0 so $\alpha_1 = e^{\delta_1} = 1$.

Step 2: Using the identified parameters α_i for $i < n$, we can solve for the transformed index parameters u_i for $i < n$. The demand shares $D_i^{\theta_0}(\mathcal{C}\backslash\{n\}))$ define the following system of equations:

$$\begin{aligned}
D_1^{\theta_0}(\mathcal{C}\backslash\{n\}) &= (1 - u_i^{\alpha_i}) + \sum_{\ell=1}^{n-1} \frac{\alpha_1}{\gamma_\ell}\left(u_\ell^{\gamma_\ell} - u_{\ell+1}^{\gamma_\ell}\right) \\
D_2^{\theta_0}(\mathcal{C}\backslash\{n\}) &= (1 - u_2^{\alpha_2})\, u_1^{\alpha_1} + \sum_{\ell=2}^{n-1} \frac{\alpha_2}{\gamma_\ell}\left(u_\ell^{\gamma_\ell} - u_{\ell+1}^{\gamma_\ell}\right) \\
&\cdots \\
D_{n-2}^{\theta_0}(\mathcal{C}\backslash\{n\}) &= \left(1 - u_{n-2}^{\alpha_{n-2}}\right) u_{n-2}^{\gamma_{n-3}} + \sum_{\ell=n-2}^{n-1} \frac{\alpha_{n-2}}{\gamma_\ell}\left(u_\ell^{\gamma_\ell} - u_{\ell+1}^{\gamma_\ell}\right) \\
D_{n-1}^{\theta_0}(\mathcal{C}\backslash\{n\}) &= \left(1 - u_{n-1}^{\alpha_{n-1}}\right) u_{n-1}^{\gamma_{n-2}} + \frac{\alpha_{n-1}}{\gamma_{n-1}}\left(u_{n-1}^{\gamma_{n-1}}\right)
\end{aligned}$$

We can solve this system and find a unique set of solutions for u_i for $i < n$. This is because each choice probability $D_i^{\theta_0}(\mathcal{C}\backslash\{n\})$ contains only terms u_k for $k \geq i$ and the choice probability of college i is strictly increasing in u_i;[40] i.e. the Jacobian of this system (with respect to u_i for $i < n$) is an upper-triangular matrix with a strictly positive diagonal.

[40]Suppose we consider the demand shares among students with a budget set $B = \{1, ..., m\}$ ordered in decreasing order of their indices. The derivative of the ith demand share with respect to u_i is:

$$\begin{aligned}
\frac{d}{du_i}\left(D_i^{\theta_0}(B)\right) &= \frac{d}{du_i}\left((1 - u_i^{\alpha_i})\, u_i^{\gamma_{i-1}} + \sum_{\ell=i}^{n}\left(u_\ell^{\gamma_\ell} - u_{\ell+1}^{\gamma_\ell}\right)\frac{\alpha_i}{\gamma_\ell}\right) \\
&= \gamma_{i-1} u_i^{\gamma_{i-1}-1} - \gamma_i u_i^{\gamma_i - 1} + \alpha_i u_i^{\gamma_i - 1} \\
&= \gamma_{i-1} \cdot u_i^{\gamma_{i-1}-1} \cdot (1 - u_i^{\alpha_i}) > 0.
\end{aligned}$$

The last inequality results from the fact that $u_i = e^{-e^{-\underline{v}_i}}$, and so $0 < u_i < 1$.

Therefore there is a unique set of solutions which can be found by using the equation for $D_i^{\theta_0}(\mathcal{C} \setminus \{n\})$ to solve for u_i, in decreasing order of i.

Step 3: It remains to find the parameters for the college with the lowest index: α_n and u_n.

Now

$$\begin{aligned} D_n^{\theta_0}(\mathcal{C} \setminus \{n\}) &= (1 - u_n^{\alpha_n}) u_n^{\gamma} + \frac{\alpha_n}{\alpha_n + \gamma}\left(u_n^{\alpha_n+\gamma}\right) = u_n^{\gamma} - \frac{\gamma}{\alpha_n+\gamma} u_n^{\alpha_n+\gamma} = d \\ \frac{\gamma_{n-1}}{\alpha_i}\left(D_i^{\theta_0}(\mathcal{C}\setminus\{n\}) - D_i^{\theta_0}(\mathcal{C})\right) &= \gamma_{n-1}\left(\frac{u_{n-1}^{\gamma_{n-1}}}{\gamma_{n-1}} - \frac{u_n^{\gamma_n}}{\gamma_n}\right) = u_{n-1}^{\gamma} - \gamma\frac{u_n^{\alpha_n+\gamma}}{\alpha_n+\gamma} = d', \end{aligned}$$

for some d, d', where $\gamma = \sum_{j \in \mathcal{C}\setminus\{n\}} \alpha_j$. Note that since we solved for $\{\alpha_i, u_i\}_{i<n}$ in steps 1 and 2, we know d, d', u_{n-1}, and γ, and the only unknowns in these two equations are α_n and u_n.

