http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
*Articles in Advance*, pp. 1–31
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

# Clearing Matching Markets Efficiently: Informative Signals and Match Recommendations

**Itai Ashlagi,[a] Mark Braverman,[b] Yash Kanoria,[c] Peng Shi[d]**

[a] Department of Management Science and Engineering, Stanford University, Stanford, California 94305; [b] Department of Computer Science, Princeton University, Princeton, New Jersey 08544; [c] Graduate Business School, Columbia University, New York, New York 10027; [d] Department of Data Sciences and Operations, Marshall School of Business, University of Southern California, Los Angeles, California 90007

**Contact:** iashlagi@stanford.edu, http://orcid.org/0000-0003-1890-3673 (IA); mbraverm@cs.princeton.edu (MB); ykanoria@columbia.edu, http://orcid.org/0000-0002-7221-357X (YK); pengshi@usc.edu, http://orcid.org/0000-0002-1057-795X (PS)

**Received:** March 24, 2018
**Revised:** August 20, 2018
**Accepted:** September 4, 2018
**Published Online in Articles in Advance:** August 7, 2019

https://doi.org/10.1287/mnsc.2018.3265



**Abstract.** We study how to reduce congestion in two-sided matching markets with private preferences. We measure congestion by the number of bits of information that agents must (i) learn about their own preferences, and (ii) communicate with others before obtaining their final match. Previous results suggest that a high level of congestion is inevitable under arbitrary preferences before the market can clear with a stable matching. We show that when the unobservable component of agent preferences satisfies certain natural assumptions, it is possible to recommend potential matches and encourage informative signals such that the market reaches a stable matching with a low level of congestion. Moreover, under our proposed approach, agents have negligible incentive to leave the marketplace or to look beyond the set of recommended partners. The intuitive idea is to only recommend partners with whom there is a nonnegligible chance that the agent will both like them and be liked by them. The recommendations are based on both the observable component of preferences and signals sent by agents on the other side that indicate interest.

**History:** Accepted by Yan Chen, decision analysis.

**Funding:** I. Ashlagi and Y. Kanoria acknowledge the research support of the National Science Foundation Division of Social and Economic Sciences [Grant SES 1254768] and Division of Civil, Mechanical and Manufacturing Innovation [Grant CMMI-1653477] and the Simons Institute for the Theory of Computing. M. Braverman acknowledges the research support of the National Science Foundation [Grant CCF-1525342], the Packard Fellowship in Science and Engineering, and the Simons Collaboration on Algorithms and Geometry. P. Shi acknowledges the support of Microsoft Research New England.

**Keywords:** marketplace and platform design • communication complexity • stable matching • match recommendations • informative signaling

## 1. Introduction

In many two-sided marketplaces, agents search to form matches with potential partners based on mutual compatibility. Examples include labor markets, college admission, online dating, and accommodation. Forming matches often requires extensive and costly communication between participants and their potential partners. For instance, on a large global freelancing platform, fewer than 10% of job applications get a response, and yet more than half of job openings remain unfilled, including when the employer invites individual freelancers (Horton 2019). We refer to a market's inability to reach a satisfactory outcome because of high communication overhead as *congestion*.[1] A major challenge for such marketplaces is to ease congestion without reducing the choice available to participants. In this paper, we explore, in a stylized setting, how matching markets can be efficiently cleared by encouraging informative signaling among users and providing good match recommendations to participants.

To study congestion in a rigorous framework, we adopt the classic notion of stability introduced by Gale and Shapley (1962) as our ideal of the market outcome without congestion. In a stable matching, no pair of agents would both prefer to match with each other over their assigned partners. This captures a key feature of such markets, which is that agents cannot simply choose their partners, but must also be chosen. Stability has been adopted as an equilibrium notion to capture real-world outcomes (Hitsch et al. 2010, Banerjee et al. 2013). Moreover, platforms that implement stable outcomes can prevent agents from looking for matches elsewhere, for instance, on other platforms.[2]

Many marketplaces use centralized clearinghouses that implement stable matchings to ease congestion, including the National Residency Matching Program

(NRMP) and many school admission systems. In these marketplaces, participants on both sides of the market submit preferences, which are then converted into a stable matching using the Gale–Shapley deferred acceptance (DA) algorithm. But even in these centralized platforms, participants may be unable to list their true preferences over all potential partners. For example, medical students participating in the NRMP are only permitted to rank hospitals with which they have interviewed, so the logistical costs of interviews limit the length of the ranking. Hence, having a centralized clearinghouse alone does not eliminate congestion.

Recent theoretical results suggest that it may be unrealistic to expect large markets to arrive at stable matchings. Using the theory of communication complexity,[3] which studies the minimum communication required to accomplish certain tasks from an information theory perspective, Segal (2007), Chou and Lu (2010), and Gonczarowski et al. (2015) prove that, for any method of finding stable matchings, there exists a class of preferences such that, to clear the market, agents must each *learn* and *communicate* their preferences for a substantial fraction of the entire market. More precisely, the worst-case amount of communication needed per agent grows linearly with the number of agents.[4] This communication requirement is implausibly large for many real markets, which have many thousands of agents on each side,[5] suggesting that there is an inevitable tension between providing choice to market participants (as captured by stability) and reducing communication overhead.

In contrast to these negative results, we show that, under natural assumptions on the distribution of preferences and prior knowledge of agents, stable matchings can be reached with limited communication even in large markets. We do this by constructing *communication protocols* that guide agents on whom to contact. The protocols help agents estimate with whom they can realistically be matched and encourage them to reach out to easier-to-get partners while waiting for harder-to-get partners to reach out to them. When every agent follows the protocol, the market reaches a stable matching with high probability and with low levels of communication and preference learning. Moreover, with high probability, it is in the best interest of each agent to comply with the protocol, assuming that others also do so.

The signaling employed by our protocol resembles features of real-world marketplaces. For instance, in the online labor market Upwork, the platform recommends suitable jobs to freelancers, who then "signal" their interest by applying; employers can also signal by inviting a suitable freelancer to apply. Both match recommendation and agent signaling play an important role in helping the market clear efficiently. Our communication protocols give platforms stylized insights on what kind of signals to allow and how to recommend matches.

### 1.1. Our Model and Results

We now describe our precise assumptions and results. Our assumptions on the preference distribution are mild. We say that a market is *separable* if the preferences of one side, say, firms, follow a latent random utility model with an additively separable structure. The preferences of workers can be arbitrary. For a firm, its latent utility for a worker is the sum of a systematic score and an idiosyncratic score,[6] both of which are heterogeneous across firm–worker pairs. The systematic score represents the worker's general level of fit based on observable characteristics, such as past experience, level of education, and test scores. This information is known to the worker and the firm but not necessarily to anyone else. The idiosyncratic score represents the idiosyncratic component of the firm's preference and is drawn independently from a certain unknown distribution. The preferences of agents, including both workers and firms, are unknown a priori to everyone, including the agents themselves. To learn their own preferences, agents have to query a choice function, which returns their most preferred partner within a given set.

Because the systematic scores can be heterogeneous across worker–firm pairs, we need the idiosyncratic scores to be sufficiently important relative to the systematic scores; otherwise, separable markets would include general markets, and the previous impossibility results of Gonczarowski et al. (2015) would hold. Precisely speaking, we assume that the range of systematic scores and the hazard rate of idiosyncratic scores are bounded above. A bounded hazard rate allows the idiosyncratic scores to take any heavy-tailed distribution, including the exponential, type I extreme value, lognormal, and Pareto distributions.

Our main result (Theorem 1) is that, in any separable market, there exists a way to find a stable matching with high probability using low levels of communication and preference learning cost. The protocol that we construct, called communication-efficient deferred acceptance (CEDA), modifies the worker-proposing DA algorithm by having workers apply only to firms at which they have a realistic chance. Workers know whether they have a chance through signals sent by firms. We show that, in any separable market, this protocol yields, with high probability, the worker-optimal stable match. Furthermore, the communication cost as measured by the number of bits agents send, on average, and the preference learning cost as measured by the number of choice function queries agents make, on average, both scale with the square root of the market size in the worst case,[7] which is much smaller than the linear scaling necessary under arbitrary preferences. In Theorem 2, we show, by constructing a plausible distribution over markets, that no stable matching protocol can have a better worst-case efficiency than CEDA for separable markets.

The source of CEDA's efficiency is the signals sent by firms, which help workers direct their applications. There are two types of signals: in the beginning, each firm bins workers based on their systematic scores and sends a *preference signal* to a certain number of its favorite workers in each bin. Receiving a preference signal indicates to a worker that the worker should apply despite having a potentially unfavorable systematic score. Although workers are applying to firms, each firm maintains and updates a systematic *qualification requirement*, which is broadcast to all workers. The qualification requirement is increased whenever the firm receives sufficiently many applications from workers whose systematic score meets the requirement. We show that, if a worker neither meets the qualification requirement nor receives a preference signal from a firm, then the worker's chance of being matched with that firm is essentially zero, so the worker should not waste time applying. Moreover, when workers apply only to firms at which their systematic scores qualify or from which they have received preference signals, then the total communication cost is small.

A drawback of the CEDA protocol is that it may require many rounds of communication as the signal an agent sends may depend on what signals the agent has received, which may, in turn, depend on the prior signals of other agents. We give an example illustrating that such sequential dependency may be necessary for any protocol under certain preference distributions. However, under stronger assumptions on preferences, we show that a two-round protocol is possible. The markets we consider are *tiered random markets*, in which agents on both sides are partitioned into tiers; an agent of a higher tier is always preferable to one of a lower tier, and preferences are uniformly random and independent within each tier.[8] For such markets, we give a two-round protocol whose preference learning and communication costs scale only polylogarithmically in the market size (Theorem 3). The protocol has the additional advantage of using only private signals, which need to be seen only by a single receiver. The protocol, called the *targeted-signaling protocol*, designates for each agent a set of easy-to-get and hard-to-get partners, based on commonly known tier information alone. In the first round, each agent signals a certain number of the agent's favorite easy-to-get partners. In the second round, each agent submits a partial preference ranking, in which the agent ranks only the subset of potential partners whom the agent signaled or who signaled the agent. The protocol outputs a matching based on these partial preferences, and we show that this matching is stable with respect to the full preferences with high probability.

Both of our protocols inherit the incentive properties of worker-proposing DA under complete information. More precisely, we show that with high probability, no worker can unilaterally deviate from the protocol and improve the worker's outcome, and no firm can unilaterally deviate and be matched to someone better than its partner in the firm-optimal stable match (Theorem 4). Recent literature has demonstrated that, in large markets, under mild assumptions, the vast majority of agents have the same match partners under the worker- and firm-optimal stable matchings (Immorlica and Mahdian 2005, Kojima and Pathak 2009, Ashlagi et al. 2017, Lee 2017, Lee and Yariv 2017). For such markets, all agents have vanishing incentives to deviate from our protocols.

### 1.2. Managerial Insights for Platforms

Our results highlight three features a platform can implement to reduce congestion. Each feature can take a variety of forms depending on the context, and we recommend implementing all three together for maximum effectiveness. The first feature is to help agents estimate their chance of obtaining a particular match. Such a feature can be as simple as publishing predicted[9] thresholds for school admissions as is currently practiced in certain universities in India, Israel, and Iran. Online platforms have the additional advantage of having a treasure trove of data on participant attributes and historical matches, and such data can be used to predict match compatibility in each direction (how much an agent may like a particular partner and vice versa.) As in CEDA, they can update their predictions by keeping track of how many favorable contacts a potential partner has received and correcting overoptimistic predictions for getting partners who are more popular than expected.

A second way for a platform to reduce congestion is to help *both* sides of the market better search for potential partners and signal interest. If only workers can initiate contact, then the top firms will receive a deluge of applications, most of which will turn out to be unfruitful. The market would be more efficient if there were a database of all workers that the top firms could use to filter for desirable worker attributes and initiate contact. Examples of such a database include class rosters that business schools post of graduating MBA students and their resumes, or an online platform, such as LinkedIn, which allows both workers and firms to search for potential partners and initiate contact. A recent empirical work that supports the importance of signaling from both sides of the market is Horton (2019), who finds that, in an online labor market in which employers recruit freelancers, the simple intervention of allowing freelancers to signal their level of availability substantially increases both match rates and market welfare.

A third helpful feature for reducing congestion is a match recommendation engine that targets potential partners for whom an agent has a realistic chance of matching. This is different from a naive "one-sided" recommendation engine that only takes into account who an agent may like, but not who may like the agent back. An online matching platform can implement such a "two-sided" engine as part of its search tool and only return potential partners who are estimated to be suitable, available, and willing to match. The value of such a recommendation engine is highlighted by the field experiment of Horton (2017) on oDesk/Upwork, which reports a 20% increase in employers' job-fill rates from adopting a recommendation engine that takes into account worker relevance, ability, and availability.

### 1.3. Relationship to Literature

As already discussed, this paper contributes to the literature on two-sided matching by defending the plausibility of stable matchings in large markets and by demonstrating the theoretical importance of match recommendations and informative signals in reducing congestion.

The conceptual framework we use to study signaling and information friction is distinct and complementary to previous works.[10] First, the role of signaling in this paper differs in that we focus on the *informative* aspect of signaling rather than the *strategic* aspect. In previous work, signaling changes the set of equilibrium outcomes, inspired by the literature on signaling games in economics. In this paper, signaling does not change the outcome, but enables the market to reach a desired outcome more quickly. Second, our approach on modeling information friction differs from previous approaches based on search theory in which agents optimize the matching outcome given costs or constraints on communication or information acquisition.[11] By contrast, we take the dual approach, in which the matching outcome is fixed, and we look for the least amount of communication needed to obtain that outcome.

The separable market model we study strictly generalizes previous models with uniformly random preferences (Wilson 1972; Pittel 1989, 1992; Coles and Shorrer 2014; Ashlagi et al. 2017), finite preference ordering on one side (Immorlica and Mahdian 2005, Kojima and Pathak 2009), and $O(\sqrt{n})$ many acceptable partners on both sides (Dagsvik 2000, Menzel 2015). This generalization is necessary for our purposes because the DA algorithm already yields low communication cost in the uniformly random case (Wilson 1972), and communication is trivially easy when agents have few acceptable partners. Our model is also distinct from the models of Lee (2017) and Lee and Yariv (2017) with bounded idiosyncratic scores and in which everyone receives the maximum possible idiosyncratic value in any large stable matching. Our model of tiered random preferences is studied in other contexts by Coles et al. (2013) and Che and Tercieux (2018a, b), and our analysis extends ideas from Pittel (1989, 1992), and Ashlagi et al. (2017).

## 2. Model

### 2.1. Review of Asymptotic Notation

This paper heavily utilizes Big O–type asymptotic notation. For clarity, we define these notations in this section.

Given two nonnegative functions $f, g : \mathbb{N} \to \mathbb{R}_+$, we say that $f(n) = O(g(n))$ if there exists $n_0$ and $M > 0$ such that $f(n) \le Mg(n)$ for all $n \ge n_0$. This indicates that up to a multiplicative constant, the tail of the function $f$ grows no faster than $g$. Similarly, we say that $f(n) = O^*(g(n))$ if there exists $n_0$ and $C > 0$ such that $f(n) \le (\log n)^C g(n)$ for all $n \ge n_0$.

We say that $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$. In other words, there exists $n_0$ and $M > 0$ such that $f(n) \ge Mg(n)$ for all $n \ge n_0$.

We say that $f(n) = o(g(n))$ if, for any $\epsilon > 0$, there exists $n_0$ such that $f(n) \le \epsilon g(n)$ for all $n \ge n_0$.

A special case is $f(n) = o(1)$, which means that $\lim_{n\to\infty} f(n) = 0$.

Intuitively speaking, $O(\cdot)$ and $O^*(\cdot)$ are used to upper bound the rate at which a function converges to infinity, and $\Omega(\cdot)$ is used to lower bound the rate; $o(\cdot)$ is used to lower bound the rate at which a function converges to zero.

### 2.2. Two-sided Matching Markets with Preference Learning

In this section, we define a generic model of two-sided matching markets with incomplete information. The markets studied in this paper, separable markets (Section 2.4) and tiered random markets (Section 4.1), are special cases of this more general model. Compared with previous work, the model has two distinguishing features: agents are allowed to have partial information on the preference distribution of others, and agents do not know their own preferences directly and must query a choice function to learn them.

A *two-sided matching market* $\mathcal{M}$ is defined by a tuple $(I, J, \omega, \mathcal{K}, \mathcal{P})$, where $I = \{1, 2, \ldots, n_I\}$ is a set of workers, and $J = \{n_I + 1, \cdots n_I + n_J\}$ is a set of firms. Both workers and firms are called *agents*. For concreteness, we refer to each worker as "she" and each firm as "it." Let $n = n_I + n_J$. This is the total number of agents, and we call this the *market size*. The parameter $\omega \in W$ represents the true *state of the world*, where $W$ is the set of possible states. (The state of world captures the distribution over agents' preferences.) $\mathcal{K}$ represents the a priori *knowledge* of agents about the state of the world.

It is indexed by each agent $i \in I \cup J$, and $\mathcal{H}_i \subseteq W$ contains the true state $\omega$. A *preference realization* $R$ is indexed by the set of agents $I \cup J$. For worker $i \in I$, $R_i$ is a permutation of $J \cup \{0\}$, which specifies her (strict) *preference ordering* over being matched with each firm and remaining unmatched, which is represented by the symbol 0. Similarly, for each firm $j \in J$, $R_j$ is a permutation of $I \cup \{0\}$, which specifies the preference ordering of firm $j$. A potential partner is said to be *acceptable* to an agent if the agent prefers being matched to the partner over being unmatched. The function $\mathcal{P}: W \rightarrow \Delta(R)$ is called the *preference function* and maps the state of the world to a probability distribution over preference realizations. Because $\omega$ is the true state of the world, $\mathcal{P}(\omega)$ is called the *true preference distribution*. We assume that each agent $i \in I \cup J$ a priori knows the function $\mathcal{P}$ and that the state of the world $\omega$ is contained in $\mathcal{H}_i$, but they do not know the true state of the world $\omega$ or the preference realization $R$.

To learn the realization of their own preference ordering, each agent must query their *choice function*, which specifies the agent's most preferred option within a given set of potential partners. Precisely speaking, for each worker $i \in I$, the choice function is $C_i : 2^{J \cup \{0\}} \rightarrow J \cup \{0\}$, which takes as input a subset of possible options, $S \subseteq J \cup \{0\}$, and outputs the highest-ranked element of $S$ according to the preference ordering $R_i$, which the agent does not directly observe. The choice function is analogously defined for every firm $j \in J$ except the set of possible options is now $I \cup \{0\}$.

The concept of choice function query is later used to quantify the preference learning cost of an agent. For the majority of the paper, we simply count the number of queries. In Section 6, we discuss an alternative formulation in which preference learning costs account for the size of the input set to each choice function query, so queries with larger inputs result in higher costs.

**2.2.1. Stable Matchings.** A collection of worker–firm pairs $\mu \subseteq I \times J$ is called a *matching* if each worker and each firm appears at most once. We refer to each $(i, j) \in \mu$ as a *matched pair*, and $i$ and $j$ as *matched partners* to one another in $\mu$. Agents who have no matched partner are said to be *unmatched* in $\mu$. Although a matching $\mu$ is, technically speaking, a set of pairs, we abuse notation and also use $\mu$ as a function that maps each agent to the agent's matched partner: if $(i, j)$ is a matched pair, then $\mu(i) = j$ and $\mu(j) = i$; if an agent $i \in I \cup J$ is unmatched, then $\mu(i) = 0$. A potential partner is said to be acceptable to an agent if the agent prefers to be matched to the partner over being unmatched. A *blocking pair* to a matching $\mu$ is a pair $(i, j) \in I \times J$ that is not a matched pair, but both agents find one another acceptable, and worker $i$ prefers firm $j$ to $\mu(i)$, and firm $j$ prefers worker $i$ to $\mu(j)$. A matching is said to be *stable* if there are no blocking pairs. For any preference realization $R$, a stable matching always exists and can be found using the DA algorithm of Gale and Shapley (1962). A version of this algorithm is stated in Section 2.3.

### 2.3. Stable Matching Protocol and Communication Cost

In this section, we formalize the concept of the communication cost of finding a stable matching, using concepts from the *communication complexity* literature [see Kushilevitz and Nisan (2006), for an overview of this literature].

A communication protocol with $n$ agents and input vector $(x_1, \ldots, x_n)$ is defined as follows. Each agent $i$ has access only to the agent's component $x_i$ of the input vector, but agents can send messages to other agents. Each message is formally represented as a sequence of zero–one bits, and we measure the length of messages by the number of bits. For now, we assume that messages are *public*, which means they are visible to all agents. (We remove this assumption in Section 4, in which we study protocols in which messages are visible only to a particular receiver.) We define the *history* of messages to be the sequence of all messages sent by any agent since the beginning of the protocol. As a function of the current history of messages, the protocol either terminates with a final *output* or chooses one agent to send the next message. The protocol also specifies what the chosen agent $i$'s message should be (including the length of the message) as a deterministic function of the agent's component $x_i$ of the input vector and the current history of messages. For every input vector $(x_1, \ldots, x_n)$, the protocol must be guaranteed to terminate with an output within a finite number of messages.

The next three definitions are essential for stating our main results. Intuitively, a matching protocol describes a way of finding a stable matching in which one can rigorously track the amount of communication. A preference learning strategy specifies how agents gather the necessary information about their own preferences when following a matching protocol, and a good strategy minimizes the expected number of choice function queries. A good matching protocol and an associated preference learning strategy together yield low communication and preference learning cost while producing stable matches with high probability.

