# Matching with Costly Information Acquisition

Justin Hadad*

Last edit: May 14, 2024

**Abstract**

We study strategic behavior in two-sided matching markets where preferences are aligned but imperfectly known, and where workers pay acquisition costs to learn their utilities from matching with different firms. When workers finish strategically obtaining match utilities, a centralized institution creates the matching by pairing successive worker-firm pairs with the highest realized surplus. We identify the class of information-acquisition mechanisms that implement the *ex-post* stable and Pareto-efficient matching, and the mechanism within the class which minimizes expected aggregate acquisition cost. Our main result proves that the number of acquisitions is minimized in expectation if the agents with the highest commonly-known values find their stable matches as early as possible.

# 1 Introduction

Most foundational work in two-sided matching theory assumes that agents have complete information about their own preferences (Roth and Sotomayor, 1990). However, in many realistic matching settings, preferences are *ex-ante* unknown and costly to learn. Consider settings where preferences are a function of match quality, and where learning qualities is costly: prospective partners must go on dates, employees must interview at firms, and potential roommates must chat on the phone. We are interested in strategic behavior and properties of equilibria in one-to-one, two-sided matching settings like these, where preferences are a function of *ex-ante* unknown match quality, and are costly to learn.

In our setting, agents derive utility equal to their match quality, which is a convex combination of four parameters: the common (publicly-known) values of the two matched agents, as well as the privately-known, idiosyncratic values that each agent has for their

*University of Oxford, Department of Economics.

match. Preferences are then unknown—agents may have priors on the quality of their match, but their *ex-post* complete preference ordering over matches is unknown. Because agents derive utility solely from their match quality, we can represent the preferences of agents on both sides via the same matrix, and so preferences are *aligned* as in Ferdowsian et al. (2023). Alignment guarantees the uniqueness of the stable match.[1]

We study a two-sided matching game with workers and firms, where workers pay a fixed homogeneous cost $\delta$ to learn their match quality (which we call the *match utility*) with a firm. Workers must pay to learn their match utility with a firm in order to match with them, and once a match utility is learned, it becomes public information to all workers. Thus the acquisition of a match utility is analogous to proposing in matching markets, except learning is nonbinding; workers can acquire match utilities with firms with whom they do not ultimately match. After some match utilities are acquired, we create final matches via a *top-top* method, by assigning matches in decreasing order of the values of the acquired match utilities.

Consider first what would happen if it were without cost ($\delta = 0$) to acquire match utilities. Then workers have a weakly dominant strategy to acquire all match utilities, and the game trivially reduces to that with complete information and the unique *ex-post* stable and Pareto-efficient matching is implemented. That the matching is unique is seen from recognizing that preferences are aligned, and that this stable matching is implemented follows from the fact that agents cannot misreport their preferences and arrive at a match superior to their unique top-top match.

Consider the *naive deferred-acceptance* game, where in each period, each worker who does not have a tentative top-top match either acquires a match utility with the firm for which they have the highest expected match utility, or exits the market. We begin by showing that, under incomplete information, naive deferred acceptance can implement an *ex-post* unstable and Pareto-inefficient matching. We also show that naive deferred acceptance can be wasteful—workers could have inferred which match utilities not to acquire based on the match utilities that had already been revealed.

We then investigate the *priority trading* game. In the priority trading game, workers are placed in some order, and sequentially either acquire match utilities until they are in the set of tentative top-top matches, or exit the market. Any worker who is displaced from their tentative top-top match is given priority, and either acquires new match utilities until she is in the set of tentative top-top matches, or exits. The game concludes when all agents are either matched in the set of tentative top-top matches or have exited the market. We show

[1] Preference alignment is a *sufficient* condition for uniqueness. The literature does not yet know a necessary and sufficient condition for preferences to admit a unique stable matching.

that this game can also implement an *ex-post* unstable and Pareto-inefficient matching—even when each worker has the weakly dominant strategy to acquire match utilities in the expected order of her true preferences. The problem, like in naive deferred acceptance, is the lack of *sufficient revelation*: workers may not acquire match utilities with firms that could have been their stable match.

We show that, so long as $\delta$ is sufficiently small, there is a class of mechanisms that implement the *ex-post* stable and Pareto-efficient matching in weakly dominant strategies.[2] These mechanisms place workers in any random order, and each worker sequentially acquires all match utilities with the firms with which they could feasibly match. In any period where a match utility is acquired such that a worker is displaced from her tentative top-top match, that displaced worker then acquires all match utilities with the new firms that could be a feasible match. We show that the complete iteration of any mechanism from this class (which we call *priority trading with multi-revelation* mechanisms) implements the *ex-post* stable and Pareto-efficient matching. We additionally find a novel mechanism, called *top-down priority trading with multi-revelation*, which implements the *ex-post* stable and Pareto-efficient matching with minimum expected aggregate cost—or analogously, the minimum number of expected proposals. This mechanism grants priority to workers in decreasing order of their commonly-known values.

Our findings demonstrate that, if interviews are of sufficiently low cost, a centralized institution can generate the *ex-post* stable and Pareto-efficient matching by ordering agents and directing them to conduct multiple interviews at once. Our main result emphasizes that an institution seeking to minimize the number of interviews should let the most desired applicants find their preferred matches first. In practice, there are many hiring platforms which centralize the interview process; take, for instance, the Econ Job Market. Our results find theoretical justification for the role of a centralized institution in interview allocation.

# 2 Literature Review

The condition that we impose on preferences—that they are *aligned*—is used elsewhere in the literature. We discuss how it relates to other preference conditions below.

Assortative Preferences $\subset$ Aligned (Acyclic) Preferences; Ordinal Potential $\subset$ SPC

[2] We are interested in *indirect* mechanisms that implement the *ex-post* stable and Pareto-efficient matching. That true preferences are unknown to the workers rules out the use of the direct-revelation principle. We frequently refer to the settings in which the indirect mechanisms are implemented as *games*, with players, payoffs, information, and moves specified in Section 3.

Preferences are aligned (as in Ferdowsian et al. (2023)) if and only if they do not have cycles. As long as the strategy set is finite, this is also the necessary and sufficient condition for an ordinal potential to exist (Hart and Mas-Colell, 1989; Monderer and Shapley, 1996; Voorneveld and Norde, 1997). Alignment is stronger than the *sequential preferences condition* (SPC) which allows for cycles in markets where agents split match utilities in fixed proportions (Eeckhout, 2000; Fernandez et al., 2021). A special case of aligned preferences is where preferences are *assortative*; that is, where agents on each side of the market share the same ranking of agents on the other side.

There are two papers which study strategic interaction in matching markets with aligned but unknown preferences; they are Fernandez et al. (2021) and Ferdowsian et al. (2023).[3]

The more closely related of the two is Ferdowsian et al. (2023), who, like us, study *decentralized* markets. In their model, agents are taxed utility proportional to the amount of time that they spend in the market before exiting. Unlike in our paper, they let agents on *both* sides behave strategically; they give examples of settings where agents on the accepting side will settle for an early match if the utility loss from staying in the market is large enough. This is a key departure from our model—we do not treat the acquisition of a match utility as an offer, and so firms have no option to exit before the centralized institution constructs the matching. Additionally, firms face no utility loss from staying in the market, so exiting early would be weakly dominated.

In each of our papers, it is immediate that the *ex-post* stable matching is not recoverable when the utility loss (either from staying in the market as in their model, or from acquiring match utilities as in ours) is sufficiently large. They prove that if this utility loss is small enough, a Bayesian Nash equilibrium exists in *weakly undominated strategies* that implements the *ex-post* stable matching. Their proof relies on agents using "reduced deferred acceptance" strategies, where they update their set of feasible matches according to what they dynamically learn from the set of offers and exits. We have a stronger result in our setting; we show that a Nash equilibrium exists in *weakly dominant strategies* when the acquisition cost is sufficiently small. Unlike their paper, we can explicitly handle correlated preferences, and in the special case where preferences are independently drawn from identical distributions, the strategies that arrive at the *ex-post* stable matching are equivalent to their reduced deferred acceptance strategies.

The other related paper—Fernandez et al. (2021)—studies *centralized* markets with

[3] We recover incomplete information in a different way from both. Fernandez et al. (2021) and Ferdowsian et al. (2023) assume preferences are drawn randomly from the set of all possible aligned preference constellations. In contrast, we have that preferences are unknown by defining match utility as the convex combination of common (publicly-known) and private (privately-known) values. We draw inspiration for this from Lee (2017), who studies asymptotic properties under uncertainty in two-sided matching markets.

aligned preferences and incomplete information. Our settings are substantially different. Unlike in our model and that of Ferdowsian et al. (2023), agents do not incur any losses in utility via in-market interaction; the only friction for arriving at the *ex-post* stable matching is that there is incomplete information. They focus on strategies that the *accepting side*, who are unaware of the preference constellation, can use when the proposing side are perfectly informed of the preferences and behave truthfully. But, so long as there are no cycles in each possible state of preferences, there is a unique Bayesian Nash equilibrium in *weakly undominated strategies* which implements the *ex-post* stable matching for each state. Our results are stronger: so long as acquisition costs are sufficiently small, the *ex-post* stable matching is implemented in *weakly dominant* strategies.

