Communication Requirements and Informative Signaling in Matching Markets

ITAI ASHLAGI, Stanford University
MARK BRAVERMAN, Princeton University
YASH KANORIA, Columbia University
PENG SHI, Microsoft Research

We study how much communication is needed to find a stable matching in a two-sided matching market with private preferences when agents have some knowledge of the preference distribution. We show that in a two-sided market with workers and firms, when workers' preferences are arbitrary and private and firms' preferences follow an additively separable latent utility model with commonly known and heterogeneous parameters, there exists a communication protocol that finds a stable matching with high probability and requires at most $O^*(\sqrt{n})$ bits of communication per agent. (We show that this is the best achievable.) The protocol, which we call communication-efficient deferred acceptance (CEDA), is a modification of the workers proposing deferred acceptance (DA) algorithm. Under CEDA, firms signal workers they privately like and thus help them to better target their applications while also broadcasting qualification requirements to discourage workers who have no realistic chances of being hired. In the special case where the preferences of both workers and firms follow a tiered structure, we show that the signaling can be done in a parallel and decentralized way. The protocols we propose inherit the incentive properties of DA in large markets. Our findings yield insights about how matching markets can be better organized to reduce congestion. Roughly, each agent should reach out to her favorite "gettables" who are likely to consider her, while waiting for her dream matches to approach her.

1 INTRODUCTION

Following the work of Gale and Shapley [1962], two-sided matching models have played an important role in the analysis of two-sided markets, including, for example, labor markets, marriage markets and college admissions. Stable matchings, which are matchings from which no pair of agents has incentive to deviate, have been used to predict equilibrium outcomes in decentralized markets [Hitsch et al., 2010a] and have been shown to be important to the success of centralized clearinghouses [Roth, 2002].

However, forming a stable matching requires agents to learn and communicate their own preferences to the market, and too much communication may result in market congestion, whereby the market cannot process information quickly enough to converge to a satisfactory outcome. For example, in a decentralized market, a stable matching may result from the equilibrium outcome of a decentralized matchmaking process; in such a process, an agent may communicate his preferences by offering to match with a partner. However, if many agents are making offers and time is constrained, agents may be forced to make decisions before receiving all offers, or to avoid sending offers to popular partners, both of which may result in a sub ptimal outcome.[1]

Congestion can also be a problem in centralized clearinghouses, which compute a match on everyone's behalf after eliciting a preference ranking from everyone. Congestion may occur because the process whereby agents learn and communicate their preferences may require time, effort, and monetary costs, all of which are constrained. For example, in the National Residency Matching Program (NRMP), a centralized clearinghouse that matches medical students with residency programs, an applicant ranks a program only if he or she travels to the program for an interview, and

[1] See [Roth, 2008], [Kagel and Roth, 2000].

the monetary and logistical cost of interviewing produces a cost to learning and communicating preferences in this market. Therefore, for the market to be efficient one must control how much preference information needs to be learned and communicated.

In practice, we observe that agents do not communicate all their preferences. For example, in the NRMP, a typical applicant submits a ranking of only about 15 programs, out of a total of more than 20,000 possible programs. Similarly, academic departments do not create complete rankings of all their job applicants, but focus only on a small subset. Hence, can we still expect the outcomes of these markets to be approximately stable? This paper addresses the following questions: What is the minimum amount of communication needed for the market to reach a stable matching? Do agents need to know their preferences over the entire market? How can a market designer help the market reach a stable matching with low communication overhead, thus limiting congestion?

To address these questions, we adopt concepts from the *communication complexity* literature.[2] In a communication protocol, agents send messages to each other, each of which may depend on the history of messages. The communication complexity of solving a problem is the minimum number of bits that must be communicated in order to find a solution. Segal [2007] studies the communication complexity of computing various types of equilibria, including stable matchings. He finds that the worst-case number of bits (i.e., the maximum number of bits across all preference profiles) needed to find a stable matching is $\Theta(n)$, where n is the number of agents on each side of the market.[3] In fact, this lower bound holds even if we allow the protocol to be randomized and to fail with a small probability [Gonczarowski et al., 2015]. These findings imply that, in the worst case, agents essentially need to know and communicate their complete preferences over the entire market. If so much communication is typically required, one may question whether stability is an appropriate solution concept.

In this paper, we study "typical markets" rather than worst-case markets. We model a typical market as one in which the preferences of agents on one side of the market follow a certain distribution, and agents possess some knowledge about this distribution. Our distributional assumptions are mild, reflecting the rich heterogeneity of preferences we observe in real markets. We also define a notion of preference learning cost by counting the number of queries to a certain preference oracle, which returns an agent's favorite option within a given set. For concreteness, we focus our exposition on labor markets, and refer to the two sides of the market as workers and firms.

Our main finding is that in the set of markets we consider, the communication complexity of finding a stable matching is much less than what the previous worst-case analysis suggests. Our assumptions on preferences are as follows. We allow the preferences of workers to be completely arbitrary and privately known, and we assume that the preferences of firms have the following *separable* form: the utility of a firm for a worker is the sum of a public score and a private score. The public score is known to both the firm and the worker concerned and can differ for every worker-firm pair,[4] thus allowing a rich heterogeneity of preferences. The private score is known only to the firm, and is assumed to be independently and identically distributed (i.i.d.) according to an arbitrary distribution with a bounded hazard rate (which includes all distributions with a sufficiently fat tail, including the exponential, the type-I extreme value, the log-normal, and the Pareto distributions).[5] We refer to such markets as *separable markets*. In such markets, we construct

[2]See Kushilevitz and Nisan [2006] for a review.
[3]Note that with n agents, one can encode a single agent using $\log_2 n$ bits of information. We write $f(n) = \Theta(g(n))$ if there exist constants $a \leq b$ and n_0 such that $ag(n) \leq f(n) \leq bg(n)$ for all $n \geq n_0$.
[4]For instance, faculty candidate X who is a promising young theorist would have a large public score with CS department Y that has a strong theory group, and is looking to hire in the area. Both X and Y would know of this compatibility beforehand.
[5]The private score of department Y for candidate X may consist of the new compatibility information they acquire during a full-day interview.

a communication protocol that finds a stable matching with only $O^*(\sqrt{n})$ bits of communication[6] and $O^*(\sqrt{n})$ preference oracle queries, and we show that this is essentially the best possible for separable markets.[7] In a market with tens of thousands of participants, one may argue that $\sqrt{n}$ is a reasonable level of communication and preference learning, thus providing a theoretical justification for using stable matchings as an approximation of the outcome in large markets.

The protocol we construct illustrates the following insight: the key to achieving low communication cost is having agents send informative signals. The protocol uses two types of signals, which we refer to as "preference" signals and "qualification" signals. In a preference signal, a firm signals to a worker that it privately likes him or her based on a high private score. In a qualification signal, a firm broadcasts[8] to the market a minimum public score that it requires for all future applicants to whom the firm did not send a preference signal. This minimum score is called the firm's "qualification" requirement. A worker is said to publicly qualify to a firm if the worker's public score for the firm meets the qualification requirement, and a worker is said to qualify to a firm if the worker either publicly qualifies or receives a preference signal from the firm. The protocol we construct, which we refer to as *communication-efficient deferred acceptance* (CEDA), embeds these two types of signals into the worker-proposing DA algorithm. Before the protocol runs the DA algorithm, it instructs firms to send preference signals. Then it applies the DA algorithm while instructing workers to apply only to firms they qualify for, and it instructs firms to increase the qualification requirement after receiving a certain number of publicly qualified applicants. The logic of the protocol is as follows: because we restrict the distribution of private scores to be sufficiently fat-tailed, after a certain number of publicly qualified applications, the firm would have probably found a worker that it sufficiently likes. Once this happens, the firm can safely increase its qualification requirement and dissuade applications from workers that have little chance of being accepted, thus reducing their communication and preference learning cost.

Moreover, we show that given additional structure on the preferences of both workers and firms, we can construct a protocol in which signals are sent in parallel and the only signals needed are preference signals. In such markets, qualification signals are no longer important, as everyone has a good understanding of which set of partners they have a chance of getting, and simply signals a certain number of their favorite partners within this set. The additional structure that we use to prove this result is as follows: agents are partitioned into tiers, and everyone prefers partners from better tiers to worse tiers and have uniformly random preferences for partners within a given tier. We call such markets *tiered random markets.* They are considerably more restrictive than general separable markets, but the results we have are considerably stronger. For tiered random markets, we construct a two-round protocol. In the first round, both workers and firms signal the partners they privately like within a certain target tier, which is defined as the tier with equal or slightly lower ranking compared to the agent's own tier (the target tier has "gettables" who are likely to consider them); in the second round, workers and firms respond to the signals they receive by ranking the senders (these are "dream matches" who reached out to them) as well as the partners they signaled to, and a matchmaker computes a matching based on these partial preference rankings. The number of partners an agent signals to in the first round depends on the agent's

[6]The O^* notation means that there exists a constant $C > 0$ such that the number of bits of communication per agent is upper bounded by a function $(C + \log^C)\sqrt{n}$.

[7]We construct a separable market in which any protocol that finds a stable matching with high probability requires at least $O(\sqrt{n})$ bits of communication per agent. We say "essentially" the best possible because there is still potentially a $\log^2(n)$ multiplicative gap between the upper and lower bounds.

[8]If interaction is via a matching platform that can estimate the preference distribution parameters (e.g., Tinder maintains an "Elo" rating for each user), *the platform can implement qualification requirements*, e.g., by filtering search/recommendation results, eliminating the need for broadcast messages.

competitiveness in the market, which can be inferred directly from the tier of the agent. We show that the protocol computes a stable matching using only $O(\log^4 n)$ bits of communication per agent and $O(\log^3 n)$ preference queries per agent.

Such a two-round protocol has the advantage of further decreased congestion, as agents can signal in parallel, without the need to wait for other agents. Moreover, such a protocol uses only private messages, which are visible only to the sender and the receiver and not to the whole market. As a result, the protocol may be plausibly implemented by a decentralized matchmaking process, without the need for a centralized platform through which firms broadcast qualification signals.

We also analyze incentive properties under all the protocols we propose and show that none of them create additional incentive issues over what is already present in DA under full preference elicitation[9]: with high probability in large markets, no agent can unilaterally deviate from any of the protocols we propose and be matched with someone better than their best stable partner under true preferences. This implies that in CEDA, the probability that a worker can profitably deviate from the protocol is vanishing in large markets. The firms, however, may profitably deviate by the same type of truncation strategies as in DA.[10] We show a stronger result for the targeted signaling protocol in tiered random markets that satisfy "general imbalance" and "large tiers" conditions: in such markets, for each agent on either side, the probability that there exists a profitable deviation goes to zero when the minimum size of any tier is large.

Finally, we briefly discuss some applications and the relevance to real markets. First note that preference and qualification signals are observed in practice. Academic departments at Tel Aviv University publish approximate acceptance thresholds for future students based on weighted SAT and grade point averages.[11] Similarly, historical threshold information is made available in college admissions in other countries such as India and Iran. Tinder, an online dating application, allows users to send one special (preference) signal per day to a potential partner beyond regular signals sent to multiple potential partners. The American Economic Association allows job market candidates to send two special signals to academic departments, and empirical evidence suggests that candidates signal departments in tiers that lower than or similar to their own [Coles et al., 2010a]. A similar form of preference signaling is documented in a field experiment on a Korean dating website where users are allocated a few virtual roses [Lee and Niederle, 2015]. Applications in colleges early admission programs can also be viewed as preference signals [Avery et al., 2009].

Our results for tiered random markets may be relevant for some applications such as college admissions and academic labor markets. Our two-round protocol for tiered random markets reflects the situation in centralized matching markets such as the NRMP, in which a first round of decentralized signaling determines the set of interviews, and a second round where preferences are submitted to a clearinghouse determines the match. Our analysis further provides predictions on how such markets will be organized and the amount of communication needed at different tiers of the market. For example, our findings may explain why we observe short preference lists in the NRMP.

Our paper relates with many strands of previous literature. For clarity of exposition, we highlight the most relevant papers in the main text and refer readers to Appendix A for a more systematic literature review.

[9]See chapter 4 of [Roth and Sotomayor, 1990] for a review of incentive issues for DA under full preference elicitation.
[10]See chapter 4 of [Roth and Sotomayor, 1990].
[11]In some departments these thresholds are not the same, in which case students with certain scores face uncertainty.

2 MODEL

2.1 Background on stable matchings

In a *two-sided matching market* there is a set of *workers* $I = \{1, 2, \ldots, n_I\}$ and a set of *firms* $J = \{1, 2, \ldots, n_J\}$. Define $n = \max\{n_I, n_J\}$. A *matching* μ is defined to be a set of pairs of the form (i, j), where $i \in I$ and $j \in J$, such that each worker and each firm appear in at most one pair. We refer to a tuple (i, j) in a matching as a *matched pair* and i and j are said to be *matched partners* of one another. Agents (firms or workers) who do not have a matched partner are said to be *unmatched.*

Each worker $i \in I$ has a *strict preference ranking* $R^I(i)$ over firms that is a permutation over $J \cup \{0\}$ where 0 denotes being unmatched. (One can interpret it as an outside option for the worker.) Similarly, each firm $j \in J$ has a *preference ranking* $R^J(j)$ over workers that is a permutation over $I \cup \{0\}$. For any agent, a partner is said to be *acceptable* if the agent prefers being matched with that partner to being unmatched.