**Definition 1** (Matching Protocol and Communication Cost). Given a two-sided matching market $\mathcal{M} = (I, J, \omega, \mathcal{H}, \mathcal{P})$, a corresponding *matching protocol* $\Pi$ is a communication protocol in which the set of agents is $I \cup J$, and for each $i \in I \cup J$, the agent's component of the input vector is $(\mathcal{H}_i, R_i)$, which includes the agent's a priori knowledge on the state of the world and the agent's preference realization. The output of the protocol is a matching $\mu \subseteq I \times J$. The *communication cost* of $\Pi$ is the expected total length of all messages (in bits) divided by the total number of agents, in which the expectation is taken over the distribution of preference realizations $\mathcal{P}(\omega)$.

**Definition 2** (Preference Learning Cost). Given a two-sided matching market $\mathcal{M} = (I, J, \omega, \mathcal{K}, \mathcal{P})$ and a matching protocol $\Pi$, a corresponding *preference learning strategy* specifies for each agent $i \in I \cup J$ how the agent should use choice function queries to obtain enough information on the agent's preference ordering $R_i$ so that the agent can follow the protocol. Precisely speaking, the strategy for an agent specifies the input to the agent's next choice function query given the history of the protocol, the sequence of choice function queries the agent made so far, and the result of these queries. (If the strategy specifies the empty set, then the agent should stop querying during this step of the protocol and send the message required of the agent.) The expected cost of a strategy is the expected total number of choice function queries across all agents, divided by the number of agents. The *preference learning cost* of matching protocol $\Pi$ is the minimum expected cost of any corresponding preference learning strategy.

**Definition 3** (Stable with High Probability). Consider a sequence of markets $(\mathcal{M}_n)_{n\in\mathbb{N}}$ indexed by the number of agents $n$ and a corresponding sequence of matching protocols $(\Pi_n)_{n\in\mathbb{N}}$. The sequence of *matching protocols* $(\Pi_n)_{n\in\mathbb{N}}$ is said to be *stable with high probability* for markets $(\mathcal{M}_n)_{n\in\mathbb{N}}$ if the probability that the matching produced is stable with respect to the realized preferences converging to one as the market size $n$ goes to $\infty$. Here, the probability is taken over the sequence of preference distributions $\mathcal{P}_n(\omega_n)$.

An example of a matching protocol that is always stable is the DA algorithm given as follows.[12] The corresponding preference learning strategy is implicitly given and uses one choice function query each step.

**Protocol 1.** *The sequential DA algorithm.*

*The protocol keeps track of a "tentative matching," which is initialized to be empty.*

(1) *Consider the set of tentatively unmatched workers who have not yet applied to all firms they find acceptable. If this set is empty, then the algorithm terminates and outputs the current tentative matching. Otherwise, one worker from this set is selected arbitrarily, say worker i, and she applies to her favorite firm j that she finds acceptable to which she has not yet applied.*

(2) *If firm j is tentatively unmatched and finds her acceptable, then it becomes tentatively matched to the worker. Otherwise, if it is already tentatively matched to some other worker, say worker i′, then it becomes tentatively matched to the more preferred worker among the two and rejects the other. The rejected worker becomes tentatively unmatched and can again be chosen in a future step. Return to step 1.*

This protocol requires high preference learning and communication cost: each step requires a choice function query to identify who to apply and who to accept. If there are $\Omega(n)$ firms, then communicating the identity of each firm requires $\Omega(\log n)$ bits of information, so each application requires a communication cost of $\Omega(\log n)$. In the worst case, a worker may apply to and be rejected from every single firm, and the average number of applications per worker may be $\Omega(n)$ (Itoga 1978). Therefore, the worst case communication cost is $\Omega(n \log n)$ per agent, and the worst case preference learning cost is $\Omega(n)$ per agent.

**Remark 1.** Although Definitions 1 and 2 define costs for a particular state of the world $\omega$, the protocol that we propose in Section 3 achieves low cost in *all* states of the world that satisfy the assumptions described in Section 2.4. Moreover, the protocol definition is the same for all states of the world.

#### 2.3.1. Impossibility of Sublinear Communication Cost for Arbitrary Markets.

The following negative result, which is a strengthening of an earlier result[13] by Segal (2007), implies that without restrictions on the preference distribution, one cannot hope to improve much upon the DA algorithm. There exists a distribution of preferences such that any matching protocol that is stable with high probability must require agents to learn and communicate, on average, at least a constant fraction of their preferences over the entire market. For real-world markets with many thousands of agents on each side, this is an unrealistically high requirement for communication and preference learning.

**Proposition 1** (Adapted from Gonczarowski et al. 2015). *There exists a sequence of two-sided matching markets* $\mathcal{M}_n = (I^n, J^n, \omega^n, \mathcal{K}^n, \mathcal{P}^n)$ *with* $\mathcal{K}_i^n = \{\omega^n\}$ *(everyone knows the state of the world and, hence, the precise distribution of preferences), such that any sequence of matching protocols that is stable with high probability requires a communication cost of at least* $\Omega(n)$ *per agent and a preference learning cost of at least* $\Omega(n/(\log n))$ *per agent.*[14]

However, this negative result requires an unrealistic preference distribution in which agent preferences are highly correlated in a contrived way.[15] The takeaway from the present paper is that, because preferences in real markets are not worst case but usually exhibit additional structure, this structure can be used to find stable matchings much more efficiently than what the previous result suggests.

### 2.4. Separable Markets

In this section, we define separable markets, a restricted class of two-sided matching markets for which we show in Section 3 that the communication requirement is much less than in arbitrary markets. After precisely stating the assumptions behind separable markets, we discuss in Section 2.4.1 why this is a reasonable model.

A separable market $\mathcal{M} = (I, J, \omega, \mathcal{K}, \mathcal{P})$ is a two-sided matching market with the following restrictions on the preference distribution $\mathcal{P}$, the true state of the world $\omega$, and the a priori knowledge $\mathcal{K}$ of the agents.

We assume no restrictions on the preferences of one side. Without loss of generality, let the unrestricted side be the workers. We assume the preferences of the firms follow a latent random utility model with the following additively separable structure (hence, the term "separable" markets). The latent utility of firm $j$ for worker $i$ is

$$u_{ji} = a_{ji} + \epsilon_{ji}, \tag{1}$$

where $a_{ji}$ is the *systematic score* of worker $i$ for firm $j$, which represents the observable characteristics for this worker–firm pair, and $\epsilon_{ji}$ is the *idiosyncratic score* of worker $i$ for firm $j$, which represents the unobserved component of the firm's preference for that worker. The utility of the firm for being unmatched, $u_{j0}$, is unrestricted. Denote this as firm $j$'s *outside utility*. We assume that all idiosyncratic scores for firm $j$ are distributed independently and identically according to a distribution $F_j$ and are independent of systematic scores, the preferences of workers, and the utility of firms for being unmatched.

Let $\mathcal{P}(\omega)$ denote the preference distribution described, in which the state of the world $\omega$ encapsulates the preferences of workers, the outside utilities of the firms, the set of systematic scores $\{a_{ji}\}$, and the idiosyncratic score distributions $\{F_j\}$. The worker preferences and firm outside utilities are completed unrestricted, and the systematic and idiosyncratic scores satisfy the following parametric assumptions, which together guarantee that the idiosyncratic scores are sufficiently important compared with the systematic scores.

**Assumption 1** (Range of Systematic Scores). *Without loss of generality,*[16] *we assume that all systematic scores $a_{ji}$ are nonnegative, and $\min_{i\in I} a_{ji} = 0$ for all $j \in J$. Further, we assume the range of systematic scores is upper bounded by a polylogarithmic function of market* size $n$, *that is,*

$$\bar{a} = O^*(1), \tag{2}$$

*where $\bar{a} = \max_{i\in I, j\in J}\{a_{ji}\}$. That is, there exists a constant $C > 0$ such that $\bar{a} \le \log^C n$.*

**Assumption 2** (Bounded Hazard Rate). *For every firm $j$, the distribution of idiosyncratic scores, $F_j$, satisfies a uniform bound on its hazard rate:*

$$h(x) := \frac{F_j'(x)}{1 - F_j(x)} \le 1 \qquad \forall x \in \mathbb{R}. \tag{3}$$

A bounded hazard rate is satisfied by any distribution with a sufficiently heavy tail, including the exponential, the type I extreme value, the log normal, and the Pareto distributions. Note also that the bound of one on the right side of (3) is without loss of generality because one can multiplicatively scale latent utilities by an arbitrary positive constant without affecting the underlying preferences, and the same multiplicative constant would become the bound on the right side of (3).

The assumptions on the a priori knowledge $\mathcal{K}$ of agents are as follows.

**Assumption 3** (Lower Bound on Prior Knowledge). *Each agent knows the systematic scores that are associated with them. (The systematic score $a_{ji}$ is a priori known to firm $j$ and worker $i$.) In addition, workers and firms know that the state of the world $\omega$ satisfies Assumptions 1 and 2.*

Note that Assumption 3 requires certain information to be included in $\mathcal{K}$, but leaves much flexibility on the remaining content. In particular, it does not specify whether agents have additional information on the state of the world, such as the systematic scores among other agents, the distribution $F_j$ for each firm, the preference of workers, or the outside utility of firms. Such knowledge assumptions are not necessary as all of our results hold regardless of how much additional knowledge agents possess beyond that specified in Assumption 3.

As described in Section 2.2, each agent does not know the agent's own preference realization a priori but can learn this via choice function queries. Agents can learn about the other agents' preferences only through communication.

#### 2.4.1. Discussion of Assumptions.

The key assumptions behind separable markets are

(1) The additively separable structure of firm latent utilities with systematic score $a_{ji}$ mutually known to firm $j$ and worker $i$.

2. The idiosyncratic scores are independent and identically distributed (i.i.d.).

3. Firm preferences are sufficiently idiosyncratic (Assumptions 1 and 2).

The additively separable structure of firm preferences is motivated as follows: in many real matching markets, the preference of at least one side depends on observable characteristics in a predictable way. For example, in hiring for a position, a firm may value an applicant's education, GPA, relevant certification, and relevant work experience. An interested worker may also have reasonable a priori knowledge of how important each of these characteristics is for a particular position and can assess her general level of fit without communicating with the firm. Likewise, the firm may observe many of a potential worker's characteristics from LinkedIn or a university's alumni database. The systematic score $a_{ji}$ represents this mutually observable general level of fit between worker $i$ and firm $j$. Note that we allow this to vary for each worker–firm pair, allowing for rich heterogeneities.

The idiosyncratic score $\epsilon_{ji}$ represents everything that is unexplained by the observable component. The assumption that they are i.i.d. is for technical convenience and is prevalent in the discrete choice literature as well as in most empirical studies in matching markets.[17] In our analysis, this assumption allows us to claim that, with high probability, after examining many workers, the firm must have found someone with a high idiosyncratic score. Furthermore, it allows us to upper bound the number of workers with a idiosyncratic score above a certain quantile. Because these are the only times we use this assumption, we expect the analysis to be generalizable to models in which the unobservable component of firm preferences exhibits mild correlations or mild variations in magnitude across workers for each firm.[18]

Regarding our assumption that firm preferences are sufficiently idiosyncratic, we note that some variant of this assumption is necessary to bypass the previous impossibility results. This is because, if idiosyncratic scores were identically zero or if the systematic scores were to have an unbounded range, then one can embed any arbitrary deterministic preferences into the separable market model, and a variant of the impossibility result of Proposition 1 would apply to any protocol acting on a particular distribution over separable markets.[19] However, Appendix B shows that Assumption 1 may be relaxed to allow the range of systematic scores to be an arbitrary sublinear function of $n$ while still obtaining a nontrivial bound.

Note also that Assumptions 1 and 2 still allow a substantial amount of systematic variation in the firm preferences. For example, suppose $F_j$ is the exponential distribution with parameter 1; then a difference in systematic scores of $3\log n$ implies that the firm will prefer the worker with the higher systematic score with a probability of $1 - 1/n^3$. Because there are only $O(n^2)$ worker–firm pairs, our assumptions allow enough systematic variation for each firm to have multiple tiers of workers, such that, with high probability, every worker in a better tier is preferred to every worker in a lower tier.[20] The number of tiers can also grow to infinity as long as it is controlled by a polynomial of $\log n$, and the tiers for each firm can be completely distinct. This allows the separable market model to be flexible enough to capture rich preference structures.

## 3. Main Results

Our main technical results for separable markets are as follows.

(1) We propose a matching protocol called the CEDA protocol, which is stable with high probability and in the worst case incurs communication and preference learning costs of $O^*(\sqrt{n})$ per agent (see Theorem 1 in Section 3.1). This is much lower than the $\Omega(n)$ cost needed for arbitrary two-sided markets (see Proposition 1).

(2) This $O^*(\sqrt{n})$ communication and preference learning cost is essentially the best possible guarantee for separable markets as we give an example of a separable market in which any matching protocol that is stable with high probability requires $O(\sqrt{n})$ bits of communication and $O(\sqrt{n}/\log n)$ bits of preference learning (see Theorem 2 in Section 3.2).

The CEDA protocol itself is a main contribution of the paper as it has an intuitively appealing structure and provides guidance on how matching platforms can facilitate efficient market clearing. We also show in Section 5 that the protocol has good incentive properties.[21]

### 3.1. CEDA

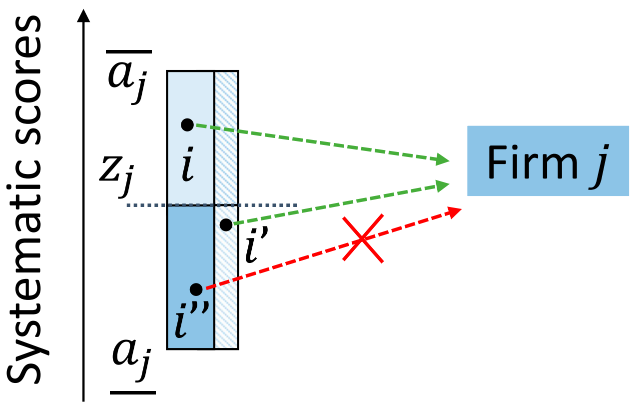
The protocol we construct, which achieves the $O^*(\sqrt{n})$ guarantee, is called the CEDA protocol. The high-level idea is to allow workers to better target their applications with the help of signals sent by firms, which help workers identify the firms at which they have a nonnegligible chance of acceptance. There are two types of signals:

- Preference signal: A firm $j$ signals to worker $i$ if it has a high idiosyncratic score $\epsilon_{ji}$ for the worker. All preference signals are sent at the outset.
- Qualification requirement signal: A firm $j$ broadcasts a qualification requirement $z_j$ to the entire market, which specifies the minimum systematic score a worker who did not receive a preference signal needs to apply to the firm. Qualification requirement updates are sent throughout the protocol.

The key property of the protocol is that if a worker $i$ does not receive a preference signal from a firm and if her systematic score is below its qualification requirement, then she would almost certainly be rejected and so should not bother applying. We explain this point in more detail in Section 3.1.1.

**Definition 4.** A worker $i$ is said to *systematically qualify* for firm $j$ if her systematic score meets its qualification requirement, $a_{ji} \geq z_j$. The worker is said to *qualify* for the firm if she either systematically qualifies for it or

**Figure 1.** (Color online) Illustration of Qualification Rules in CEDA for Applying to a Firm




*Notes.* The vertical axis shows the systematic score of workers for firm $j$. The long rectangle on the left represents the set of all workers with the dashed regions to the right representing workers who received a preference signal from the firm. There are three workers in the figure: $i$, $i'$, and $i''$. Worker $i$ qualifies for firm $j$ because her systematic score exceeds the threshold $z_j$, worker $i'$ qualifies because she received a preference signal, but worker $i''$ does not qualify because she neither meets the qualification requirement $z_j$ nor received a preference signal.

receives a preference signal from it. See Figure 1 for an illustration.

The protocol is defined formally as follows and also implies a preference learning strategy.

**Protocol 2** (CEDA). *Initialize the qualification requirement for each firm j to the minimum possible systematic score*[22] $z_j = 0$. *There are two phases in CEDA.*

(1) *Preference signaling: Each firm j places workers based on their systematic scores into unit-ranged bins* $[0, 1)$, $[1, 2)$, $\cdots$, $[\bar{a} - 1, \bar{a}]$.

*The firm sends a preference signal to its top* $2e\sqrt{n}$ *most preferred workers from each bin.*[23]

(2) *Deferred acceptance with qualification requirement: Run the sequential worker-proposing deferred acceptance algorithm (Protocol 1 in Section 2.3) with the following modifications*[24]:

(a) *Workers apply only to firms for for which they qualify (see Definition 4).*

(b) *Each firm j, after every* $3e \log n\sqrt{n}$ *applications received from systematically qualified applicants, increases its qualification requirement* $z_j$ *by one and broadcasts this increase to the market by sending a qualification requirement signal.*

In practice, a platform can make it easier for agents to follow the protocol by estimating the systematic scores, automating the qualification requirement updates, and only showing workers the firms for which they currently qualify. This would also eliminate the need for broadcast messages. We highlight also that the CEDA protocol is exactly the same for any market with set of workers $I$, set of firms $J$, and range of systematic scores $\bar{a}$; thus, it demands only minimal knowledge on the part of the protocol designer; compare with Remark 1 in Section 2.3.

The following is our main result about CEDA.

**Theorem 1.** *In any sequence of separable markets satisfying Assumptions 1–3, the CEDA protocol (Protocol 2) is a matching protocol that is stable with high probability. Its communication cost and preference learning cost are both at most* $O^*(\sqrt{n})$ *per agent in the worst case.*

Before sketching the proof, we present the high-level intuition on what drives the reductions in communication overhead in CEDA. Worker-proposing DA (Protocol 1) yields high communication cost if there are many applications before the algorithm converges, and this can happen if either

(i) most applications by worker are rejected by firms or

(ii) firms' utility only improves slightly with every acceptance, implying that a firm may tentatively match with many candidates before reaching its final match.

CEDA overcomes the first problem by only allowing applications from workers who have a significant chance of being accepted. The second problem does not occur in separable markets because idiosyncratic scores are heavy tailed (Assumption 2), implying that each acceptance is likely to improve a firm's utility by a significant amount.

**3.1.1. Sketch of the Proof.** The formal proof of Theorem 1 is in Appendix A. The main reason behind this result is what we call CEDA's *no-false-negatives property*.

**Lemma 1** (No-False-Negatives Property). *With high probability, throughout the running of CEDA, if a worker does not qualify for a firm at a certain time, then she will not be accepted if she had applied to the firm at that time.*

This property implies that, with high probability, CEDA prevents only applications that would have been rejected anyway. So the sequence of acceptances in CEDA exactly matches that in the DA protocol, and CEDA succeeds in finding the worker-optimal stable match.

The high-level intuition of why the property holds is as follows. At least one of two things must occur for a worker to be (tentatively) accepted by a firm: she must have either a high systematic score or a high idiosyncratic score. The definition of "not qualifying" (the inverse of Definition 4) rules out both these possibilities, so workers who do not qualify would not have been accepted anyway.

More precisely, define $\gamma_j$ to be the $(1 - 1/\sqrt{n})$th fractile of the idiosyncratic scores for firm $j$, $\gamma_j := F_j^{-1}(1 - 1/\sqrt{n})$. Define $x_j$ to be firm $j$'s latent utility for the current tentative match in CEDA. (This is initialized to $u_{j0}$ and increases whenever the firm tentatively accepts a new worker.). We show in Appendix A that CEDA satisfies the following two properties with high probability.

(1) Every worker $i$ whose idiosyncratic score for firm $j$ is higher than $\gamma_j$ receives a preference signal. (This follows from the i.i.d. property of idiosyncratic scores, the definition of $\gamma_j$, and the number of preference signals sent.)

(2) For every firm $j$, the tentative match value $x_j$ increases sufficiently quickly relative to $z_j$ so that the following invariant holds throughout the running of the CEDA protocol:

$$z_j \leq \max(0, x_j - \gamma_j), \tag{4}$$

which implies that, if a worker does not systematically qualify for firm $j$, then her systematic score satisfies $a_{ji} < x_j - \gamma_j$. This invariant holds because of the following concentration bound: the maximum idiosyncratic score among $3e \log n\sqrt{n}$ independent draws of $F_j$ is at least $\gamma_j + 1$ with probability at least $1 - 1/n^3$. Intuitively, the maximum of many independent draws of a heavy-tailed distribution is large with high probability, and we can make the probability as high as $1 - 1/n^k$ by taking $ke \log n\sqrt{n}$ draws.

These two points together imply the no-false-negatives property because, when worker $i$ does not qualify for firm $j$, then the first property implies that $\epsilon_{ji} < \gamma_j$, and the second implies that $a_{ji} < x_j - \gamma_j$, so the firm's utility for the worker satisfies $u_{ji} = a_{ji} + \epsilon_{ji} < x_j$, and thus, the worker would not have been accepted anyway.

Having explained the intuition as to why CEDA successfully computes the worker-optimal stable match with high probability, we now explain the reasoning behind the $O^*(\sqrt{n})$ bounds in communication and preference learning cost. This follows from the observation that, in CEDA, both communication and preference learning can be upper bounded by the number of signals sent and the number of applications. By construction, each firm sends only $O^*(\sqrt{n})$ preference signals and $O^*(1)$ qualification requirement signals. This implies that it can receive only $O^*(\sqrt{n})$ applications from workers who do not systematically qualify, and $O^*(\sqrt{n})$ applications from workers who do systematically qualify. This last bound follows from the observation that there are at most $O^*(1)$ updates to the qualification requirement for each firm because Assumption 1 implies that, after this many updates, the qualification requirement will have exceeded the highest possible systematic score, and no worker would systematically qualify. Finally, for each update to the qualification score, there are—by construction—only $O^*(\sqrt{n})$ applications from systematically qualified workers.