In our model, agents decide with which firms they acquire match utilities. For an analysis of two-sided matching markets where costly search randomly acquires unknown match utilities, see Choi (2022).

The inspiration for our indirect implementation of information-acquisition mechanisms comes from the implementation of top trading cycles, as originally discussed in Shapley and Scarf (1974).

# 3 Model

There is a set of workers $W$ and a set of firms $F$; let $N = \max\{|W|, |F|\}$. Each pair of $w \in W$ and $f \in F$ is associated with a match utility $\zeta_{w,f}$, which is the sum of two components:

$$\zeta_{w,f} := u_{w,f} + \ell_{w,f}. \tag{1}$$

We then define

$$u_{w,f} := \lambda c_f + (1-\lambda) v_{w,f},$$

where $\lambda \in [0,1]$ is a publicly-known weight, $c_f$ is the value of firm $f$ that is commonly known by all agents in the market, and $v_{w,f}$ is the private value that worker $w$ has for firm $f$. The private value $v_{w,f}$ is realized *ex ante*, known only to worker $w$, and drawn from a publicly-known distribution along $[\underline{v}_{w,f}, \bar{v}_{w,f}]$. Note that $\lambda$ functions as the "certainty weight"; a larger $\lambda$ means that $u_{w,f}$ is determined less by the idiosyncratic, privately-known component $v_{w,f}$. We likewise define

$$\ell_{w,f} := \lambda r_w + (1-\lambda) s_{w,f},$$

where $\lambda \in [0,1]$ is the same publicly-known weight as above, $r_w$ is the value of worker $w$ that is commonly known by all agents in the market, and $s_{w,f}$ is the private value that firm

$f$ has for worker $w$. The private value $s_{w,f}$ is realized *ex ante*, known only to firm $f$, and drawn from a publicly-known distribution along $[\underline{s}_{w,f}, \bar{s}_{w,f}]$. We separated contributions to $\zeta_{w,f}$ in Equation (1) as such because each worker $w$ makes decisions on which match utilities to acquire based on their values of $u_{w,f}$.

We define an *economy* as a tuple

$$\mathcal{E} := (W, F, \lambda, (r_w)_{w \in W}, (c_f)_{f \in F}, (v_{w,f}, s_{w,f})_{w,f \in W \times F}),$$

which is identified by sets of workers and firms, the values for each agent in the market, and the weight $\lambda$.

Each period is denoted by some $t \in \{1, 2, ..., T\}$. In each period, some worker $w$ either acquires a match utility $\zeta_{w,f}$ with some firm $f$, or exits the market. Each worker suffers a fixed, uniform, and homogeneous cost $\delta \geq 0$ to acquire a match utility, and matches between a worker and firm can only form if the worker has acquired the match utility for that pair. When a final match is made between worker $w$ and firm $f$, firm $f$ gets utility $\zeta_{w,f}$, meanwhile worker $w$ gets utility

$$\mathbf{u}_{w,f} := \zeta_{w,f} - \eta_w \delta, \tag{2}$$

where $\eta_w$ is the total number of match utilities that worker $w$ has acquired.[4] Define the set of pairs of workers and firms for which match utilities have been acquired through period $t$ as

$$R_t := \{(w, f) \mid \zeta_{w,f} \text{ acquired through period } t\}.$$

Then $\eta_w := |\{f \mid (w, f) \in R_T\}|$. We further enforce that the outside option for both workers and firms grants zero utility, and note that the quasilinearity of utilities immediately implies risk neutrality.

Let $R_t^f$ be the set of workers who have acquired their match utilities with firm $f$ through period $t$,

$$R_t^f := \{w \mid (w, f) \in R_t\},$$

and $R_t^w$ be the set of firms with which worker $w$ has acquired her match utility through period $t$,

$$R_t^w := \{f \mid (w, f) \in R_t\}.$$

Because match utilities are split in fixed proportion within worker-firm pairs, we have that

[4] We use a special case of alignment where match utilities are the same for both the worker and the firm in a given match. Without loss of generality, we could say each matched worker $w$ and firm $f$ receive utility split in some other fixed proportion of $\zeta_{w,f}$; consider each worker $w$ receives $\alpha\zeta_{w,f} - \eta_w\delta$ from matching with firm $f$, while firm $f$ receives $\beta\zeta_{w,f}$. This merely scales the conditions for $\delta$ that we derive throughout our paper by a factor of $\alpha$.

preferences are *aligned* as in Ferdowsian et al. (2023). Preference alignment implies that the rejection of any match cannot generate a superior match for either the worker or the firm, and so we can assign matches in decreasing order of acquired match utilities. Define the set of *tentative top-top matches* as the matches constructed by the successive pairing of workers and firms with the highest acquired match utilities by period $t$; we denote this set of tentative top-top matches as $D_t$. Formally, in each period $t$, we find $(w', f') = \arg\max_{(w,f)\in R_t} \zeta_{w,f}$, which constitutes a top-top match from the set of acquired match utilities. Then define the *modified set of acquisitions* $\tilde{R}_t = \{(w, f) \in R_t \mid w \neq w' \text{ and } f \neq f'\}$, and find $(w'', f'') \in \arg\max_{(w,f)\in\tilde{R}_t} \zeta_{w,f}$ which again constitutes a top-top match. We iterate until no pairs of workers and firms remain in the modified set of acquisitions $\tilde{R}_t$, and set $D_t = \{(w', f'), (w'', f''), ...\}$. At times we refer to $D_T$ as the matching $\mu$.

We construct games—defined by naive deferred acceptance in Examples 1 and 2, priority trading in Section 3.1, and what we call *priority trading with multi-revelation* in Section 3.2—with specific players, payoffs, information, and moves. The **players** in each game are the workers, who acquire match utilities to maximize their **payoffs**, which is their final match utility discounted by a factor proportional to their number of acquisitions. The **information** set for each worker $w$ in each period $t$ includes the distributions along $[\underline{v}_{w,f}, \bar{v}_{w,f}]$ and $[\underline{s}_{w,f}, \bar{s}_{w,f}]$, their private values $v_{w,f}$ for all $f \in F$, the commonly-known values $r_w$ and $c_f$ for all $w \in W$ and $f \in F$, the set of acquired match utilities $R_t$, and the set of tentative top-top matches $D_t$. We describe the set of **moves** for each game in each relevant section.

We are interested in the ability for indirect mechanisms (that is, the construction of particular games) to implement *ex-post* stable and Pareto-efficient matchings in weakly dominant strategies. We define a *weakly dominant strategy* to be the set of moves that grant weakly greater payoff than any other set of moves for any profile of other workers' actions. We say that a matching $\mu$ is *ex-post* stable if there exists no pair $(w, f) \notin \mu$ for which $\zeta_{w,f} > \zeta_{w,f'}$ and $\zeta_{w,f} > \zeta_{w'',f}$ where $(w, f'), (w'', f) \in \mu$. We say that a matching is *ex-post* Pareto-efficient if there exists no pair $(w, f) \in \mu$ for which $\zeta_{w',f} > \zeta_{w,f}$ and $\zeta_{w,f''} > \zeta_{w,f}$ while every other worker receives weakly greater payoff. Because preferences are aligned, *ex-post* stability guarantees Pareto-efficiency; if the matching were *ex-post* Pareto-inefficient then there is at least one unmatched worker-firm pair who could receive a strict increase in match utility from matching together, hence the matching is *ex-post* unstable. We frequently shorten *ex-post* stability and Pareto-efficiency to "*ex-post* stability." We say that a mechanism is *ex-post* stable and Pareto-efficient if the matching it produces in weakly dominant strategies is guaranteed to be *ex-post* stable and Pareto-efficient.

We now illustrate that the mechanism described by the naive deferred-acceptance algorithm is *ex-post* unstable and *wasteful*, meaning that workers acquire match utilities with firms that could never have been their stable match. The deferred-acceptance game works as follows: In each period, each worker does not acquire a new match utility if she is already matched in the set of tentative top-top matches. Otherwise, she either acquires a new match utility or exits the market. In *naive* deferred acceptance, each worker acquires the match utility with the firm for which she has the highest expected match utility. We show that naive deferred acceptance can implement *ex-post* unstable matchings for some realizations of firms' private values (Example 1). Additionally, when workers could have inferred which match utilities not to acquire based on match utilities that had already been acquired, naive deferred acceptance can elicit unnecessary acquisitions (Example 2).

