

# Two-Sided Time-Independent Regret for Matching Markets with Limited Interviews

Amirmahdi Mirfakhar, Xuchuang Wang, Mengfan Xu, Hedyeh Beyhaghi, Mohammad Hajiesmaili

*University of Massachusetts Amherst*

[arxiv.org/abs/2602.12224](https://arxiv.org/abs/2602.12224)

Consider a recruiting platform with two kinds of learning agents: job seekers and employers, matched over repeated hiring cycles. Each cycle, a candidate screens a small number of employers—through introductory calls or technical screens—before committing to one application; the employer then accepts its most preferred applicant or holds the position open if it remains uncertain. Neither side knows upfront how good a match will be, and both form noisy impressions over time. A platform designer faces a natural question: *how many low-cost screenings per cycle does each participant need before the platform reliably settles into stable, mutually preferred matches?*

We study this through the lens of bandit learning in two-sided matching markets, capturing two features missing from prior work. First, screenings act as *queried hints* that reveal noisy preference signals to both sides while restricting who can apply. This creates a fundamental tension: more screenings improve information but impose heavy evaluation costs on employers and, as everyone converges on the same top choices, drive congestion; fewer screenings ease that burden but leave candidates competing over a very limited pool with little information to guide them, sharply raising the risk of going unmatched. Second, unlike all prior bandit-matching work, we allow *employers to be uncertain about their own preferences*. We introduce *strategic deferral*—allowing an employer to hold a position open rather than commit prematurely—as a primitive that lets the market self-correct without explicit coordination.

Our main result is that **a constant number of screenings per cycle suffices for horizon-independent regret**, improving over the  $O(\log T)$  guarantees of all prior work under strictly weaker assumptions. In the centralized setting, we show that *two screenings per round suffice* using empirical mean estimators, resolving an open conjecture on single-agent hinted bandits along the way. In decentralized settings—where candidates observe only coarse anonymous signals such as whether a position is still open or whether an employer’s hire changed—we design both a coordinated and a coordination-free algorithm that achieve the same horizon-independence under employer uncertainty and strategic deferral, all converging to an actual stable matching rather than merely a zero-regret outcome.

This model applies more broadly to modern multi-agent platforms such as creator-brand sponsorship markets, online dating apps, freelance marketplaces, and AI-assisted mentorship or collaboration networks—wherever low-cost pre-match interactions between learning agents shape who applies and who accepts.