The argument is formalized in Appendix A. The proof shows that the probability that the resulting matching is not stable is bounded above by $(\bar{a}(n))/n^2 + ne^{-\sqrt{n}/3}$, where $\bar{a}(n) = O^*(1)$ is the bound on the range of systematic scores in Assumption 1.

### 3.2. The Optimality of the $O^*(\sqrt{n})$ Guarantee

The following example shows that the $O^*(\sqrt{n})$ guarantee of CEDA is asymptotically near optimal: consider a market with $n$ workers and $n$ firms. The preferences of both workers and firms follow a separable structure with all systematic scores being zero and all idiosyncratic scores being drawn from an exponential distribution with rate parameter one. Each agent has an outside option of value $(\log n)/2$, which implies that each agent finds a uniformly random subset of about $\sqrt{n}$ partners acceptable. Agent preferences among acceptable partners are also uniformly random.

Intuitively, it is hard to find a stable matching in this market without at least $\Omega(\sqrt{n})$ bits of communication because otherwise it is difficult even to identify which pairs of agents find each other mutually acceptable: out of $n$ possible partners, an agent has, on average, only one who the agent finds acceptable and who reciprocally finds the agent acceptable, so finding mutually acceptable partners is like finding needles in a haystack. The most obvious way of finding all mutually acceptable partners is for every agent to communicate the set of agents that the agent finds acceptable, which requires $\Theta(\log n\sqrt{n})$ bits per agent. Theorem 2 shows that no protocol can do much better.

**Theorem 2.** *There exists a sequence of separable markets for which any sequence of matching protocols that is stable with high probability requires a communication cost of $\Omega(\sqrt{n})$ per agent and a preference learning cost of $\Omega(\sqrt{n}/\log n)$ per agent.*[25]

The sequence of markets in Theorem 2 is exactly the example described. An outline of the proof in Appendix C is as follows. First, we show that any matching protocol that is stable with probability at least 90% must approximately identify the set of worker–firm pairs who find each other mutually acceptable. (The precise definition of "approximately identify" is technical and is presented in the formal proof in the appendix.) The reason is that there are $O(n)$ such mutually acceptable pairs, and a significant fraction of them must be matched in all stable matchings because $O(n)$ agents are matched in total. There are $n^2$ worker–firm pairs in total. For each pair, we use ideas from information theory based on Shannon's entropy and Shannon's mutual information (see Braverman 2015) to show that the protocol must use $\Omega(1/\sqrt{n})$ bits of communication, on average, to approximately determine mutual acceptability.

This implies that $n^2 \times \Omega(1/\sqrt{n}) = \Omega(n^{3/2})$ total bits of communication are needed, which implies the $\Omega(\sqrt{n})$ per agent bound.

Finally, we show that this $\Omega(\sqrt{n})$ lower bound on communication cost immediately implies a $\Omega(\sqrt{n}/\log n)$ lower bound on preference learning cost. The reason is that any protocol that uses $Q$ choice function queries can be modified into a communication protocol with a communication cost of $O(Q \log n)$ bits.

## 4. Simultaneous Two-Round Protocol with Private Communication

The CEDA protocol in Section 3.1 is sequential: a worker's decision about to which firms to apply depends on the firms' current qualification requirements, which, in turn, depend on other workers' application decisions. Implementing such a protocol may result in slow market convergence as agents need to wait for other agents to act before knowing what preference information to learn next and how to act next. In this section, we explore the possibility of *simultaneous* protocols, in which the dependence of each agent's action on prior actions is minimized. In particular, we consider *two-round protocols*.[26] In the first round, agents simultaneously signal to various partners, and in the second round, everyone reports a partial preference list to a central matchmaker based on signals received during the first round. One interpretation of the central matchmaker is a centralized clearinghouse as in the NRMP. Another interpretation is that it is a proxy for a decentralized matching process after which agents arrive at a matching in which no pair of agents who contacted one another form a blocking pair.

It turns out that one can construct a separable market in which any two-round protocol of this form that is stable with high probability requires $\Omega(n)$ bits of communication per agent (see Appendix H). To obtain a positive result, we introduce a simpler model of matching markets, which we call tiered random markets. In such markets, agents are partitioned into tiers, and each agent prefers better tiers to worse tiers and has uniformly random preferences among agents in a given tier (see Section 4.1). This model still allows both vertical and horizontal differentiation, and it yields clean insights about what kind of signals are most informative. Moreover, there is a highly efficient two-round stable matching protocol, which we present in Section 4.4.

This protocol has an additional advantage: it *uses only private messages*, which are messages visible only to a sender and a receiver and not to anyone else. Requiring messages to be private models markets in which there is no efficient way to broadcast a particular message to all agents simultaneously. The formal definition is as follows.

**Definition 5.** A communication protocol $\Pi$ is said to use only private messages if every message specifies the identity of a receiving agent and is visible only to that agent. Furthermore, each message an agent sends can depend only on the history of messages that the agent has seen and not on messages the agent has not seen.

As in the definition of a communication protocol in Section 2.3, the protocol still observes all messages and chooses when to terminate with an output and who should send the next message.[27]

### 4.1. Tiered Random Markets

A *tiered market* is a two-sided matching market $(I, J, \omega, \mathcal{H}, \mathcal{P})$ in which agents are partitioned into commonly known tiers, and every agent prefers partners from a better tier to those from a worse tier and has uniformly random preferences for partners within a given tier.[28]

Precisely speaking, there are $K \geq 1$ tiers of workers and $L \geq 1$ tiers of firms. The state of the world $\omega = (s, t)$, where $s$ is a $K$-dimensional vector of positive integers and $t$ is an $L$-dimensional vector of positive integers, satisfying

$$0 =: s_0 < s_1 < \cdots < s_K = n_I,$$
$$\text{and } 0 =: t_0 < t_1 < \cdots < t_L = n_J.$$

Let $I_k = \{s_{k-1}+1, s_{k-1}+2, \ldots, s_k\}$. This denotes the $k$th tier of workers. $I_1$ is the best tier and $I_K$ the worst. Given worker $i \in I$, let $k(i)$ denote the tier of the worker. This is the unique $k$ such that $s_{k-1} < i \leq s_k$. Similarly, let $J_l = \{n_I + t_{l-1} + 1, \ldots, n_I + t_l\}$ denote the $l$th tier of firms. (We add $n_I$ because we index the firms from $n_I + 1$ to $n_I + n_J$.) For every firm $j$, let $l(j)$ denote the tier of the firm. The a priori knowledge $\mathcal{H}_i$ of every agent $i \in I \cup J$ is $\{\omega\}$. The true preference distribution $\mathcal{P}(\omega)$ is such that the preference realization $R_i$ of each worker $i \in I$ is a uniformly random permutation of $J_1$, followed by a uniformly random permutation of $J_2$ and so on. We assume that every firm is acceptable to the worker. The preferences of firms are defined analogously.

Tiered random markets model markets in which the vertical differentiation is coarse and the horizontal differentiation is idiosyncratic. For example, in the academic job market for certain subfields, one may argue that departments are clustered into quality-differentiated tiers, and every applicant prefers better tiers, but preferences within each tier are driven by personal preferences that can be modeled as essentially random. Furthermore, applicants may also be clustered into tiers based on publication record and school of origin with different departments having essentially random preferences within each tier based on their particular needs at the moment.

### 4.2. Better and Worse Positions

We define a notion of relative competitive position, which intuitively indicates who has more market power in a tiered random market. We say that a worker $i$ is in a *weakly better position* than firm $j$ if $s_{k(i)} \le t_{l(j)}$. In other words, there are weakly fewer workers in tiers as good as $i$ as there are firms in tiers as good as $j$ (workers as good as $i$ are in short supply). Similarly define this for firms. Define *weakly worse position* in an analogous way with the inequality reversed.

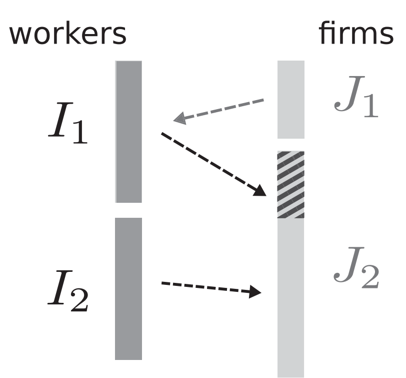
These definitions are motivated by the following result from previous work. In a *uniformly random market*, which is a special case of tiered random markets with one tier on each side, the DA algorithm terminates quickly when the proposing side has weakly fewer agents but not when it has strictly more agents. (For example, Ashlagi et al. (2017) show that, in a market with $n-1$ workers and $n$ firms, the average number of applications in the worker-proposing DA algorithm is about $\log n$ per agent, whereas the average number of applications in the firm-proposing DA is about $n/(\log n)$ per agent.) This implies that it is more efficient in terms of communication to have the side with fewer agents do the proposing.

The protocol we propose has agents only initializing contact with potential partners in weakly worse positions than themselves.

### 4.3. Example and Intuition

Before giving the full protocol, we consider a simple numerical example, which illustrates the main insights. Suppose that there are two tiers of workers of 50 workers each. We call these the top workers and average workers, respectively. Similarly, there are two tiers of firms, which we call the top and average firms. There are 20 top firms and 90 average firms as illustrated in Figure 2.

**Figure 2.** An Example of a Tiered Random Market (See Section 4.3)




*Notes.* There are two tiers of workers, $I_1$ and $I_2$, and two tiers of firms, $J_1$ and $J_2$. The height of each rectangle corresponds to the number of agents in that tier. The arrows show the direction of the most informative signaling. The top workers $I_1$ should signal to the average firms $J_2$ and wait for the top firms to signal to them. The bottom workers $I_2$ should signal to the bottom firms as well, but they need to amplify the number of signals they send because some of the bottom firms would already be taken by top workers.

The first observation is that the tier structure precludes certain matches in a stable matching. For example, in every stable matching, every top firm must be matched with a top worker. The reason is that there are more top workers than top firms, and any top worker who is not matched with a top firm would like to be matched to one. Thus, the average workers in this example have no chance whatsoever of being matched with a top firm. More generally, in a tiered random market, worker $i$ can be matched with firm $j$ only if their tiers overlap: $s_{k(i)-1} < t_{l(j)}$ and $s_{k(i)} > t_{l(j)-1}$.

The second observation is that it is more efficient to have the agents signal partners with (weakly) worse positions than themselves. In the example, it is more efficient for top firms to signal top workers than the reverse because there are fewer top firms than top workers. Similarly, it is more efficient for average firms to wait for signals from both top and average workers because, when we take out the top workers who will be matched to the top firms, there are $100 - 20 = 80$ workers left and 90 average firms. So the previous results for uniformly random markets suggest that it is better for the workers to signal. These directions of signaling are shown in the arrows in Figure 2.

The third observation is that because preference signals are sent in parallel, certain agents may have to send extra signals to account for the fact that potential partners may already be taken up by competitors from better tiers. For example, for each of the average workers, there are 90 average firms she can signal, but 30 of these will end up matching with a top worker against whom she has no chance. So, for every three signals she sends, in expectation, one would be wasted. This means that she should amplify the number of signals she sends by a factor of 3/2.

For certain agents, this amplification effect may be large. For example, consider the case in which there is a single tier of $n$ firms and there are $n$ tiers of one worker each so that the workers are completely vertically differentiated. In this example, the last ranked workers need to signal order $n$ firms because of the amplification effect. However, one can show that the *average* amplification needed is only $O(\log n)$ in this case.[29] For arbitrary tier structures, one can show that the average amplification needed is always small.

### 4.4. The Targeted Signaling Protocol

In this section, we describe a two-round protocol for tiered random markets that is stable with high probability and uses only private messages. We begin with a high-level summary. The protocol designates for each agent a *target tier* based on the tier structure alone (Definition 6). In the first round, every agent signals a certain number of favorite partners within the agent's target tier. A signal intuitively represents initiating contact in a decentralized job market. In the second round, agents submit a partial preference ranking of partners among those they signaled or received a signal from, and the protocol runs the DA algorithm using the partial preferences collected.

**Definition 6.** The target tier of a worker $i$ is the best tier of firms that is in a weakly worse position than she is. In other words, this is $J_l$, where $l = \min\{l : t_l \geq s_{k(i)}\}$. Similarly define the target tier of each firm $j$.

In the example in Figure 2, the target tiers are indicated by the arrows so that the target tier of top firms is top workers, the target tier of top workers is bottom firms, and so on. The bottom firms do not have a target tier because they are in the worst position possible. As explained in Section 4.3, it is most efficient for agents to each limit signals to their target tier. However, the number of signals sent needs to be amplified in inverse proportion to the fraction of target-tier agents left over after competitors in better tiers have found matches. This is precisely stated.

**Definition 7.** For each worker $i$ in tier $k$ with target tier $l$, define her *relative competitiveness for her target tier*

$$\rho(i) = \min\left(1, \frac{t_l - s_{k-1}}{t_l - t_{l-1}}\right).$$

This is the proportion of the target tier that will not be matched to workers from better tiers. Similarly define $\rho(j)$ for firms.

**Definition 8.** Define the *target number* of an agent $i \in I \cup J$ as

$$r(i) = \frac{24 \log^2 n}{\rho(i)}.$$

In the example, the relative competitiveness of top firms and top workers is one, whereas the relative competitiveness of the bottom workers is the proportion of bottom firms that are unshaded in Figure 2, which is equal to $60/90 = 2/3$. The most important feature of the target number is that it is inversely proportional to $\rho(a)$, which matches the level of amplification in the example. The $\log^2 n$ multiplier accounts for competition from the agent's own tier.[30] The constant 24 is an artifact of the analysis in Appendix E and is not the minimum possible constant.

The protocol is precisely defined as follows.

**Protocol 3** (Targeted Signaling Protocol). *The protocol has two rounds and uses only private messages.*

(1) *Signaling round: Every agent $i \in I \cup J$ signals to the agent's favorite $r(i)$ partners in the agent's target tier.*

(2) *Matching round: Every agent submits to the protocol a partial preference ranking of partners with the ranking restricted to partners who either signaled to the agent or to whom the agent signaled. The protocol then outputs the worker-optimal stable matching with respect to the submitted preferences (computed by running worker-proposing DA off-line).*

**Theorem 3.** *For any sequence of tiered random markets, the targeted signaling protocol is a matching protocol that uses only private messages and is stable with high probability. Its average communication cost is $O(\log^4 n)$ per agent, and its average preference learning cost is $O(\log^3 n)$ per agent.*

The proof of Theorem 3 is in Appendix E. The main steps are as follows. First, we show that, with high probability, a certain subset of signals sent in the signaling round contains a stable matching. This result, stated in Lemma 5, is obtained by proving a new bound on the average rank of agents in any stable matching in unbalanced uniform random markets.[31] Second, we show that whenever this subset of signals contains a stable matching, running DA on the partial preferences, as in the matching round, returns a stable matching. This result uses the structure of tiered markets and the definition of target tier. (In general, even if a set of partial preferences contains a stable matching with respect to full preferences, running DA on these partial preferences may result in a matching that is stable only with respect to the partial preferences.) Third, we count the total number of signals and show that it is no more than $O(\log^3 n)$ per agent, and the communication cost naturally has an additional logarithmic factor.[32] The proof shows that the probability that the resultant matching is not stable is no more than $18/n$, and the average number of signals sent per agent is no more than $60 \log^3 n$.

**Remark 2.** An immediate corollary of Theorem 3 is that, in the targeted signaling protocol, with high probability, the number of agents who experience more than $\sqrt{n}$ bits of communication is at most[33] $O^*(\sqrt{n})$. Intuitively speaking, some agents may experience relatively more communication, but the number of such agents is a vanishing fraction of the population. In Theorem 8 of Appendix I, we demonstrate a tiered random market in which some agents *must* experience $\Omega(\sqrt{n})$ bits of communication under any protocol that uses private messages and is stable with high probability.

**Remark 3.** We also have a variant of the targeted signaling protocol that saves a factor of $\log n$: its communication cost is $O(\log^3 n)$ bits per agent and its preference learning cost is $O(\log^2 n)$ per agent. It involves a much more complicated formula for the target number than in Definition 8. In the interest of a clean exposition, we show the simpler version here.

**Remark 4.** The targeted signaling protocol presented above achieves near optimal communication cost. In Theorem I.1, we give an example of a tiered random market in which any matching protocol that uses only private communication and is stable with high probability must use at least $\Omega(\log^2 n)$ bits of communication per agent. This lower bound includes protocols that allow an arbitrary number of rounds of communication. The targeted signaling protocol achieves near optimal cost using only two rounds.

## 5. Incentive Compatibility

The DA protocol, even in the setting with full preference elicitation, is not completely incentive compatible: an agent on the nonproposing side may profitably deviate from truthful reporting by truncating the agent's preference ranking.[34] Nevertheless, the DA mechanism is known to be strategy-proof for the proposing side.[35] Moreover, assuming truthful reporting by other agents, an agent on the nonproposing side cannot unilaterally deviate (misreport the agent's preferences) and be matched with someone better than the agent's best stable partner.[36]

We show that, with high probability, all the matching protocols proposed in this paper are in a certain sense "as incentive compatible" as DA with full preference elicitation. By theorem 4.11 in Roth and Sotomayor (1990) (originally resulting from Demange et al. 1987), being as incentive compatible as DA is the best one can hope for in any mechanism that returns stable matchings.

**Definition 9.** For a given sequence of two-sided matching markets, a sequence of matching protocols is as incentive compatible as DA with high probability if there exists a function $\delta(n)$ with $\delta(n) \to 0$ as $n \to \infty$, such that for any fixed agent with probability at most $\delta(n)$, the agent can unilaterally deviate from the protocol and be matched with someone better than the agent's best stable partner (under complete preferences). The probabilities are defined with respect to the distribution of preferences $\mathcal{P}(\omega)$.

**Theorem 4.** *For any sequence of separable markets, the CEDA protocol is as incentive compatible as DA with high probability. For any sequence of tiered random markets, the targeted signaling protocol is as incentive compatible as DA with high probability.*

The "with high probability" caveat is needed because, with a small probability, the protocols may fail to find a stable matching. The proof of Theorem 4 is in Appendix G. The proof is based on applying a "blocking lemma" by Gale and Sotomayor (1985) to a market in which preferences are restricted to the *subgraph of signals*, which is the set of worker–firm pairs in which at least one member of each pair sent a signal to the other (i.e., a preference signal or an application in the CEDA protocol, or a signal in the targeted signaling protocol). The same lemma is used to prove the incentive properties of the original DA algorithm.[37] The new ingredient here is to use properties of the subgraph of signals generated by our protocols. In particular, we make use of the fact that a single deviating agent has little control over the edges of the subgraph between other agents.

The proof shows that the respective protocols are as incentive compatible as DA whenever they succeed in finding a stable matching. By the proofs of Theorems 1 and 3, we can set the function in Definition 9 to $\delta(n) = (\bar{a}(n))/n^2 + ne^{-\sqrt{n}/3}$ for CEDA, where $\bar{a}(n) = O^*(1)$ is the upper bound to the range of systematic scores in Assumption 1 and to $\delta(n) := 18/n$ for the targeted signaling protocol.

Because both protocols we propose are based on the worker-proposing DA, they are also, with high probability, strategy-proof for workers.

**Corollary 1.** *For any sequence of separable markets, the CEDA protocol is strategy-proof for all workers with high probability. For any sequence of tiered random markets, the targeted signaling protocol is strategy-proof for all workers with high probability.*[38]

Theorem 4 also implies that, whenever an agent has a unique stable partner, the probability that the agent can profitably deviate from the protocol vanishes in large markets. Thus, if each agent's probability of having multiple stable partners is small, then following the protocol is an $\epsilon$-Bayes Nash equilibrium. In "typical" tiered random markets, we can prove that agents have a vanishing probability of having multiple stable partners, thus implying approximate incentive compatibility of the targeted signaling protocol for everyone. This is made precise as follows.

**Definition 10.** A tiered market satisfies *general imbalance* if the sets $\{t_1, t_2, \cdots\}$ and $\{s_1, s_2, \cdots\}$ are disjoint.

**Theorem 5.** *There exists a function $\delta : \mathbb{N} \to \mathbb{R}$ satisfying $\delta(y) \to 0$ as $y \to \infty$ such that the following holds. For any $y \in \mathbb{N}$, consider any tiered random market satisfying general imbalance in which the number of agents in each tier is at least $y$. Then, for each agent, the probability that the agent can profitably deviate from the targeted signaling protocol is at most $\delta(y)$.*

The proof is in Appendix G. Besides applying Theorem 4, the proof uses ideas from Ashlagi et al. (2017) to show that the proportion of agents with multiple stable partners converges to zero under the general imbalance condition when the number of agents in each tier is large. The convergence rate shown in the proof is $\delta(y) = o(1/(\sqrt{\log y}))$ and is close to the best possible.[39]

## 6. Discussion

Although we study particular mathematical models in this paper, the general message is broader: a market clearing outcome (stable matching) can be achieved in large matching markets by informative signaling, despite real-world limitations in communication. The types of informative signals we uncover for matching markets include the following: signaling partners for whom one has a particular idiosyncratic preference, broadcasting to the market a requirement on observable characteristics (the platform can do this) and focusing on initiating contact with easy-to-get partners while waiting for hard-to-get partners to initiate contact. These types of signaling are already present in many real matching markets, and our results highlight their importance to market efficiency.