Notation for the match-utility matrices in all examples is as follows:

| | $\underline{c_{f_1}}$<br>$f_1$ | $\underline{c_{f_2}}$<br>$f_2$ |
|---|---|---|
| $r_{w_1}$ \| $w_1$ | $v_{w_1,f_1}$ $s_{w_1,f_1}$<br>$\zeta_{w_1,f_1}$ | $v_{w_1,f_2}$ $s_{w_1,f_2}$<br>$\zeta_{w_1,f_2}$ |
| $r_{w_2}$ \| $w_2$ | $v_{w_2,f_1}$ $s_{w_2,f_1}$<br>$\zeta_{w_2,f_1}$ | $v_{w_2,f_2}$ $s_{w_2,f_2}$<br>$\zeta_{w_2,f_2}$ |

**Example 1** (Ex-Post Instability)**.** Consider match utilities drawn as follows, where it is known that $v_{w,f}, s_{w,f} \in [0, 100]$ and $\lambda = \frac{1}{2}$.

| | $\underline{10}$<br>$f_1$ | $\underline{0}$<br>$f_2$ |
|---|---|---|
| 10 \| $w_1$ | 0 100<br>**60** | 100 0<br>55 |
| 0 \| $w_2$ | 80 0<br>45 | 0 50<br>**25** |

The unique stable matching is $\{(w_1, f_1), (w_2, f_2)\}$. But $u_{w_1,f_2} > u_{w_1,f_1}$ and $u_{w_2,f_1} > u_{w_2,f_2}$ so in the first period the acquisitions are $R_1 = \{(w_1, f_2), (w_2, f_1)\}$. No more match utilities are acquired so the matching $\mu = \{(w_1, f_2), (w_2, f_1)\}$ is realized, which is *ex-post* unstable.

**Example 2** (Waste)**.** Consider the same setting from Example 1 but with $v_{w_1,f_1} = 100$.

Match utilities are then given by the following matrix.

| | 10<br>$f_1$ | 0<br>$f_2$ |
|---|---|---|
| 10 \| $w_1$ | 100 100<br>**110** | 100 0<br>55 |
| 0 \| $w_2$ | 80 0<br>45 | 0 50<br>**25** |

In Example 1, the realization of $\zeta_{w_1,f_1}$ could not have precluded $(w_2, f_1)$ from being in the *ex-post* stable matching. But in this example, there is no draw of $s_{w_2,f_1}$ for which $\zeta_{w_2,f_1} > \zeta_{w_1,f_1}$ once $\zeta_{w_1,f_1}$ has been acquired, and so acquiring $\zeta_{w_2,f_1}$ is wasteful. Still, worker $w_2$ acquires match utility $\zeta_{w_2,f_1}$ in naive deferred acceptance because $u_{w_2,f_1} > u_{w_2,f_2}$, hence naive deferred acceptance is wasteful.

The preceding examples demonstrate that naive deferred acceptance cannot guarantee *ex-post* stability or minimize wasteful acquisitions. We now construct a game using the *priority trading* mechanism, where workers sequentially acquire match utilities, and where any worker who is displaced from a tentative top-top match is given priority to acquire match utilities again (or exit the market). We will show that this mechanism also cannot guarantee *ex-post* stability.

## 3.1 Priority Trading

We now introduce the game characterized by the priority trading mechanism (which we call the *PT game*), the outcomes of which we find using the *You Request My House, I Get Your Turn* algorithm (Abdulkadiroglu and Sonmez, 1999). The game works as follows: In the first period, a worker $w$ is randomly chosen to acquire match utilities until she is matched in the set of tentative top-top matches, or she can exit the market.[5] In the next period, another worker $w'$ is randomly chosen to acquire match utilities until she is in the set of tentative top-top matches, or she can exit the market. If worker $w'$ displaces worker $w$ in the set of tentative top-top matches, then worker $w$ acquires match utilities or exits the market, else some worker $w''$ is randomly chosen to acquire match utilities or exit the market. The game continues until all workers either have a match in the set of tentative top-top matches or have exited the market, at which point the top-top matches are made final.

[5] If match utilities are weakly positive, then the first worker to acquire match utilities will only have to make one acquisition.

We show that there exists a sufficiently small $\delta$ for which workers have the weakly dominant strategy to acquire match utilities only with the firms for which they could profitably compete, in decreasing order of the expected match utility. We then show that games with this small-enough $\delta$ can still generate *ex-post* unstable matchings.

We now construct the set of firms $A_w$ with which worker $w$ would ever *reasonably* acquire match utilities, meaning the set of firms with which acquisitions could be profitable. We hold that worker $w$ will not acquire match utility $\zeta_{w,f}$ with firm $f$ if there is some other worker $w'$ who is guaranteed to acquire match utility $\zeta_{w',f}$ where $\zeta_{w',f} > \zeta_{w,f}$ for any draws of private values. If firm $f$ satisfies this, we say firm $f$ has a commonly-known value that is *too high*. Additionally, worker $w$ will not acquire match utility $\zeta_{w,f}$ with firm $f$ if there is some firm $f'$ with which worker $w$ could definitely match, where $\zeta_{w,f'} > \zeta_{w,f}$ for any draws of private values. If firm $f$ satisfies this, we say firm $f$ has a commonly-known value that is *too low*. The set $A_w$, which we frequently call the set of *reasonable firms* for worker $w$, contains all the firms with commonly-known values that are neither too high nor too low; that is, the reasonable firms are the set of firms for which worker $w$ could profitably compete.

Define $w_n$ and $f_n$ to be the worker and firm respectively with the $n$-th highest commonly-known value. Let $k \leq m < n$, so $c_{f_k} \geq c_{f_m} > c_{f_n}$ and $r_{f_k} \geq r_{f_m} > r_{f_n}$. Worker $w_n$ will not acquire match utility $\zeta_{w_n,f_k}$ with *upward* firm $f_k$ (1) if there is some worker $w_m$ whose match utility $\zeta_{w_m,f_k}$ with firm $f_k$ is greater than her match utility $\zeta_{w_m,f_n}$ with firm $f_n$ for all draw of private values; *and* (2) if firm $f_k$ receives greater match utility with worker $w_m$ than with worker $w_n$ for all draws of private values. Formally, we can then define the set of firms that are *too high* for worker $w_n$ as

$$\begin{aligned} A_{w_n}^{\text{high}} := \{f_k \in F \,|\exists w_m \in W \text{ where } & c_{f_k} \geq c_{f_m} > c_{f_n}, r_{w_k} \geq r_{w_m} > r_{w_n} \text{ s.t.} \\ & \lambda c_{f_k} + (1-\lambda)(\underline{v}_{w,f} + \underline{s}_{w,f}) \geq \lambda c_{f_n} + (1-\lambda)(\bar{v}_{w,f} + \bar{s}_{w,f}) \text{ and} \\ & \lambda r_{w_m} + (1-\lambda)(\underline{v}_{w,f} + \underline{s}_{w,f}) \geq \lambda r_{w_n} + (1-\lambda)(\bar{v}_{w,f} + \bar{s}_{w,f})\}. \end{aligned}$$

Now let $k > m \geq n$, so $c_{f_k} < c_{f_m} \leq c_{f_n}$ and $r_{f_k} < r_{f_m} \leq r_{f_n}$. Worker $w_n$ will not acquire match utility $\zeta_{w_n,f_k}$ with *downward* firm $f_k$ (1) if there is some firm $f_m$ where the match utility $\zeta_{w_n,f_m}$ with firm $f_m$ is greater than match utility $\zeta_{w_n,f_k}$ with firm $f_k$ for all draws of private values; *and* (2) if firm $f_m$ gets greater match utility with worker $w_n$ than with worker $w_m$ for all draws of private values. Formally, we can then define the set of firms that are *too* *low* for worker $w_n$ as

$$A_{w_n}^{\text{low}} := \{f_k \in F \mid \exists f_m \in F \text{ where } c_{f_k} < c_{f_m} \leq c_{f_n}, r_{w_k} < r_{w_m} \leq r_{w_n} \text{ s.t.}$$
$$\lambda c_{f_m} + (1-\lambda)(\underline{v}_{w,f} + \underline{s}_{w,f}) \geq \lambda c_{f_k} + (1-\lambda)(\bar{v}_{w,f} + \bar{s}_{w,f}) \text{ and}$$
$$\lambda r_{w_n} + (1-\lambda)(\underline{v}_{w,f} + \underline{s}_{w,f}) \geq \lambda r_{w_m} + (1-\lambda)(\bar{v}_{w,f} + \bar{s}_{w,f})\}.$$

Then the set of *reasonable firms* for worker $w$ is the set

$$A_w := F \setminus (A_w^{\text{low}} \cup A_w^{\text{high}}).$$

Note the construction of $A_w$ is independent of the *draws* of the private values, and takes as parameters just the commonly-known values and certainty weight $\lambda$.