For a matching μ define $\mu^I : I \to J \cup \{0\}$ to be a function that maps each worker to his/her matched partner; if the worker i is unmatched, then $\mu^I(i) = 0$. Similarly define $\mu^J : J \to I \cup \{0\}$ to be a mapping from the set of firms to their matched partners. A *blocking pair* to the matching is a pair (i, j) which is not a matched pair, but worker i prefers firm j to $\mu^I(i)$, and firm j prefers worker i to $\mu^J(j)$. A matching is called *stable* if there is no blocking pair.

A stable matching always exists and can be found using the deferred acceptance (DA) algorithm by [Gale and Shapley, 1962]. The *worker-proposing* DA finds a stable matching works as follows.

Protocol 1. The deferred acceptance (DA) algorithm

The algorithm begins without any matches and builds up a matching through a series of rounds.

(1) *In each round, an unmatched worker (if any such remains) is selected arbitrarily and applies to the worker's favorite firm that the worker has not yet applied to.*
(2) *If that firm is unmatched, then it accepts the application and is said to be tentatively matched to the worker. If that firm is already tentatively matched to someone else, then it becomes matched to the more preferred of these two workers, and rejects the less preferred of these two workers. The rejected worker becomes unmatched and can again be chosen in a future round to apply to their next favorite firm.*

The algorithm ends when every worker is either matched to a firm or has already been rejected by all firms the worker finds acceptable.

It can be shown that the result of this algorithm does not depend on the order in which unmatched workers are chosen to apply. Furthermore, the worker-proposing DA algorithm gives every worker the best possible outcome under any stable matching. This optimal stable matching for workers is called the *worker-optimal stable match.*

2.2 Separable markets

A two-sided matching market is called a *separable* market if for at least one side of the market, preferences follow a latent random utility model with the following additive structure. (The preferences of the other side can be completely arbitrary and privately known.) Without loss of generality, label the side with preference structure as *the firms.* Assume that the latent utility of firm j for worker i is

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

where a_{ji} is called the *public score* of worker i for firm j and is known to this firm and this worker. ϵ_{ji} is called the *private score* of worker i for firm j, and is only privately known to firm j. For each firm j, we assume that the private scores ϵ_{ji} are independently drawn from a continuous distribution with CDF F_j with a bounded hazard rate. Precisely speaking, there exists a constant $B > 0$ such that the hazard rate

$$h_j(x) := \frac{F_j'(x)}{1 - F_j(x)} \leq B.$$

This assumption is satisfied by distributions whose tail is as fat as the exponential distribution, including the exponential distribution, the type-I extreme value distribution commonly used in discrete choice modeling, the log-normal distribution, and the Pareto distributions. The latent utility of firm j for the outside option, u_{j0}, is privately known.

Assume that for each firm j, the public scores have an additive range of at most *polylog*(n). Formally, we assume that there exists a constant $C > 0$, such that $\max_{i \in I}\{a_{ji}\} - \min_{\in I}\{a_{ji}\} \leq C \log^C n$ for sufficiently large n.[12] Although some bound on the range of public scores is necessary for our result, the precise expression of $C \log^C(n)$ allows for simplicity of exposition, and can be relaxed to a general bound $g(n) = o(n)$ with slight modification to our results.[13]

2.3 Communication and preference learning cost

We are interested in measuring the amount of communication and preference learning required to construct a stable matching. In this section, we formally define these measures, using terminology from the communication complexity literature (see [Kushilevitz and Nisan, 2006] for an overview of this literature).

In a *communication protocol*, agents communicate with each other through messages. For now, we assume that messages are visible to all other agents, and protocols that satisfy this assumption are called protocols *with broadcast*. (We remove this assumption in Section 4, allowing for protocols without broadcast.) As a function of all messages sent so far, the protocol either terminates with a final output, or uniquely determines the identity of the next agent who is supposed to send a message. The agent's message is also uniquely determined by the agent's private information as well as the history of messages. For concreteness, the message space is $\{0, 1\}$, so the amount of communication is measured in bits.[14] In this paper, all protocols are deterministic.

Given a separable market and a communication protocol, the sequence of messages is uniquely determined given the preference realizations of workers and firms. Define the *communication cost* of agent i to be the total number of bits sent by agent i, given the preference realizations. Since the preferences of firms are random and independent of the preferences of workers, we can define the *expected communication cost* of agent i as the expectation of the above communication cost taken over the randomness in the preferences. Define the *average communication cost per agent* as the across-agent average of the expected communication cost for each agent. This is the main quantity we use to measure communication cost in this paper.

To quantify preference learning cost, we assume that agents do not know their preferences a priori, but can uncover them only through querying a preference oracle. In each query to the oracle, an agent gives as input a subset of partners and receives from the oracle the identity of the

[12] Note that this still allows for significant ex-ante vertical differentiation among workers. For example, if the private scores are exponentially distributed with rate parameter B, then a difference in public scores of $\frac{3}{B}\log(n)$ means that firm j would prefer the worker with the higher public score with probability of at least $1 - \frac{1}{n^3}$. Thus, allowing for public scores to be as large as $\frac{3k}{B}\log(n)$ allows us to create k tiers of workers, so that with probability of at least $1 - \frac{1}{n}$, every firm will always a prefer a worker from a better tier to a worker from a worse tier.

[13] Under a bound of $g(n)$, we can modify parameters in our protocols and yield a version of Theorem 3.1 with every bound multiplied by a factor of $O(\sqrt{g(n)})$. This is shown in Appendix C.

[14] Any other constant sized message space affects only constant factors in our result.

agent's most preferred partner within this subset. (We do not count the communication with the oracle as part of the communication cost.) As before, since protocols are deterministic in this paper, the sequence of preference oracle queries is fixed given the realization of preferences. Define the *preference learning cost* for agent i as the number of such queries to the oracles that the agent makes before the communication protocol terminates. Define similarly the *average preference learning cost per agent* as the across-agent average of the expectation of the preference learning cost per agent, when the expectation is taken over the randomness in the preferences.

Define a matching protocol to be a communication protocol that, upon termination, outputs a matching. The protocol is said to *succeed* if the produced matching is stable with respect to the true preferences. We are interested in finding matching protocols that succeed with high probability, where the randomness is again induced by the randomness of the preferences.[15]

A simple example of a matching protocol that always succeeds is the naive protocol that elicits from every agent all of their private information and finds a stable matching using the deferred acceptance algorithm. Since every agent can be represented using $O(\log n)$ bits, this can be done by eliciting $O(n \log n)$ bits of information from every agent. Segal [2007] shows that against an adversarial distribution of preferences (R^I, R^J), any communication protocol that always succeeds requires an average communication cost per agent of $\Omega(n)$, which is only a logarithmic factor less than the naive protocol.[16] This negative result was strengthened to protocols that need only succeed with high probability.

Proposition 2.1 ([Gonczarowski et al., 2015]). *There exists a distribution on (R^I, R^J) such that any matching protocol that succeeds with probability of at least $2/3$ requires at least an average communication cost per agent of $\Omega(n)$.*[17]

Observe that in any protocol with non trivial messages, for every $\log n$ bits of communication there is at least one query to the preference oracle. Therefore, Proposition 2.1 implies the following corollary, which states that against an adversarial distribution of preferences, the average preference learning cost per agent is close to the maximum possible value of n. (Note that an agent can learn her complete preference lists with n sequential oracle calls.)

Corollary 2.2. *There exists a distribution on (R^I, R^J) such that any matching protocol that succeeds in finding a stable matching with probability of at least $2/3$ must have an average preference learning cost per agent of $\Omega(n/\log n)$.*

However, these negative results requires fairly contrived distribution of preferences. In this paper, we restrict preferences to follow those of separable markets: one side can be completely arbitrary, while the other side follows an additive structure allowing rich heterogeneities. As shown in the next section, we find that under these assumptions, both the communication and the preference learning costs can be reduced to $O^*(\sqrt{n})$, which is much less than the $O^*(n)$ result suggested by the above negative results.[18]

3 MAIN RESULTS

For separable markets (defined in Section 2.2, we construct a matching protocol called the communication-efficient deferred acceptance (CEDA) protocol, which succeeds to find a stable matching with high

[15] Formally, given a sequence of markets of n agents and given the preferences of workers, we want a sequence of protocols such that the measure of firm preferences for which the protocol fails tends to 0 as $n \to \infty$.

[16] We say that a function $f(n) = \Omega(g(n))$ if there exists a constant $C > 0$ such that $f(n) \geq Cg(n)$ for all sufficiently large n.

[17] The finding of [Gonczarowski et al., 2015] is originally stated in terms of communication protocols that allows randomization, but their result implies this proposition by Yao's minimax theorem.

[18] We say that a function $f(n) = O^*(g(n))$ if there exists a constant $C > 0$ such that $f(n) \leq C \log^C(n) g(n)$ for all sufficiently large n.

probability and that has provably small communication and preference learning costs (costs that are bounded by $O(\sqrt{n})$ up to logarithmic factors). This suggests that a stable matching can be found (or reached via an equilibrium process) in an efficient manner.

Theorem 3.1. *The communication-efficient deferred acceptance (CEDA) protocol (Protocol 3 in Section 3.1.2) succeeds in finding the worker-optimal stable matching with high probability in any separable market. Its average communication cost per agent is at most $O^*(\sqrt{n})$ and its average preference learning cost per agent is at most $O^*(\sqrt{n})$.*

The protocol, as presented in Section 3.1, has an intuitive two-phase structure: In the first phase, firms signal to workers they privately like. The second-phase efficiently simulates the worker proposing deferred acceptance algorithms, with workers applying down their preference list but skipping applications to firms where they have no realistic chance of being accepted.

The CEDA protocol also achieves approximately optimal worst-case communication and preference learning costs.

Theorem 3.2. *There exists a separable market such that any communication protocol with broadcast that computes a stable matching with high probability requires an average communication cost of $\Omega(\sqrt{n})$ per agent and an average preference learning cost of $\Omega(\sqrt{n}/\log n)$.*

In the market constructed to establish Theorem 3.2, agents' preferences on both sides of the market are drawn from a latent random utility model. The construction leading to this result and intuition on the proof are given in Section 3.2.

3.1 Communication-efficient deferred acceptance (CEDA)

The key idea in the communication-efficient deferred acceptance (CEDA) protocol is to allow workers to better target their applications with the help of signals sent by firms. There are two types of signals:

- **Preference signal:** A firm j signals to worker i, which intuitively speaking, indicates that firm j has high private score ϵ_{ji} for worker i.
- **Qualification signal:** A firm j broadcasts a qualification requirement z_j to the entire market, which specifies the minimum public score for a worker who did not receive a preference signal to apply to the firm.

Although we formally allow every message to be public (see Section 2.3), this is not necessary. The protocol needs only qualification signals to be broadcast to every worker; preference signals need to be seen only by the worker to whom it is addressed.

The purpose of these signals is to help workers estimate at what firms they have a non-negligible chance of being accepted. For clarity of exposition, define the following terms.

Definition 3.3. A worker i is said to *publicly qualify* for firm j if the worker's public score meets the qualification requirement, $a_{ji} \geq z_j$. The worker is said to *qualify* for firm j if the worker either publicly qualifies or received a preference signal from the firm.

Moreover, define q_j to be the top $\frac{1}{\sqrt{n}}$ quantile of private scores, $q_j = F^{-1}(1 - \frac{1}{\sqrt{n}})$. Define $\underline{a}_j$ to be the minimum public score for firm j, $\underline{a}_j = \min_i\{a_{ji}\}$. Define μ_j to be firm j's latent utility for the current worker the firm is tentatively matched with. (This is initialized to u_{j0} and monotonically increases with each acceptance.)

Before describing the actual protocol, we give a simplified version (Protocol 2 in Section 3.1.1) to provide the main intuition. This version assumes that firms know μ_j and q_j. The actual CEDA protocol (Protocol 3 in Section 3.1.2) will be a modification that uses only ordinal preference information, which agents obtain via queries to the preference oracle (defined in Section 2.3). It also requires knowledge only of the bound B on the hazard rate, but no other information about the distribution of private scores.

3.1.1 Simplified protocol with cardinal utilities. The simplified version of the protocol is as follows.

Protocol 2. Simplified CEDA

For firm j, initialize the qualification requirement to $z_j = \underline{a}_j$.

(1) **Preference signaling:** *Each firm j sends a preference signal to each worker i that it privately likes, specifically, if $\epsilon_{ji} \geq q_j$.*
(2) **Deferred acceptance with qualification requirement:** *Run the worker-proposing deferred acceptance algorithm (Protocol 1 in Section 2.1) with the following two modifications:*
   (a) *The worker applies only to firms for which she qualifies. (So in Step 1 of Protocol 1, the chosen worker applies to the worker's favorite firm for which the worker both qualifies and has not yet been rejected.)*
   (b) *Whenever a firm tentatively accepts an application, let the latent utility for the tentatively matched option be μ_j. If the qualification requirement z_j is less than $\mu_j - q_j$, then send a qualification signal to update it to $z_j = \mu_j - q_j$.*[19]

Lemma 3.4. *Protocol 2 always finds the worker-optimal stable matching and, with high probability, the average communication cost per agent is at most $O^*(\sqrt{n})$.*[20]

The proof of Lemma 3.4 is given in Appendix C. We describe the main ideas below. Observe that, by construction, if a worker i is not qualified for firm j, then the firm would not accept the worker even if the worker applies (because the worker's not qualifying means that her public and private scores are both low). Thus, the above protocol implements the DA worker-proposing DA algorithm, except that it skips workers' proposals that would have never have been accepted. Therefore, the protocol always yields the worker-optimal stable match.