One potential criticism of our model is that preference learning costs do not distinguish between finding one's best option among many or few alternatives. However, if we alter the cost model so each choice function query with input set $S$ incurs cost $c(|S|)$ for some increasing function $c(\cdot)$, then Theorem 1 still holds with the guarantees on preference learning cost weakened to $O^*(\sqrt{n}c(n))$ bits per agent. This is still an asymptotic improvement over previous results if $c(n) = o(\sqrt{n})$. Another model of preference learning would be to only allow each agent to query for the agent's idiosyncratic score for one partner at a time, but in that model, it is hopeless to achieve sublinear preference learning cost because there is always the chance that one's favorite partner was not queried.

Another potential criticism is that agents in our model know the systematic scores associated with them perfectly. However, if agents know their systematic scores up to an additive error of no more than a constant $\delta > 0$, then Theorems 1 and 4 still hold with the following variant of CEDA: (1) an agent systematically qualifies for firm $j$ if the agent's estimate of the agent's systematic score is no less than $z_j - \delta$; (2) increase by a multiplicative factor of $e^\delta$ the number of preference signals sent by firms to each bin as well as the number of applications from systematically qualified agents before increasing the qualification requirement by one.

This paper raises several questions for future research. One question is how many rounds of communications are required to keep the overall level of communication low.[40] In Appendix H, we give a simple example of a separable market in which no two-round protocol can reach a stable matching with low communication cost. In that example, the problem would be solved by having an additional "aftermarket" round.[41]

Another question is, "How heterogeneous do incurred communication and preference learning costs need to be across agents?" In the targeted-signaling protocol, although most agents incur a very low communication cost (polylogarithmic in the market size), some agents require more communication effort. To see this, consider the following example: there are $n$ tiers of one worker each, and one tier of $n$ firms. In this example, the top workers can essentially choose whichever firm they like, but the bottom workers can be matched only with leftover firms after workers from higher tiers have already made their choices. In the targeted-signaling protocol, the bottom worker in this example incurs a linear communication cost. Moreover, we show in Appendix I that, for this market, any protocol that uses only private messages requires some agent to incur a communication cost on the order of the square root of the market size. It would be interesting to study, either theoretically or empirically, the relationship between the market position of an agent and the minimum communication effort required from the agent.

### Acknowledgments

This paper has greatly benefited from conversations with Avinatan Hassidim, Ilya Segal, and Assaf Romm. The authors also thank Federico Echenique, Marina Halac, Scott Kominers, Anqi Li, Jacob Leshno, and Marek Pycia for very helpful comments.

## Appendix A. Proof of Correctness and Efficiency of CEDA (Theorem 1)

In this appendix, we prove Theorem 1 from Section 3.1, which states that in any separable market, the CEDA protocol is a matching protocol that is stable with high probability (see Definition 1) and that it uses only communication and preference learning costs of $O^*(\sqrt{n})$ per agent. After proving this, we give a generalization of CEDA that works when the range of systematic scores is upper bounded by a general function $g(n)$ instead of the $O^*(1)$ upper bound assumed in Section 2.4.

**Proof of Theorem 1.** We first prove that CEDA is stable with high probability in separable markets. (In other words, it succeeds in finding a stable matching with high probability.) Define $\gamma_j$ to be the $(1 - 1/\sqrt{n})$th fractile of idiosyncratic score for firm $j$, $\gamma_j = F_j^{-1}(1 - 1/\sqrt{n})$. Define $x_j$ to be firm $j$'s latent utility for the current tentative match. (This is initialized to $u_{j0}$ and increases with every tentative acceptance by the firm.) The proof relies on the two following lemmas, which we prove later:

**Lemma A.1.** *With probability at least $1 - o(1/n)$, for every firm $j$, every worker $i$ whose idiosyncratic score for the firm is at least $\gamma_j$ receives a preference signal from the firm.*

**Lemma A.2.** *With probability at least $1 - o(1/n)$, throughout the running of the CEDA protocol, we have the following invariant for every firm $j$,*

$$z_j \leq \max(0, x_j - \gamma_j). \tag{A.1}$$

Given Lemmas A.1 and A.2, we show that CEDA successfully computes the worker-optimal stable match with high probability. The argument is that, by showing what is referred to as the no-false-negatives property in Section 3.1, which is that with high probability, CEDA never prevents an application that would have been accepted. The usefulness of this property comes from the fact that when CEDA does not prevent any application that would have been accepted, then CEDA reproduces the outcome of the standard DA algorithm (Protocol 1), which is the worker-optimal stable match.

The no-false-negatives property follows from these two lemmas. This is because the only applications CEDA prevents are from worker–firm pairs $(i,j)$ in which worker $i$ does not qualify to firm $j$. By definition, this means that the worker does not systematically qualify for the firm and has not received a preference signal from the firm. Suppose that we are in the $1-o(1/n)$ fraction of preferences in which the statements in both Lemmas A.1 and A.2 are true. Then the first clause implies that $a_{ji} < z_j$, which implies that $z_j > 0$, so by inequality (A.1), $a_{ji} < z_j < x_j - \gamma_j$. The second clause implies that $\epsilon_{ji} < \gamma_j$. Together, these two clauses imply that

$$u_{ji} = a_{ji} + \epsilon_{ji} < x_j - \gamma_j + \gamma_j = x_j,$$

so worker $i$ would not have been accepted by firm $j$ even if the worker applied. This implies that with probability $1 - o(1/n)$ (with probability defined on the randomness in the preferences), CEDA produces the worker-optimal stable match.

Having established that CEDA is stable with high probability in separable markets, we now prove the $O^*(\sqrt{n})$ bounds on the average communication cost per agent as well as on the average preference learning cost per agent.

First, we bound the number of signals sent per firm. For preference signals, this is, at most, $O^*(\sqrt{n})$ because there are $O^*(1)$ bins and $2e\sqrt{n}$ signals per bin. For qualification signals, this is at most $O^*(1)$ because this is the maximum range of systematic scores for any firm, and the qualification requirement increases by one for every signal. (When the qualification requirement exceeds the maximum systematic score, then the increases in qualification requirement would also stop because there are no systematically qualified applicants.)

The bounds on the number of signals sent per firm imply a $O^*(\sqrt{n})$ bound on the number of applications received by each firm. This is because the number of applications from workers who do not systematically qualify is upper bounded by the number of preference signals. The number of applications from workers who do systematically qualify is upper bounded by $3e\log n\sqrt{n}$ times the number of qualification requirement updates. Both of these are $O^*(\sqrt{n})$.

To bound preference learning cost, observe that the only choice function queries needed in CEDA are for preference signals, applications, and responses to an application, and each of these requires exactly one choice function query. This proves the $O^*(\sqrt{n})$ per agent bound on preference learning cost. For communication cost, observe that each preference signal, application, and response to an application can be communicated in $O(\log n)$ bits as this suffices to encode the identity of an agent. Each qualification requirement signal can also be represented by $O^*(1)$ bits because each qualification requirement $z_j$ only takes integer values and has a range of $O^*(1)$ bits.

To complete the proof of Theorem 1, we prove Lemmas A.1 and A.2. Both of these use Lemma A.3, which we derive first.

**Lemma A.3.** *A random variable $Z$ is distributed according to CDF $F$ with bounded hazard rate*

$$h(x) = \frac{F'(x)}{1-F(x)} \le 1 \qquad \forall x \in \mathbb{R}.$$

*Then, for all $x \in \mathbb{R}$,*

$$\frac{\mathbb{P}(Z \ge x+1)}{\mathbb{P}(Z \ge x)} \ge \frac{1}{e}.$$

**Proof of Lemma A.3.** Define $\phi(x) = \log(1-F(x))$. Note that the bounded hazard rate implies the derivative $\phi'(x) \ge -1$. The result follows from the fact that the desired conditional probability is simply $\exp(\phi(x+1) - \phi(x))$. □

**Proof of Lemma A.1.** For a given firm $j$, suppose, on the contrary, that there exists a worker $i$ whose idiosyncratic score exceeds $\gamma_j$ but who does not get a preference signal from the firm. Then there must be at least $2e\sqrt{n}$ other workers in the same bin as worker $i$ who do receive a preference signal. This, in turn, implies that there are $2e\sqrt{n}$ other workers with an idiosyncratic score at least $\gamma_j - 1$.

Let random variable $Y$ denote the number of workers in the bin with an idiosyncratic score of at least $\gamma_j - 1$. Define $y = E[Y]$. By Lemma A.3 (with $x$ in the lemma being $\gamma_j - 1$), we get $y \le e\sqrt{n}$. Furthermore, $y \ge \sqrt{n}$ by definition of $\gamma_j$. Let $M = 2e\sqrt{n}$. Note that $2y \le M$. By Chernoff bound,

$$\mathbb{P}(Y \ge M) \le \mathbb{P}(Y \ge 2y) \le \exp\left(-\frac{1}{3}y\right) \le \exp\left(-\frac{1}{3}\sqrt{n}\right) = o\left(\frac{1}{n^2}\right). \tag{A.2}$$

Thus, after a union bound on the $O^*(1)$ bins per firm and $O(n)$ firms, we get the desired result. □

**Proof of Lemma A.2.** Fix firm $j$, the invariant on $z_j$ (see inequality (A.1)) is initially satisfied by definition as $z_j = 0 \le \max(0, x_j - \gamma_j)$. There are, at most, $O^*(1)$ increases to $z_j$ throughout the running of CEDA. Because the right-hand side of the invariant can only increase, it suffices to upper bound the probability that $z_j > x_j + \gamma_j$ after each of these increases in $z_j$.

Now, suppose that the qualification requirement $z_j$ is equal to $y$ at some point in CEDA. We show that, with probability at least $1 - 1/n^3$, after $3e\log n\sqrt{n}$ applications from systematically qualified workers, the tentative match value $x_j$ is at least $y + \gamma_j + 1$. This is because, if $x_j$ is less than this after so many applications from systematically qualified workers, then it must be that all of those workers had an idiosyncratic score less than $\gamma_j + 1$. By Lemma A.3, the chance that any worker has an idiosyncratic score at least $\gamma_j + 1$ is at least $1/(e\sqrt{n})$. (Lemma A.3 applies because of the restriction in footnote 24 regarding how the next worker applicant is picked: as a result, as analysts, we can pretend we don't know which preference signals were sent, and when a worker $i$ applies to firm $j$, we can first check if the worker is systematically qualified for that firm. If so, $\epsilon_{ji}$ remains independent of the information we know so far and is, thus, a fresh draw from distribution $F_j$.) So the chance that not one of $3e\log n\sqrt{n}$ workers had such a high idiosyncratic score is at most

$$\left(1-\frac{1}{e\sqrt{n}}\right)^{3e\log n\sqrt{n}} \le \frac{1}{n^3}.$$

This statement implies that, regardless of what the value of $y = z_j$ is at a certain time, after $3e\log n\sqrt{n}$ applications from systematically qualified workers, we have that, with probability at least $1-1/n^3$, the qualification requirement at that time is

$$z_j = y + 1 \le x_j - \gamma_j \le \max(0, x_j - \gamma_j).$$

Using this argument and counting from the first application from a systematically qualified applicant, we get that, with probability at least $1-(O^*(1))/n^3 = 1-o(1/n^2)$, the invariant is satisfied after every qualification requirement update. A union bound on the $O(n)$ firms yields the desired result. □

This completes the proof of Theorem 1. □

## Appendix B. Extension of CEDA to Larger Range of Systematic Scores

Suppose now that the range in systematic scores for a firm is upper bounded by a general function $g(n)$ instead of by the $O^*(1)$ bound assumed in Assumption 1. Finding a stable match in such a market is a more difficult problem because there is a greater variation of possible unknowns. In fact, if we allow $g(n)$ to be as high as $O^*(n)$, then one can show that, even with idiosyncratic scores satisfying the bounded hazard rate, it is possible to embed the worst-case examples of Gonczarowski et al. (2015) into our model so that any matching protocol that is stable with high probability must use at least $\Omega(n)$ bits of communication per agent.

Nevertheless, we can generalize CEDA to cases in which $g(n) = o(n)$ and still provide a $O^*(\sqrt{ng(n)})$ guarantee on communication and preference learning costs. This is nontrivial because $O^*(\sqrt{ng(n)}) = o(n)$, so this bypasses the impossibility result for arbitrary markets (Proposition 1). The protocol is as follows.

**Protocol B.1** (Generalized CEDA). *Initialize the qualification requirement for each firm $j$ to $z_j = 0$. There are two phases.*

(1) *Preference signaling: Each firm $j$ bins workers according to systematic scores into unit-ranged bins $[0,1)$, $[1,2)$, $\cdots$. For each bin, let the number of workers be $l$. The firm sends a preference signal to its top*

$$M(n,l) = \max\left((\log n)^2, 2el\sqrt{\frac{g(n)}{n}}\right) \tag{B.1}$$

*most preferred workers from the bin.*

(2) *Deferred acceptance with qualification requirement: Proceed as in step 2 of the CEDA protocol (see Section 3.1) except that we change the number of systematically qualified applications received before increasing the qualification requirement by one to*

$$3e\log n\sqrt{\frac{n}{g(n)}}.$$

Note that the only changes when generalizing CEDA are the number of preference signals sent in each bin and the number of applications to wait for before each update of the qualification requirement. The underlying framework is the same.

**Theorem B.1.** *In any separable market in which the range of systematic scores is at most $g(n)$, the generalized CEDA protocol (Protocol B.1) is a matching protocol that is stable with high probability. In the worst case, its communication cost and preference learning cost are $O^*(\sqrt{ng(n)})$ per agent.*

**Proof of Theorem B.1.** Define $\gamma_j$ to be now the $(1-\sqrt{g(n)/n})$th fractile of idiosyncratic scores. By the argument in the proof of Theorem 1, to show that the generalized CEDA protocol is stable with high probability, it suffices to prove that Lemmas A.1 and A.2 hold in this new context with this new definition of $\gamma_j$.

For Lemma A.1, the same argument works except we need to modify the Chernoff bound (A.2) to handle bins with few workers. As in the previous proof of Lemma A.1, for a particular bin with $l$ workers, define $Y$ to be the number of workers in the bin with idiosyncratic score at least $\gamma_j - 1$, and $y = E[Y]$. By Lemma A.3, $y \le e\sqrt{(g(n))/n}l$. Define $\beta = M(n,l)/y - 1$. By definition of $M$ (Equation (B.1)), $\beta \ge 1$, and $\beta y \ge 1/2M(n,l) = \Omega(\log^2 n)$. By Chernoff bound,

$$\mathbb{P}(Y \ge M(n,l)) \le \max\left(e^{-\frac{\beta^2 y}{3}}, e^{-\frac{\beta^2 y}{2+\beta}}\right) \le \exp\left(-\frac{1}{2}\frac{\beta y}{3}\right) \le \exp(-\Omega(\log^2(n))) = o\left(\frac{1}{n^3}\right).$$

As before, a union bound over all $g(n)$ bins for each firm and the $O(n)$ firms yields the desired result.

For Lemma A.2, the same argument as in the previous proof applies. Concretely speaking, suppose that the tentative match value $z_j$ is equal to some value $y$ at some point. With probability at least $1-1/n^3$, after seeing $3e\log n\sqrt{n/(g(n))}$ applications from systematically qualified workers, the tentative match value $x_j$ is at least $y+\gamma_j+1$. This is because the chance that any worker has an idiosyncratic score of at least $\gamma_j+1$ is at least $1/e\sqrt{(g(n))/n}$, and the chance that none of these systematically qualified applicants have such a high idiosyncratic score is at most

$$\left(1-\frac{1}{e}\sqrt{\frac{g(n)}{n}}\right)^{3e\log n\sqrt{\frac{n}{g(n)}}} \le \frac{1}{n^3}.$$

After establishing Lemmas A.1 and A.2 for this new setting with range of systematic score being $g(n)$, we get from the argument in the proof of Theorem 1 that the generalized CEDA protocol is stable with high probability for such markets.

As before, to prove the bound on communication and preference learning costs, it suffices to bound the number of signals and applications per agent. By definition and by the assumption that $g(n) = o(n)$, each firm sends at most $O^*(\sqrt{ng(n)})$ preference signals. This also bounds the number of applications from workers who are not systematically qualified. Each firm also sends at most $g(n)$ qualification requirement signals and, for each such signal, receives at most $O^*(n/(g(n)))$ applications from systematically qualified workers. So the total number of signals and applications is at most $O^*(\sqrt{ng(n)})$ per agent, which is what we needed to prove. □

## Appendix C. Proof of Near Optimality of CEDA (Theorem 2)

We obtain the communication lower bound in a model that is even stronger than the broadcast model. Consider the following two-party relaxation of the problem of finding a stable matching. Alice controls all the workers, and Bob controls all the firms. The workers' and firms' preferences are generated according to the model. Alice and Bob want to figure out a stable matching by communicating with each other. Note that if there is a distributed broadcasting protocol that uses a total of $B$ bits of communication, then Alice and Bob can simulate it using $B$ bits of communication: Alice will simulate all the workers' messages, and Bob will simulate all the firms' messages. Note that the converse is not true because Alice's messages are allowed to depend on the preferences of all workers simultaneously (which amounts to having "free" communication among workers). We show an example in which $\Omega(n^{3/2})$ communication between Alice and Bob is necessary, which immediately implies an $\Omega(n^{3/2}/n) = \Omega(\sqrt{n})$ lower bound on average communication per agent needed to solve the original (harder) problem.

In fact, we show our lower bound under a further restriction of the model with the workers' preferences being stochastic and similar to the firms' preferences. The construction is as follows. There are $n$ workers and $n$ firms. Let $v_{ij}$ be worker $i$'s latent utility for firm $j$ and $u_{ji}$ be firm $j$'s latent utility for worker $i$. Let both be distributed independently according to $Exp(1)$, which is the exponential distribution with rate parameter one (i.e., the systematic scores are zero, and the idiosyncratic scores are exponentially distributed). Let the value of the outside option be $(\log n)/2$ for every agent. Note that $\mathbb{P}(v_{ij} \geq (\log n)/2) = 1/\sqrt{n}$. Therefore, we expect every agent to have around $\sqrt{n}$ acceptable partners.

Let Alice be given all the workers' preferences and Bob be given all the firms' preferences. Let $\pi$ be the communication protocol, at the end of which Alice and Bob output a matching $\mu_\pi$ that is stable with probability at least 0.9 (that is, the matching protocol is successful with a high constant probability). We claim that it must be the case that the length of the protocol (i.e., the number of bits of communication) is bounded as $|\pi| = E[|\Pi|] = \Omega(n^{3/2})$, where $\Pi$ is the realization of the protocol $\pi$.

We focus only on whether a pair of agents find each other mutually acceptable (ignoring other ordinal information), such mutually acceptable pairs being the only pairs that can be matched under any stable matching (this leads to Claim C.2). If worker $i$ and firm $j$ are a mutually acceptable pair, and moreover, this is the only mutually acceptable pair of which each of $i$ and $j$ is a member, then $(i, j)$ *must* be a matched pair under any stable matching. This yields Claim C.1. We draw on ideas from information complexity theory (see, e.g., Braverman 2015), together with Claims C.1 and C.2, to establish Claim C.3 and, hence, our lower bounds. Note that our proof is short and self-contained, using only basic facts from information theory (Cover and Thomas 2012).

We define the following Boolean random variables for each worker–firm pair $(i, j)$:

$$A_{ij} := \begin{cases} 1 & \text{if } v_{ij} \geq \frac{\log n}{2}, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \qquad \text{(C.1)}$$

$$B_{ij} := \begin{cases} 1 & \text{if } u_{ji} \geq \frac{\log n}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $A_{ij} = 1$ means that worker $i$ likes firm $j$ more than the outside option, and $B_{ij} = 1$ means that firm $j$ likes worker $i$ more than the outside option. Note that all $A_{ij}$, $B_{ij}$ are distributed as Bernoulli$(\alpha)$, where $\alpha = 1/\sqrt{n}$, and are independent of each other. In addition, let $M_{ij}$ be the indicator random variable of whether worker $i$ is matched to firm $j$ under $\mu_\pi$.

**Claim C.1.** For sufficiently large $n$, we have $\mathbb{P}[M_{ij} = 1 | A_{ij} = B_{ij} = 1] > 10^{-2}$.

**Proof of Claim C.1.** Assume $A_{ij} = B_{ij} = 1$, so worker $i$ and firm $j$ find one another acceptable. By a standard Chernoff bound, the probability that worker $i$ has more than $2\sqrt{n}$ acceptable partners is at most $\exp(-\sqrt{n}/3)$. Because the probability that each of these firms finds worker $i$ to be acceptable is exactly $1/\sqrt{n}$, and the probability that another firm out of these other than $j$ finds worker $i$ acceptable is at most $1 - (1 - 1/\sqrt{n})^{2\sqrt{n}}$, which converges to $1 - e^{-2}$ for large $n$. For sufficiently large $n$, the sum of these two probabilities is no more than $1 - e^{-2.1}$, so the chance that $j$ is the unique firm that both finds $i$ acceptable and also is acceptable to $i$ is at least $e^{-2.1}$. Because the preferences of workers and firms are independent, the chance that both $i$ and $j$ are the unique mutually acceptable partners for one another is at least $e^{-4.2}$ for sufficiently large $n$. Therefore, for sufficiently large $n$,

$$\begin{aligned} &\mathbb{P}[M_{ij} = 1 | A_{ij} = B_{ij} = 1] \\ &\quad \geq \mathbb{P}[\pi \text{ outputs a stable match}] \\ &\qquad \cdot \mathbb{P}[i \text{ is with } j \text{ in all stable matches} | A_{ij} = B_{ij} = 1] \\ &\quad > (.9) \cdot e^{-4.2} \\ &\quad > 10^{-2}. \end{aligned}$$ □

**Claim C.2.** $\mathbb{P}[M_{ij} = 1 | A_{ij} = 1, B_{ij} = 0] = 0$, $\mathbb{P}[M_{ij} = 1 | A_{ij} = 0, B_{ij} = 1] = 0$, and $\mathbb{P}[M_{ij} = 1 | A_{ij} = 0, B_{ij} = 0] = 0$.