Let the set of workers who would reasonably compete for firm $f$ be defined

$$A_f := \{w \in W \mid f \in A_w\}.$$

Then define the probability that worker $w$ acquiring $\zeta_{w,f}$ results in the match $(w, f)$ given acquisitions from all $w' \in A_f$ as

$$\gamma(w, A_f) := \prod_{w' \in A_f, w' \neq w} \Pr(\zeta_{w,f} > \zeta_{w',f}). \tag{3}$$

Because we assume that preferences are aligned, no two match utilities will be exactly equivalent. Without this assumption, we lose the guarantee that the *ex-post* stable matching is unique (Ferdowsian et al., 2023).[6] We then slightly abuse the definition for the utility that a worker receives from a match (Equation (2)), and write the expected utility that worker $w$ receives from acquiring $\zeta_{w,f}$ when facing competition from all $w' \in A_f$ as

$$\mathrm{E}[\mathbf{u}_{w,f}] = \gamma(w, A_f) \cdot \mathrm{E}[\zeta_{w,f}] - \delta. \tag{4}$$

Note that the expected utility from acquiring a match utility is distinct from the *expected match utility*, which is given by $\mathrm{E}[\zeta_{w,f}]$ alone.

We construct an upper bound on $\delta$ such that acquisition behavior always replicates that of the game *without* acquisition costs—that is, workers have the weakly dominant strategy to acquire match utilities with their respective reasonable firms in their expected preference

[6] We could additionally enforce that if two match utilities take equivalent values—either $\zeta_{w,f} = \zeta_{w',f}$ where $w \neq w'$, or $\zeta_{w,f} = \zeta_{w,f'}$ where $f \neq f'$—that one of the two match utilities is randomly chosen as more preferred. And if either of the distributions along $[\underline{v}_{w,f}, \bar{v}_{w,f}]$ or $[\underline{s}_{w,f}, \bar{s}_{w,f}]$ is continuous, then the probability measure of acquiring identical match utilities is zero.

ordering (Proposition 1). The maximum size of $\delta$ is such that the expected loss that a worker suffers from skipping down her preferences is weakly greater than the cost of the acquisition itself. Then, we show that these weakly dominant strategies—where workers acquire match utilities in their expected preference ordering, while still skipping the firms with which they could never reasonably match—still cannot guarantee *ex-post* stability (Proposition 2). This is because there are draws of private values such that blocking pairs exist, for which the match utilities were never acquired.

**Proposition 1.** *Suppose* $\delta \leq \bar{\delta}$ *where*

$$\bar{\delta} \leq \frac{\mathrm{E}[\boldsymbol{u}_{w,f^*}] - \mathrm{E}[\boldsymbol{u}_{w,f}]}{1 - \gamma(w, A_{f^*})} + \gamma(w, A_f)\, \mathrm{E}[\zeta_{w,f}], \tag{5}$$

*for all* $f \neq f^*$ *where* $f^* = \arg\max_{f \in A_w \setminus R_t^w} u_{w,f}$. *Then the PT game admits weakly dominant strategies where workers acquire match utilities in decreasing order of* $u_{w,f}$ *from the set of reasonable firms, until she is matched in the set of tentative top-top pairs.*

*Proof.* First consider any match utility $\zeta_{w,f}$ that worker $w$ could acquire with any firm $f \notin A_w \cup R_t^w$. If $f \notin A_w$, then there are two cases. The first is that $f \in A_w^{\text{high}}$, where acquiring $\zeta_{w,f}$ will surely lead to rejection, because firm $f$ is guaranteed greater utility from matching with some worker $w'$ who will certainly acquire match utility $\zeta_{w',f}$. The second is that $f \in A_w^{\text{low}}$, but then worker $w$ is guaranteed greater utility from matching with some firm $f'$ with a higher commonly-known value than firm $f$, and with which she could be guaranteed a final match. And because no worker $w'$ will abandon a tentative top-top match $(w', f) \in D_t$ unless some $\zeta_{w'',f}$ is acquired such that $\zeta_{w'',f} > \zeta_{w',f}$, the match utility that firm $f$ receives weakly increases over time, and so worker $w$ acquiring any $\zeta_{w,f}$ where $f \in R_t^w$ necessarily leads to another rejection. Hence in period $t$, it is weakly dominant for each worker $w$ to acquire match utility $\zeta_{w,f}$ with firm $f$ where $f \in A_w \setminus R_t^w$.

Let $f^* = \arg\max_{f \in A_w \setminus R_t^w} u_{w,f}$ be the most-preferred reasonable firm for worker $w$ who has not yet rejected her. For acquiring match utility $\zeta_{w,f^*}$ to be the weakly dominant strategy in period $t$, the expected utility gain from acquiring match utility $\zeta_{w,f^*}$ must be weakly greater than acquiring $\zeta_{w,f}$ immediately. But the match $(w, f^*)$ is not guaranteed; with probability $1 - \gamma(w, A_{f^*})$, worker $w$ will have to acquire match utility with some firm $f$ after acquiring match utility $\zeta_{w,f^*}$. Thus we need the condition

$$\mathrm{E}[\mathbf{u}_{w,f^*}] + (1 - \gamma(w, A_{f^*}))\, \mathrm{E}[\mathbf{u}_{w,f}] \geq \mathrm{E}[\mathbf{u}_{w,f}],$$

for all firms $f \neq f^*$, where $f^* = \arg\max_{f \in A_w \setminus R_t^w} u_{w,f}$. Rearranging and plugging in the definition of $\mathrm{E}[\mathbf{u}_{w,f}]$ from Equation (4), we can isolate $\delta$:

$$\delta \leq \frac{\mathrm{E}[\mathbf{u}_{w,f^*}] - \mathrm{E}[\mathbf{u}_{w,f}]}{1-\gamma(w, A_{f^*})} + \gamma(w, A_f)\,\mathrm{E}[\zeta_{w,f}].$$

This is Condition (5). □

We now show that even with $\delta \leq \bar{\delta}$ satisfying Condition (5), the resulting matching is not guaranteed to be *ex-post* stable. The proof in Proposition 2 formalizes the following intuition: Consider a worker $w$ who acquires match utility $\zeta_{w,f^*}$ with firm $f^*$ where $f^* = \arg\max_{f \in A_w \setminus R_t^w} u_{w,f}$, and ultimately $(w, f^*)$ are matched in $D_T$. That $(w, f^*)$ ultimately match does not rule out the existence of some firm $f'$ with which worker $w$ has not acquired match utility $\zeta_{w,f'}$, where $\zeta_{w,f'} > \zeta_{w,f^*}$. Then a matching with $(w, f')$ could have been stable.

**Proposition 2.** *Suppose that the PT game admits weakly dominant strategies as in Proposition 1. The resulting matching is not guaranteed to be ex-post stable and Pareto-efficient.*

*Proof.* Suppose at some period $t$, worker $w$ acquired match utility $\zeta_{w,f^*}$ where $f^* = \arg\max_{f \in A_w \setminus R_t^w} u_{w,f}$, and say $(w, f^*) \in D_T$. Because $(w, f^*) \in D_T$, no worker $w'$ acquired match utility $\zeta_{w',f^*}$ where $\zeta_{w',f^*} > \zeta_{w,f^*}$. Then let $(w', f') \in D_T$ where at some period $t' \neq t$, firm $f'$ satisfied $f' = \arg\max_{f \in A_{w'} \setminus R_{t'}^{w'}} u_{w',f}$. This acquisition ordering follows from weakly dominant strategies as in Proposition 1 because worker $w$ acquired match utility $\zeta_{w,f^*}$ where $f^* = \arg\max_{f \in A_w \setminus R_t^w} u_{w,f}$ and worker $w'$ acquired match utility $\zeta_{w',f'}$ where $f' = \arg\max_{f \in A_{w'} \setminus R_{t'}^{w'}} u_{w',f}$.

Then let $f' \in (A_w \setminus R_t^w)$ and $f^* \in (A_{w'} \setminus R_{t'}^{w'})$, and so both $\zeta_{w,f'}$ and $\zeta_{w',f^*}$ were never acquired. But because $f' \in (A_w \setminus R_t^w)$ and $f^* \in (A_{w'} \setminus R_{t'}^{w'})$ then for some draws of private values ($v_{w,f'}$ and $v_{w',f^*}$, and $s_{w,f'}$ and $s_{w',f^*}$), we could have $\zeta_{w,f'} > \zeta_{w,f^*}$ and $\zeta_{w',f^*} > \zeta_{w',f'}$. But if $\zeta_{w,f'} > \zeta_{w,f^*}$ and $\zeta_{w',f^*} > \zeta_{w',f'}$, then any matching with $(w, f^*), (w', f') \in D_T$ is *ex-post* Pareto-inefficient. Then the matching is *ex-post* unstable because workers $w, w'$ and firms $f^*, f'$ improve from new matches $(w, f'), (w', f^*) \in D_T$. □

We have shown that there are draws of private values such that blocking pairs could exist between workers and firms whose match utilities had not been acquired. As a result, the matching resulting from the weakly dominant strategies in Proposition 1 is not guaranteed to be *ex-post* stable. The missing element is *sufficient revelation*; that is, workers need to sufficiently acquire their match utilities to ensure that they have arrived at their *ex-post* stable match. We now propose a class of mechanisms (in that we construct a new game with different possible moves) that guarantee sufficient revelation in weakly dominant strategies, given that $\delta$ is sufficiently small. We then characterize the mechanism within the class which minimizes the number of expected acquisitions.