The main part of the proof of the lemma is to establish the bound on communication cost. This can be done by bounding the total number of preferences signals and the total number of applications. By construction, the number of preference signals per firm is $O(\sqrt{n})$ with high probability, and this bounds the number of applications from workers who do not publicly qualify. It suffices then to bound the number of applications from workers who publicly qualify.

For firm j, define Phase 1 as the period before the tentative match value μ_j reaches the threshold $\underline{a}_j + q_j$. Note that during Phase 1, no acceptance results in a qualification update, but after Phase 1, every acceptance results in a qualification update. Now, using the fact that q_j is defined to be the top $\frac{1}{\sqrt{n}}$th quantile, we show that Phase 1 ends with high probability after $2\log(n)\sqrt{n}$ applications. After Phase 1, we see that by the bounded hazard rate assumption, every application from a publicly qualified applicant will cause the qualification requirement to increase by at least 1 with probability at least $\frac{1}{\exp(B)\sqrt{n}}$. Since the total increase in qualification requirement is bounded by the range of public scores, which is at most $C\log^C(n)$, we obtain that with high probability, the total number of publicly qualified applications to firm j is at most $O^*(\sqrt{n})$.

[19]In practice, when updating the qualification requirement, we update it to the minimum public score that is above the current qualification requirement, as this allows us to communicate each qualification signal in $O(\log(n))$ bits by indicating the identity of the worker whose public score we set as the threshold.

[20]Note that one can terminate the above protocol with a failure message when the communication cost reaches a threshold of $O^*(\sqrt{n})$ to yield a protocol that is always bounded in communication and succeeds in finding a stable match with high probability.

3.1.2 Complete protocol with preference oracle only. The complete protocol that we propose is a modification of the previous protocol that can be implemented without firms having knowledge of their private scores. Furthermore, agents do not know their preferences a priori, but learn them by querying the preference oracle defined in Section 2.3 (*i.e.*, with each query, they learn their most preferred partner within a certain set). Finally, we no longer need any information about the distribution of private scores other than the bound B on the hazard rate.

Protocol 3. Communication-efficient deferred acceptance (CEDA)

Initialize the qualification requirement for each firm j to be $z_j = \underline{a}_j$.

(1) **Preference signaling:** *Each firm j bins workers according to public scores into bins $[\underline{a}, \underline{a}+1)$, $[\underline{a}+1, \underline{a}+2), \cdots$. In each bin, let the number of workers be l. The firm sends a preference signal to its top $k_1(n,l) := 2\exp(B)\max(\log^2(n), \frac{l}{\sqrt{n}})$ workers from that bin. (If there are fewer than this number of workers in the bin, then the firm sends a preference signal to all of them.)*

(2) **Deferred acceptance with qualification requirement:** *Run the worker-proposing deferred acceptance algorithm (Protocol 1 in Section 2.1) with the following two modifications:*
   (a) *Workers apply only firms for which they are qualified.*
   (b) *Each firm j updates the qualification requirement z_j as follows: it keeps z_j equal to the minimum value of $\underline{a}_j$ for the first $k_2(n) := 3\exp(B)\sqrt{n}$ applications that it receives. Call this Phase 1. After Phase 1, the firm sends a qualification signal to increase z_j by 1 after every $k_3(n) := 3\exp(B)\log(n)\sqrt{n}$ applications it receives from workers that publicly qualify.*[21]

The above protocol eliminates firms' need to know their cardinal utilities using two techniques: signaling by binning and updating qualification signals z_j based on receiving many applications. The binning in the preference signaling phase is designed so that, with high probability, a firm sends a preference signal to every worker that it would have signaled under the simpler protocol. The updating of qualification signals is designed so that the signal z_j is always lower than what it would have been under the simpler protocol. The idea is that because private scores are relatively fat-tailed, if a firm receives many applications, then the private score of at least one of them must be fairly high. And since these are applicants who publicly qualify, their public scores are also reasonably high, so the firm likely have found a good candidate. When this happens, workers who are not as competitive in terms of public score and who did not receive a preference signal have negligible chance of being accepted, so they might as well not apply.

3.2 The lower bound: Theorem 3.2

We use a natural distribution over preferences to prove the lower bound. For simplicity, we assume that preferences of both sides follow a separable structure, with the private component distributed according to an exponential distribution with rate parameter 1. There are n workers, ex ante identical to each other, and n firms, ex ante identical to each other. Each agent has an outside option of value $\frac{\log n}{2}$, and the public score between every worker and every firm is 0. We show that in this market, any matching protocol that succeeds with high probability requires $\Omega(\sqrt{n})$ bits of communication and $\Omega(\sqrt{n}/\log n)$ preference queries per agent.

Note that this is close to what is achieved under the naive protocol in which firms and workers simply communicate their entire preference lists. The reason is that in this market, each

[21]Note that if workers can observe the number of applications a firm receives from publicly qualified applicants, then the firms does not need to send qualification signal. Either the workers themselves or the platform can use the number of applications already received to infer that certain workers do not have a chance.

agent finds roughly $O(\sqrt{n})$ partners acceptable (whom they prefer over their outside option), and communicating the entire preference list over acceptable partners requires only $\sqrt{n}\log n$ bits.

The formal proof is deferred to Appendix B. The main ideas are as follows. First, we make the problem simpler and allow workers to communicate with each other without cost (effectively creating a "super-worker" for communication purposes, who possesses all information known to workers), and allow firms to communicate with each other without cost (creating a "super-firm").[22] In this simplified communication problem, we prove an $\Omega(n^{3/2})$ lower bound on the number of bits of communication needed between the super-worker and the super-firm, to compute a stable matching with high probability. This implies an $\Omega(\sqrt{n})$ lower bound on the average communication cost per agent. We infer an $\Omega(\sqrt{n}/\log n)$ lower bound on the preference learning cost per agent, since any protocol that uses K preference learning oracle calls can be modified into a communication protocol that has a communication cost of $O(K\log(n))$.

To show the above bounds, we first show that for a matching protocol to succeed with high probability, it must identify approximately if each worker-firm pair finds each other mutually acceptable. The reason is that a significant fraction of these mutually acceptable pairs must be matched in all stable matchings. There are n^2 worker-firm pairs in total. For each pair, we use ideas from information complexity theory (see [Braverman, 2015]) to show that the protocol must use $\Omega(1/\sqrt{n})$ bits of communication on average to approximately determine mutual acceptability,[23] leading to a bound of $\Omega(n^2/\sqrt{n}) = \Omega(n^{3/2})$ on the total amount of communication.

4 TWO-ROUND PROTOCOL WITH PRIVATE COMMUNICATION

The CEDA protocol in Section 3.1 is sequential: a worker's decision about which firm to apply to depends on the firms' current qualification requirements, which in turn depends on other workers' application decisions. Implementing such a protocol may create undesirable congestion, 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*[24]: 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 from the first round. One interpretation of the central matchmaker is a centralized clearinghouse such as in the National Residency Matching Program (NRMP). Another interpretation is that it is a proxy for decentralized matchmaking. Signaling then would intuitively correspond to initiating contact, and we assume that after a decentralized matchmaking process, agents arrive at a matching with no blocking pair of agents who are in contact with one another.[25]

It turns out one can construct a separable market in which any two-round protocol that finds a stable matching with high probability requires $\Omega(n)$ bits of communication per agent (see Appendix D). As a result, we settle for a simpler model of matching markets, which we call *tiered random markets*. In such markets, agents are partitioned into tiers, and everyone prefers better tiers to worse tiers and has uniformly random preferences among partners of a given tier (see Section 4.1). This model still allows both vertical and horizontal differentiation, and it yields clean insights into

[22]Further, we allow the public scores to be common knowledge among agents.

[23]Note that in this setup, i finds j acceptable with probability $1/\sqrt{n}$, and the same with i and j swapped. Each of i and j can communicate this exactly using $O(\log n/\sqrt{n})$ bits (on average). In Appendix B, we show that one cannot do much better, in particular, $\Omega(1/\sqrt{n})$ bits of communication are *required* to approximately determine mutual compatibility.

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

[25]In the labor market, either the worker needs to reach out to the firm, or the firm needs to reach out to the worker (perhaps through a headhunter) before the formal screening process can commence.

the relative competitive positions of agents and what kinds of signals are the most informative (see Section 4.2 for an example). We present our two-round protocol in Section 4.3.

An additional advantage of the protocol we present is that it uses only *private messages*, which are messages that are visible only to a sender and a receiver, but not to anyone else. (This is in contrast with the messages defined in Section 2.3, which are visible to everyone.) Requiring messages to be private models communication in decentralized markets, in which there is no central platform to filter for each agent communication relevant for that agent.[26] Requiring private messages automatically controls for the amount of information agents receive through the protocol.

4.1 Tiered random markets

A *tiered market* is a two-sided matching market with certain restrictions on preferences. Let agents be partitioned into disjoint *tiers*, and restrict the preferences of agents so that each agent strictly prefers partners of better tiers to worse tiers. For general tiered markets, the preference of agents for partners of the same tier can be arbitrary. Let I_1 denote the first (best) tier of workers, I_2 the second tier, and so on. We have that the set of workers, $I = \bigcup_k I_k$, where index k is called the *tier index* of workers. Similarly define J_k as the kth tier of firms, with $J = \bigcup_k J_k$. For each k, let $m_k = |I_k|$ and $n_k = |J_k|$. Let $m = \sum_k m_k$ be the total number of workers and $n = \sum_k n_k$ be the total number of firms. (In Section 2, we denoted m and n by n_I and n_J respectively, but we alter notation slightly for simplicity in the tiered random case.)

For any tiered market, define s_k to be the number of workers in tiers 1 through k, $s_k = \sum_{i=1}^{k} m_k$. Similarly, define $t_k = \sum_{i=1}^{k} n_k$. We say that tiers I_k and J_l are *overlapping* if $s_k > t_{l-1}$ and $t_l > s_{k-1}$. Note that only agents from overlapping tiers can be matched in any stable matching.

A *uniformly random market* is a two-sided matching market in which the preferences of agents are uniformly random and independent, and every agent finds every partner acceptable.[27] The uniformly random market is the most well-studied stochastic model of matching in the literature.[28]

A *tiered random market* is a tiered market in which the preferences of agents between any individual tier I_k of workers and J_l of firms follow that of a uniformly random market. In other words, agents strictly prefer partners of better tiers, but have uniformly random and independent preferences for partners within a given tier. Furthermore, everyone prefers being matched to anyone over being unmatched.

4.2 Example and intuition

Before giving the full protocol, we consider a simple numerical example, which will illustrate the main insights. Suppose that there are two tiers of workers, of 50 workers each. There are also two tiers of firms, the top tier with 20 firms, and the second with 90 firms. This is illustrated in Figure 1.

The first observation is that the tier structure precludes certain matches in stable matching. For example, firms in J_1 are guaranteed to be able to get workers in I_1, so workers in lower tier I_2 cannot hope to be matched to the top tier firms in J_1.

The second observation is that we can infer whose preferences matter more. One insight from the theory of uniformly random markets is that it is most efficient in terms of communications for the short side to propose to the long side. (For example, [Ashlagi et al., 2016] show that in a market

[26]In CEDA, it may be realistic to think that a central platform notifies workers whenever they are no longer qualified for a firm, or that there is a central place where workers can see which firms they are currently qualified for. But it would be unrealistic to expect every agent to keep checking the website of each firm individually for updates to the qualification requirements.

[27]Precisely speaking, for every worker $i \in I$, the preference ranking $R^I(i)$ is a uniformly random and independent permutation of J. An analogous statement holds for firms.

[28]See for example [Wilson, 1972], [Knuth, 1976], [Pittel, 1989], [Ashlagi et al., 2016].

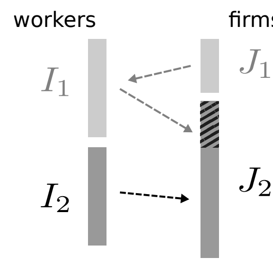

Fig. 1. An example of a tiered random market (see Section 4.2). There are two tiers of workers, I_1 and I_2, and two tiers of firms, J_1 and J_2. The height of each rectangle corresponds to the number of agents in that tier. The arrows show the direction of the most informative signaling. The diagram highlights the situation of worker tier I_2: these workers should signal to the firms in J_2 since they are in a better competitive position. However, they should amplify the number of signals they send to account for the fact that some signals would be "wasted" on the firms (represented by stripes) that received offers from better workers from I_1.

with $n-1$ workers and n firms, the number of applications in the worker-proposing DA algorithm is about $\log(n)$ per agent, whereas the number of applications in the firm-proposing DA is about $\frac{n}{\log(n)}$ per agent.) Hence, it makes the most sense for the top firms in J_1 to signal to the workers in I_1 rather than the other way around. (See the arrows in Figure 1.)

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 many partners they signal would be taken up by agents from better tiers, against whom they do not have a chance. For example, for the second tier of workers I_2, there are 90 firms in J_2 they can signal, but 30 of them would receive better offers from workers in I_1, so for every three signals they send, one would be essentially wasted. This means that they should amplify the number of signals they send by a factor of $\frac{3}{2}$.

For certain agents under certain tier structures, this amplification effect may be large. For example, consider the case in which there is one 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 bottom workers may need to signal up to $O(n)$ number of firms, because of the amplification effect. However, one can show that the *average* amplification needed is only $O(\log n)$.[29] For arbitrary tier structures, one can show that this is always the case: the average amplification needed is small.