**Proof of Claim C.2.** This follows from the fact that two agents can be matched with one another in a stable matching only if both find one another acceptable (more preferable than the outside option). □

Note that Claims C.1 and C.2 together imply that $\pi$ must approximately compute the value of the AND function $A_{ij} \wedge B_{ij}$ in the sense that knowing that $M_{ij} = 1$ implies that $A_{ij} \wedge B_{ij} = 1$, and when $A_{ij} \wedge B_{ij} = 1$, we know that $M_{ij} = 1$ with at least a constant probability. Next, we make the following information-theoretic claim, which quantifies the information complexity of approximately computing the Boolean AND function (this line of reasoning is similar to Braverman et al. 2013).

**Claim C.3.** We have

$$I(A_{ij}B_{ij};\Pi) = \Omega(\alpha) = \Omega(1/\sqrt{n}). \tag{C.2}$$

Here $\Pi$ is again the random variable representing the realization of the protocol $\pi$, and $I(X;Y)$ is Shannon's mutual information, which, informally, measures the amount of information a random variable $X$ contains about a variable $Y$ (and vice versa). In terms of Shannon's entropy $H(\cdot)$, the mutual information is $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$. In other words, Alice and Bob cannot hope to even approximate the value of $A_{ij} \wedge B_{ij}$ without revealing a substantial amount of information about themselves. Note that fully revealing the values of $A_{ij}, B_{ij}$ corresponds to Shannon's entropy $H(A_{ij}, B_{ij}) = \Theta(\log n/\sqrt{n})$.[42] Let us prove Claim C.3.

**Proof of Claim C.3.** We rely on the following basic facts about protocols and about mutual information.

(1) If $A_{ij}$ and $B_{ij}$ are independent and independent of the players' other inputs, then, for each transcript realization $\pi$ of $\Pi$, the variables $(A_{ij}|\Pi = \pi)$ and $(B_{ij}|\Pi = \pi)$ are also independent. The reason is that one player speaking at a time cannot introduce a dependence between these variables (the formal proof is by induction on protocol rounds).

(2) Therefore, we have by the chain rule (see, e.g., Cover and Thomas (2012) for information theory basics):

$$\begin{aligned} I(A_{ij}B_{ij};\Pi) &= I(A_{ij};\Pi) + I(B_{ij};\Pi|A_{ij}) \\ &= I(A_{ij};\Pi) + I(B_{ij};\Pi A_{ij}) - I(B_{ij};A_{ij}) \\ &= I(A_{ij};\Pi) + I(B_{ij};\Pi A_{ij}) \\ &= I(A_{ij};\Pi) + I(B_{ij};\Pi) + I(B_{ij};A_{ij}|\Pi) \\ &= I(A_{ij};\Pi) + I(B_{ij};\Pi). \end{aligned}$$

Therefore, it will be enough to lower bound $I(A_{ij};\Pi) + I(B_{ij};\Pi)$.

(3) We can write the mutual information expression we are interested in in terms of KL-divergence as follows:

$$I(A_{ij};\Pi) = \mathbf{E}_{\pi\sim\Pi}\mathbf{D}_{KL}(A_{ij}|_{\Pi=\pi}\|A_{ij}). \tag{C.3}$$

A similar expression holds for $B_{ij}$. Again, a proof and further discussion can be found in information theory texts, such as Cover and Thomas (2012).

(4) It can be shown by direct calculation that for any constant $c < 1$, and $x < 1/2$, it is the case that for $c' < c$

$$\mathbf{D}_{KL}(\text{Bernoulli}(c' \cdot x)\|\text{Bernoulli}(x)) = \Omega_c(x), \tag{C.4}$$

where the Bernoulli random variable Bernoulli$(x)$ takes the value one w.p. $x$, and the value zero w.p. $1 - x$.

By Claims C.1 and C.2, we have that

$$\mathbb{P}[M_{ij} = 0] \geq \mathbb{P}[(A_{ij}, B_{ij}) \neq (1,1)] = (1 - \alpha^2), \tag{C.5}$$

and

$$\mathbb{P}[M_{ij} = 0, (A_{ij}, B_{ij}) = (1,1)] < \alpha^2 \cdot (1 - 10^{-2}). \tag{C.6}$$

Therefore, for a sufficiently large $n$ (and, thus, a sufficiently small $\alpha$),

$$\mathbb{P}[(A_{ij}, B_{ij}) = (1,1)|M_{ij} = 0] < \frac{\alpha^2(1 - 10^{-2})}{1 - \alpha^2} \leq \alpha^2 \cdot (1 - 9 \cdot 10^{-3}). \tag{C.7}$$

Let $\Pi_{M_{ij}=0}$ be the distribution of the history of the protocol, conditional on $M_{ij} = 0$, we have by observation 1 that

$$\begin{aligned} &\mathbf{E}_{\pi\sim\Pi_{M_{ij}=0}}\mathbb{P}[(A_{ij}, B_{ij}) = (1,1)|\Pi = \pi] \\ &\quad = \mathbb{P}[(A_{ij}, B_{ij}) = (1,1)|M_{ij} = 0], \end{aligned}$$

and, thus, by Markov's inequality

$$\begin{aligned} &\mathbb{P}_{\pi\sim\Pi_{M_{ij}=0}}[\mathbb{P}[(A_{ij}, B_{ij}) = (1,1)|\Pi = \pi] < \alpha^2 \cdot (1 - 2 \cdot 10^{-3})] \\ &\quad \geq 1 - \frac{0.991}{0.998} > 7 \cdot 10^{-3}. \end{aligned} \tag{C.8}$$

Note that $\mathbb{P}[(A_{ij}, B_{ij}) = (1,1)|\Pi = \pi] < \alpha^2 \cdot (1 - 2 \cdot 10^{-3})$ implies that either $\mathbb{P}[A_{ij} = 1|\Pi = \pi] < \alpha \cdot (1 - 10^{-3})$ or $\mathbb{P}[B_{ij} = 1|\Pi = \pi] < \alpha \cdot (1 - 10^{-3})$, and by (C.4), for any $\pi \in \Pi_{M_{ij}=0}$,

$$\mathbf{D}_{KL}(A_{ij}|_{\Pi=\pi}\|A_{ij}) + \mathbf{D}_{KL}(B_{ij}|_{\Pi=\pi}\|B_{ij}) = \Omega(\alpha). \tag{C.9}$$

By (C.5) and (C.8), the probability of such a $\pi$ is at least $7 \cdot 10^{-3} \cdot (1 - \alpha^2) > 6 \cdot 10^{-3}$ for sufficiently large $n$. The contribution of such $\pi$s to the expectation of $\mathbf{D}_{KL}(A_{ij}|_{\Pi=\pi}\|A_{ij}) + \mathbf{D}_{KL}(B_{ij}|_{\Pi=\pi}\|B_{ij})$ is, therefore, at least $\Omega(\alpha)$ because $\mathbf{D}_{KL}$ is always nonnegative, which implies by (C.3) that

$$\begin{aligned} I(A_{ij};\Pi) + I(B_{ij};\Pi) = \mathbf{E}_{\pi\sim\Pi}[&\mathbf{D}_{KL}(A_{ij}|_{\Pi=\pi}\|A_{ij}) \\ &+ \mathbf{D}_{KL}(B_{ij}|_{\Pi=\pi}\|B_{ij})] = \Omega(\alpha), \end{aligned}$$

concluding the proof. □

**Fact C.1.** If $X$ and $Y$ are independent, then $I(X,Y;Z) \geq I(X;Z) + I(Y;Z)$.

We can now conclude the proof of Theorem 2. Because $A_{ij}, B_{ij}$ are mutually independent for the different values of $(i,j)$, we get using the preceding fact and the fact that entropy is always an upper bound on mutual information,

$$\begin{aligned} H(\Pi) \geq I(A_{11}B_{11}\dots A_{nn}B_{nn};\Pi) &\geq \sum_{i=1}^{n}\sum_{j=1}^{n} I(A_{ij}B_{ij};\Pi) \\ &= n^2 \cdot \Omega(1/\sqrt{n}) = \Omega(n^{3/2}). \end{aligned}$$

(We have used Claim C.3 here.) Observing that $|\pi| \geq H(\Pi) = \Omega(n^{3/2})$ concludes the proof of the lower bound on communication cost.

The preference learning lower bound follows because, if there is a protocol of preference learning cost $R$, each time a preference oracle call is made, Alice (or Bob) can share the learned preference with the other player at cost $O(\log n)$, yielding a communication protocol with cost $C = O(R\log n)$. Therefore, $R = \Omega(n^{3/2}/\log n)$.

## Appendix D. A Generic Algorithm for Stable Matchings in Tiered Markets

In this section, we present a generic algorithm for computing stable matchings in tiered markets based on the ability to compute stable matchings between single tiers of workers and firms. One corollary of this is that the CEDA protocol in Section 3.1 can be generalized to *tiered separable* markets, which are tiered markets in which the preferences between any tier of workers and any tier of firms follow assumptions of the separable market as in Section 2.4. This implies that the $O^*(\sqrt{n})$ average communication cost of computing a stable matching can also be attained for generalizations of separable markets that allow for arbitrarily many tiers. (Recall that the assumptions on separable markets in Section 2.4 restrict it to having only a constant number of tiers.)

Before presenting the algorithm (Theorem D.1), we first define the concept of submatching. For any matching $\mu$, any subset $A \subseteq I$ of workers, and $B \subseteq J$ of firms, let $\mu(A, B)$ be the *submatching* restricted to agents $A \times B$, which is defined as $\{(i, j) \in \mu : i \in A, j \in B\}$. We say that the submatching $\mu(A, B)$ is stable if everyone prefers to be matched to their partner over being unmatched and there are no blocking pairs in $A \times B$.

**Theorem D.1.** *A stable matching in a tiered market can be constructed as follows. Initialize $\mu = \emptyset$.*

(1) *Construct a stable submatching between the top tiers $I_1$ and $J_1$. Add these matches to $\mu$.*

(2) *Remove any matched agent in the submatching as well as any unmatched agent in $I_1$ or $J_1$ who finds someone in $J_1$ or $I_1$ unacceptable. (These agents find all agents in worse tiers unacceptable.) After this, either $I_1$ or $J_1$ would have been completely removed.*

(3) *If either all of the workers or all of the firms have been removed, then return $\mu$. Otherwise repeat step 1 for the top remaining tiers on both sides.*

Moreover, every stable matching can be constructed in this way.

**Proof of Theorem D.1.** The theorem follows from the following claim, which implies that the set of stable matchings in tiered markets can be decomposed into the Cartesian product of stable submatchings for the top tiers and stable submatchings for the rest of market.

**Claim D.1.** In a tiered market, a matching $\mu$ is stable if and only if

(1) Submatching $\mu(I_1, J_1)$ is stable.

(2) Submatching $\mu(I_1 \backslash (I_1^m \cup I_1^u), J_1 \backslash (J_1^m \cup J_1^u))$ is stable, where $I_1^m \subseteq I_1$ denotes the matched workers in $\mu(I_1, J_1)$, and $I_1^u \subseteq I_1$ denotes the unmatched workers who find someone in $J_1$ unacceptable. The sets of firms $J_1^m$ and $J_1^u$ are similarly defined.

Given this claim, both directions of Theorem D.1 follow from straightforward induction. To show the first direction of this claim, assume that $\mu$ is a stable matching for the tiered market. Note that, in any stable matching $\mu$, for a fixed set $A$ of workers and fixed set $B$ of firms, the submatching $\mu(A, B)$ must be stable. Therefore, submatching $\mu(I_1, J_1)$ must be stable. Apply the rural hospital theorem on this submarket, we have that the sets $I_1^m$, $I_1^u$, $J_1^m$, and $J_1^u$ are fixed in all possible stable matchings $\mu$. Hence, the sets $I_1 \backslash (I_1^m \cup I_1^u)$ and $J_1 \backslash (J_1^m \cup J_1^u)$ are fixed and the submatching $\mu(I_1 \backslash (I_1^m \cup I_1^u), J_1 \backslash (J_1^m \cup J_1^u))$ must be stable.

For the second direction of the claim, suppose that the designated submatchings are stable; we show that $\mu$ must be stable. To do this, we need to show that workers in $I_1^u$ and firms in $J_1^u$ cannot be matched in any stable matching and that there can be no blocking pairs between workers $I_1^m$ and any firm in $J$ and no blocking pairs between firms $J_1^m$ and any worker in $I$.

First, observe that workers $I_1^u$ are unmatched in the stable submatching $\mu(I_1, J_1)$, which implies that these workers cannot be matched to anyone in $J_1$ in a stable matching for the whole market. However, because they find certain firms in the top tier $J_1$ unacceptable, they must find every firm in worse tiers unacceptable, so they cannot be matched to them either. A similar statement can be made for firms in $J_1^u$.

Now, there can be no blocking pairs between workers in $I_1^m$ and firms in $J_1$ by the fact that submatching $\mu(I_1, J_1)$ is stable. Moreover, there cannot be any blocking pairs between workers in $I_1^m$ and firms in $J \backslash J_1$ because the worker is already matched to someone of a better tier. This implies that there cannot be blocking pairs between workers in $I_1^m$ and any firm. A similar statment can be made for firms in $J_1^m$. This implies that $\mu$ is stable as desired. □

**Corollary D.1** (Generalization of the Rural Hospital Theorem). *In a tiered market, if we define the partner tier of a given agent in a matching as the tier index of the agent's matched partner (and zero if the agent is unmatched), then, for every agent, the partner tier of that agent is the same in every stable matching.*

## Appendix E. Proof of Correctness and Efficiency of the Targeted Signaling Protocol (Theorem 3)

In this section, we prove Theorem 3, which claims that the targeted signaling protocol succeeds with high probability and bounds its communication and preference learnings costs. The proof is based on studying the properties of the following mathematical object.

**Definition E.1.** In the targeted-signaling protocol, define the subgraph of signals as the collection of tuples $(i, j)$ for which either

(1) the worker $i$ signaled to firm $j$ in the signaling round or

(2) The firm $j$ signaled to the worker $i$ and $s_{k(i)} \neq t_{l(j)}$. (We do not count signals from firms that are at an equal position with their target tier, $s_{k(i)} = t_{l(j)}$.)

We break the proof of Theorem 3 into three claims.

• Claim E.1: With probability at least $1 - 18/n$, the subgraph of signals contains a stable matching.

• Claim E.2: Whenever the subgraph of signals contains a stable matching, the matching returned by the targeted signaling protocol is stable (with respect to complete preferences).

• Claim E.3: The total number of signals is at most $\Theta(n \log^3 n)$.

Claims E.1 and E.2 imply that the targeted signaling protocol succeeds with probability at least $1 - 18/n$. Claim E.3 implies the desired bounds on communication and preference learning costs. This is because, in the signaling round, sending each signal requires $O(\log n)$ communication cost and $O(1)$ preference learning cost. In the matching round, the sum of the length of everyone's partial rankings is exactly twice the total number of signals, and producing each ranking of length $k$ requires $O(k \log n)$ communication cost and $O(k)$ preference learning cost.

### Appendix E.1. Proof of Claim E.1

**Definition E.2.** Define the *tiered DA matching* as the matching produced when running the algorithm from Theorem D.1 on the tiered random market with the stable submatching between top tiers $I_1$ and $J_1$ in step 1 being produced by the following algorithm: if $|I_1| \le |J_1|$, run the worker-proposing DA algorithm in the submarket with only tiers $I_1$ and $J_1$; otherwise, run the firm-proposing DA algorithm in this submarket.

We prove Claim E.1 by proving that, with probability $1 - 18/n$, the tiered DA matching is contained in the subgraph of signals. The crux is proving the following lemma, which gives a bound on the rank obtained by a given agent in a uniformly random matching market with certain partners being unavailable. (Intuitively, the unavailable partners represent those that have been matched to better tiers in the tiered DA matching). As in the whole paper, log here denotes the natural logarithm.

**Lemma E.1.** *Consider a matching market with $m$ workers, $n \ge m$ available firms, and $u$ unavailable firms. The preferences of workers for the $n+u$ firms are uniformly random, and the preference of available firms for workers are uniformly random. The unavailable firms prefer to be unmatched. Let $N \ge 2$ be such that $N \ge n+u$. For any given worker with probability at least $1 - 9/N^2$, we have that, in the worker-proposing DA algorithm, the given worker is matched to one of the worker's top $r$ firms, where*

$$r = 24\frac{n+u}{n}\log^2(N).$$

**Proof of Lemma E.1.** First, note that, without loss of generality, $N \ge 100$ because, otherwise, $r > N$.

Label the fixed worker to be worker 1. Label the available firms 1 through $n$ and the unavailable firms $n+1$ through $n+u$. Consider the firm-optimal stable match in the submarket without worker 1 and without the unavailable firms and call this matching $\mu_1$. Note that $\mu_1$ does not depend on worker 1's preferences, nor does it depend on the preference of firms for worker 1. Define $E_1$ as the event that the total rank of workers in matching $\mu_1$ (ignoring the unavailable firms) does not exceed $R = 4e(m-1)\log N$. We lower bound the probability of $E_1$ using a proposition we prove in Appendix F about the average rank of workers in this setting (Proposition F.1). By plugging $z = 2\log N$ into Proposition F.1, we have that the probability of event $E_1$ is at least $1 - 8/N^2$.

Let $\mu_2$ be the matching formed by running the worker-proposing DA algorithm from initialization $\mu_1$. In other words, suppose that we start with everyone else matched according to $\mu_1$ and have worker 1 propose to worker 1's top choice as in the DA algorithm. This may cause a previously matched worker to be rejected from a firm, and we will have this worker apply to the worker's next choice, which may result in a chain of rejections leading to someone applying to one of the $n-m+1$ unmatched available firms. $\mu_2$ is a stable matching (with respect to the entire market). Because the rank of worker 1 in $\mu_2$ is no better than in the worker-optimal stable match, it suffices to upper bound the rank obtained by worker 1 in $\mu_2$.

First, let us make a few structural observations on $\mu_1$ and $\mu_2$ under event $E_1$. For each firm $j \le n$, let $B_j$ be the set of workers who weakly prefer firm $j$ to their partner in $\mu_1$. (We use the letter "B" because this is the set of workers who want to block with $j$ in $\mu_1$.) In $\mu_1$, firm $j$ is matched to the firm's favorite worker in $B_j$. Furthermore, the sum $\sum_{j=1}^n B_j \le R = 4e(m-1)\log N$ as this sum always equals the total rank obtained by workers in $\mu_1$ (ignoring the unavailable firms). Now, consider running the DA algorithm with initialization $\mu_1$, drawing only as needed the preference of firms for worker 1. When worker 1 applies to an available firm $j$, the probability that worker 1 is accepted is exactly $p_j = 1/(1+|B_j|)$ because this is worker 1's chance of being the firm's favorite worker among a set of size $1 + |B_j|$. If the worker is rejected, then the worker applies to the worker's next choice, and the same formula for acceptance probability applies. If the worker is accepted, then this triggers a rejection chain that ends with one of the $n-m+1$ unmatched available firms. Note that this rejection chain can never circle back to firm $j$ and cause worker 1 to be rejected because that would contradict the assumption that $\mu_1$ is the firm-optimal stable match in the submarket without worker 1. For the unavailable firms $j \ge n$, define $p_j = 0$. We have that by Jensen's inequality,

$$\sum_{j=1}^{n+u} p_j = \sum_{j=1}^{n} p_j \ge \frac{n}{4e\log N + 1}. \tag{E.1}$$

Now, conditional on event $E_1$, let $P$ be the probability that worker 1 is not matched to worker 1's top $r$ firms in the worker-proposing DA algorithm. Let $A$ be all subsets of the $n+u$ firms of cardinality $\lfloor r \rfloor$. We have,

$$P \le \frac{1}{|A|}\sum_{S\in A}\prod_{j\in S}(1-p_j) \tag{E.2}$$

$$\le \left(1 - \frac{\sum_{j=1}^n p_j}{n+u}\right)^r \tag{E.3}$$

$$\le \exp\left(-\frac{nr}{(n+u)(4e\log N + 1)}\right) \tag{E.4}$$

$$< \frac{1}{N^2}. \tag{E.5}$$

Inequality (E.2) follows from the independence between the preference of worker 1, the event $E_1$, and each firm $j$'s preference for worker 1. Inequality (E.3) follows from the fact that the sum of product in the inner part of Equation (E.2) increases if we replace any two different $p_i$, $p_j$ with their average $(p_i + p_j)/2$, so the maximum is attained when all of them are equal. Inequality (E.4) follows from inequality (E.1) and the bound $(1-x) \le \exp(-x)$, and inequality (E.5) follows from the fact that when $N \ge 100$, we have $(24\log^2 N)/(4e\log N + 1) > 2\log N$.

Because the probability that $E_1$ does not occur is at most $8/N^2$, we have that the total probability that worker 1 does not get one of worker 1's top $r$ choices is at most $8/N^2 + 1/N^2 = 9/N^2$, which is what we needed. □

Claim E.1 follows from Lemma E.1 and taking a union bound over the $m + n \le 2n$ agents, observing that, in the tiered DA matching, every agent is in a scenario described by Lemma E.1. So, with probability at least $1 - 18/N$, the tiered DA matching is contained in the subgraph of signals.

### Appendix E.2. Proof of Claim E.2

We break the proof of Claim E.2 into two parts.