### 3.2 Priority Trading With Multi-Revelation

We propose a game called *priority trading with multi-revelation* (PTMR). The game works as follows: In the first period, a worker $w$ is randomly chosen to acquire match utilities until she is matched in the set of tentative top-top matches, or she exits the market. After she is matched in the set of tentative top-top matches, then starting in the subsequent period, worker $w$ can then acquire any additional number of match utilities. In the first period after which worker $w$ no longer acquires match utilities, another worker $w'$ is randomly chosen to acquire match utilities until she is matched in the set of tentative top-top matches, or exits the market. Worker $w'$ can then acquire subsequent match utilities after she is matched in the set of tentative top-top matches if she chooses. If worker $w'$ displaces worker $w$ from the set of tentative top-top matches, then worker $w$ acquires match utilities again or exits the market; else some worker $w''$ acquires match utilities or exits the market. The game continues until all workers stop acquiring match utilities, at which point the top-top matches are made final.

We define $S_t^w$ to be the *feasible stable matches* (or the *feasible firms*) for worker $w$ as determined by the set of acquired match utilities by period $t$—that is, $S_t^w$ is the set of firms with which worker $w$ could feasibly match in the *ex-post* stable matching, given her information in period $t$. We show that for sufficiently small $\delta$, it is weakly dominant for worker $w$ in period $t$ to acquire match utility with firm $f^* = \arg\max_{f \in S_t^w} u_{w,f}$, update her set of feasible stable matches $S_{t+1}^w$, then acquire match utilities with all firms in that set (Proposition 3). If some match utility $\zeta_{w,f'}$ is acquired in some period $\tilde{t} > t$ such that $\zeta_{w,f'} > \zeta_{w,f^*}$ and $(w, f') \in D_{\tilde{t}}$, then worker $w$ updates her set of feasible stable matches $S_{\tilde{t}+1}^w$, and acquires match utilities with all firms in that set. We show that when $\delta$ is small enough such that these are the weakly dominant strategies, the game admits a matching that is guaranteed to be *ex-post* stable and Pareto-efficient (Proposition 4).

We have so far not specified the initial priority ordering of workers. Define the *PTMR class of mechanisms* to be the set of PTMR games, each with a unique priority ordering of workers. We show that prioritizing the workers in decreasing order of commonly-known values is *cost-minimizing*, meaning the expected number of acquisitions is minimized. We call the PTMR mechanism that places workers in decreasing order of commonly-known values *top-down PTMR* (Proposition 5). The intuition for the cost-minimization aspect is that as the game proceeds, less firms with higher commonly-known values remain feasible stable matches, and so there are less acquisitions made with firms already paired with their stable match.

We formally define the set of feasible stable matches. In any period $t$, each worker $w$ has an interim set of feasible firms, called $S_t^w$, for which it is potentially profitable to acquire match utilities:

$$S_t^w := F \setminus (\chi_t^w \cup \Psi_t^w \cup R_t^w),$$

with sets to be defined. The interim set of feasible firms for worker $w$ is the total set of firms less the sets of firms that are unreachable ($\chi_t^w$) by worker $w$ in period $t$, the firms with commonly-known values that are too small to possibly constitute an optimal match ($\Psi_t^w$) for worker $w$ in period $t$, and the firms with which worker $w$ has already acquired match utility ($R_t^w$).

We define the interim set of firms $\chi_t^w$ for which the already-acquired match utility of some other tentative top-top match is *too high* as

$$\chi_t^w := \{f \in F \mid \lambda(c_f + r_w) + (1-\lambda)(v_{w,f} + \bar{s}_{w,f}) \leq \zeta_{w',f} \quad \exists (w', f) \in D_t\}. \tag{6}$$

The interpretation of this is as follows: $\chi_t^w$ are the firms for which there is no draw of private values such that these firms could prefer to match with worker $w$ over their tentative top-top matches. Then worker $w$ will not acquire match utilities with those firms in period $t$.

We additionally define the interim set of firms $\Psi_t^w$ for which the highest possible realized match utility is *too low* relative to the tentative top-top match of worker $w$:

$$\Psi_t^w := \{f \in F \mid \lambda(c_f + r_w) + (1-\lambda)(v_{w,f} + \bar{s}_{w,f}) \leq \zeta_{w,f'} \quad \exists (w, f') \in D_t\}. \tag{7}$$

The set $\Psi_t^w$ contains the firms for which no draws of private values could remove $(w, f')$ from the set of tentative top-top matches. Then worker $w$ will not acquire match utilities with those firms in period $t$.

Now we show that we can find a sufficiently small $\delta$ for which each worker $w$ has the weakly dominant strategy to acquire match utility $\zeta_{w,f^*}$ in period $t$, where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$. Specifically, $\delta$ must be small enough such that in every period where a worker acquires match utilities or exits, the expected loss from missing on a most-preferred feasible stable match is weakly greater than the cost of an acquisition. There are two scenarios: The first is where worker $w$ is not yet matched in the set of tentative top-top matches. We derive a condition for $\delta$ under which missing on a most-preferred feasible stable match is always weakly more costly than immediately acquiring the match utility with some less-preferred firm. The condition for $\delta$ when worker $w$ is unmatched in the set of tentative top-top matches follows the same construction as that of Condition (5).

The second scenario is where there exists some firm $f'$ such that $(w, f') \in D_t$. Again, let firm $f^*$ satisfy $f^* = \arg\max_{f \in S_t^w} u_{w,f}$. For worker $w$ to have the weakly dominant strategy to acquire match utility $\zeta_{w,f^*}$ with firm $f^*$, worker $w$ must gain more utility in expectation from acquiring match utility $\zeta_{w,f^*}$ than she suffers in incurred cost. But the expected utility gain, whenever worker $w$ is already matched in the set of tentative top-top matches, is conditional on the probability that $\zeta_{w,f^*}$ is acquired such that $\zeta_{w,f^*} > \zeta_{w,f'}$, *as well as* the probability that firm $f^*$ becomes a top-top match, given by $\gamma(w, A_{f^*})$. We derive a sufficient condition for $\delta$ where the expected utility gain from acquiring match utility $\zeta_{w,f^*}$—which is additionally dependent on the probability that worker $w$ and firm $f^*$ result in a final match—is always a weak improvement in expectation than merely remaining in the match $(w, f')$.

**Proposition 3.** *Suppose $\delta \leq \bar{\delta}$ at all $t$ where*

$$\bar{\delta} = \begin{cases} \frac{\mathrm{E}[\boldsymbol{u}_{w,f^*}] - \mathrm{E}[\boldsymbol{u}_{w,f}]}{1-\gamma(w,A_{f^*})} + \gamma(w, A_f)\,\mathrm{E}[\zeta_{w,f}] & \text{if } \nexists f' \text{ s.t. } (w,f') \in D_t \\ \Pr(\zeta_{w,f^*} > \zeta_{w,f'}) \cdot [\gamma(w, A_{f^*}) \cdot \mathrm{E}[\zeta_{w,f^*} \mid \zeta_{w,f^*} > \zeta_{w,f'}] - \\ \qquad \gamma(w, A_{f'}) \cdot \zeta_{w,f'}] & \text{if } \exists f' \text{ s.t. } (w,f') \in D_t \end{cases} \tag{8}$$

*for all $f \neq f^*$ where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$. Then the PTMR game admits weakly dominant strategies where workers acquire match utilities in decreasing order of $u_{w,f}$ until no firms remain in the set of feasible stable matches.*

*Proof.* We first show that it is not profitable for worker $w$ to acquire match utility $\zeta_{w,f}$ in period $t$ with any firm $f \notin S_t^w$. If $f \notin S_t^w$, then there are three cases. First, if $f \in \chi_t^w$, then there exists some $(w', f) \in D_t$ where there are no such private values where $\zeta_{w,f} > \zeta_{w',f}$, so worker $w$ will surely be rejected by firm $f$ in period $t$. Second, if $f \in \Psi_t^w$, then worker $w$ is already matched with firm $f'$ in the set of tentative top-top matches, and there are no such private values where $\zeta_{w,f'} > \zeta_{w,f}$. Third, if $f \in R_t^w$, then match utility $\zeta_{w,f}$ has already been acquired and an additional acquisition will only incur additional cost. Hence it is only profitable for worker $w$ to acquire match utility $\zeta_{w,f}$ with some $f \in S_t^w$ in period $t$.