These insights are illustrated in Figure 1, and they form the basis of our protocol in Section 4.3.

4.3 The targeted signaling protocol

In this section, we derive a two round protocol for tiered random markets called the *targeted signaling protocol.* It elicits information in a simultaneous way, having agents signal to those they idiosyncratically like via private messages in parallel.

We first present a high-level summary of the protocol. First, we designate for each agent a *target tier* based on the tier structure alone, which intuitively represents the best partners for whom the agent has high chances of getting. The protocol requires every agent to signal his favorite partners within his target tier, with the number of signals being determined by a formula that incorporates the agent's competitiveness in the market (see Definition 4.3). After all the signals have been sent, the protocol then collects partial preferences among agents between whom a signal was sent, and runs the DA algorithm on the partial preferences collected. (The DA algorithm is for concreteness; any process that results in a stable matching restricted to the partial preferences can be used.)

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

More precisely, we define the following notions for a tiered random market. For any tier of workers I_k, we say that the tier of firms J_l is in a *(weakly) worse competitive position* if $t_l \geq s_k$.

Definition 4.1. For each worker i in tier I_k, define the *target tier* of the worker as best tier of firms that is in a (weakly) worse competitive position. In other words, this is J_l, where

$$l = \min\{l : t_l \geq s_k\}.$$

Similarly define the target tier of firms by switching the role of I and J, s and t.

Definition 4.2. For each worker $i \in I$ in tier I_k with target tier J_l, define the worker's *target tier competitiveness* as

$$\rho(i) = \frac{t_l - s_{k-1}}{n_l}.$$

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

In the example in Figure 1, the target tiers are indicated by the arrows, so that the target tier of firm-tier J_1 is worker-tier I_1, the target tier of worker-tier I_1 is firm-tier J_2 and so on. Firms J_2 do not have a target tier because they are in the worst competitive position possible. The target tier competitiveness of the worker-tier I_1 and firm-tier J_1 are both 1, since these are the best tiers on each side. The target tier competitiveness of I_2 is the proportion of agents in J_2 who are not shaded, and is equal to $\frac{60}{90} = \frac{2}{3}$ in this example.

Definition 4.3. Define the *target number* of an agent $a \in I \cup J$ as

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

Note that the target number is higher for agents with lower target tier competitiveness.

Definition 4.4. A *redundant firm-tier* is a firm-tier l whose target tier k satisfies $s_k = t_l$.

The targeted signaling protocol is defined precisely as follows. Beside the workers and firms, there is an additional agent called the matchmaker, who synchronizes the market into two rounds and runs the DA algorithm on partial preference submissions. (This is a theoretical construct and can be interpreted as a proxy for a decentralized matching process.)

(1) **Signaling round:** Each agent $a \in I \cup J$, except those in redundant firm-tiers, signals to its favorite $r(a)$ partners in the agent's target tier. After this, the agent sends a notification to the matchmaker indicating that the signaling is done.
(2) **Matching round:** Once the matchmaker receives the notification from everyone, he elicits from each agent 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 matchmaker then runs the DA algorithm (either worker-proposing or firm-proposing) based on the partial rankings. If the result contains fewer than $\min(m, n)$ matches, then output a failure message. Otherwise, output the resultant matching.

THEOREM 4.5. *For any tiered random market, the targeted signaling protocol is a stable matching protocol that uses only private messages and succeeds with high probability.*[30] *Without loss of generality, let $n \geq m$. The protocol's average communication cost is $\Theta(\log^4(n))$ per agent and its average preference learning cost is $\Theta(\log^3(n))$ per agent.*

[30]The protocol is deterministic, so the probability is taken over the randomness in the preferences.

The proof of Theorem 4.5 is in Appendix F. The main steps are as follows: First, we show that, with high probability, the set of signals sent in the signaling round contains a stable matching. This result, stated in Lemma F.2, can be interpreted as a generalization of Theorem 6.1 b) of [Pittel, 1992] to markets with richer structures. The proof of this uses a new bound on the average rank of agents in any stable matching in unbalanced uniform random markets.[31] Second, we show that whenever the set of signals contains a stable matching, running DA on the partial preferences as in the matching round would also return a stable matching.[32] Third, we count the total number of signals and show that it is no more than $O(\log^3(n))$ per agent.

Remark 1. *We also have a variant of the targeted signaling protocol with a lower average communication cost of* $\Theta(\log^3(n))$ *per agent and a preference learning cost of* $\Theta(\log^2(n))$ *per agent, but that involves a much more complicated formula for the target number than Definition 4.3. We think that the gain is minimal, so we show the simpler and more intuitive version here.*

Remark 2. *In Appendix H, we show that any stable matching protocol that uses only private communication and succeeds with high probability uses at least* $\Omega(\log^2(n))$ *bits of communication per agent. This lower bound can be attained by a protocol that uses sequential signaling, which means it is not a two-round protocol. The targeted signaling protocol presented above achieves near optimal communication cost, while having the added advantage of being a two-round protocol.*

5 INCENTIVE COMPATIBILITY

The deferred acceptance (DA) algorithm, even in the setting with full preference elicitation, is not completely incentive compatible: an agent on the non-proposing side may profitably deviate from truthful reporting by truncating his or her preference ranking.[33] Nevertheless, the DA algorithm is known to be strategyproof for the proposing side.[34] Moreover, assuming truthful reporting by other agents, an agent on the non-proposing side cannot unilaterally deviate and be matched with someone better than her best stable partner,[35], which implies that DA is incentive compatible for these agents if they have a unique stable partner. In summary, under DA agents can never unilaterally deviate from truthful reporting and get someone better than their best stable partner, and under full information of the preferences of others, they can always unilaterally deviate to get matched to one of their stable partners.

In a communication protocol, we say that the agent *complies* with the protocol if she truthfully participates according to what is prescribed, and that the agent *deviates* from the protocol otherwise. We show that all the communication protocols proposed in this paper are in a certain sense “as incentive compatible” as DA with full preference elicitation, in large markets.

Definition 5.1. A stable-matching protocol is *as incentive compatible as DA* if no agent can unilaterally deviate from the protocol and be matched with someone better than their best stable partner (under true preferences). Similarly, a protocol 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,

[31] The bound, stated in Proposition G.1, is analogous to Theorem 1(i) of [Ashlagi et al., 2016], except that the constant is worse but the notion of “high probability” is stronger. The proof of Proposition G.1 is based on the integral formula technique of [Pittel, 1992] and is therefore much cleaner than the analysis in [Ashlagi et al., 2016].

[32] This result uses the structure of tiered markets and the definition of target tier, and is not true in general: even if a set of partial preferences contains a stable matching, running DA on these partial preferences may result in a matching that is stable only with respect to the partial preferences but not with respect to the whole market.

[33] See Chapter 4 of [Roth and Sotomayor, 1990] for a thorough exposition of all these incentive properties of DA.

[34] See [Roth, 1982] and [Dubins and Freedman, 1981].

[35] *I.e.*, their most preferred partner in all stable matchings. For a proof, see Theorem 4.11 in [Roth and Sotomayor, 1990] which is originally credited to [Demange et al., 1987].

the probability that the agent can unilaterally deviate from the the protocol and be matched with someone better than her best stable partner (under the true underlying preferences) is at most $\delta(n)$.

Theorem 5.2. *The simple version of the CEDA protocol (from Section 3.1.1) is as incentive compatible as DA. The modified version of the CEDA protocol (from Section 3.1.2) and the targeted signaling protocol (from Section 4.3) are as incentive compatible as DA with high probability.*

The "high probability" qualifications are needed for the latter two protocols because, with a small probability, they may fail to find a stable matching. The proof of Theorem 5.2 is in Appendix I with an outline of the main argument here: we first observe that all these protocols can be described as running the worker-proposing DA on a suitable "subgraph of signals" G. When an agent unilaterally deviates from the protocol, she may influence the final subgraph to be some other G' and induce a matching μ' outputted by the protocol. We then apply a blocking lemma[36] as in [Gale and Sotomayor, 1985] on μ' and G and argue that there must be an edge $(i, j) \in G$ that blocks μ', and we argue that this same edge must be present in G', which contradicts μ' being the output of the protocol.

Corollary 5.3. *The simple version of the CEDA protocol is strategyproof for workers. In the modified version of the CEDA protocol, for each worker, with high probability she cannot profitably deviate from the protocol with high probability.*

Theorem 5.2 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 for each agent, the probability that the agent has multiple stable partners is small, then complying with the protocol is an ϵ-Bayes Nash equilibrium for all parties. In tiered-random markets, we can prove this second condition for "typical" markets when the size of each tier is large. This is made precise as follows.

Definition 5.4. A tiered market satisfies *general imbalance* if the sets $\{t_1, t_2, \cdots\}$ and $\{s_1, s_2, \cdots\}$ are disjoint.

Theorem 5.5. *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 $\delta(y)$.*

6 DISCUSSION

This paper studies the communication requirements in two-sided matching markets. We find that the communication cost for reaching a stable matching is typically very low. For markets in which firms have separable preferences and workers have arbitrary preferences, we construct a communication protocol (CEDA), under which each agent only needs to send $O^*(\sqrt{n})$ messages on average, which vastly reduces congestion compared to the $O(n)$ benchmark of a naive implementation of deferred acceptance. With additional structure over preferences (as in tiered-random markets), we identify a two-round protocol, which allows for parallel signaling and requires only $\Theta(\log^4 n)$ bits of communication per agent. In each of these protocol, the main idea is to encourage agents to send signals that help other agents understand with whom they have a realistic chance of matching with.

[36]The same blocking lemma can be used to prove the incentive results for DA. See pages 92–93 of [Roth and Sotomayor, 1990].

Our results provide some theoretical justification for applying the solution concept of stability to real world matching markets. They can also help guide improved designs for matching markets. For example, in school choice a student may find it very useful to know which set of schools she is likely to be accepted at, as this can reduce her search costs. So our results motivate school choice systems providing more information about chances of matching, similar to what is currently done for university admission in Israel. Online matching markets (labor markets, dating platforms) may also use prior knowledge to guide match recommendations, facilitate appropriate signalling, and so on.

Our paper suggests directions for future research on the nature of congestion. For example, in Appendix D, 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 round of an aftermarket (as done, for example, in the *scramble* after the main match conducted by the NRMP). This raises the question of how many agents may remain unmatched after a two-round protocol, or alternatively, how many rounds are needed to find a stable matching.

REFERENCES

Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda. 2015. Expanding "choice" in school choice. *American Economic Journal: Microeconomics* 7, 1 (2015), 1–42.

Itai Ashlagi, Yash Kanoria, and Jacob D. Leshno. 2016. Unbalanced Random Matching Markets: The Stark Effect of Competition. *Journal of Political Economy (to appear)* (2016).

Christopher Avery, Andrew Fairbanks, and Richard J. Zeckhauser. 2009. *The early admissions game.* Harvard University Press.

Christopher Avery, Christine Jolls, Richard A. Posner, and Alvin E. Roth. 2001. The market for federal judicial law clerks. *The University of Chicago Law Review* 68, 3 (2001), 793–902.

Laszlo Babai, Peter Frankl, and Janos Simon. 1986. Complexity classes in communication complexity theory. In *Proceeding of the 27th Annual Symposium on Foundations of Computer Science.* IEEE, 337–347.

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

Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. 2013. From information to exact communication. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing.* ACM, 151–160.

Jen-Hou Chou and Chi-Jen Lu. 2010. Communication requirements for stable marriages. In *Algorithms and Complexity.* Springer, 371–382.

Peter Coles, John Cawley, Phillip B. Levine, Muriel Niederle, Alvin E. Roth, and John J. Siegfried. 2010a. The job market for new economists: A market design perspective. *The Journal of Economic Perspectives* 24, 4 (2010), 187–206.

Peter Coles, Alexey Kushnir, and Muriel Niederle. 2010b. *Preference signaling in matching markets.* Technical Report. National Bureau of Economic Research.

Thomas M. Cover and Joy A. Thomas. 2012. *Elements of information theory.* John Wiley & Sons.

Gabrielle Demange, David Gale, and Marilda Sotomayor. 1987. A further note on the stable matching problem. *Discrete Applied Mathematics* 16, 3 (1987), 217–222.

William Diamond and Nikhil Agarwal. 2016. Latent Indices in Assortative Matching Models. *Quantitative Economics* (2016). forthcoming.

Lester E. Dubins and David A. Freedman. 1981. Machiavelli and the Gale-Shapley algorithm. *The American Mathematical Monthly* 88, 7 (1981), 485–494.

Federico Echenique, Alistair J. Wilson, and Leeat Yariv. 2016. Clearinghouses for two-sided matching: An experimental study. *Quantitative Economics* 7, 2 (2016), 449–82.

Federico Echenique and Leeat Yariv. 2012. An experimental study of decentralized matching. (2012). working paper.

David Gale and Lloyd L. Shapley. 1962. College admissions and the stability of marriage. *Amer. Math. Monthly* 69 (1962), 9–15.

David Gale and Marilda Sotomayor. 1985. Some remarks on the stable matching problem. *Discrete Applied Mathematics* 11, 3 (1985), 223–232.

Yannai A. Gonczarowski, Noam Nisan, Rafail Ostrovsky, and Will Rosenbaum. 2015. A stable marriage requires communication. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms.* SIAM, 1003–1017.