(1) Claim E.2(a). Whenever the subgraph of signals contains a stable matching, running worker-proposing DA with preferences restricted to this subgraph returns a matching $\mu$ that is stable with respect to complete preferences.

(2) Claim E.2(b). Whenever this happens, the targeted signaling protocol returns the same matching $\mu$.

We first show Claim E.2(b). Let $\mu$ denote the matching that arises from running worker-proposing DA with preferences restricted to the subgraph of signals. If $\mu$ is stable (with respect to complete preferences), then it must match every worker $i$ who has a target tier: that is, there exists firm tier $l$ such that $t_l \ge s_{k(i)}$. Now, observe that the output of the targeted signaling protocol is simply running the worker-proposing DA with respect to the subgraph of signals plus certain additional edges. Precisely speaking, these edges are the ones ruled out in Definition E.1, which are tuples $(i, j)$, where $j$ signaled to $i$ and they $s_{k(i)} = t_{l(j)}$. Because the DA algorithm without these edges ended up matching every one of these agents $i$ incident to one of these edges, and the favorite partner of agent $i$ are already included in the subgraph of signals, running DA with these edges would not change the result. Therefore, the targeted signaling protocol returns $\mu$.

Claim E.2(a) follows from the following structural result (Lemma E.2) on the subgraph of signals. First, let us give a few definitions. For any subgraph (defined as a collection of worker-firm tuples $(i, j)$), we say that a matching $\mu$ is *stable with respect to the subgraph* if every matched agent in $\mu$ prefers to be matched than unmatched and there are no blocking pairs to $\mu$ within the subgraph. Define the *complete graph* as the Cartesian product $I \times J$. (The original definition of stability is equivalent to stability with respect to the complete graph.) Define a matching to be *full* if it has cardinality $\min(n_I, n_J)$, which corresponds to matching all agents of at least one of the sides.

**Lemma E.2.** *Any matching $\mu$ that is full and stable with respect to the subgraph of signals is stable with respect to the complete graph.*

To see why Claim E.2(a) follows from Lemma E.2, note that, when the subgraph of signals contains a stable matching, then running worker-proposing DA with preferences restricted to the subgraph returns a matching that is full and that is stable with respect to the subgraph. Lemma E.2 implies that this matching is stable with respect to the complete graph.

**Proof of Lemma E.2.** Let the subgraph of signals be $G$ and let $\mu$ be a full matching stable with respect to $G$; we show that $\mu$ is stable with respect to the complete graph. It suffices to show that $\mu$ does not contain any blocking pairs.

We show that no tuple $(i, j)$ in the complete graph can be a blocking pair. Let worker $i$ be in worker-tier $k$ and firm $j$ be in firm-tier $l$. First, any $(i, j) \in G$ cannot be a blocking pair by the definition of $\mu$ being stable in $G$.

Suppose first that $s_k \ne t_l$. Without loss of generality, let $s_k < t_l$. In this case, worker $i$ must be matched in $\mu$, say, to firm $j'$. Let $\tilde{l}$ be the target tier of worker $i$ and let $l'$ be the tier of firm $j'$. Note that $l' \le \tilde{l}$ because $G$ can only contain edges between $i$ and weakly better tiers than $\tilde{l}$. Moreover, we have $\tilde{l} \le l$ because $\tilde{l} = \min\{l : t_l \ge s_k\}$ by definition. Combining the two inequalities, we have $l' \le \tilde{l} \le l$. Suppose that $l' < l$, then $i$ would prefer $j'$ to $j$, so $(i, j)$ cannot be a blocking pair. Suppose that $l' = l$, then both must equal $\tilde{l}$, which is the target tier of workers $I_k$. This means that the target tier of firms $J_l$ must be a strictly worse tier of workers than $I_k$. (Otherwise, we would need $s_k = t_l$.) So the only reason that $(i, j') \in \mu$ is in the subgraph of signals $G$ is that $i$ signals to $j'$. But $i$ does not send a signal to $j$, and both $j$ and $j'$ are in the same tier, so $i$ must prefer $j'$ to $j$, so $(i, j)$ cannot be a blocking pair.

The only remaining case is $s_k = t_l$. In this case, the subgraph of signals (Definition E.1) only includes signals from $I_k$ to $J_l$. Thus, $i$ must be matched in $\mu$, say to $j'$, and we have that $j'$ is either in a better tier than $j$ or is signaled to by $i$. In either cases, $i$ must prefer $j'$ to $j$, so $(i, j)$ cannot be a blocking pair.

Because there are no blocking pairs, $\mu$ must be stable with respect to the complete graph. □

### Appendix E.3. Proof of Claim E.3

We complete the proof of Theorem 3 by proving Claim E.3.

**Lemma E.3.** *The total number of signals sent during the signaling round of the targeted signaling protocol is at most $60n\log^3 n$.*

**Proof of Lemma E.3.** The result is trivially true if $n \le e^4 < 60$ because the average number of signals is at most $n$. Assume now that $\log n \ge 4$.

Define a *grouping* of tiers as the set of all tiers of agents that share the same target tier. For example, if the shared target tier is firm-tier $J_l$, then the grouping is all worker-tiers $I_k$ such that $t_{l-1} < s_k \le t_l$.

Consider an arbitrary grouping. Without loss of generality, let the shared target tier be $J_l$ as before. Let $\mathcal{A}_k = \{k : t_{l-1} < s_k \le t_l\}$ as before. Let the minimum and maximum element of $\mathcal{A}_k$ be $k_0$ and $k_1$, respectively. For each $k \in \mathcal{A}_k$, define $r_k$ as the target number of workers in tier $I_k$ (see Definition 8).

Let the total number of signals sent by this grouping be $\sigma = \sum_{k=k_0}^{k_1} m_k r_k$. Note first that $r_{k_0} = 24\log^2 n$ because the competitiveness of workers (see Definition 7) in $I_{k_0}$ must be one. For $\mathcal{A}_k \backslash \{k_0\}$, we have

$$
\begin{aligned}
\sum_{k=k_0+1}^{k_1} m_k r_k &= 24 n_l \log^2 n \sum_{k=k_0+1}^{k_1} \frac{m_k}{t_l - s_{k-1}} \\
&\le 24 n_l \log^2 n \sum_{k=k_0+1}^{k_1} \left( \frac{1}{t_l - s_{k-1}} + \cdots + \frac{1}{t_l - s_k + 1} \right) \\
&\le 24 n_l \log^2 n \left( \frac{1}{t_l - s_{k_0}} + \frac{1}{t_l - s_{k_0} - 1} + \cdots + \frac{1}{2} + \frac{1}{1} \right) \\
&\le 24 n_l \log^2 n \log(t_l - s_{k_0} + 1) \\
&\le 24 n_l \log^3 n
\end{aligned}
$$

This shows that

$$\sigma \le 24 m_{k_0} \log^2 n + 24 n_l \log^3 n \le (6 m_{k_0} + 24 n_l) \log^3 n.$$

From this, the desired result follows because the sum of all possible $m_{k_0}$ is at most $2n$, and the same holds for the sum of all possible $n_l$. □

## Appendix F. Average Rank in Unbalanced Matching Market

The proof of Lemma E.1 requires a concentration bound for the average rank obtained by workers in any stable matching in a uniformly random matching market with strictly more firms than workers. When there are $n-1$ workers and $n$ firms, Ashlagi et al. (2017) implies that this is asymptotically $\log n$ as $n \to \infty$. Compared with their result, the following proposition gives up a constant factor of $2e$ in the asymptotics but obtains a stronger probabilistic guarantee that holds for any $n$. The proof builds on the techinques from Pittel (1989, 1992), which are based on an integral formula from Knuth (1976).

**Proposition F.1.** *Consider a uniformly random matching market with $n-1$ workers and $n$ firms. For any $z \geq 2$, we have that, with probability at least $1-8\exp(-z)$, the average rank obtained by workers in any stable matching is no more than*

$$\bar{r} = e(2\log n + z).$$

**Proof of Proposition F.1.** Define $m = n-1$. Let $P_0$ be the probability that a uniformly random matching market with $m$ workers and $n$ firms has a stable matching in which the average rank of workers is greater than $\bar{r}$. We upper bound $P_0$ by $8\exp(-z)$. Observe that this is trivially true if $m \leq 7$ because $\bar{r} \geq 3e > 8$, so we assume from now on that $m \geq 8$.

As in Pittel (1989), define the standard matching $\mu_0$ as the matching in which for each $1 \leq i \leq m$, worker $i$ is matched to firm $i$. The unmatched firms are denoted by indices $j > m$. For each tuple $(i,j)$, $1 \leq i \leq m$, $1 \leq j \leq n$, let $X_{ij}$ and $Y_{ij}$ be i.i.d. draws from Uniform$[0,1]$. These values induce the preferences of workers and firms as follows: the smaller the value of $X_{ij}$, the more worker $i$ prefers firm $j$. Similarly, the smaller the value of $Y_{ij}$, the more firm $j$ prefers worker $i$. For simplicity, define $x_i = X_{ii}$ and $y_i = Y_{ii}$ for $1 \leq i \leq m$. We call the matrices $X$ and $Y$ the cardinal preferences of workers and firms and the vectors $x$ and $y$ the matching values of workers and firms.

Adapting equation 2.2 of Pittel (1989), we have that given the matching values $x$ and $y$, the probability that the standard matching is stable and the total rank of workers equals $R$ is exactly

$$[\xi^{R-m}]\left\{\prod_{i=1}^{m}(1-x_i)\prod_{1\leq i\neq j\leq m}(1-x_i+\xi x_i(1-y_j))\right\}, \tag{F.1}$$

where $[\xi^a]\{f(\xi)\}$ denotes the coefficient of $\xi^a$ in the expansion of polynomial $f(\xi)$. This formula is analogous to equation 2.2 of Pittel (1989), and we provide a brief explanation. The rank obtained by each worker is exactly one plus the number of firms the worker wants to block with, so the total rank of workers is $R$ if and only if the total number of firms workers want to block with is $R-m$, counting with multiplicity. The expression in the braces computes the probability that the standard matching $\mu_0$ is stable while keeping track of who wants to block with whom using dummy variable $\xi$. The expression is a product of various terms, and is based on the i.i.d. assumptions of entries of $X$ and $Y$. In the first product, $(1-x_i)$ is the probability that worker $i$ does not want to block with the unmatched firm as $P(X_{in} > X_{ii}) = 1-x_i$. In the second product, we have the linear combination of two terms: $1-x_i$ is the probability that worker $i$ does not want to block with firm $j$, and $x_i(1-y_j)$ is the probability that worker $i$ wants to block with $j$ but $j$ does not reciprocate. Expanding the product as a polynomial in $\xi$ and examining the coefficient of $\xi^{R-m}$ obtains exactly the probability that the matching is stable and the total rank of workers is $R$.

Define $A = \{(x,y) : 0 \leq x_i, y_i \leq 1, 1 \leq i \leq m\}$. We have

$$P_0 \leq n! \int_A \sum_{R=\lfloor m\bar{r}\rfloor+1}^{\infty}[\xi^{R-m}]\left\{\prod_{i=1}^{m}(1-x_i)\prod_{1\leq i\neq j\leq m}(1-x_i + \xi x_i(1-y_j))\prod_{i=1}^{m}\right\}dx\,dy \tag{F.2}$$

$$\leq n!\int_A \inf_{\xi\geq 1}\left\{\xi^{m(1-\bar{r})}\exp(-m\sum_{i=1}^{m}x_i + \xi\sum_{1\leq i\neq j\leq m}x_i(1-y_j))\right\}dx\,dy. \tag{F.3}$$

Inequality (F.2) follows from integrating Equation (F.1) over the uniform distribution of matching values; then using a union bound over all $n!$ matchings based on symmetry. Inequality (F.3) comes from the Chernoff method of bounding the tail of power series[43] and using the fact that $\lfloor m\bar{r}\rfloor + 1 \geq m\bar{r}$.

Define $\mu = \bar{r}/e = 2\log n + z$. Partition the region of integration $A$ into $A_1 = \{(x,y) \in A : \sum_{i=1}^{m} x_i \geq \mu\}$, and $A_2 = A\backslash A_1$. Define this integral in regions $A_1$ and $A_2$ to be $P_1$ and $P_2$, respectively. It suffices to bound $P_1$ and $P_2$. For convenience, define $s = \sum_{i=1}^{m} x_i$.

To bound $P_1$, we set $\xi = 1$. Let $s_i = s - x_i$, $\Psi(x) = \int_0^1 \exp(-xy)dy = (1-\exp(-x))/x$, and $A_1^x = \{x : 1 \leq x_i \leq 1, \sum_{i=1}^{m} x_i \leq \mu\}$. For clarity, we write a series of inequalities and explain them one by one afterward.

$$P_1 \leq n!\int_{A_1}\exp(-s - \sum_{i=1}^{m}s_i y_i)\,dx\,dy \tag{F.4}$$

$$= n!\int_{A_1^x}\exp(-s)\prod_{i=1}^{m}\Psi(s_i)\,dx \tag{F.5}$$

$$\leq e^2 n!\int_{A_1^x}\exp(-s)\frac{1}{s^m}\,dx \tag{F.6}$$

$$\leq e^2 n!\int_{\mu}^{m}\exp(-s)\frac{1}{s^m}\frac{s^{m-1}}{(m-1)!}\,ds \tag{F.7}$$

$$\leq e^2 n(n-1)\int_{\mu}^{m}\exp(-s)\,dx. \tag{F.8}$$

$$\leq e^2\exp(-z) \tag{F.9}$$

Inequality (F.4) follows from plugging in $\xi = 1$ into inequality (F.3). In Equation (F.5), we integrate with respect to each $y_i$. In inequality (F.6), we make use of the fact that for each $a > 0$,

$$(\log\Psi(a)) = \frac{1}{\exp(a)-1} - \frac{1}{a} \geq -\frac{2}{a+2}.$$

So

$$\begin{aligned}\log\left(\prod_{i=1}^{m}\Psi(s_i)\right) &\le 2+\sum_{i=1}^{m}\left(\log(\Psi(s_i))-\frac{2x_i}{s+1}\right)\\ &\le 2+\sum_{i=1}^{m}\left(\log(\Psi(s_i))-\frac{2x_i}{s_i+2}\right)\\ &\le 2+\sum_{i=1}^{m}\log(\Psi(s_i+x_i))=2+m\log(\Psi(s))\\ &\le 2-m\log(s).\end{aligned}$$

In inequality (F.7), we use a standard change of variable from $x$ to $s$ (see equation 3.2 and inequality 3.3 of Pittel 1989). In inequality (F.8), we use the fact that $1/s$ is decreasing in $s$ and that $\mu \ge 1$. In inequality (F.9), we integrate out $s$ and make use of the definition of $\mu$ and $z \ge 1$.

To bound $P_2$, we set $\xi = e\mu/s$. As before, we state a series of inequalities and explain them afterward.

$$\begin{aligned}P_2 &\le n!\int_{A_2}\left(\frac{e\mu}{s}\right)^{m(1-e\mu)}\exp(-ms+e(m-1)\mu)\,dx\,dy &\text{(F.10)}\\ &\le n^2\int_0^{\mu}\left(\frac{e\mu}{s}\right)^{m(1-e\mu)}\exp(-ms+e(m-1)\mu)s^{m-1}\,dx\,dy &\text{(F.11)}\\ &\le n^2\int_0^{\mu}\exp(-(m+e)\mu+m)\mu^{m-1}\,ds &\text{(F.12)}\\ &= [n^2\exp(-\mu)][\exp(-(m+e-1)\mu)\mu^m]\exp(m) &\text{(F.13)}\\ &\le \frac{1}{e}\exp(-z). &\text{(F.14)}\end{aligned}$$

Equation (F.10) substitutes in $\xi = e\mu/s$ to inequality (F.3) and uses the fact that $\sum_{1\le i\ne j\le m} x_i(1-y_j) \le s(m-1)$. Inequality (F.11) again integrates out each $y_i$ and uses the change of variable from $x$ to $s$ as in inequality 3.3 of Pittel (1989). Inequality (F.12) uses the fact that the function $f(s) = \exp(-ms)s^{em\mu-1}$ is increasing in $[0,\mu]$. Equation (F.13) simplifies the formula and arranges into groups, denoted by square brackets. Inequality (F.14) comes from bounding each group. We bound the first group by $\exp(-z)$ using the formula for $\mu$. We bound the second group using the observation that the function $f(\mu) = \exp\cdot(-(m+e-1)\mu)\mu^m$ is maximized when $\mu = m/(m+e-1)$, so $f(\mu)\exp(m) \le (m/(m+e-1))^m = (1-(e-1)/(m+e-1))^m \le \exp(-((e-1)m)/(m+e-1)) < \exp(-1)$ because $m \ge 8$.

Combining inequalities (F.9) and (F.14), we have

$$P_0 = P_1 + P_2 \le \left(e^2 + \frac{1}{e}\right)\exp(-z) < 8\exp(-z),$$

which completes the proof. □

**Remark F.1.** The preceding proof can be modified to show the following statement. Consider a uniformly random matching market with $m$ men and $n = m + d$ women, in which $1 \le d \le (e-1)m$. For any $z \ge 2$, with probability at least $1 - 8\exp(-z)$, the average rank of men in any stable matching is no more than

$$\bar{r} = e\left(\log\left(\frac{n}{d}\right) + \frac{\log m + z}{d} + 1\right).$$

## Appendix G. Proofs of Incentive Properties (Theorems 4 and 5)

In a matching protocol, we say that the agent *complies* with the protocol if the agent truthfully participates according to what is prescribed, and that the agent *deviates* from the protocol otherwise.

Before proving that both protocols in this paper are as incentive compatible as DA with high probability (Theorem 4), we establish two lemmas. This allows us to prove the desired properties for both protocols simultaneously under one framework.

Define the subgraph of signals in the targeted-signaling protocol as in Definition E.1. For CEDA, define the subgraph of signals as follows.

**Definition G.1.** In the CEDA protocol, for a given realization of preferences, define the subgraph of signals as the set of tuples $(i, j)$ for which worker $i$ applies to $j$ at some point during the protocol.

Let $G$ be the subgraph of signals in either protocol, assuming everyone complies. In either protocol, define the worker-optimal stable match *restricted to subgraph G* as the result of running the worker-proposing DA algorithm using only preference information within pairs of agents in $G$. Similarly define the firm-optimal stable match restricted to $G$. Furthermore, we say that a matching is stable with respect to $G$ if it is individually rational[44] and there are no blocking pairs $(i, j) \in G$ with respect to true preferences. A matching is stable with respect to complete preferences if it is stable with respect to the complete graph $I \times J$.

**Lemma G.1.** *In either CEDA (Protocol 2) or targeted signaling (Protocol 3), let the subgraph of signals be G. The produced matching $\mu$ is not blocked by any edge $(i, j) \in G$. Then we have, with high probability,*

(1) *The worker-optimal stable match restricted to G (defined assuming everyone complies) is stable with respect to complete preferences.*

(2) *The firm-optimal stable match restricted to G is also a stable with respect to complete preferences.*

**Proof of Lemma G.1.** In CEDA, the outputted matching $\mu$ is simply the result of running the worker-proposing DA on $G$, so $\mu$ is not blocked by any edge $G$. In the targeted signaling protocol, the outputted matching is the result of running the worker-proposing DA on a graph that contains[45] $G$, so $\mu$ must not be blocked by any edge of $G$.

In CEDA, with high probability, the outputted matching is stable with respect to complete preferences (Theorem 1). When this happens, the worker-optimal stable match restricted to $G$, which is $\mu$, is stable. The firm-optimal stable match restricted to $G$ is also $\mu$. (To see this, note that, by definition of $G$, every firm in $\mu$ gets their favorite partner in $G$.)

In the targeted signaling protocol, with high probability, the subgraph of signals contains a stable matching by Claim D.1. When this happens, the result of the worker-optimal and firm-optimal DA must be full[46] and stable with respect to the subgraph of signals. Therefore, by Lemma E.2, both matchings are stable with respect to complete preferences. □

We are now ready to prove Theorems 4 and 5.

**Proof of Theorem 4.** We prove that both protocols are as incentive compatible as DA with high probability. In either protocol, let $G$ be the subgraph of signals when everyone complies.

Suppose, to the contrary, that a certain agent can unilaterally deviate from the protocol and cause the resultant matching to be $\mu'$, which gives the deviating agent someone other than the agent's best stable partner under complete, true preferences. By Lemma G.1, we have that, with high probability, the agent obtains a better partner in $\mu'$ than what the agent gets in either the worker- or firm-optimal stable match restricted to $G$ (defined previously with respect to true preferences). Note that regardless of the agent's deviation, $\mu'$ must be individual rational for all agents. We apply the following blocking lemma from Gale and Sotomayor (1985) (the version here is from Roth and Sotomayor 1990, lemma 3.5).

**Lemma G.2** (Blocking Lemma, Gale and Sotomayor 1985). *Let $\mu_W$ be the worker-optimal stable match. Let $\mu$ be any individually rational matching with respect to strict preferences $R$ and let $I'$ be all workers who prefer $\mu$ to $\mu_W$. If $I'$ is nonempty, there is a worker–firm pair $(i, j)$ that blocks $\mu$ such that $i \notin I'$ and $\mu(j) \in I'$.*

The strict preferences $R$ we use for this lemma are the true preferences restricted to $G$, treating everything else as unacceptable. In other words, worker $i$'s preference is induced by her true preferences over $\{j : (i, j) \in G\}$, and if $(i, j) \notin G$, then she treats firm $j$ as unacceptable. Similarly define preferences of firms. The matching we use is the $\mu'$.

An implication of the lemma is that there exists $(i, j) \in G$ such that

(1) neither $i$ or $j$ is the same as the deviating agent, so they are assumed to be complying to the protocol;

(2) $(i, j)$ blocks $\mu'$ (under the true preferences of $i$ and $j$).