Define $f^* = \arg\max_{f \in S_t^w} u_{w,f}$. When there does not exist some firm $f'$ such that $(w, f') \in D_t$, it is a weakly dominant strategy for worker $w$ to acquire $\zeta_{w,f^*}$ in all periods $t$ if the expected utility loss from acquiring match utility with some other firm $f$ is weakly greater than $\delta$. We closely follow the proof of Proposition 1: With probability $\gamma(w, A_{f^*})$, worker $w$ acquiring match utility $\zeta_{w,f^*}$ results in match $(w, f^*) \in D_T$, but with probability $1 - \gamma(w, A_{f^*})$, worker $w$ will make some other acquisition(s) or will exit the market. So for acquiring match utility $\zeta_{w,f^*}$ to be the weakly dominant strategy in period $t$, the expected utility gain from acquiring match utility $\zeta_{w,f^*}$ must be weakly greater than acquiring $\zeta_{w,f}$ immediately. Hence when there does not exist some firm $f'$ such that $(w, f') \in D_t$, workers have the weakly dominant strategy to acquire match utilities in decreasing order of $u_{w,f}$ so long as

$$\mathrm{E}[\mathbf{u}_{w,f^*}] + (1 - \gamma(w, A_{f^*}))\,\mathrm{E}[\mathbf{u}_{w,f}] \geq \mathrm{E}[\mathbf{u}_{w,f}],$$

for all firms $f \neq f^*$, where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$. Rearranging and plugging in the definition of $\mathrm{E}[\mathbf{u}_{w,f}]$ from Equation (4), we recover:

$$\delta \leq \frac{\mathrm{E}[\mathbf{u}_{w,f^*}] - \mathrm{E}[\mathbf{u}_{w,f}]}{1 - \gamma(w, A_{f^*})} + \gamma(w, A_f)\,\mathrm{E}[\zeta_{w,f}].$$

We now examine the case where $(w, f') \in D_t$ for some firm $f'$. We show that, for sufficiently small values of $\delta$, there exists a weakly dominant strategy for worker $w$ to acquire match utilities with firms in the set of feasible stable matches in decreasing order of her expected match utility—that is, with the firm $f^*$ that sequentially satisfies $f^* = \arg\max_{f \in S_t^w} u_{w,f}$.[7]

We prove that worker $w$ should acquire match utilities from the set $S_t^w$ in decreasing order of the expected value of the match utility. First consider if worker $w$ acquired match utility $\zeta_{w,f}$ with some firm $f \neq f^*$ where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$. Recall that the size of the set of *too-low* firms $\Psi_{t'}^w$ at some period $t' > t$ is increasing in the match utility $\zeta_{w,f'}$, where $(w, f') \in D_{t'}$. But the match utility $\zeta_{w,f}$ is smaller in expectation than $\zeta_{w,f^*}$ for all $f \neq f^*$. Then any order of acquisitions that is not decreasing in commonly-known values may be wasteful—worker $w$ may acquire some $\zeta_{w,f}$ with some firm $f$ that would have otherwise entered the set of *too-low* firms $\Psi_{t'}^w$ at some period $t' > t$.

Now we find a condition for $\delta$ where worker $w$ has the weakly dominant strategy to acquire all match utilities with firms in her set of feasible stable matches $S_t^w$. With probability $\gamma(w, A_{f^*}) \cdot \Pr(\zeta_{w,f^*} > \zeta_{w,f'})$, the acquisition of $\zeta_{w,f^*}$ where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$ will result in a tentative top-top match with firm $f^*$; note we now consider the probability that firm $f'$ is displaced from the tentative top-top matches *alongside* the probability that $(w, f^*) \in D_T$. Then with probability $1 - \Pr(\zeta_{w,f^*} < \zeta_{w,f'})$, worker $w$ merely remains with firm $f'$ in the set of tentative top-top matches after acquiring match utility $\zeta_{w,f^*}$.

Then worker $w$ receives weakly greater utility when acquiring $\zeta_{w,f^*}$ when $(w, f') \in D_t$ if the following condition holds:

$$\Pr(\zeta_{w,f^*} > \zeta_{w,f'}) \cdot \mathrm{E}[\mathbf{u}_{w,f^*} \mid \zeta_{w,f^*} > \zeta_{w,f'}] + (1 - \Pr(\zeta_{w,f^*} > \zeta_{w,f'})) \cdot (\gamma(w, A_{f'}) \cdot \zeta_{w,f'} - \delta) \geq \gamma(w, A_{f'}) \cdot \zeta_{w,f'}, \quad (9)$$

where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$ and $(w, f') \in D_t$. The expected utility from worker $w$ acquiring

[7] We sacrifice some notation in favor of clarity. Formally, we mean that each worker $w$ has the weakly dominant strategy to acquire match utility with firm $f^* = \arg\max_{f \in S_t^w} u_{w,f}$, then $f^{**} = \arg\max_{f \in S_{t+1}^w} u_{w,f}$, and so on, until the set of feasible stable matches is empty. We continue this slight abuse of notation throughout the proof.

match utility $\zeta_{w,f^*}$ conditional on $\zeta_{w,f^*} > \zeta_{w,f'}$ is given by

$$\mathrm{E}[\mathbf{u}_{w,f^*} \mid \zeta_{w,f^*} > \zeta_{w,f'}] = \gamma(w, A_{f^*}) \cdot \mathrm{E}[\zeta_{w,f^*} \mid \zeta_{w,f^*} > \zeta_{w,f'}] - \delta,$$

which follows from Equation (4). Then we can rewrite Inequality (9) and isolate $\delta$:

$$\delta \leq \Pr(\zeta_{w,f^*} > \zeta_{w,f'}) \cdot \left[\gamma(w, A_{f^*}) \cdot \mathrm{E}[\zeta_{w,f^*} \mid \zeta_{w,f^*} > \zeta_{w,f'}] - \gamma(w, A_{f'}) \cdot \zeta_{w,f'}\right],$$

where $f^* = \arg\max_{f \in S_t^w} u_{w,f}$ and $(w, f') \in D_t$. Note worker $w$ updates her set of feasible stable matches if she acquires some $\zeta_{w,f^{**}}$ where $\zeta_{w,f^{**}} > \zeta_{w,f'}$ such that $(w, f^{**}) \in D_{t'}$ at some $t' > t$.

□

We note that Condition (8) is a sufficient, but not a necessary, condition for the game to exhibit these weakly dominant strategies, where each worker $w$ acquires match utilities in decreasing order of $u_{w,f}$ until no firms remain in the set of feasible stable matches. Consider the second case within Condition (8), where there is some firm $f'$ such that $(w, f') \in D_t$. The probability that a worker-firm pair results in a top-top match dynamically updates according to the set of match utilities already acquired. We compute the expected utility that worker $w$ receives from her current match as $\gamma(w, A_{f'}) \cdot \zeta_{w,f'}$, which is the product of the probability that worker $w$ acquires match utility greater than those acquired by all other workers for whom firm $f$ is reasonable, and the match utility $\zeta_{w,f'}$ which worker $w$ has already acquired.[8] It may seem natural, though, to compute $\gamma(w, A_{f'} \setminus R_t^{f'})$ instead of $\gamma(w, A_{f'})$, because worker $w$ has beaten out some competition already if $(w, f') \in D_t$. But there may be some workers who have acquired a higher match utility with firm $f'$, but who have a more preferred top-top match elsewhere—then it is *still possible* that $(w, f') \in D_t$.

Indeed, the top-top construction of final matches implies that the change in one match can lead to a "cascading effect" where the set of tentative top-top matches changes for multiple workers and firms. Deriving a *necessary* condition for the weakly dominant strategies in Proposition 3 that incorporates the dynamic state of the game (which updates the probabilities of matching with a firm according to the set of already-acquired match utilities) is still an open question. We derive a sufficient condition for $\delta$ based on the expected utility of acquiring match utility $\zeta_{w,f}$ with firm $f$, given competition from all workers for whom firm $f$ is a reasonable match.

The implication of the weakly dominant strategies in Proposition 3 is that workers sufficiently acquire all match utilities with firms that could feasibly be their stable match. Now,

[8] See the right-hand side of Inequality (9).

unlike our result for PT games in Proposition 2, the PTMR game guarantees an *ex-post* stable matching given these are the weakly dominant strategies. We prove that the matching is guaranteed *ex-post* stable by constructing the unique *ex-post* stable matching independently of PTMR, and then showing that the matching from PTMR (under these weakly dominant strategies) coincides with the unique *ex-post* stable matching.

**Proposition 4.** *Suppose that the PTMR game admits weakly dominant strategies as in Proposition 3. The matching is then guaranteed to be ex-post stable and Pareto-efficient.*

*Proof.* Define a *submarket* as a nonempty set of workers and a nonempty set of firms, possibly without the complete set of workers and firms from $W$ and $F$ respectively. Consider the construction of the matching $\mu$ as follows: Find $(w', f') = \arg\max_{(w,f)\in(W\times F)} \zeta_{w,f}$, then define the submarket $(\tilde{W}, \tilde{F})$ where $\tilde{W} = W \setminus w'$ and $\tilde{F} = F \setminus f'$. Then find $(w'', f'') = \arg\max_{(w,f)\in(\tilde{W}\times\tilde{F})} \zeta_{w,f}$. Because preferences are aligned, a top-top match is found in each iteration. Continue until no submarkets remain, and let $\mu = \{(w', f'), (w'', f''), ...\}$.