Dan Gusfield. 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16, 1 (1987), 111–128.

Guillaume Haeringer and Myrna Wooders. 2011. Decentralized job matching. *International Journal of Game Theory* 40, 1 (2011), 1–28.

John William Hatfield, Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp. 2015. *Chain stability in trading networks.* Technical Report. Working paper.

Friedrich August Hayek. 1945. The use of knowledge in society. *The American Economic Review* (1945), 519–530.

Günter J Hitsch, Ali Hortaçsu, and Dan Ariely. 2010a. Matching and sorting in online dating. *The American Economic Review* 100, 1 (2010), 130–163.

Günter J. Hitsch, Ali Hortaçsu, and Dan Ariely. 2010b. What makes you click? Mate preferences in online dating. *Quantitative marketing and Economics* 8, 4 (2010), 393–427.

Leonid Hurwicz. 1973. The design of mechanisms for resource allocation. *The American Economic Review* 63, 2 (1973), 1–30.

Nicole Immorlica and Mohammad Mahdian. 2005. Marriage, honesty, and stability. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms.* 53–62.

John H. Kagel and Alvin E. Roth. 2000. The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *Quarterly Journal of Economics* (2000), 201–235.

Bala Kalyanasundaram and Georg Schnitger. 1992. The Probabilistic Communication Complexity of Set Intersection. *SIAM Journal on Discrete Mathematics* 5, 4 (Nov. 1992), 545–557.

Alexander S. Kelso Jr. and Vincent P. Crawford. 1982. Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society* (1982), 1483–1504.

Donald E. Knuth. 1976. *Marriages stables et leurs relations avec d'autres problemes combinatoires.* Les Presses de l'Universite de Montreal.

F. Kojima and P. A. Pathak. 2009. Incentives and Stability in Large Two-Sided Matching Markets. *American Economic Review* 99 (2009), 608–627.

Eyal Kushilevitz and Noam Nisan. 2006. *Communication Complexity.* Cambridge University Press.

Robin S. Lee and Michael Schwarz. 2009. *Interviewing in two-sided matching markets.* Technical Report. National Bureau of Economic Research.

SangMok Lee. 2016. Incentive compatibility of large centralized matching markets. *Review of Economic Studies* (2016). forthcoming.

Soohyung Lee and Muriel Niederle. 2015. Propose with a rose? signaling in internet dating markets. *Experimental Economics* 18, 4 (2015), 731–755.

Konrad Menzel. 2015. Large Matching Markets as Two-Sided Demand Systems. *Econometrica* 83, 3 (2015), 897–941.

Kenneth Mount and Stanley Reiter. 1974. The informational size of message spaces. *Journal of Economic Theory* 8, 2 (1974), 161–192.

Cheng Ng and Daniel S. Hirschberg. 1991. Three-Dimensional Stable Matching Problems. *SIAM Journal on Discrete Mathematics* 4, 2 (1991), 245–252.

Muriel Niederle and Leeat Yariv. 2009. *Decentralized matching with aligned preferences.* Technical Report. National Bureau of Economic Research.

Joana Pais. 2008. Incentives in decentralized random matching markets. *Games and Economic Behavior* 64, 2 (2008), 632–649.

Marcin Peski. 2015. Large Roommate Problem with Non-Transferable Random Utility. (2015). working paper.

Boris Pittel. 1989. The average number of stable matchings. *SIAM Journal on Discrete Mathematics* 2, 4 (1989), 530–549.

Boris Pittel. 1992. On Likely Solutions of a Stable Marriage Problem. *The Annals of Applied Probability* 2, 2 (1992), 358–401.

Alexander Razborov. 1992. On the Distributed Complexity of Disjointness. *TCS: Theoretical Computer Science* 106 (1992).

Alvin E. Roth. 1982. The economics of matching: Stability and incentives. *Mathematics of operations research* 7, 4 (1982), 617–628.

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

Alvin E. Roth. 2008. What have we learned from market design? *The Economic Journal* 118, 527 (2008), 285–310.

Alvin E. Roth and Oliveira Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Cambridge University Press.

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

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

Lloyd S. Shapley and Martin Shubik. 1971. The assignment game I: The core. *International Journal of game theory* 1, 1 (1971), 111–130.

L. B. Wilson. 1972. An analysis of the stable marriage assignment algorithm. *BIT Numerical Mathematics* 12, 4 (1972), 569–575.

A RELATED LITERATURE

This paper relates to several strands of literature. First is the literature on communication complexity in economic models. The amount of communication required to reach efficient outcomes in markets with prices was first studied formally by Hurwicz [1973] and Mount and Reiter [1974].[37] Segal [2007] studies more generally the communication requirements for social choice rules, including stable matching, and shows that minimal information is captured by "budget sets," which can be viewed as analogs of budget sets in convex economies.[38]

A number of papers analyze the communication requirements for reaching or verifying stability in two-sided matching markets under different models of communication (Chou and Lu [2010], Gonczarowski et al. [2015], Ng and Hirschberg [1991]).[39] These studies, including Segal [2007], make no assumptions about agents' preferences and find that $\Omega(n)$ bits of information are needed per agent in the worst case.

The next strand of literature concerns large two-sided matching markets. Various papers investigate the size of the core and incentive compatibility of stable mechanisms (Ashlagi et al. [2016], Immorlica and Mahdian [2005], Kojima and Pathak [2009], Lee [2016]). Knowing that the core is small does not imply, however, that the communication cost for finding a stable matching is small. Close to our paper are Ashlagi et al. [2016] and Pittel [1992] who study random markets with a single tier on each side. Ashlagi et al. [2016] find that in slightly unbalanced markets, agents on the short side get matched, on average, to their $O(\log n)$-th ranked partner. Their results imply that on average the amount of information needed is $O(\log^2 n)$ per agent, since a communication protocol can query preferences of agents on the short side and send application messages to agents on the long side. Their analysis can be extended to tiered markets with a constant number of large tiers on each side, but it is unclear what happens when the number of tiers grows with n. Our protocol for tiered markets achieves $O(\log^4 n)$ bits of communication per agent for an *arbitrary tier structure*, and this protocol has the additional advantage of allowing signals in parallel. The analysis builds on techniques from Pittel [1992] and thus can be viewed as a generalization of his work to more complex preferences. Various papers ask under what conditions a decentralized market is likely to reach a stable outcome (Echenique and Yariv [2012], Haeringer and Wooders [2011], Niederle and Yariv [2009], Pais [2008]). These papers find that complete information enables stable matchings to arise in equilibrium, but more structure over preferences is required under incomplete information.

Structural assumptions on agents' preferences in large matching market is not novels. Numerous papers take this approach for the purposes of identification and analysis of equilibrium properties (see Diamond and Agarwal [2016], Hitsch et al. [2010b], Lee [2016], Menzel [2015], Peski [2015]). It is worth noting that Menzel [2015] and Peski [2015] assume agents have only $O(\sqrt{n})$ acceptable partners, so agents in their model have only $O^*(\sqrt{n})$ communication cost. We show that the same low level of communication is achievable under much more general assumptions about preferences.

Our paper also connects with the literature on congestion and signaling in matching markets, and we contribute to this literature by giving a mathematically precise measure of congestion in terms of communication complexity. Previous work document congestion and how agents deal with it in practice, but do not give a mathematically precise measure. Roth and Xing [1994] and Avery

[37] These studies stemmed from Hayek [1945], who asked how decentralized information can be utilized to reach economic efficiency.

[38] Various papers implicitly study budget sets to characterize core outcomes in two-sided markets (Kelso Jr. and Crawford [1982], Shapley and Shubik [1971]) and in trading networks (Hatfield et al. [2015]).

[39] Prior to these studies, Knuth [1976] asked whether stability can be reached in faster than n^2 steps as required by DA in the worst-case scenario and Gusfield [1987] asked whether verifying stability can be done in faster than quadratic time.

et al. [2001] document congestion in labor markets, initiating a large literature on unraveling.[40] Echenique et al. [2016] study the performance of a centralized deferred acceptance mechanism in the lab and find that many agents on the proposing side tend to skip potential partners, which resembles our protocol if such skipping arises from agents' pessimism about the likelihood of being accepted. Previous papers that study signaling in matching markets include Lee and Schwarz [2009], Coles et al. [2010b], Lee and Niederle [2015], and Abdulkadiroğlu et al. [2015]. However, in all these papers, signaling improves the cardinal efficiency over stable matchings, whereas in our work, signaling still leads to a stable matching but the communication overhead is decreased.

Finally, our work builds on tools developed in the computer science literature on communication complexity. Our proof for the lower bound in separable markets uses the theory of information complexity, which is a relatively recent approach that uses Shannon's information theory to derive tight communication complexity bounds.[41] The $\sqrt{n}$ expression in the lower bound is related to results established for the communication complexity of set disjointness, one of the most studied problems in communication complexity theory (see [Kalyanasundaram and Schnitger, 1992] and [Razborov, 1992]).

B PROOF OF LOWER BOUND WITH BROADCAST

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, since Alice's messages are allowed to depend on the preferences of all workers simultaneously (which amounts to having "free" communication among workers). We will show an example where $\Omega(n^{3/2})$ communication between Alice and Bob is necessary, which will immediately imply 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 $Exp(1)$, which is the exponential distribution with rate parameter 1. (*i.e.*, the public scores are zero, and the private scores are exponentially distributed.) Let the the value of the outside option be $\frac{\log n}{2}$ for every agent. Note that $\mathbb{P}(v_{ij} \geq \frac{\log n}{2}) = \frac{1}{\sqrt{n}}$. 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 π be the communication protocol, at the end of which Alice and Bob output a matching μ_π that is stable with probability at least .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

[40]In the the market for clinical psychologists, which was organized in the early 1990's in a decentralized match day, numerous program directors skipped their rankings and made offers to applicants that indicated they would accept the offer (Roth and Xing [1994]).

[41]A more detailed account of information complexity and its connections to communication complexity can be found in [Braverman, 2015].

bits of communication), is bounded as $|\pi| = \mathrm{E}[|\Pi|] = \Omega(n^{3/2})$, where Π is the realization of the protocol π.

We will 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 will lead to Claim B.2 below). If worker i and firm j are a mutually acceptable pair, and moreover, this is the only mutually acceptable pair that each of i and j is a member of, then (i, j) *must* be a matched pair under any stable matching. This will yield Claim B.1 below. We will draw on ideas from information complexity theory (see, e.g., [Braverman, 2015]), together with Claims B.1 and B.2, to establish Claim B.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):

$$A_{ij} := \begin{cases} 1 & \text{if } v_{ij} \geq \frac{\log n}{2} \\ 0 & \text{otherwise} \end{cases} \quad \text{and } B_{ij} := \begin{cases} 1 & \text{if } u_{ji} \geq \frac{\log n}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

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(α), 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 μ_π.

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

Proof. 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 $\exp(-\frac{\sqrt{n}}{3})$. Since the probability that each of these firms find worker i to be acceptable is exactly $\frac{1}{\sqrt{n}}$, the probability that another firm out of these other than j finds worker i acceptable is at most $1 - (1 - \frac{1}{\sqrt{n}})^{2\sqrt{n}}$, 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 $e^{-2.1}$. Since 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,

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

□

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

Note that Claims B.1 and B.2 together imply that π 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 B.3. *We have*

$$I(A_{ij}B_{ij};\Pi) = \Omega(\alpha) = \Omega(1/\sqrt{n}). \tag{3}$$

Here Π is the again the random variable representing the realization of the protocol π, 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 them. Note that fully revealing the values of A_{ij}, B_{ij} corresponds to Shannon's entropy $H(A_{ij}, B_{ij}) = \Theta((\log n)/\sqrt{n})$. Let us prove Claim B.3.

Proof. 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 π of Π, the variables $(A_{ij}|\Pi = \pi)$ and $(B_{ij}|\Pi = \pi)$ are also independent. This is because 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):

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

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:

$$I(A_{ij};\Pi) = \mathbf{E}_{\pi\sim\Pi}\mathbf{D}_{KL}(A_{ij}|_{\Pi=\pi}\|A_{ij}). \tag{4}$$

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 $c' < c$

$$\mathbf{D}_{KL}(\text{Bernoulli}(c' \cdot x)\|\text{Bernoulli}(x)) = \Omega_c(x), \tag{5}$$

where the Bernoulli random variable Bernoulli(x) takes the value 1 w.p. x, and the value 0 w.p. $1 - x$.

By Claims B.1 and B.2 we have that

$$\mathbb{P}[M_{ij} = 0] \geq \mathbb{P}[(A_{ij}, B_{ij}) \neq (1,1)] = (1 - \alpha^2), \tag{6}$$

and

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

Therefore, for a sufficiently large n (and thus a sufficiently small α),

$$\mathbb{P}[(A_{ij}, B_{ij}) = (1,1)|M_{ij} = 0] < \frac{\alpha^2(1 - 10^{-2})}{1 - \alpha^2} \leq \alpha^2 \cdot (1 - 9 \cdot 10^{-3}). \tag{8}$$

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

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

and thus by Markov's inequality

$$\mathbb{P}_{\pi\sim\Pi_{M_{ij}=0}}[\mathbb{P}[(A_{ij},B_{ij})=(1,1)|\Pi=\pi]<\alpha^2\cdot(1-2\cdot 10^{-3})]\geq 1-\frac{0.991}{0.998}>7\cdot 10^{-3}. \tag{9}$$

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 (5) above, for any $\pi\in\Pi_{M_{ij}=0}$,

$$\mathbf{D}_{KL}(A_{ij}|_{\Pi=\pi}\|A_{ij})+\mathbf{D}_{KL}(B_{ij}|_{\Pi=\pi}\|B_{ij})=\Omega(\alpha). \tag{10}$$

By (6) and (9), the probability of such a π is at least $7\cdot 10^{-3}\cdot(1-\alpha^2)>6\cdot 10^{-3}$ for sufficiently large n. The contribution of such π'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 $\Omega(\alpha)$, since $\mathbf{D}_{KL}$ is always non-negative, this implies by (4) that

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

concluding the proof. □

Fact 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 3.2. Since A_{ij},B_{ij} are mutually independent for the different values if (i,j), we get using the above fact, and the fact that entropy is always an upper bound on mutual information,

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

(We have used Claim B.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, since 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 $R=\Omega(n^{3/2}/\log n)$.