Let $G'$ be the subgraph of signals under the deviating action by the agent. (Regardless of the deviation, this is well defined at the end of both protocols). Because $(i, j)$ blocks $\mu'$, $(i, j) \notin G'$ by Lemma G.1. However, $(i, j) \in G$ by construction. We show that this discrepancy can happen with vanishing probability in either protocol, thus proving the desired result. The common idea is that the deviating agent has limited control over the edges of the subgraph of signals involving complying agents only.

In the targeted signaling protocol, $(i, j) \in G$ implies that $(i, j) \in G'$ because both $i$ and $j$ send the same set of signals in the signaling round regardless of what the deviating agent does. This proves that targeted signaling is as incentive compatible as DA with high probability.

In the CEDA protocol, $(i, j) \in G$ implies that $(i, j) \in G'$ with high probability. This is because, by construction, both $i$ and $j$ prefer each other to their partners in $\mu'$. Suppose, on the contrary, that $(i, j) \notin G'$, then $i$ skipped firm $j$ in her application decision, which implies that the firm must not have sent her a preference signal and that she must not have systematically qualified for the firm, $a_{ji} < z_j$, where $z_j$ is the qualification requirement of firm $j$ at the end of the protocol. Let $x_j$ be the matched utility of firm $j$ at the end of the protocol. The fact that $j$ prefers $i$ to its matched partner in $\mu'$ implies

$$x_j < u_{ji} = a_{ji} + \epsilon_{ji}.$$

This means that either $\epsilon_{ji} > \gamma_j$ or $z_j > a_{ji} > x_j - \gamma_j$. In other words, either it's the case that $\epsilon_{ji} > \gamma_j$ but firm $j$ does not send a preference signal to $i$, or it's the case that $z_j > \max(0, x_j - \gamma_j)$ for some firm $j$ at the end CEDA. But because firm $j$ (by construction) complies to the protocol, the first case happens with vanishing probability by Lemma A.1. The second also happens with vanishing probability by the proof Lemma A.2. The key is to note that, for a compliant firm, the invariant proved in Lemma A.2 that $z_j \le \max(0, x_j - \gamma_j)$ holds with high probability by a statistical argument based on the randomness of idiosyncratic scores and is independent of the application decisions of agents and also independent of whatever strategic actions that influence applications. This proves that CEDA is as incentive compatible as DA with high probability. □

**Proof of Theorem 5.** By Theorem 4, it suffices to show that, under general imbalance and when the number of agents of every tier is large, the probability that any given agent has multiple stable partners is vanishing. By the decomposition result in Theorem D.1, the desired result follows from the following lemma, which is based on techniques from Ashlagi et al. (2017). □

**Lemma G.3.** *There exists a nonincreasing function $\delta : \mathbb{N} \to \mathbb{R}$ such that $\delta(y) \to 0$ as $y \to \infty$ with the following property. In a matching market with a single tier of $m$ workers and a single tier of $n \ge m+1$ firms, the probability that any given agent, conditional on being matched, has multiple stable partners is upper bounded by $\delta(n)$.*

**Proof of Lemma G.3.** Consider two cases. Suppose that $m \ge n/2$, then by lemma B.1.(ii) in the online appendix of Ashlagi et al. (2017), there exists $m_0$ such that, for all $m > m_0$, with probability at least $1 - \exp(-\log^{0.4} n)$, the number of workers with multiple stable partners is no more than $m/(\log^{0.5} m)$, and the same statement holds for firms. This implies that, for $n \ge 2m_0$, the probability that any given worker has multiple stable partners is at most

$$\exp(-\log^{0.4} n) + \log^{-0.5}\left(\frac{n}{2}\right).$$

Similarly, the probability that a firm, conditional on being matched, has multiple stable partners is also upper bounded by this. Define function $\delta_1 : \mathbb{N} \to \mathbb{R}$ that has $\delta(n) = 1$ for $n < 2m_0$ and as the previous quantity when $n \ge 2m_0$. Then function $\delta_1$ satisfies the desired result in the region when $m \ge n/2$.

Suppose now that $m < n/2$. Then we show that the probability that a given matched agent has multiple stable partners is still small. Consider the outcome of the worker-proposing DA algorithm. Consider any firm that is matched in this worker-optimal stable match (WOSM). Let $u = n - m \ge n/2$. As in Ashlagi et al. (2017), suppose that the firm has multiple stable partners; then the chain of proposals triggered by the firm rejecting its current partner must come back to this firm before going to the $u$ unmatched firms. The chance that this happens is at most $1/u$. Moreover, if the firm has more than two stable partners, then when the firm rejects the second stable partner, the chain of proposals triggered also needs to come back to the firm rather than go to the $u$ unmatched firms. So the number of stable partners of this firm is stochastically dominated by Geometric($2/n$). For $n \geq 11$, the expectation of this is less than $3/n$. Thus, the expected total number of worker–firm pairs that can be in a stable match and that is not already in the WOSM is at most $3m/n$ for any $n \geq 11$. By symmetry, for any agent, conditional on being matched, the chance the agent has multiple stable partners is no more than $3/n$ when $n \geq 11$. Define $\delta_2 : \mathbb{N} \to \mathbb{R}$ to be $\delta(n) = 1$ for $n < 11$ and $\delta(n) = 3/n$, and we have that $\delta_2$ satisfies the desired result in the region when $m < n/2$.

Finally, we note that $\delta = \max(\delta_1, \delta_2)$ satisfies the desired property for any $n \geq m + 1$. □

## Appendix H. Impossibility of Two-Round Protocol for General Separable Markets

In this section, we show a simple example of a separable market for which no two-round protocol that uses $o(n)$ bits of communication per agent computes a stable matching with high probability.

For clarity of exposition, the example is described with certain agents preferring any partner of a specific type over any partner of another type. This can be approximated with high probability with idiosyncratic scores distributed as Exp(1) by having a systematic score difference of $3 \log n$ for the type that one prefers.

**Example H.1.** Consider an academic job market with two types of departments and two types of applicants. The types are teaching-focused departments and applicants, and research-focused departments and applicants. An agent prefers to be matched with a partner of the agent's own focus, and otherwise, preferences are drawn uniformly at random. Research-focused departments and teaching-focused candidates are in short supply: there are $n$ research-focused departments, $n + 2$ teaching-focused departments, $n + 2$ research-focused applicants, and $n$ teaching-focused applicants.

In this example, in any stable matching, there are two research-focused applicants and two teaching-focused departments who are matched with each other. However, a priori, the chance that any two agents of different focus are matched in a stable matching is $2/n$. So, in a two-round protocol, it's never worthwhile for agents to signal across their own focus. The communication-efficient method is to first let research-focused departments and teaching-focused applicants pick their partners in a two-round protocol and then run an additional aftermarket to match the remaining agents.

## Appendix I. Optimal Communication Cost in Tiered Random Markets

The targeted signaling protocol in Section 4.4 uses $\Theta(\log^4 n)$ bits of communication per agent. In this section, we show that the best possible is $\Theta(\log^2 n)$.

Consider the protocol based on the generic algorithm for tiered markets in Theorem D.1, in which, in step 1, if $|I_1| \leq |J_1|$, we simulate the worker-proposing DA algorithm, and if $|I_1| > |J_1|$, then we simulate the firm proposing. Because we only allow private messages, each time an agent proposes to a partner, the agent does not know whether the partner is taken by agents from better tiers. This results in wasted proposals. Nevertheless, one can show that the number of wasted proposals is not too high. In fact, one can show that, with high probability, this protocol terminates with a stable matching using $\Theta(n \log n)$ proposals. This shows that there exists a stable matching protocol that only uses private messages and that succeeds with high probability, using communication cost $\Theta(\log^2 n)$ per agent and preference learning cost of $\Theta(\log n)$ per agent.

The following shows that this bound on average communication cost is the best possible.

**Theorem I.1** (Lower Bound on Communication Cost with Private Messages). *Consider a uniformly random market with $n - 1$ workers and $n$ firms. In such a market, any matching protocol that is stable with high probability and only uses private messages must incur a communication cost of $\Omega(\log^2 n)$ bits per agent.*

The following theorem says that, in certain tiered random markets, some agents must experience up to $\sqrt{n}$ bits of communication even though the average is only polylogarithmic in $n$.

**Theorem I.2** (Lower Bound on Agent-Specific Communication). *Consider a market consisting of $n - 1$ workers, each worker in a tier of her own, and $n$ firms all in a single tier. Consider the worker in the $\sqrt{n}$th worst tier. In any matching protocol that is stable with high probability and only uses private messages, this worker must incur a communication cost of at least $\Omega(\sqrt{n})$ bits.*

We now prove the theorems in turn.

**Proof of Theorem I.1.** The market in question is the same as that in Ashlagi et al. (2017). It suffices to prove the lower bound for any one worker because all workers are ex ante the same. Some features of this market that we know from Ashlagi et al. (2017) are

- In all stable matchings, the average rank of workers for their matched firms is very close to $\log n$, and a vanishing fraction of workers (firms) have multiple stable partners. (The average rank of workers for their matched firms is at least $0.99n/\log n$.)
- Fix a worker $i$. With high probability (whp), the worker has a unique stable partner. Conditioned on the stable partner being unique, the distribution of worker $i$'s rank of her stable partner is asympototically close to Geometric($1/\log n$). In particular, the conditional probability that the unique stable partner is one of worker $i$'s top $\log n$ most preferred firms is $p \in (1 - 1/e - 0.01, 1 - 1/e + 0.01)$ for large enough $n$.
- Run the worker-proposing deferred acceptance algorithm with worker $i$ excluded. Whp, all but $n^{0.99}$ firms receive between $0.9 \log n$ and $1.1 \log n$ proposals from workers.

Our proof approach is as follows: we consider some worker $i$ and an oracle who knows the preferences of all other agents and seek to find, whp, a stable partner of $i$. A slight complication here is that workers may have multiple stable partners in these markets. We work around this by defining a communication problem $P_1$ as follows: the correct answer is "yes" if the unique stable partner of $i$ occurs in her top $\log n$ most preferred firms, and the correct answer is "no" if the unique stable partner of $i$ does not occur in her top $\log n$ most preferred firms. In the case that $i$ does not have a unique stable partner, either a "yes" or a "no" is considered correct. Note that, given a candidate stable partner of $i$, one can output "yes" if the stable partner is among $i$'s top $\log n$ most preferred firms and a "no" if not. If the candidate stable partner is truly a stable partner, the output is a correct answer. Hence, the problem of producing an output that is correct whp for problem $P_1$ is no harder than the problem of finding an agent $j$ who, whp, is a stable partner of $i$.

Call a firm $j$ "accessible" if it satisfies the following: fix preferences of all agents except worker $i$. Suppose worker $i$ moves agent $j$ to the top of her preference list, keeping the rest of her preferences unchanged. Then worker $i$ will be matched to agent $j$ under the worker-optimal stable matching. Denote by $J_a$ the set of firms accessible to worker $i$. Note that whether $j \in J_a$ or not *does not depend on the preferences of worker i*. This follows immediately from the fact that the worker-optimal stable matching can be computed using the worker-proposing deferred acceptance algorithm, and this algorithm makes no use of worker $i$'s preferences unless she is rejected by firm $j$, in which case we already know that $j \notin J_a$. Finally, note that when $i$ has a unique stable partner, this is her most preferred firm in $J_a$.

We now control the size of set $J_a$, showing that its size is close to $n/\log n$. We make use of an analysis resembling that in Ashlagi et al. (2017) (with revelation of preferences as needed) and using the third bullet stated. Consider the worker-optimal stable matching with worker $i$ excluded. Let $J'$ be the set of firms that have each received between $0.9\log n$ and $1.1\log n$ proposals. Using the third bullet, we have

$$|J - J'| \le n^{0.99}$$

For each of these firms, independently, the probability that they prefer $i$ over their currently matched worker is at least $1/(1 + 1.1\log n) \ge 0.8/\log n$ and at most $1/(1 + 0.9\log n) \le 1.2/\log n$. Let $J''$ be the set of firms in $J'$ who prefer $i$ over their currently matched worker. It follows using a standard concentration bound that, whp, we have

$$0.7n/\log n \le |J''| \le 1.3n/\log n\,.$$

The probability that a rejection chain starting at a firm in $J''$, compare with Ashlagi et al. (2017), will return to the firm with a proposal that the firm prefers over worker $i$ before it terminates by going to the unmatched firm is at most $1/(2 + 0.9\log n) \le 1.2/\log n$. Call the set of firms for which the rejection chain returns $\hat{J}$. These firms may or may not be in $J''$. With high probability, using Markov's inequality, we have $|\hat{J}| \le f_n |J''| 1.2/\log n$ for any $f_n = \omega(1)$. Using $f_n = \sqrt{\log n}$, we obtain a bound of

$$|\hat{J}| \le \sqrt{n} 1.3n/\log n \cdot 1.2/\log n \le 2n/(\log n)^{3/2}\,.$$

All firms in $J''\backslash\hat{J}$ (here the rejection chain terminates without returning to the firm) are, for sure, a part of $J_a$. Thus, we have, whp,

$$\begin{aligned}|J_a| &\ge |J''\backslash\hat{J}| = |J''| - |\hat{J}| \ge 0.7n/\log n - 2n/(\log n)^{3/2}\\ &\ge 0.5n/\log n.\end{aligned}$$

On the other hand, we have $J_a \subseteq J'' \cup (J - J')$, leading to, whp,

$$|J_a| \le |J''| + |J - J'| \le 1.3n/\log n + n^{0.99} \le 1.5n/\log n\,.$$

Let the set of the worker's most preferred $\log n$ firms be $J_p$. Now again consider the communication problem $P_1$ and take any protocol that solves it. In cases in which worker $i$ has a unique stable partner and the protocol finds a correct answer, the answer is exactly $\mathbb{I}(J_a \cap J_p \ne \phi)$. Recalling that, whp, worker $i$ has a unique stable partner, and because the output of the protocol matches $\mathbb{I}(J_a \cap J_p \ne \phi)$ correctly whp in these cases, the protocol finds $\mathbb{I}(J_a \cap J_p \ne \phi)$ correctly with high probability overall.

We are now close to obtaining a lower bound on the expected number of bits needed for the protocol using the second part of Proposition I.1. Suppose we gave the worker $i$ access to $|J_a|$. Recall that $J_a$ is a uniformly random subset of $J$ and independent of $J_p$, conditioned on $|J_a|$. Let the lower bound in Proposition I.1 be $C(\log n)^2$ (i.e., we just named the constant factor $C$). We prove our result by contradiction. Suppose the protocol requires less than $(C/2)(\log n)^2$ bits in expectation. We found that, whp, we have

$$|J_a| \in (0.5n/\log n, 1.5n/\log n)\,.$$

(Notice that these are the same bounds that are needed in Proposition I.1.) Combining this fact and Markov's inequality on the expected number of bits used by the protocol conditioned on $|J_a|$, we deduce that, with probability at least 1/3,[47] the protocol is faced with $|J_a|$ bounded as required and such that the expected number of bits the protocol uses for that "bad" $|J_a|$ is at most $2(C/2)(\log n)^2 = C(\log n)^2$ bits. For each such bad $|J_a|$, the protocol must output something different from $\mathbb{I}(J_a \cap J_p \ne \phi)$ with probability at least $\epsilon$, using Proposition I.1. Because such bad $|J_a|$s occur with probability at least 1/3, the overall probability of outputting something different than $\mathbb{I}(J_a \cap J_p \ne \phi)$ is at least $\epsilon/3$. This contradicts that the protocol finds $\mathbb{I}(J_a \cap J_p \ne \phi)$ correctly whp. Thus, we have a contradiction. We conclude that any protocol solving problem $P_1$ must use at least $(C/2)(\log n)^2$ bits in expectation. Returning to the fact that problem $P_1$ is at least as hard as finding $j$ who is a stable partner of $i$ with high probability, we conclude that the agent-specific communication complexity for worker $i$ has the same lower bound. Finally, using symmetry over workers (all are in the same tier), we obtain the bound of $\Omega((\log n)^2)$ on the average agent-specific communication complexity. □

**Proof of Theorem I.2.** This market always has a unique stable matching, which can be constructed by serial dictatorship: workers choose their most preferred unmatched firm in the order of worker tiers. Let $i$ be the worker in question, whose tier is the $(n - \sqrt{n} + 1)$th from the top. When it is $i$'s turn to choose under serial dictatorship, there is a uniformly random subset of $\sqrt{n}$ unmatched firms remaining. Call this subset of firms $J'$. Worker $i$'s unique stable partner is the firm in subset $J'$ that appears highest in her preference list. Let $j$ be $i$'s unique stable partner. We prove our lower bound by showing that an even easier problem requires $\Omega(\sqrt{n})$ bits of communication: the problem is that of determining, with high probability, if $i$'s unique stable partner is one of the top $\sqrt{n}$ entries in her preference list. Now this occurs if and only if the set $I'$ consisting of the $\sqrt{n}$ firms that worker $i$ most prefers, intersects with set $J'$. Note that $I'$ and $J'$ are independent, uniformly random subsets of the set of $n$ firms, each of size $\sqrt{n}$, and $I'$ is known only to agent $i$, whereas $J'$ is known only to the oracle. The lower bound of $\Omega(\sqrt{n})$ follows from first part of Proposition I.1, which is implied by theorem 8.3 of Babai et al. (1986). Note that, for the tier structure described, the same lower bound (up to constant factors) can be obtained for each worker who has rank $n - \Theta(\sqrt{n})$. $\square$

## Appendix I.1. Communication Complexity of Set Disjointness

In this section, we prove the technical result that is used in the proof of Theorem I.1.

Consider a set $N$ such that $|N| = n$.[48] Suppose there is a uniformly random subset $A \subset N$ with $|A| = l_a$ known to agent $a$ and an independent uniformly random subset $B \subset N$ with $|B| = l_b$ known to agent $b$. Agents $a$ and $b$ are able to interactively communicate with each other, and the goal is to determine whether $A$ and $B$ have a nontrivial intersection or not. We are interested in lower bounds for the communication complexity of determining the correct answer with a probability of error that vanishes as $n$ grows although the bounds below hold even for a constant positive error independent of $n$. Note that the prior probability of intersection between the sets is bounded away from zero and one for any $l_a$ and $l_b$ such that $l_a, l_b = o(n)$ and $l_a \cdot l_b / n \in [0.1, 10]$.

**Proposition I.1.** *There exists $\epsilon > 0$ such that the following holds:*

- *With $l_a = l_b = \sqrt{n}$, the communication complexity of finding $\mathbb{I}(A \cap B \neq \phi)$ correctly with probability $1 - \epsilon$ is $\Omega(\sqrt{n})$ bits.*
- *With $l_a = \log n$ and $l_b = cn/\log n$ for some $c \in [1/2, 2]$, the communication complexity of finding $\mathbb{I}(A \cap B \neq \phi)$ correctly with probability $1 - \epsilon$ is $\Omega((\log n)^2)$ bits, uniformly over $c$ in the specified range.*

**Proof of Proposition I.1.** The first part of the proposition is just the lower bound part of theorem 8.3 of Babai et al. (1986).

The second part can be deduced with some work using general results about the communication complexity of disjointness (Braverman et al. 2013). We include a simpler direct proof, inspired by the proof of Babai et al. (1986) for the first part of the proposition. Because the communication setting in the proposition is distributional (i.e., the inputs come from a specified distribution of inputs), it suffices to consider deterministic protocols (because a randomized protocol over a distribution of inputs can be converted into a deterministic one by fixing the random seed that gives the lowest error over the input distribution).

Let $\mathcal{A}$ and $\mathcal{B}$ denote the sets of inputs to Alice and Bob. Thus, $|\mathcal{A}| = \binom{n}{l_a}$ and $|\mathcal{B}| = \binom{n}{l_b}$. Assume that there is a protocol $\pi$ of communication cost $d$. The randomized protocol $\pi$ induces a partition of $\mathcal{A} \times \mathcal{B}$ into at most $2^d$ combinatorial rectangles on each of which the output is either zero or one. For all values of $c$, the probability that the output is zero (i.e., that the sets are disjoint) is a constant, and therefore, for a sufficiently small $\epsilon$, a constant fraction of the mass is covered by zero rectangles on each of which the error rate is $< c_1\epsilon$ for an absolute constant $c_1 > 0$. We show that the maximum possible mass of each such rectangle is at most $2^{-\Omega(\log^2 n)}$, and therefore, there must be at least $2^{\Omega(\log^2 n)}$ such rectangles, and thus, $d = \Omega(\log^2 n)$.

Let $R_1 = \mathcal{A}_1 \times \mathcal{B}_1$ be a combinatorial rectangle in $\mathcal{A} \times \mathcal{B}$ such that at most a $(c_1\epsilon)$ fraction of the elements of $R$ are not disjoint (and, thus, the zero output is wrong). We need to show that the size of $R$ is relatively small. Let $\mathcal{A}_2$ denote the elements in $\mathcal{A}_1$ that intersect at most a $(2c_1\epsilon)$ fraction of the elements in $\mathcal{B}_1$. Note that we must have $|\mathcal{A}_2| \geq |\mathcal{A}_1|/2$. Let $\mathcal{B}_2 := \mathcal{B}_1$, and $R_2 := \mathcal{A}_2 \times \mathcal{B}_2$. It suffices to show that $R_2$ has mass at most $2^{-\Omega(\log^2 n)}$.