We show that $\mu$ is *ex-post* stable and Pareto-efficient. Say some $(w', f')$ is a blocking pair. But because preferences are aligned then $(w', f') = \arg\max_{(w,f)\in(\tilde{W}\times\tilde{F})} \zeta_{w,f}$ for some submarket $(\tilde{W}, \tilde{F})$, so $(w', f')$ must be matched, hence a matching from this construction is *ex-post* stable. Consider that the matching is *ex-post* Pareto-inefficient and some $(w, f)$ are matched, so $\zeta_{w,f'} > \zeta_{w,f}$ and $\zeta_{w'',f} > \zeta_{w,f}$, and all workers and firms besides worker $w$ and firm $f$ weakly improve from $(w, f'), (w'', f) \in \mu$. But because the matching is *ex-post* stable then $(w, f')$ and $(w'', f)$ could not possibly be matched, so the matching is *ex-post* Pareto-efficient.

Now we show that the matching $\mu'$ constructed by PTMR in the weakly dominant strategies from Proposition 3 coincides with the above, by showing that the set of acquisitions $R_T^w$ generated by the PTMR game contains the set of stable matches. Consider that there is some firm $f$ such that $f \notin S_t^w$ in all periods $t$, but $(w, f)$ is in the *ex-post* stable matching. If $f \in \chi_t^w$ for all periods $t$, then there always exists some $(w', f) \in D_t$ such that no match utility $\zeta_{w,f}$ is possible such that $\zeta_{w,f} > \zeta_{w',f}$. Hence $(w, f)$ is not a feasible stable match. Then consider that $f \in \Psi_t^w$ for all periods $t$. But then the acquisition of $\zeta_{w,f}$ is never profitable because there are no draws of private values where worker $w$ prefers firm $f$ to her current match, so $(w, f)$ cannot be a stable match. Hence if $f \in \chi_t^w \cup \Psi_t^w$ for all periods $t$ then firm $f$ cannot be a stable match for worker $w$. By the weakly dominant strategies, if $f \notin \chi_t^w \cup \Psi_t^w$ in some period $t$ then the match utility $\zeta_{w,f}$ is acquired, so $(w, f) \in R_T$. And so $R_T^w$ contains the set of feasible stable matches for each worker $w$ at time $T$. Hence matching $(w', f') = \arg\max_{(w,f)\in R_T} \zeta_{w,f}$, coincides with $(w', f') = \arg\max_{(w,f)\in(W\times F)} \zeta_{w,f}$, and the top-top match coincides for every submarket that elides all previously matched members. Hence the matching is *ex-post* stable and Pareto-efficient. □

We have shown that weakly dominant strategies from Proposition 3 guarantee *ex-post* stability. The matching could be *ex-post* unstable if workers deviate from these strategies—for instance, if a worker exits prematurely or neglects acquiring a match utility with a firm from her set of feasible stable matches. So we have identified a sufficient condition for $\delta$ for which sufficient revelation is guaranteed, and the *ex-post* stable matching is implemented.

We now show that the *top-down PTMR* mechanism, where agents are given priority in decreasing order of their commonly-known values, is the cost-minimizing mechanism from the class of PTMR mechanisms—meaning that this mechanism minimizes the expected number of acquisitions. We prove this by showing that, whenever a worker with a smaller commonly-known value is given priority over a worker with a higher commonly-known value, the expected average size of the set of feasible stable matches increases.

**Proposition 5.** *Suppose that the PTMR game admits weakly dominant strategies as in Proposition 3. Then top-down PTMR is the cost-minimizing mechanism from the class of PTMR mechanisms that guarantees the ex-post stable and Pareto-efficient matching.*

*Proof.* Consider a PTMR game where worker $w_m$ acquires match utilities before worker $w_n$ where $r_{w_m} > r_{w_n}$. Say there is some firm $f$ such that $(w_m, f) \in D_t$. The probability that firm $f \in S_t^{w_n}$ decreases in match utility $\zeta_{w_m,f}$ by definition of the *too-high* set of firms $\chi_t^{w_n}$ in Equation (6). Because $\zeta_{w_m,f} > \zeta_{w_n,f}$ in expectation, flipping the priority of $w_m$ and $w_n$ increases the expected average size of the set of feasible stable matches. Because this holds for every $w_m, w_n \in W$ where $r_{w_m} > r_{w_n}$, it must be that top-down PTMR is cost-minimizing. □

In the examples that follow, we illustrate that top-down PTMR is the cost-minimizing mechanism guaranteeing the *ex-post* stable matching. First, in Example 3, we show that the top-down PTMR game can admit weakly dominant strategies where the *ex-post* stable matching is realized in one acquisition for each worker.

**Example 3** (Replicating the Singleton)**.** Consider match utilities drawn as follows, where it is known that $v_{w,f}, s_{w,f} \in [0, 10]$ and $\lambda = \frac{1}{2}$.

| | $\underline{+100}$<br>$f_1$ | $\underline{+50}$<br>$f_2$ | $\underline{+0}$<br>$f_3$ |
|---|---|---|---|
| +100 \| $w_1$ | 0 0<br>**100** | 5 5<br>80 | 10 9<br>58 |
| +50 \| $w_2$ | 10 10<br>85 | 5 5<br>**55** | 10 9<br>34 |
| +0 \| $w_3$ | 10 10<br>60 | 10 10<br>35 | 0 0<br>**0** |

In $t = 1$, worker $w_1$ acquires match utility $\zeta_{w_1,f_1} = 100$ with firm $f_1 = \arg\max_{f \in S_1^{w_1}} u_{w_1,f}$. Then $\Psi_2^{w_1} = \{f_2, f_3\}$ because private values $s_{w_1,f_2}$ and $s_{w_1,f_3}$ cannot be large enough such that $\zeta_{w_1,f_2} > \zeta_{w_1,f_1}$ or $\zeta_{w_1,f_3} > \zeta_{w_1,f_1}$. Then $S_2^{w_1} = \emptyset$ because $R_1^{w_1} = \{f_1\}$ so no more match utilities are acquired.

In $t = 2$, $\chi_2^{w_2} = \{f_1\}$ because there do not exist draws of $s_{w_2,f_1}$ such that $\zeta_{w_2,f_1} > \zeta_{w_1,f_1}$. Then worker $w_2$ acquires match utility $\zeta_{w_2,f_2} = 55$ with firm $f_2 = \arg\max_{f \in S_2^{w_2}} u_{w_2,f}$. Then $\Psi_2^{w_2} = \{f_3\}$ because private value $s_{w_2,f_3}$ cannot be large enough such that $\zeta_{w_2,f_3} > \zeta_{w_2,f_2}$. Then $S_3^{w_2} = \emptyset$ so no more match utilities are acquired.

In $t = 3$, $\chi_3^{w_3} = \{f_1, f_2\}$ because there do not exist draws of $s_{w_3,f_1}$ and $s_{w_3,f_2}$ such that $\zeta_{w_3,f_1} > \zeta_{w_1,f_1}$ or $\zeta_{w_3,f_2} > \zeta_{w_2,f_2}$. Then $S_3^{w_3} = \{f_3\}$ and so worker $w_3$ acquires match utility $\zeta_{w_3,f_3} = 0$ with firm $f_3 = \arg\max_{f \in S_3^{w_3}} u_{w_3,f}$. Hence $D_3 = \{(w_1, f_1), (w_2, f_2), (w_3, f_3)\}$ and is *ex-post* stable. Thus the minimum number of match utilities is acquired and the acquisitions replicate the singleton preference submission for each worker.

The next example demonstrates that even when values are drawn such that many acquisitions may be anticipated, top-down PTMR minimizes the number necessary. Example 4 has a game with clustered private values and match utilities, where the *ex-post* stable matching is found in the minimum number of acquisitions.

**Example 4** (Minimum Necessary Checks)**.** Consider match utilities drawn as follows, where it is known that $v_{w,f}, s_{w,f} \in [0, 10]$ and $\lambda = \frac{1}{2}$.

| | +10<br>$f_1$ | +5<br>$f_2$ | +0<br>$f_3$ |
|---|---|---|---|
| +10 \| $w_1$ | 10 0<br>15 | 4 10<br>**14.5** | 8 10<br>14 |
| +5 \| $w_2$ | 10 10<br>**17.5** | 9 10<br>17 | 0 0<br>2.5 |
| +0 \| $w_3$ | 10 6<br>13 | 10 10<br>12.5 | 0 0<br>**0** |

We spare some notation and discuss how top-down PTMR arrives at the stable matching in the minimum number of acquisitions. In $t = 1$, worker $w_1$ acquires match utility $\zeta_{w_1,f_1} = 15$ with firm $f_1 = \arg\max_{f \in S_1^{w_1}} u_{w_1,f}$. She does not acquire any other match utilities because even with $s_{w_1,f_j} = \bar{s}_{w,f}$, we have that $\zeta_{w_1,f_1} > \zeta_{w_1,f_2}$ and $\zeta_{w_1,f_1} > \zeta_{w_1,f_3}$.