C PROOF OF CORRECTNESS AND EFFICIENCY OF CEDA

In this appendix, we prove more general versions of Lemma 3.4 and Theorem 3.1 in Section 3.1. The generality allows us to relax the $C\log^C(n)$ bound on the range of public scores to a general function $g(n)$. The more general results are stated below.

Let $s(n)$ and $g(n)$ be functions of n. In the main text of the paper, $s(n)=\sqrt{n}$ and $g(n)=C\log^C(n)$. In this section, let them be general functions with $s(n)=\Omega(\sqrt{n})$, $s(n)=o(\frac{n}{\log n})$, $g(n)=\Omega(\log n)$, and $g(n)=o(n)$.[42] Let B be the upper bound on the hazard rate of the distribution of private scores. We have the following results which subsume Lemma 3.4 and Theorem 3.1.

Theorem C.1. *Assume that the public scores satisfy* $\max_i a_{ji}\leq\min_i a_{ji}+g(n)$ *for all firms j. The simplified CEDA protocol (Protocol 2 in Section 3.1.1) using $q_j=F_j^{-1}(1-\frac{1}{s(n)})$ always finds the worker-optimal stable match. Its communication cost is at most*

$$O(\max(g(n)s(n),\frac{n}{s(n)})\log n)$$

with probability at least $1-O(\frac{1}{n})$.

[42] We say that a function $f(n)=\Omega(g(n))$ if there exists constant $c>0$ such that $f(n)\geq cg(n)$ for all sufficiently large n. We say that a function $f(n)=o(g(n))$ if $\lim_{n\to\infty}\frac{f(n)}{g(n)}\to 0$. We say that $f(n)=O(g(n))$ if there is a constant $c>0$ such that $f(n)\leq cg(n)$ for all sufficiently large n.

Theorem C.2. *Assume that the public scores satisfy* $\max_i a_{ji} \leq \min_i a_{ji} + g(n)$ *for all firms* j*. The full CEDA protocol (Protocol 3 in Section 3.1.2) with the parameters* $k_1(n,l)$*,* $k_2(n)$ *and* $k_3(n)$ *in the protocol replaced by* $k_1(n,l) = 2\exp(B)\max(log^2(n), \frac{l}{s(n)})$*,* $k_2(n) = 3\log(n)s(n)$*, and* $k_3(n) = 3\exp(B)\log(n)s(n)$ *achieves communication cost at most*

$$O(\max(\log(n)g(n)s(n), \frac{n}{s(n)})\log(n))$$

and preference learning cost at most

$$O(\max(\log(n)g(n)s(n), \frac{n}{s(n)}))$$

and succeeds to find the worker-proposing stable match with probability at least $1 - O(\frac{1}{n})$.

When $s(n) = \sqrt{n}$ and $g(n) = C\log^C(n)$, these imply Lemma 3.4 and Theorem 3.1. More generally, setting $s(n) = \sqrt{\frac{n}{g(n)}}$, we have the when public scores have an additive range of $g(n)$, the bounds in the above results become $O^*(\sqrt{ng(n)})$ for both simplified and full CEDA.[43]

One key assumption that both proofs use is that the private scores have bounded hazard rate, and the key implication of this assumption is as follows.

Lemma C.3. *If random variable* z *is distributed according to CDE* F *with hazard rate*

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

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

$$\mathbb{P}(z \geq x+1 | z \geq x) \geq \exp(-B).$$

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

Proof of Theorem C.1. We first show that the simplified CEDA protocol always produces a stable matching. We then bound the total number of preference signals and applications, which will yield the desired result.

Let μ_j be firm j's utility for the partner the firm is currently tentatively matched with. This is initialized to u_{j0} and steadily increases as the firm accepts each application. The key of the protocol is the interplay between μ_j with the qualification requirement z_j, as well as the threshold on private score for sending preference signals q_j (which is defined to be the top $\frac{1}{s(n)}$ quantile of private scores). Define $\underline{a}_j = \min_i a_{ji}$. Recall also the definitions for "qualify" and "publicly qualify" in the beginning of Section 3.1.

Claim C.1. *The simplified CEDA protocol produces the worker-optimal stable match.*

Proof of Claim C.1. This follows from the observation that if a worker i does not qualify for firm j, then i would not have been accepted anyway. This is because by not qualifying, the public score $a_{ji} < \mu_j - q_j$ and the private score $\epsilon_{ji} < q_j$, so

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

□

[43]Recall that $f(n) = O^*(g(n))$ if there exists constant $C > 0$ such that $f(n) \leq C\log^C(n)g(n)$ for all sufficiently large n.

Claim C.2. *With probability at most* $\exp(-\frac{n}{3s(n)})$*, a firm* j *sends more than* $\frac{2n}{s(n)}$ *preference signals.*

Proof of Claim C.2. This follows from a simple Chernoff bound, since the chance that a firm will send a certain worker a preference signal is exactly $\frac{1}{s(n)}$, so the expected total number sent by the firm is $\frac{n}{s(n)}$, and signals are independent. □

Define the event S_j as the first time that μ_j is at least $\underline{a}_j + q_j$. In the simplified CEDA protocol, this is the first time a qualification signal is sent by firm j.

Claim C.3. *The chance that a firm* j *receives more than* $2\log(n)s(n)$ *applications before event* S_j *is at most* $\frac{1}{n^2}$*.*

Proof of Claim C.3. Before S_j, everyone publicly qualifies to firm j, so preference signals from the firm would not have made any difference to the application decisions of workers. S_j will happen whenever the firm receives an application from a worker whose private score is at least q_j. By independence of private scores, the chance that this does not happen for $2\log(n)s(n)$ applications is

$$(1 - \frac{1}{s(n)})^{2\log(n)s(n)} \leq \frac{1}{n^2}.$$

□

Claim C.4. *After* S_j*, every application from a worker who publicly qualifies will cause* z_j *to increase by* 1 *with probability at least* $\frac{1}{\exp(B)s(n)}$*, independent of whatever has happened before in the protocol.*

Proof of Claim C.4. If worker i publicly qualifies for firm j, then the worker would have applied regardless of whether the worker received a preference signal. Now, the worker will cause z_j to increase by 1 whenever $\epsilon_{ji} \geq q_j + 1$, which happens with probability at least $\frac{1}{\exp(B)s(n)}$ by Lemma C.3, since q_j is the top $\frac{1}{s(n)}$ quantile by definition. Furthermore, by independence of private scores, this is independent of everything that has happened before. □

Claim C.5. *The total number of application received by a firm* j *from workers who publicly qualify is more than* $\alpha(n) := 2\exp(B)s(n)\max(g(n), 8\log(n))$ *with probability at most* $\frac{1}{n^2}$*.*

Proof of Claim C.5. This follows from a simple Chernoff bound by Claim C.4, using the fact that the chance that the sum of $\alpha(n)$ *Bernoulli*$(\frac{1}{\exp(B)s(n)})$ random variables is more than half the mean, $\frac{\alpha(n)}{2\exp(B)s(n)} \geq g(n)$, with probability at least $1 - \exp(-\frac{\alpha(n)}{8\exp(B)s(n)}) \geq 1 - \frac{1}{n^2}$. However, when this happens, the qualification requirement would have exceeded the maximum possible public score $\max_i a_{ji}$ and no one publicly qualifies any more. □

To finish the proof, define a failure event as one of

- The number of preference signal of some firm is more than $\frac{2n}{s(n)}$.
- The number of applications from workers who publicly qualify is more than $2\log(n)s(n) + \alpha(n) = O(g(n)s(n))$.

By Claims C.2, C.3, and C.5, the probability that no failure event occurs is at least

$$1 - [n\exp(-\frac{n}{3s(n)}) + n(\frac{1}{n^2} + \frac{1}{n^2})] \geq 1 - O(\frac{1}{n}).$$

Since applications from workers that do not publicly qualify can be bounded by the number of preference signals, when there are no failures, the total number of preference signals and applications per agent is at most

$$\frac{4n}{s(n)} + 2\log(n)s(n) + \alpha(n) = O(\max(g(n)s(n), \frac{n}{s(n)})).$$

Since the number of qualification signals is bounded by the number of applications, and since each application, application decision, preference signal, and qualification signal can be sent using $O(\log(n))$ bits, this proves the desired bound.[44] □

Proof of Theorem C.2. The proof of Theorem C.2 is analogous to C.1. The difference is that while the simplified CEDA protocol always yields a stable matching, but only achieves low communication cost with high probability, the full CEDA protocol always achieves low communication cost, but only yields a stable matching with high probability.

To see that CEDA always achieves low preference learning and communication cost, observe that the total number of preference signals is at most $O(\frac{n}{s(n)} + \log(n)g(n))$, and the total number of applications is at most $O(\log(n)s(n)g(n))$. Therefore, the average number of preference signals or applications per agent is at most $O(\max(\frac{n}{s(n)}, \log(n)s(n)g(n)))$, which upper-bounds the number of preference oracle queries. The communication cost is upper-bounded by $O(\log(n))$ this amount, because each signal or application can be communicated using $O(\log(n))$ bits.

We now prove that CEDA yields the worker-optimal stable match with high probability. As in the proof of Claim C.1, it suffices to show that with high probability, throughout the running of CEDA, there never exists a time in which a worker i is not qualified to firm j, but the firm would have accepted i if the worker applied at that instant, or $u_{ji} > \mu_j$. If such a scenario arises, then it must be that either

(1) worker i did not get a preference signal from firm j, but $\epsilon_{ji} \geq q_j$;
(2) or worker i did not publicly qualify, $z_j > a_{ji}$, but $a_{ji} \geq \mu_j - q_j$.

It suffices to show that both of these events occur with vanishing probability in CEDA.

Claim C.6. *The chance that there exists a worker i and firm j such that $\epsilon_{ji} \geq q_j$ but i does not get a preference signal from j is at most $O(\frac{1}{n})$.*

Proof of Claim C.6. For each worker j and each bin of public scores or range 1, suppose that there are l workers in this bin, if there is a worker i in this bin with $\epsilon_{ji} \geq q_j$ but the worker does not get a preference signal from j, then there must be at least $k_1(n) = 2\exp(B)\max(\log^2(n), \frac{l}{s(n)})$ other workers in that bin with private score of at least $q_j + 1$. By Lemma C.3 and the definition of q_j as the top $\frac{1}{s(n)}$th quantile of private scores, the chance that any given worker in that bin has such a high private score is at most $\frac{\exp(B)}{s(n)}$, independent of everything else. If $\frac{l}{s(n)} < \log^2(n)$, then this mean is less than $\exp(B)\log^2(n)$, so by a Chernoff bound, the number of such workers is as high as $k_1(n)$ with probability at most $\exp(-\frac{\exp}{3}\log^2(n)) = \exp(-O(\log^2 n))$. If $\frac{l}{s(n)} \geq \log^2(n)$, then a similar Chernoff bound yields that the number of such workers is as high as $k_1(n)$ with probability at most $\exp(-\frac{k_1}{3})$, which is also bounded by $\exp(-O(\log^2 n))$. Using a union bound over the n workers and $g(n) = O(n)$ bins, the total probability is at most $O(n^2\exp(-O(\log^2(n)))) = O(\frac{1}{n})$. □

Claim C.7. *The chance that at some point in the running of CEDA, we have that there exists a firm such that $z_j > \underline{a}_j$, $z_j > \mu_j - q_j$ is at most $O(\frac{1}{n})$*

Proof of Claim C.7. For each firm j, there are by definition at most $g(n)$ increase to the qualification requirement z_j. It suffices to upper-bound the probability that $z_j > \mu_j + q_j$ after each of these increases.

[44] Qualification signals can also be sent using $O(\log(n))$ bits because each qualification signal z_j can be round up to the smallest public score of workers for this firm that is at least z_j, and the identity of this worker can be specified using $O(\log(n))$ bits.