Subtracting the two equations yields

$$u_n = (d - d')^{1/\gamma}.$$

Finally, α_n satisfies

$$\frac{\gamma}{\alpha_n + \gamma} u_n^{\alpha_n} = 1 - \frac{d}{u^{\gamma}}$$

and since $\frac{\partial}{\partial \alpha}\left(\frac{\gamma}{\alpha_n+\gamma} u_n^{\alpha_n}\right) = \gamma\left(\frac{(\alpha_n+\gamma)(\ln u_n) u_n^{\alpha_n} - u^{\alpha_n}}{(\alpha_n+\gamma)^2}\right) = \frac{\gamma u_n^{\alpha_n}}{(\alpha_n+\gamma)^2}\left((\alpha_n+\gamma)(\ln u_n) - 1\right) < 0$ (as $u_n = e^{-e^{-\underline{v}_n}} < 1$) it follows that there is a unique solution for α_n.

Proof of Lemma 6. Since the indices $\underline{v}_i$ are decreasing in i, a student will inspect a set of colleges $\{1, 2, \ldots, \ell\}$ for some ℓ. Let college i be such that student ω chooses $D^{\omega}(B) = i$. It must hold that $\ell \geq i$, and that $v_i = \max_{j \leq \ell} v_j$. If $\ell > i$ then it must be the case that $v_i \in \left(\underline{v}_{\ell+1}, \underline{v}_{\ell}\right]$. If $i = \ell$ it must be the case that $v_{\ell} = \max_{j \leq \ell} v_j > \underline{v}_{\ell+1}$, and so either $v_i = v_{\ell} \in \left(\underline{v}_{\ell+1}, \underline{v}_{\ell}\right]$ or $v_{\ell} > \underline{v}_{\ell}$.

Consider the case where $v_i \in \left(\underline{v}_{\ell+1}, \underline{v}_\ell\right]$ for $\ell \geq i$. This occurs with probability

$$
\begin{aligned}
\int_{\underline{v}_{\ell+1}}^{\underline{v}_\ell} g\left(x-\delta_i\right) \prod_{j \leq \ell, j \neq i} G\left(x-\delta_j\right) dx &= \int_{\underline{v}_{\ell+1}}^{\underline{v}_\ell} e^{-(x-\delta_i)} e^{-e^{-x} \sum_{j \leq \ell} e^{\delta_j}} dx \\
&= e^{\delta_i} \int_{\underline{v}_{\ell+1}}^{\underline{v}_\ell} e^{-e^{-x} \sum_{j \leq \ell} e^{\delta_j}} e^{-x} dx \\
&= e^{\delta_i} \int_{e^{-\underline{v}_\ell}}^{e^{-\underline{v}_{\ell+1}}} e^{-y \sum_{j \leq \ell} e^{\delta_j}} dy \\
&= \frac{e^{\delta_i}}{\sum_{j \leq \ell} e^{\delta_j}} \left(e^{-e^{-\underline{v}_\ell} \sum_{j \leq \ell} e^{\delta_j}} - e^{-e^{-\underline{v}_{\ell+1}} \sum_{j \leq \ell} e^{\delta_j}}\right) \\
&= \frac{e^{\delta_i}}{\sum_{j \leq \ell} e^{\delta_j}} \left(\prod_{j \leq \ell} G\left(\underline{v}_\ell - \delta_j\right) - \prod_{j \leq \ell} G\left(\underline{v}_{\ell+1} - \delta_j\right)\right),
\end{aligned}
$$

where $G(x) = e^{e^{-x}}$ and $g(x) = e^{-\left(x+e^{-x}\right)}$ are the cdf and pdf respectively of the extreme value distribution $EV[0,1]$.

Consider the case where $v_i = v_\ell > \underline{v}_\ell > v_j$ for all $j < \ell$. This occurs with probability

$$
\left(1 - G\left(\underline{v}_i - \delta_i\right)\right) \prod_{j<i} G\left(\underline{v}_i - \delta_j\right).
$$

Summing the two probabilities over all possible values of ℓ gives

$$
D_i^{\theta_0}(B) = \left(1 - G\left(\underline{v}_i - \delta_i\right)\right) \prod_{j<i} G\left(\underline{v}_i - \delta_j\right) + \sum_{\ell=i}^{m} \frac{e^{\delta_i}}{\sum_{j \leq \ell} e^{\delta_j}} \left(\prod_{j \leq \ell} G\left(\underline{v}_\ell - \delta_j\right) - \prod_{j \leq \ell} G\left(\underline{v}_{\ell+1} - \delta_j\right)\right).
$$

$\square$