Construct a sequence of elements $S_1, \ldots, S_k$ of $\mathcal{A}_2$ with the following property: for each $i$, $|S_i \setminus \cup_{j=1}^{i-1} S_j| \geq l_a/2$. We continue constructing this sequence incrementally until one of two things happens: (1) we cannot add another element to the sequence, or (2) we have $|\cup_{j=1}^{k} S_j| \geq \sqrt{n}$. We consider each of these cases separately:

*Case* 1. There are sets $S_1, \ldots, S_k$ of $\mathcal{A}_2$ such that $|\cup_{j=1}^{k} S_j| < \sqrt{n}$, and each element $S \in \mathcal{A}_2$ satisfies $|S \setminus \cup_{j=1}^{k} S_j| \leq l_a/2$. Then this gives the following upper bound on the size of $\mathcal{A}_2$:

$$|\mathcal{A}_2| \leq \binom{\sqrt{n}}{l_a/2} \cdot \binom{n}{l_a/2} < n^{-\Omega(l_a)} \cdot \binom{n}{l_a} = 2^{-\Omega(\log^2 n)} \cdot \binom{n}{l_a},$$

and thus, $\mathcal{A}_2$ is small in this case, and we are done.

*Case* 2. There are sets $S_1, \ldots, S_k$ of $\mathcal{A}_2$, such that $\sqrt{n} \leq |\cup_{j=1}^{k} S_j| \leq \sqrt{n} + l_a$, and for each $i$, $|S_i \setminus \cup_{j=1}^{i-1} S_j| \geq l_a/2$. Each of the $S_i$s intersects at most a $(2c_1\epsilon)$ fraction of the elements in $\mathcal{B}_2$, and thus, at least half the elements in $\mathcal{B}_2$ intersect at most $4c_2\epsilon k$ of these sets. Denote these elements in $\mathcal{B}_2$ by $\mathcal{B}_3$. We have $|\mathcal{B}_3| \geq |\mathcal{B}_2|/2$. Each element $T$ in $\mathcal{B}_3$ can now be described as follows: first, specify the $S_i$s that $T$ intersects; there are at most $k \cdot \binom{k}{4c_2\epsilon k}$ ways of doing this. Notice that the union of the $S_i$s that $T$ does not intersect is at least $(k - 4c_2\epsilon k) l_a/2 > k l_a/3$ because each set contributes at least $l_a/2$ new elements to the union. Therefore, there are, at most, $\binom{n - kl_a/3}{l_b}$ ways to select the elements of $T$ from the remaining elements. Putting these together we get

$$\begin{aligned}\frac{|\mathcal{B}_2|}{\binom{n}{l_b}} &\leq \frac{2|\mathcal{B}_3|}{\binom{n}{l_b}} \leq \frac{2k \cdot \binom{k}{4c_2\epsilon k} \cdot \binom{n - kl_a/3}{l_b}}{\binom{n}{l_b}} \leq 2k \cdot \left(\frac{e}{4c_2\epsilon}\right)^{4c_2\epsilon k} \cdot \left(1 - \frac{kl_a}{3n}\right)^{l_b} \\ &\leq_1 2k \cdot e^{k/12} \cdot e^{-kl_a l_b/3n} \leq_2 2k \cdot e^{k/12} \cdot e^{-k/6} = 2k \cdot e^{-k/12} \\ &\ll_3 2^{-\Omega(\log^2 n)}.\end{aligned}$$

Here $\leq_1$ holds for a sufficiently small $\epsilon$ because, when $4c_2\epsilon < e^{-5}$, $(e/(4c_2\epsilon))^{4c_2\epsilon} < e^{1/12}$; $\leq_2$ holds because $l_a l_b \geq n/2$, and $\ll_3$ holds because $k > \sqrt{n}/l_a \gg \log^2 n$. Thus, $\mathcal{B}_2$ is very small in this case, and so is $R_2$, concluding the proof. $\square$

## Endnotes

[1] Congestion in matching markets has been studied in laboratory experiments and in the field. See Roth and Xing (1994, 1997), Roth (2008), and Avery et al. (2001) for empirical studies of unravelling in labor markets because of congestion, and Kagel and Roth (2000) for related laboratory experiments. More recently, Fradkin (2017) documented large reductions on the number of eventual bookings on Airbnb because the initial contact made by searchers went to hosts who rejected the offer.

[2] Roth finds stability to be an important requirement in the success of centralized matching markets (Roth 2002, Kagel and Roth 2000).

[3] See Kushilevitz and Nisan (2006) for a review of the communication complexity literature. The importance of studying communication in economic models is highlighted in the seminal essay Hayek (1945). This research direction was first formalized in Hurwicz (1973) and Mount and Reiter (1974).

[4] An agent can specify the agent's preference list over all $n$ agents on the other side using $n \log n$ bits. Segal (2007) proves that the communication needed per agent may grow linearly in the number of agents, assuming deterministic preferences, deterministic communication protocols, and exact stability. Chou and Lu (2010) extend this result to approximate stability. Gonczarowski et al. (2015) extend it to protocols that succeed with probability at least 2/3 on the worst-case distribution of preferences. We give a precise restatement in Section 2.3.1.

[5] In 2017, the number of applicants in the NRMP was about 43,000, and the number of positions was 31,000 (National Residency Matching Program 2017).

[6] This structure of preferences is analogous to many discrete choice models, such as the multinomial logit or probit.

[7] The square root comes from the communication complexity of set disjointness under independent inputs (Babai et al. 1986).

[8] Such preferences are called "block correlated" in Coles et al. (2013).

[9] If the market is stable across years, then publishing historical thresholds would suffice.

[10] See Coles et al. (2010), Chade et al. (2017), Lee and Niederle (2015), Abdulkadiroğlu et al. (2015), Lee and Schwarz (2017), and Jagadessan and Wei (2018).

[11] See Eeckhout (1999), Adachi (2003), Das and Kamenica (2005), Lauermann and Nöldeke (2014), Arnosti et al. (2014), and Kanoria and Saban (2017). For a review of the use of search theory in matching, see Chade et al. (2017).

[12] In particular, this is the sequential version of the worker-proposing DA algorithm developed by McVitie and Wilson (1971).

[13] The earlier result of Segal (2007) implies that, for any matching protocol that always produces a stable matching, there exists a preference realization $R$ such that the protocol requires a communication cost of $\Omega(n)$ per agent.

[14] The original results in Gonczarowski et al. (2015) concern communication cost only. However, any lower bound on communication cost automatically implies a bound on preference learning because any protocol that uses $Q$ choice function queries can be made into a communication protocol with $Q \log n$ bits of communication because the only information relevant to computing a stable matching is the result of choice function queries.

[15] The construction of Gonczarowski et al. (2015) is based on embedding a worst-case input distribution of the unique-disjointness problem (Razborov 1992) into the preference distribution of a submarket of $m = \Omega(n)$ workers and firms. The preferences are such that (1) every agent within the submarket values a uniformly random subset of $m/4$ partners within the submarket highly; (2) with constant probability, no pair of agents within this submarket value each other highly reciprocally; (3) with remaining probability, *exactly one* uniformly random pair of agents value each other highly; and (4) the structure of the unique stable matching hinges on the existence of such a pair of agents. Any matching protocol that is stable with high probability can be used to determine the existence of such a pair of agents who value each other highly, and so the protocol requires $\Omega(n)$ bits of communication per agent as implied by the hardness result of Razborov (1992) on the unique-disjointness problem.

[16] Simple translation of $a_{ji}$'s and $u_{j0}$ yields this property for each $j$ given that $u_{j0}$ is unrestricted.

[17] See, for example, Choo and Siow (2006), Agarwal (2015), Dagsvik (2000), Peski (2017), Menzel (2015), and Galichon et al. (2016) as well as the review article Chiappori and Salanié (2016).

[18] However, if the correlations in preferences are allowed to be large and complex, then the negative result of Gonczarowski et al. (2015) applies, and no communication efficient protocol exists.

[19] Without the idiosyncratic component of firm preferences, worker $i$ can infer nothing from her systematic score for a firm because she cannot meaningfully compare systematic scores across firms as any increasing nonlinear transformation of systematic scores for a firm would perfectly preserve preference orderings.

[20] In comparison, the preference distributions in Kojima and Pathak (2009) correspond to a $O(1)$ bound on the range of systematic scores and cannot incorporate enough systematic variation to have multiple tiers.

[21] The incentive properties for CEDA are as follows: with high probability, no worker can unilaterally deviate and improve her matching, and no firm can unilaterally deviate and obtain someone better than its best stable partner. Previous results have found that, in a variety of large-market models, the difference between the worker-optimal and firm-optimal stable matches is small for the vast majority of agents (Immorlica and Mahdian 2005, Kojima and Pathak 2009, Ashlagi et al. 2017, Lee 2017, Lee and Yariv 2017). In such cases, CEDA is approximately incentive compatible for everyone.

[22] The qualification requirement is an integer throughout the protocol, ensuring that it can be communicated with $O(\log n)$ bits.

[23] If there are fewer than $2e\sqrt{n}$ workers in the bin, then the firm sends a preference signal to all of them.

[24] In addition, we require that the identity of the worker selected to make the next application in step 1 of Protocol 1 cannot depend on any information outside the history of application and acceptance decisions of DA, such as the preference signals received by workers from firms to which they have not applied yet.

[25] Our lower bound allows the communication protocol to be adaptive, to be randomized, and to broadcast messages to every agent. Moreover, the bound holds even if all workers can communicate among themselves for free, and all firms can communicate among themselves for free.

[26] We need at least two rounds for agents to be able to respond to signals from others.

[27] One can interpret this as the existence of an additional "operator" through which all messages are routed, so the operator sees all messages and can choose who should send the next message.

[28] This preference structure is called "block-correlated" in Coles et al. (2013).

[29] The reason is that the worker ranked number $k$ needs to send only $O(n/(n-k+1))$ signals.

[30] A factor of $\log n$ is necessary because, even when there is one tier on each side, agents on the short side of the market obtain an expected rank of $\Omega(\log n)$ (Pittel 1989). (Obtaining a rank of one means being matched with one's favorite partner; a rank of two means one's second favorite, and so on.) The extra logarithmic factor in $24 \log^2 n$ is an artifact of the analysis (see Remark 3 after Theorem 3.)

[31] The bound, stated in Proposition F.1, is analogous to theorem 2 of Ashlagi et al. (2017) except that the constant is worse but the notion of "high probability" is stronger as per our requirement.

[32] This is because communicating the identity of each particular agent requires $\log n$ bits.

[33] This follows because the total amount of communication is $O(n \log^3 n) = O^*(n)$.

[34] See chapter 4 of Roth and Sotomayor (1990) for a thorough exposition of the incentive properties of DA. See Coles and Shorrer (2014) for a discussion of the manipulability of DA in a balanced uniformly random market.

[35] See Roth (1982) and Dubins and Freedman (1981).

[36] The best stable partner of an agent is the agent's most preferred partner in all stable matchings.

[37] See pp. 92–93 of Roth and Sotomayor (1990).

[38] Here, "with high probability" is defined as in Definition 9: the probability space is $\mathcal{P}(\omega)$, and the probability of not being strategy-proof is no more than $\delta(n) = o(1)$.

[39] It is impossible to show a convergence rate faster than $O(1/(\log n))$ because, in a uniformly random market with $n$ workers and $n-1$ firms in which every partner is acceptable to every agent, the probability that an agent has multiple stable partners is approximately $1/(\log n)$, following the analysis of Ashlagi et al. (2017).

[40] One paper that studies the number of rounds of communication needed in a matching model is Dobzinski et al. (2014). However, they do not consider the stability of the match.

[41] An analogous aftermarket is implemented in the NRMP after the main match takes place.

[42] The $\Theta$ notation represents that two functions grow to infinity at similar rates: $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$.

[43] For any power series $f(\xi)$ with positive coefficients, $\sum_a^\infty [\xi^a]\{f(\xi)\} \leq \xi^{-a} \inf_{\xi \geq 1}\{f(\xi)\}$.

[44] A matching is individually rational if agents are only matched to partners they find acceptable.

[45] Specifically, the graph is the collection of tuples $(i, j)$ in which at least one signaled to the other during the signaling round. By Definition E.1, this is $G \cup \{(i, j) : j \text{ signaled to } i, s_{k(i)} = t_{l(j)}\}$.

[46] A matching is full if it matches $\min(n_I, n_J)$ pairs of agents.

[47] We use 1/3 instead of 1/2 to accommodate that, with small probability, $|J_a|$ may not fall in the desired range.

[48] We redefine $n$ here. For the purposes of this section, there are no workers or firms.

## References


Abdulkadiroğlu A, Che YK, Yasuda Y (2015) Expanding "choice" in school choice. *Amer. Econom. J. Microeconom.* 7(1):1–42.

Adachi H (2003) A search model of two-sided matching under nontransferable utility. *J. Econom. Theory* 113(2):182–198.

Agarwal N (2015) An empirical model of the medical match. *Amer. Econom. Rev.* 105(7):1939–1978.

Arnosti N, Johari R, Kanoria Y (2014) Managing congestion in decentralized matching markets. Babaioff M, Conitzer V, Easley D, eds. *Proc. 15th ACM Conf. Econom. Comput.* (ACM, New York), 451.

Ashlagi I, Kanoria Y, Leshno JD (2017) Unbalanced random matching markets: The Stark effect of competition. *J. Political Econom.* 125(1):69–98.

Avery C, Jolls C, Posner RA, Roth AE (2001) The market for federal judicial law clerks. *Univ. Chicago Law Rev.* 68(3):793–902.

Babai L, Frankl P, Simon J (1986) Complexity classes in communication complexity theory. Hartmanis J, ed. *Proc. 27th Annual Sympos. Foundations Comput. Sci.* (ACM, New York), 337–347.

Banerjee A, Duflo E, Ghatak M, Lafortune J (2013) Marry for what? Caste and mate selection in modern India. *Amer. Econom. J. Microeconom.* 5(2):33–72.

Braverman M (2015) Interactive information complexity. *SIAM J. Comput.* 44(6):1698–1739.

Braverman M, Garg A, Pankratov D, Weinstein O (2013) From information to exact communication. Boneh D, Roughgarden T, Feigenbaum J, eds. *Proc. 45th Annual ACM Sympos. Theory Comput.* (ACM, New York), 151–160.

Chade H, Eeckhout J, Smith L (2017) Sorting through search and matching models in economics. *J. Econom. Literature* 55(2):493–544.

Che YK, Tercieux O (2018a) Efficiency and stability in large matching markets. Working paper, Columbia University, New York.

Che YK, Tercieux O (2018b) Payoff equivalence of efficient mechanisms in large matching markets. *Theoret. Econom.* 13(1):239–271.

Chiappori PA, Salanié B (2016) The econometrics of matching models. *J. Econom. Literature* 54(3):832–861.

Choo E, Siow A (2006) Who marries whom and why. *J. Political Econom.* 114(1):175–201.

Chou JH, Lu CJ (2010) Communication requirements for stable marriages. Diaz J, Calamoneri T, eds. *Proc. 7th Internat. Conf. Algorithms Complexity* (Springer-Verlag, Berlin, Heidelberg), 371–382.

Coles P, Shorrer R (2014) Optimal truncation in matching markets. *Games Econom. Behav.* 87:591–615.

Coles P, Kushnir A, Niederle M (2013) Preference signaling in matching markets. *Amer. Econom. J. Microeconom.* 5(2):99–134.

Coles P, Cawley J, Levine PB, Niederle M, Roth AE, Siegfried JJ (2010) The job market for new economists: A market design perspective. *J. Econom. Perspect.* 24(4):187–206.

Cover TM, Thomas JA (2012) *Elements of Information Theory* (John Wiley & Sons, Hoboken, NJ).

Dagsvik JK (2000) Aggregation in matching markets. *Internat. Econom. Rev.* 41(1):27–58.

Das S, Kamenica E (2005) Two-sided bandits and the dating market. *Proc. 19th Internat. Joint Conf. Artificial Intelligence* (Morgan Kaufmann Publishers Inc., San Francisco), 947–952.

Demange G, Gale D, Sotomayor M (1987) A further note on the stable matching problem. *Discrete Appl. Math.* 16(3):217–222.

Dobzinski S, Nisan N, Oren S (2014) Economic efficiency requires interaction. Shmoys D, ed. *Proc. 46th Annual ACM Sympos. Theory Comput.* (ACM, New York), 233–242.

Dubins LE, Freedman DA (1981) Machiavelli and the Gale-Shapley algorithm. *Amer. Math. Monthly* 88(7):485–494.

Eeckhout J (1999) Bilateral search and vertical heterogeneity. *Internat. Econom. Rev.* 40(4):869–887.

Fradkin A (2017) Search, matching, and the role of digital marketplace design in enabling trade: Evidence from Airbnb. Working paper, MIT, Cambridge, MA.

Gale D, Shapley LL (1962) College admissions and the stability of marriage. *Amer. Math. Monthly* 69(1):9–15.

Gale D, Sotomayor M (1985) Some remarks on the stable matching problem. *Discrete Appl. Math.* 11(3):223–232.

Galichon A, Kominers SD, Weber S (2016) Costly concessions: An empirical framework for matching with imperfectly transferable utility. Working paper, New York University, New York.

Gonczarowski YA, Nisan N, Ostrovsky R, Rosenbaum W (2015) A stable marriage requires communication. Indyk P, ed. *Proc. 26th Annual ACM-SIAM Sympos. Discrete Algorithms* (SIAM, Philadelphia), 1003–1017.

Hayek FA (1945) The use of knowledge in society. *Amer. Econom. Rev.* 34(4):519–530.

Hitsch GJ, Hortaçsu A, Ariely D (2010) Matching and sorting in online dating. *Amer. Econom. Rev.* 100(1):130–163.

Horton JJ (2017) The effects of algorithmic labor market recommendations: Evidence from a field experiment. *J. Labor Econom.* 35(2):345–385.

Horton JJ (2019) Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution. *Management Sci.* Forthcoming.

Hurwicz L (1973) The design of mechanisms for resource allocation. *Amer. Econom. Rev.* 63(2):1–30.

Immorlica N, Mahdian M (2005) Marriage, honesty, and stability. Buchsbaum A, ed. *Proc. 16th Annual ACM-SIAM Sympos. Discrete Algorithms* (SIAM, Philadelphia), 53–62.

Itoga SY (1978) The upper bound for the stable marriage problem. *J. Oper. Res. Soc.* 29(8):811–814.

Jagadessan M, Wei A (2018) Varying the number of signals in matching markets. Christodoulou G, Harks T, eds. *Web and Internet Economics—WINE 2018*, Lecture Notes in Computer Science, vol. 11316 (Springer, Cham, Switzerland), 232–245.

Kagel JH, Roth AE (2000) The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *Quart. J. Econom.* 115(1):201–235.

Kanoria Y, Saban D (2017) Facilitating the search for partners on matching platforms: Restricting agent actions, Working paper, Columbia University, New York.

Knuth DE (1976) *Mariages Stables et leurs Relations avec d'autres Problèmes Combinatoires: Introduction à l'analyse Mathématique des Algorithmes* [*Stable Marriage and Its Relation to Other Combinatorial Problems: An Introduction to the Mathematical Analysis of Algorithms*] (Presses de l'Université de Montréal, Montréal).

Kojima F, Pathak PA (2009) Incentives and stability in large two-sided matching markets. *Amer. Econom. Rev.* 99(3):608–627.

Kushilevitz E, Nisan N (2006) *Communication Complexity* (Cambridge University Press, Cambridge, UK).

Lauermann S, Nöldeke G (2014) Stable marriages and search frictions. *J. Econom. Theory* 151:163–195.

Lee RS, Schwarz M (2017) Interviewing in two-sided matching markets. Technical report, National Bureau of Economic Research, Cambridge, MA.

Lee S (2017) Incentive compatibility of large centralized matching markets. *Rev. Econom. Stud.* 84(1):444–463.

Lee S, Niederle M (2015) Propose with a rose? Signaling in internet dating markets. *Experiment. Econom.* 18(4):731–755.

Lee S, Yariv L (2017) On the efficiency of stable matchings in large markets. Working paper, Washington University, St. Louis.

McVitie DG, Wilson LB (1971) The stable marriage problem. *Comm. ACM* 14(7):486–490.

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

Mount K, Reiter S (1974) The informational size of message spaces. *J. Econom. Theory* 8(2):161–192.

National Residency Matching Program (2017) Press release: 2017 NRMP main residency match the largest match on record. Accessed July 13, 2018, http://www.nrmp.org/press-release-2017-nrmp-main-residency-match-the-largest-match-on-record/.

Peski M (2017) Large roommate problem with non-transferable random utility. *J. Econom. Theory* 168:432–471.

Pittel B (1989) The average number of stable matchings. *SIAM J. Discrete Math.* 2(4):530–549.

Pittel B (1992) On likely solutions of a stable marriage problem. *Ann. Appl. Probab.* 2(2):358–401.

Razborov A (1992) On the distributed complexity of disjointness. *TCS: Theoret. Comput. Sci.* 106(2):385–390.

Roth AE (1982) The economics of matching: Stability and incentives. *Math. Oper. Res.* 7(4):617–628.

Roth AE (2002) The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica* 70(4):1341–1378.

Roth AE (2008) What have we learned from market design? *Econom. J.* 118(527):285–310.

Roth AE, Sotomayor O (1990) *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis* (Cambridge University Press, Cambridge, UK).

Roth AE, Xing X (1994) Jumping the gun: Imperfections and institutions related to the timing of market transactions. *Amer. Econom. Rev.* 84(4):992–1044.

Roth AE, Xing X (1997) Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists. *J. Political Econom.* 105(2):284–329.

Segal I (2007) The communication requirements of social choice rules and supporting budget sets. *J. Econom. Theory* 136(1): 341–378.

Wilson L (1972) An analysis of the stable marriage assignment algorithm. *BIT Numer. Math.* 12(4):569–575.