First note that with probability at least $1 - \frac{1}{n^3}$, $\mu_j \geq \underline{a}_j + q_j$ at the end of Phase 1. This follows from the fact that everyone publicly qualifies at this point, so every application has private score $\epsilon_{ji} \geq q_j$ with probability exactly $\frac{1}{s(n)}$, independent of everything else, so the chance that none of the first $k_2(n) = 3\log(n)s(n)$ applicants have such a high private score to firm j is at most

$$(1 - \frac{1}{s(n)})^{3\log(n)s(n)} \leq \frac{1}{n^3}.$$

Now, suppose that $\mu_j \geq z + q_j$ and z is the current qualification requirement for firm j. We show that with high probability, after the next $k_3(n) = 3\exp(B)\log(n)s(n)$ publicly qualified applications, the latent tentative match value μ_j must have increased to at least $z + q_j + 1$. This is because if μ_j has not increased to this amount, it means that none of the $k_3(n)$ publicly qualified applicants had a private score at least $q_j + 1$. Now, for publicly qualified applicants, their private score from the perspective of an outside observer is still distributed according to F_j, since they qualify for firm j regardless of whether they receive a preference signal. Thus, the probability that their private score is at least $q_j + 1$ is at least $\frac{1}{\exp(B)s(n)}$ by Lemma C.3. So the chance that none of the $k_3(n)$ applicants had such a high private score is at most

$$(1 - \frac{1}{\exp(B)s(n)})^{3\exp(B)s(n)\log(n)} \leq \frac{1}{n^3}.$$

Thus, with probability at least $1 - \frac{g(n)+1}{n^3}$, firm j never experiences $z_j > \mu_j + q_j$ after Phase 1. A union bound over the n firms shows the desired result.

□

Combining Claims C.6 and C.7, we get that with probability at least $1 - O(\frac{1}{n})$, CEDA succeeds to yield the worker-optimal stable match.

□

D IMPOSSIBILITY OF TWO-ROUND PROTOCOL

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 private scores distributed as $Exp(1)$ by having a public score difference of $3\log(n)$ for the type that one prefers.

Example D.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 his or her own focus, and otherwise preferences are drawn uniformly at random. Research-focused departments and teaching focused candidates are in short supply: here 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 $\frac{2}{n}$. 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.

E 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 2.2. 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 2.2 restrict it to having only a constant number of tiers.)

Before presenting the algorithm (Theorem E.1), we first define the concept of sub-matching. For any matching μ, any subset $A \subseteq I$ of workers and $B \subseteq J$ of firms, let $\mu(A, B)$ be the *sub-matching* restricted to agents $A \times B$, which is defined as $\{(i, j) \in \mu : i \in A, j \in B\}$. We say that the sub-matching $\mu(A, B)$ 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 E.1. *A stable matching in a tiered market can be constructed as follows. Initialize $\mu = \emptyset$.*

(1) Construct a stable sub-matching between the top tiers I_1 and J_1. Add these matches to μ.
(2) Remove any matched agent in the sub-matching found above, as well as any unmatched agent in I_1 or J_1 who finds someone in J_1 or I_1 unacceptable. (These agents will 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 μ. Otherwise repeat step 1 for the top remaining tiers on both sides.

Moreover, every stable matching can be constructed in the above way.

Proof. 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 sub-matchings for the top tiers and stable sub-matchings for the rest of market.

Claim E.1. *In a tiered market, a matching μ is stable if and only if*

(1) Sub-matching $\mu(I_1, J_1)$ is stable.
(2) Sub-matching $\mu(I_1 \backslash (I_1^m \cup I_1^u), J_1 \backslash (J_1^m \cup J_1^u))$ 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 E.1 follow from straightforward induction. To show the first direction of this claim, assume that μ is a stable matching for the tiered market. Note that in any stable matching μ, for a fixed set A of workers and fixed set B of firms, the sub-matching $\mu(A, B)$ must be stable. Therefore, sub-matching $\mu(I_1, J_1)$ must be stable. Apply the Rural Hospital Theorem on this sub-market, 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 μ. 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 sub-matching $\mu(I_1 \backslash (I_1^m \cup I_1^u), J_1 \backslash (J_1^m \cup J_1^u))$ must be stable.

For the second direction of the claim, suppose that the designated sub-matchings are stable, we show that μ 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 sub-matching $\mu(I_1, J_1)$, 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 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 sub-matching $\mu(I_1, J_1)$ 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 μ is stable, as desired. □

Corollary E.2 (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.*

F PROOF OF CORRECTNESS AND EFFICIENCY OF THE TARGETED SIGNALING PROTOCOL

In this section, we prove Theorem 4.5, which claims that the targeted signaling protocol succeeds with high probability and bounds its communication and preference learnings costs. Without loss of generality, let $n \geq m$. Define the subgraph of signals as the collection of tuples (i, j) for which either worker i signals to j or firm j signals to i during the signaling round of the targeted signaling protocol. We break the proof of Theorem 4.5 into 3 claims.

- **Claim 1:** With probability at least $1 - \frac{18}{n}$, the subgraph of signals contains a stable matching.
- **Claim 2:** Whenever the subgraph of signals contains a stable matching, running DA with preferences restricted to the subgraph (as in the matching round of the targeted signaling protocol) returns a stable matching (for the whole market).
- **Claim 3:** The total number of signals is at most $\Theta(n \log^3(n))$.

Claims 1 and 2 imply that the targeted signaling protocol succeeds with probability at least $1 - \frac{18}{n}$. Claim 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.

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

We prove claim 1 by proving that with probability $1 - \frac{18}{n}$, 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 F.2. *Consider a matching market with m workers, $n \geq m$ 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 $N \geq 2$ be such that $N \geq n + u$. For any given worker, with probability at least $1 - \frac{9}{N^2}$, we have that in* *the worker-proposing DA algorithm, the given worker is matched to one of his top r firms, where*

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

Proof of Lemma F.2. First, note that without loss of generality, $N \geq 100$ 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 sub-market without worker 1 and without the unavailable firms, and call this matching μ_1. Note that μ_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 μ_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 G about the average rank of workers in this setting (Proposition G.1). By plugging in $z = 2\log(N)$ into Proposition G.1, we have that the probability of event E_1 is at least $1 - \frac{8}{N^2}$.

Let μ_2 be the matching formed by running the worker-proposing DA algorithm from initialization μ_1. In other words, suppose that we start with everyone else matched according to μ_1 and have worker 1 propose to his 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 his next choice, which may result in a chain of rejections leading to someone applying to one of the $n-m+1$ unmatched available firms. μ_2 is a stable matching (with respect to the entire market). Because the rank of worker 1 in μ_2 is no better than in the worker-optimal stable match, it suffices to upper-bound the rank obtained by worker 1 in μ_2.

First, let us make a few structural observations on μ_1 and μ_2 under event E_1. For each firm $j \leq n$, let B_j be the set of workers who weakly prefer firm j to their partner in μ_1. (We use the letter B because this is the set of workers who want to block with j in μ_1.) In μ_1, firm j is matched to the firm's favorite worker in B_j. Furthermore, the sum $\sum_{j=1}^{n} B_j \leq R = 4e(m-1)\log(N)$, as this sum always equals the total rank obtained by workers in μ_1 (ignoring the unavailable firms). Now, consider running the DA algorithm with initialization μ_1, drawing only as needed the preference of firms for worker 1. When worker 1 applies to an available firm j, the probability that he is accepted is exactly $p_j = \frac{1}{1+|B_j|}$, because this is his chance of being the firm's favorite worker among a set of size $1 + |B_j|$. If the worker is rejected, then the worker apply to his next choice, and the same formula for acceptance probability will apply. If the worker is accepted, then this will trigger 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 μ_1 is the firm-optimal stable match in the sub-market without worker 1. For the unavailable firms $j \geq n$, define $p_j = 0$. We have that by Jensen's inequality,

$$\sum_{j=1}^{n+u} p_j = \sum_{j=1}^{n} p_j \geq \frac{n}{4e\log(N)+1}. \tag{11}$$

Now, conditional on event E_1, let P be the probability that worker 1 is not matched to his 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,

$$P \leq \frac{1}{|A|} \sum_{S \in A} \prod_{j \in S} (1 - p_j) \tag{12}$$

$$\leq \left(1 - \frac{\sum_{j=1}^{n} p_j}{n+u}\right)^r \tag{13}$$

$$\leq \exp\left(-\frac{nr}{(n+u)(4e\log(N)+1)}\right) \tag{14}$$

$$< \frac{1}{N^2} \tag{15}$$

Inequality (12) follows from the independence between the preference of worker 1, the event E_1, and each firm j's preference for worker 1. Inequality (13) follows from the fact that the sum of product in the inner part of equation 12 increases if we replace any two different p_i, p_j with their average $\frac{p_i+p_j}{2}$, so the maximum is attained when all of them are equal. Inequality (14) follows from inequality 11 and the bound $(1-x) \leq \exp(-x)$, and inequality (15) follows from the fact that when $N \geq 100$, we have $\frac{24\log^2(N)}{4e\log(N)+1} > 2\log(N)$.

Since the probability that E_1 does not occur is at most $\frac{8}{N^2}$, we have that the total probability that worker 1 does not get one of his top r choices is at most $\frac{8}{N^2} + \frac{1}{N^2} = \frac{9}{N^2}$, which is what we needed. ◻

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

Claim 2 follows from the following structural result (Lemma F.3) 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 μ is *stable with respect to the subgraph* if every matched agent in μ prefers to be matched than unmatched, and there are no blocking pairs to μ 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 $\min(m,n)$, which corresponds to matching all agents of at least one of the sides.

Lemma F.3. *Any matching μ that is full and stable with respect to the subgraph of signals is stable with respect to the complete graph.*

To see why Claim 2 follows from Lemma F.3, note that when the subgraph of signals contains a stable matching, then running DA (or any other stable matching algorithm) with preferences restricted to the subgraph returns a matching that is full and that is stable with respect to the subgraph. Lemma F.3 implies that this matching is stable with respect to the complete graph.

Proof of Lemma F.3. Let the subgraph of signals be G. let μ be a full matching stable with respect to G, we show that μ is stable with respect to the complete graph. It suffices to show that μ 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 be j in firm-tier l. Firstly, any $(i,j) \in G$ cannot be a blocking pair by the definition of μ being stable in G.

Suppose first that $s_k \neq t_l$. Without loss of generality, let $s_k < t_l$. In this case, worker i must be matched in μ, 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 $l' \leq \tilde{l}$ because G can only contain edges between i and weakly better tiers than $\tilde{l}$. Moreover, we have $\tilde{l} \leq l$, since $\tilde{l} = \min\{l : t_l \geq s_k\}$ by definition. Combining the two inequalities, we have $l' \leq \tilde{l} \leq l$. 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, firm-tier l is one of the redundant tiers (see Definition 4.4). As before, i must be matched in μ, 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.

Since there are no blocking pairs, μ must be stable with respect to the complete graph. □

We complete the proof of Theorem 4.5 by proving Claim 3.

Lemma F.4. *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 F.4. The result is trivially true if $n \leq e^4 < 60$, since the average number of signals is at most n. Assume now that $\log n \geq 4$.

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 $t_{l-1} < s_k \leq t_l$.

Consider an arbitrary grouping. Without loss of generality, let the shared target tier be J_l as above. Let $K = \{k : t_{l-1} < s_k \leq t_l\}$ as above. Let the minimum and maximum element of K be k_0 and k_1 respectively. For each $k \in K$, define r_k as the target number of workers in tier I_k (see Definition 4.3).

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 $r_{k_0} = 24 \log^2 n$ since the competitiveness of workers (see Definition 4.2) in I_{k_0} must be 1. For $K \backslash \{k_0\}$, we have

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

This shows that

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

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. □

G AVERAGE RANK IN UNBALANCED MATCHING MARKET

The proof of Lemma F.2 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., 2016] implies that this is asymptotically $\log(n)$ as $n \longrightarrow \infty$. Compared to 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] and [Pittel, 1992], which are based on an Integral formula due to Knuth.

Proposition G.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. 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 μ_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 the standard matching is stable and the total rank of workers equals R is exactly

$$[\xi^{R-m}]\{\prod_{i=1}^{m}(1-x_i)\prod_{1\leq i\neq j\leq m}(1-x_i+\xi x_i(1-y_j))\}, \tag{16}$$

where $[\xi^a]\{f(\xi)\}$ denotes the coefficient of ξ^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 below. 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 firms workers want to block with is $R-m$, counting with multiplicity. The expression in the braces computes the probability the standard matching μ_0 is stable while keeping track of who wants to block with whom using dummy variable ξ. 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 ξ and examining the coefficient of ξ^{R-m} obtains exactly the probability that the matching is stable and the total rank of workers is R.

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

$$P_0 \le n! \int_A \sum_{R=\lfloor m\bar{r} \rfloor+1}^{\infty} [\xi^{R-m}] \left\{ \prod_{i=1}^{m} (1 - x_i) \prod_{1 \le i \ne j \le m} (1 - x_i + \xi x_i (1 - y_j)) \right\} dx\, dy \tag{17}$$

$$\le n! \int_A \inf_{\xi \ge 1} \left\{ \xi^{m(1-\bar{r})} \exp(-m \sum_{i=1}^{m} x_i + \xi \sum_{1 \le i \ne j \le m} x_i (1 - y_j)) \right\} dx\, dy \tag{18}$$

Inequality (17) follows from integrating equation (16) over the uniform distribution of matching values, then using a union bound over all $n!$ matchings based on symmetry. Inequality 18 comes from the Chernoff method of bounding the tail of power series[45], and using the fact that $\lfloor m\bar{r} \rfloor + 1 \ge m\bar{r}$.