Then in $t = 2$, worker $w_2$ recognizes that private value $s_{w_2,f_1}$ could be drawn such that $\zeta_{w_2,f_1} > \zeta_{w_1,f_1}$ thus blocking the pair $(w_1, f_1)$. So worker $w_2$ acquires match utility $\zeta_{w_2,f_1}$ and recognizes that there is no draw of private value $s_{w_2,f_2}$ or $s_{w_2,f_3}$ such that $\zeta_{w_2,f_2} > \zeta_{w_2,f_1}$ or $\zeta_{w_2,f_3} > \zeta_{w_2,f_1}$.

Now worker $w_1$ is given priority again and recognizes that she can no longer match with firm $f_1$, so she acquires another match utility. Even though $v_{w_1,f_3} > v_{w_1,f_2}$, she acquires match utility $\zeta_{w_1,f_2}$ because $u_{w_1,f_2} > u_{w_1,f_3}$. And indeed, there is no draw of $s_{w_1,f_3}$ such that $\zeta_{w_1,f_3}$ is acquired where $\zeta_{w_1,f_3} > \zeta_{w_1,f_2}$, which ultimately saves her an acquisition of $\zeta_{w_1,f_3}$.

Then worker $w_3$ acquires match utilities. Even though her values for firms $f_1$ and $f_2$ are especially high—both $v_{w_1,f_1} = \bar{v}_{w,f}$ and $v_{w_1,f_2} = \bar{v}_{w,f}$—worker $w_3$ does not acquire match utilities with them because there is no draw of private values such that $\zeta_{w_3,f_2} > \zeta_{w_1,f_2}$ or $\zeta_{w_3,f_1} > \zeta_{w_2,f_1}$. Then even in this game with clustered private values, the *ex-post* stable matching $\{(w_1, f_2), (w_2, f_1), (w_3, f_3)\}$ is realized by acquiring the minimum number of match utilities; the only match utility acquired where the associated worker-firm pair is not in the *ex-post* stable matching is $\zeta_{w_1,f_1}$. This match utility would have been necessary to check even if worker $w_2$ or worker $w_3$ were given priority first, because for some $s_{w_1,f_1}$, both $\zeta_{w_1,f_1} > \zeta_{w_2,f_1}$ and $\zeta_{w_1,f_1} > \zeta_{w_3,f_1}$ could hold.

We have identified a class of mechanisms—the PTMR games—where the *ex-post* stable matching can be implemented in weakly dominant strategies, so long as $\delta$ is small enough. Now we show that, when $\delta$ is sufficiently *large*, noncooperative behavior in both PT and PTMR mechanisms generates equilibria where agents do not acquire match utilities (Proposition 6). In every economy, there exists a $\delta$ large enough where no acquisitions are profitable.

**Proposition 6.** *In every economy where $\delta \geq \underline{\delta}$ for all $w \in W, f \in F$ where*

$$\underline{\delta} = \lambda(c_f + r_w) + (1 - \lambda)(\bar{v}_{w,f} + \mathrm{E}[s_{w,f}]), \tag{10}$$

*both PT and PTMR games admit weakly dominant strategies where no workers acquire any match utilities.*

*Proof.* Define $\bar{\zeta}_{w,f} = \lambda(c_f + r_w) + (1 - \lambda)(\bar{v}_{w,f} + \mathrm{E}[s_{w,f}])$ which is the maximum match utility that a worker $w$ could receive from matching with firm $f$. Then no match utilities are acquired if $\delta \geq \bar{\zeta}_{w,f}$ because no match utility is greater in expectation than the expected cost $\delta$, even with probability 1 of the match being final. We now show this is the minimum $\delta$ to satisfy this for all economies. Say $\delta < \bar{\zeta}_{w,f}$, then if $w', f'$ exist such that $w' = \arg\max_{w \in W} r_w$, $f' = \arg\max_{f \in F} c_f$, and $v_{w',f'} = \bar{v}_{w,f}$, then $\zeta_{w',f'} > \delta$ in expectation. Then it would be weakly dominant for $w'$ to acquire match utility $\zeta_{w',f'}$ with $f'$. □

We remark that $\delta$ satisfying Condition (10) is not necessary for both PT and PTMR games to admit weakly dominant strategies where no workers acquire any match utilities. Indeed, consider a game where each worker $w$ has private value $\underline{v}_{w,f}$ for each firm $f$. Then for any $\epsilon > 0$, if

$$\delta \geq \lambda(c_f + r_w) + (1 - \lambda)(\underline{v}_{w,f} + \mathrm{E}[s_{w,f}]) + \epsilon,$$

then each worker immediately exiting the market is a weakly dominant strategy.

Now we discuss the range of $\delta$ values in the results for PTMR. For sufficiently small $\delta$ satisfying Condition (8), there exists an indirect mechanism (top-down PTMR) that generates the *ex-post* stable and Pareto-efficient matching in the minimum number of expected acquisitions. For sufficiently large $\delta$ satisfying Condition (10), workers have the weakly dominant strategy to immediately exit the market.

In PTMR where $\delta$ satisfies neither Condition (8) nor Condition (10), workers acquire match utilities with firms to maximize their expected utility conditional on the probability of eventually matching with that firm, and may not acquire all match utilities which may be stable. Indeed, neither of the aforementioned strategies from Proposition 3 and Proposition 6 would be weakly dominant in PTMR games where neither condition is satisfied. Any $\delta$ outside of Condition (8) would imply that at least one worker $w$ will skip acquiring at least one match utility $\zeta_{w,f}$, as the expected utility gain from acquiring match utility $\zeta_{w,f}$ with some firm $f$ is too small relative to $\delta$. Still, match $(w, f)$ could be stable so long as firm $f$ remains in the set of feasible stable matches for worker $w$. And we showed in Proposition 6 that $\delta$ satisfying Condition (6) was the minimum $\delta$ to incentivize against all acquisitions. Any other $\delta$ cannot guarantee that all workers immediately exit.

# 4 Conclusion

We study matching games where workers strategically acquire their match utilities with firms, and where acquisitions are costly and necessary to find a match. We operate in a setting of *aligned preferences*, where workers and firms split utility in fixed proportion. We construct a class of mechanisms—which we call *priority trading with multi-revelation* (PTMR) mechanisms—where workers sequentially acquire match utilities or exit the market, and whenever one worker replaces another in the set of tentative matches, the newly-replaced worker is given priority to acquire match utilities again. There exists a sufficiently small acquisition cost for workers to have the weakly dominant strategy to acquire match utilities with all firms with which they could feasibly match. These strategies guarantee that the matching is *ex-post* stable and Pareto-efficient. We then show that the PTMR mechanism that orders workers in decreasing order of their commonly-known values is the mechanism that minimizes the expected number of acquisitions.

Our findings highlight the impact of costly preference revelation on equilibrium outcomes. The implications of this research extend to job markets where centralized institutions can strategically direct interviews, and social situations where interactions must be had to determine mutual compatibility. Future research can consider heterogeneous costs to agents and alternative preference structures which lack the form of the ordinal potential.

# References


Abdulkadiroglu, A. and Sonmez, T. (1999). House Allocation with Existing Tenants. *Journal of Economic Theory*, 88(2):233–260.

Choi, A. (2022). Matching with Searching.

Eeckhout, J. (2000). On the uniqueness of stable marriage matchings. *Economics Letters*, 69(1):1–8.

Ferdowsian, A., Niederle, M., and Yariv, L. (2023). Strategic Decentralized Matching: The Effects of Information Frictions.

Fernandez, M. A., Rudov, K., and Yariv, L. (2021). Centralized Matching with Incomplete Information.

Hart, S. and Mas-Colell, A. (1989). Potential, Value, and Consistency. *Econometrica*, 57(3):589–614. Publisher: [Wiley, Econometric Society].

Lee, S. (2017). Incentive Compatibility of Large Centralized Matching Markets. *The Review of Economic Studies*, 84(1 (298)):444–463. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

Monderer, D. and Shapley, L. S. (1996). Potential Games. *Games and Economic Behavior*, 14(1):124–143.

Roth, A. E. and Sotomayor, M. A. O. (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monographs. Cambridge University Press, Cambridge.

Shapley, L. and Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37.

Voorneveld, M. and Norde, H. (1997). A Characterization of Ordinal Potential Games. *Games and Economic Behavior*, 19(2):235–242.