Define $\mu = \frac{\bar{r}}{e} = 2\log(n) + z$. Partition the region of integration A into $A_1 = \{(x, y) \in A : \sum_{i=1}^{m} x_i \ge \mu\}$, and $A_2 = A \backslash A_1$. Define the above 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 $s_i = s - x_i$, $\Psi(x) = \int_0^1 \exp(-xy)dy = \frac{1-\exp(-x)}{x}$, and $A_1^x = \{x : 1 \le x_i \le 1, \sum_{i=1}^{m} x_i \le \mu\}$. For clarity, we write a series of inequalities and explain them one by one afterward.

$$P_1 \le n! \int_{A_1} \exp(-s - \sum_{i=1}^{m} s_i y_i)\, dx\, dy \tag{19}$$

$$= n! \int_{A_1^x} \exp(-s) \prod_{i=1}^{m} \Psi(s_i)\, dx \tag{20}$$

$$\le e^2 n! \int_{A_1^x} \exp(-s) \frac{1}{s^m}\, dx \tag{21}$$

$$\le e^2 n! \int_{\mu}^{m} \exp(-s) \frac{1}{s^m} \frac{s^{m-1}}{(m-1)!}\, ds \tag{22}$$

$$\le e^2 n(n-1) \int_{\mu}^{m} \exp(-s)\, dx \tag{23}$$

$$\le e^2 \exp(-z) \tag{24}$$

Inequality (19) follows from plugging in $\xi = 1$ into inequality (18). In equation (20), we integrate with respect to each y_i. In inequality (21), we make use of the fact that for each $a > 0$,

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

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

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 (22), we use a standard change of variable from x to s (see equation (3.2) and inequality (3.3) of [Pittel, 1989]). In inequality (23), we use the fact that $\frac{1}{s}$ is decreasing in s and that $\mu \ge 1$. In inequality (24), we integrate out s and make use of the definition of μ and $z \ge 1$.

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

$$
\begin{aligned}
P_2 &\le n!\int_{A_2}\left(\frac{e\mu}{s}\right)^{m(1-e\mu)}\exp(-ms+e(m-1)\mu)\,dx\,dy && (25)\\
&\le n^2\int_0^{\mu}\left(\frac{e\mu}{s}\right)^{m(1-e\mu)}\exp(-ms+e(m-1)\mu)s^{m-1}\,dx\,dy && (26)\\
&\le n^2\int_0^{\mu}\exp(-(m+e)\mu+m)\mu^{m-1}\,ds && (27)\\
&= [n^2\exp(-\mu)][\exp(-(m+e-1)\mu)\mu^m]\exp(m) && (28)\\
&\le \frac{1}{e}\exp(-z) && (29)
\end{aligned}
$$

Equation (25) substitutes in $\xi = \frac{e\mu}{s}$ to inequality (18) and uses the fact that $\sum_{1\le i\ne j\le m} x_i(1-y_j) \le s(m-1)$. Inequality (26) 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 (27) uses the fact that the function $f(s) = \exp(-ms)s^{em\mu-1}$ is increasing in $[0,\mu]$. Equation (28) simplifies the formula and arrange into groups, denoted by square brackets. Inequality 29 comes from bounding each group. We bound the first group by $\exp(-z)$ using the formula for μ. We bound the second group using the observation that the function $f(\mu) = \exp(-(m+e-1)\mu)\mu^m$ is maximized when $\mu = \frac{m}{m+e-1}$, so $f(\mu)\exp(m) \le (\frac{m}{m+e-1})^m = (1-\frac{e-1}{m+e-1})^m \le \exp(-\frac{(e-1)m}{m+e-1}) < \exp(-1)$, since $m \ge 8$.

Combining inequalities (24) and (29), we have

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

which completes the proof. □

Remark 3. *The above proof can be modified to show the following statement. Consider a uniformly random matching market with m men and $n = m + d$ women, where $1 \le d \le (e-1)m$. 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*

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

H OPTIMAL COMMUNICATION COST IN TIERED RANDOM MARKETS

The targeted signaling protocol in Section 4.3 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 E.1, in which in Step 1, if $|I_1| \le |J_1|$, 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 propose to a partner, the agent does not know whether or not the partner is taken by agents from better tiers. This results in wasted proposals. Nevertheless, one can show as in [Knuth, 1976] 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 succeed 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 H.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, if a stable matching protocol that only uses private messages succeeds with high probability, then the protocol requires at least $\Omega(\log^2 n)$ bits of communication per agent on average.*

Proof of Theorem H.1. The market in question is the same as that in [Ashlagi et al., 2016]. It suffices to prove the lower bound for any one worker, since all workers are ex ante the same. Some features of this market that we know from [Ashlagi et al., 2016] 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. Whp (with high probability), 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, cf. Section 2.1, 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 his 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., 2016] (with revelation of preferences as needed) and using the third bullet stated above. 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

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

For each of these firms, independently, the probability that they prefer i over their currently matched worker is at least $1/(1 + 1.1 \log n) \geq 0.8/\log n$ and at most $1/(1 + 0.9 \log n) \leq 1.2/\log n$. 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

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

The probability that a rejection chain starting at a firm in J'', cf. [Ashlagi et al., 2016], 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 $1/(2 + 0.9 \log n) \leq 1.2/\log n$. 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 $|\hat{J}| \leq f_n |J''| 1.2/\log n$, for any $f_n = \omega(1)$. Using $f_n = \sqrt{\log n}$ we obtain a bound of

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

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,

$$|J_a| \geq |J'' \backslash \hat{J}| = |J''| - |\hat{J}| \geq 0.7n/\log n - 2n/(\log n)^{3/2} \geq 0.5n/\log n\,.$$

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

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

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 where worker i has a unique stable partner and the protocol finds a correct answer, the answer is exactly $\mathbb{I}(J_a \cap J_p \neq \phi)$. Recalling that whp, worker i has a unique stable partner, and since the output of the protocol matches $\mathbb{I}(J_a \cap J_p \neq \phi)$ correctly whp in these cases, the protocol finds $\mathbb{I}(J_a \cap J_p \neq \phi)$ 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 Proposition H.2. 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 H.2 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 H.2.) 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,[46] the protocol is faced with problem 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 $\mathbb{I}(J_a \cap J_p \neq \phi)$ with probability at least ϵ, using Proposition H.2. Since such bad $|J_a|$'s occur with probability at least 1/3, the overall probability of outputting something different than $\mathbb{I}(J_a \cap J_p \neq \phi)$ is at least $\epsilon/3$. This contradicts that the protocol finds $\mathbb{I}(J_a \cap J_p \neq \phi)$ 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. □

H.1 Complexity Foundations

Consider a set N such that $|N| = n$.[47] 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 non-trivial intersection or not. We are interested in lower bounds for the communication complexity of determining the correct answer with 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 0 and 1 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 H.2. *There exists $\epsilon > 0$ such that 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. The proposition 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. Since 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 (since 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 inputs 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 π of communication cost d. The randomized protocol π induces a partition of $\mathcal{A} \times \mathcal{B}$ into at most 2^d combinatorial rectangles, on each of which the output is either 0 or 1. For all values of c, the probability that the output is 0 (i.e. that the sets are disjoint) is a constant, and therefore, for a sufficiently small ϵ, a constant fraction of the mass is covered by 0-rectangles, on each of which the error rate is $< c_1\epsilon$ for an absolute constant $c_1 > 0$. We will 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 0 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

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

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

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 uper 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 which 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 which T does not intersect is at least $(k - 4c_2\epsilon k)l_a/2 > kl_a/3$, since 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 ϵ, since when $4c_2\epsilon < e^{-5}$, $\left(\frac{e}{4c_2\epsilon}\right)^{4c_2\epsilon} < e^{1/12}$; $\leq_2$ 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. □

I PROOFS OF INCENTIVE COMPATIBILITY PROPERTIES

Before proving that all the protocols in this paper are as incentive compatible as DA with high probability (Theorem 5.2), we establish two lemmas. This allows us to prove the desired properties for all protocols simultaneously under one framework.

For either versions of CEDA, define the *subgraph of signals* G as the set of tuples (i, j) for which worker i applies to firm j at some point during the protocol, assuming that everyone complies to the protocol. (The subgraph G is deterministic given the preference realizations.) In the targeted signaling protocol, define the subgraph of signals G as in Appendix F: the set of tuples (i, j) for which either worker i signals to firm j or vice versa during the signaling round. 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 (similar to in the matching round of the targeted signaling protocol). Similarly define the firm-optimal stable match restricted to G.

LEMMA I.1. *Whenever CEDA or the targeted signaling protocol succeeds to find a stable matching:*

(1) The resultant match μ is the same as the worker-optimal stable match restricted to the subgraph of signals G.

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

Proof of Lemma I.1. The first point follows immediately from the definition of the protocols.

The second point holds in CEDA because the firm-optimal stable match restricted to G is also μ. (To see this, note that by definition of G, every firm in μ gets their favorite partner in G.) The second point holds in the targeted signaling protocol because the protocol succeeding implies that μ is full (matching $\min(m, n)$ pairs of agents), which implies that the firm-optimal stable match restricted to G is full. The desired result follows from Lemma F.3 in Appendix F. □

We are now ready to prove Theorem 5.2 and 5.5.

Proof of Theorem 5.2. Let G be the subgraph of signals (which we defined for all three protocols in question). Suppose that a certain agent can unilaterally deviate from the protocol and cause the resultant matching to be μ', which gives the deviating agent someone than the agent's best stable partner (under complete, true preferences). By Lemma I.1, whenever the respective protocols succeed, the agent obtains a better partner in μ' than what the agent gets in either the worker-optimal or firm-optimal stable match restricted to G (defined with respect to true preferences). Note that regardless of the agent's deviation, μ' must be individual rational for all agents. Applying Lemma 3.5 (as known as the blocking lemma) in [Roth and Sotomayor, 1990], which is originally due to [Gale and Sotomayor, 1985], we have that there must exist a worker i and a firm j such that

(1) neither i or j are the same as the deviating agent, so are assumed to be complying to the protocol;
(2) $(i, j) \in G$ and (i, j) blocks μ' (under the true preferences of i and j).

Let the subgraph of signals that resulted from the agent's deviant behavior be G' (this is well defined as well at the end of each protocol). It must be that $(i, j) \notin G'$, otherwise μ' would not be stable restricted to G' since both i and j are reporting true preferences with respect to one another regardless of what the deviant agent does. However, this is impossible in the targeted signaling protocol since $(i, j) \in G$ implies that $(i, j) \in G'$, since the signaling happens simultaneously in the first round of the protocol and the deviating agent cannot affect signals between other compliant agents. This proves that targeted signaling is as incentive compatible as DA for every agent whenever the protocol succeeds under compliant behavior for everyone (which happens with high probability by the Theorem 4.5).

For either version of CEDA, suppose that $(i, j) \notin G'$, but worker i and firm j prefer one another to their final matched partner in μ'. Let the final utility obtained by firm j in μ' be u_j^*, and the final qualification requirement be z_j^* (allowing whatever behavior from the deviant agent). The fact that $(i, j) \notin G'$ and worker i gets matched to a less preferred firm in μ' means that worker i skipped firm j in his/her application decision, which implies that 1) firm j must not have sent a preference signal to i and 2) the qualification requirement z_j^* at the end of the run of the protocol exceeds a_{ji}. Whenever CEDA succeeds under compliant behavior from everyone, the first statement above implies that $\epsilon_{ji} < q_j$. Combining with the second statement, we get

$$z_j^* > a_{ji} = u_{ji} - \epsilon_{ji} > u_j^* - q_j.$$

However, this never happens under the simple version of CEDA by definition of z_j, and it happens with vanishing probability at most $O(\frac{1}{n})$ under the modified version of CEDA by the proof of Claim C.7 (the key is to notice that the proof of Claim C.7 is a statistical argument, and does not require compliant behavior by workers or other firms.). Thus, the simple version of CEDA is always as incentive compatible as DA. For the modified version of CEDA, for any fixed agent, the chance that the agent can deviate and obtain someone better than his/her best stable partner is at most $O(\frac{1}{n})$.[48] □

Proof of Theorem 5.5. By Theorem 5.2, 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 E.1 in Appendix E, the desired result follows from the following Lemma, which is based on techniques from [Ashlagi et al., 2016]. □

Lemma I.2. *There exists a non-increasing 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 $n \geq m + 1$ firms, the probability that any given agent, conditional on being matched, has multiple stable partners is upper-bounded by $\delta(n)$.*

Proof of Lemma I.2. Consider two cases, suppose that $m \geq \frac{n}{2}$, then by Theorem 3 ii) of [Ashlagi et al., 2016], 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 $\frac{m}{\log^{0.5} m}$, and the same statement holds for firms. This implies that for $n \geq 2m_0$, the probability that any given worker has multiple stable partners is at most

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

Similarly, the probability that a firm, conditional on being matched, has multiple stable partners is also upper-bounded by this. Define function $\delta_1 : \mathbb{N} \to \mathbb{R}$ has $\delta(n) = 1$ for $n < 2m_0$ and as the above quantity when $n \geq 2m_0$. Then function δ_1 satisfies the desired result in the region when $m \geq \frac{n}{2}$.

Suppose now that $m < \frac{n}{2}$, 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 $u = n - m \geq \frac{n}{2}$. As in [Ashlagi et al., 2016], 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 $\frac{1}{u}$. 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($\frac{2}{n}$). For $n \geq 11$, the expectation of this is less than $\frac{3}{n}$. 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 $\frac{3m}{n}$ 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 $\frac{3}{n}$ when $n \geq 11$. Define $\delta_2 : \mathbb{N} \to \mathbb{R}$ to be $\delta(n) = 1$ for $n < 11$ and $\delta(n) = \frac{3}{n}$, we have that δ_2 satisfies the desired result in the region when $m < \frac{n}{2}$.

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

[48] This also uses the observation that the modified CEDA succeeds under compliant behavior with probability at least $1 - O(\frac{1}{n})$.