Learning through Transient Matching in Congested Markets*

Andrew Ferdowsian†

This version: November 1, 2022, New Version

Abstract

I introduce a framework for studying transient matching in decentralized markets where workers learn about their preferences through their experiences. Limits on the number of available positions force workers to compete over matches. Each capacity-constrained firm employs workers whose match value exceeds a threshold. Since employment offers both payoff and information benefits, workers effectively face a multi-armed bandit problem: to them each firm acts as a bandit, where the probability of "success" at the firm is driven by market competition. In such markets, I show that aggregate demand for firms satisfies the gross substitutes condition, yielding equilibrium existence. The resulting search patterns match a variety of stylized facts from labor market data: high-quality workers search less, and tenure increases with age. In general, equilibria are inefficient because competition depresses the level of search. From a market design perspective, I find that the utilization of headhunters has differential effects depending on workers' quality, conclusively improving both outcomes for low-quality workers and overall efficiency. Reducing congestion through unemployment benefits, however, can depress search and may ultimately reduce match efficiency.

Keywords: Dynamic Matching, Market Design, Congestion.

JEL: C72, C73, C78, D47, D83, J64.

*I thank Dorisz Albrecht, Robert Bell, Sylvain Chassang, Navin Kartik, Kwok Hao Lee, Lukas Mann, Daniel McGee, Xiaosheng Mu, Pietro Ortoleva, Richard Rogerson, Kirill Rudov, Evgenii Safonov, João Thereze, Tim Wang, and audiences at Princeton and Stony Brook for their helpful advice. Special thanks go to Alessandro Lizzeri, Can Urgun, and Leeat Yariv for their continued support. I gratefully acknowledge financial support from the Dietrich Economic Theory Center.

†Department of Economics, Princeton University.

1 Introduction

1.1 Overview

In many markets, participants learn about potential matches by pairing with them. In labor markets, employees often learn about a company's culture on the job; in residency markets, aspiring doctors learn about their preferred specialties through apprenticeships; and in marriage markets, individuals learn about prospective spouses through dating. These markets are often congested. Firms hire a limited number of employees; hospitals have federal funding limits capping the number of residents they can hire; and many relationships are monogamous. Congestion limits the ability of agents to learn: a worker who is not hired cannot learn her match value with a firm. How do agents strategically search in congested matching markets? When agents learn through matching, who ultimately is matched with whom? How do common interventions—hiring intermediaries or increasing unemployment benefits—change the quality of matches?

This paper develops a novel model of learning through matching when there is a limited number of positions available. I extend techniques from the operations literature on multi-armed bandits to determine workers' equilibrium search patterns, when the rewards from search are endogenous. I characterize the set of equilibria in congested markets with transient matching. In general, equilibria are inefficient due to competition depressing the level of search. In line with empirical work on tenure, in equilibrium, workers with higher than anticipated match values search less, as do older workers. I proceed by considering the impact of two common policy interventions: introducing informed intermediaries—headhunters—and increasing unemployment benefits. I show that revealing information about a firm's match values unambiguously improves equilibrium welfare, though benefits are unequally distributed among workers. When there is a commonly known top firm, increasing unemployment benefits increases competition at that firm, decreasing equilibrium welfare. In contrast, when markets are uncongested, unemployment benefits always improve equilibrium welfare.

In the model, a continuum of workers repeatedly search for jobs at a finite number of firms. Workers differ in their observable characteristics, but do not know their match value at a given firm until they are hired by that firm.[1] Each period, every worker chooses a single firm to apply to. Firms interview their set of applicants and hire the workers with the best fit, subject to capacity constraints. Hired workers learn about the quality of their match, while rejected workers only learn of their rejections.[2] Workers and firms split the surplus from matching.

[1]For instance, a young graduate of computer science would know her grades and the school she attended, but might not know that a position at Google would feature her best fit.

[2]Results in the empirical literature on search motivate the modelling choice that workers only learn upon being matched (Menzio, Telyukova, and Visschers 2016). In the appendix, I show that the qualitative results are similar

At the end of the period, workers retire with a fixed probability, exiting the market. At the beginning of the next period, a proportional mass of workers is born.

A novel technical contribution of the paper is the translation of the set of firms to a multi-armed bandit problem. I show that firms act as *endogenous* bandits. In the standard multi-armed bandit problem, a single decision-maker explores a fixed set of bandits. An important difference in this setting is that multiple workers simultaneously compete over a limited number of positions at firms. The reward from a given firm depends on the probability of hire at that firm. However, the probability of hire is endogenously determined by workers' strategies. *Hiring thresholds* suffice as a manner of describing competitive forces and the ability to learn in equilibrium. Despite the fact that workers simultaneously learn and optimize, thresholds fill a role similar to that of prices in competitive equilibrium, and are able to guarantee an equilibrium exists. While prices and thresholds fill a similar role, the two are not interchangeable. In particular, a worker's demand for a firm amounts to a lottery in the presence of hiring thresholds with incomplete preference information.

Firms act as endogenous bandits whose rewards are determined in equilibrium. Utilizing the multi-armed bandit characterization, I show that aggregate demand over firms satisfies the gross substitutes condition of Kelso and Crawford (1982). This condition enables the development of a tâtonnement algorithm, wherein thresholds converge to a fixed point equilibrium in which firms' hiring thresholds are consistent with workers' behavior. To the best of my knowledge, my model is the first to endogenously determine the rewards of experimentation through competition in a market setting. Additionally, the techniques easily extend to other markets. I show in the appendix that gradual learning, heterogeneous discounting, and flexible firm capacities can all be incorporated into the model, without affecting qualitative results, as none cause the gross substitutes condition to be violated.

In tenure data, two facts emerge. Workers' transition rates between jobs decrease with age (Menzio, Telyukova, and Visschers 2016), and higher quality workers are more likely to be satisfied with any given match (Network 2017) compared to lower quality workers. I show that both facts are a natural consequence of transient matching with incomplete preference information. Older workers have had more chances to investigate firms, and so are likely to have found a satisfactory match. Then, an older worker's incentive to test out other firms is lower than that of a young worker with far more unexplored options to choose from. Separately, when two workers share a prior, the higher quality worker will search less. Any of her matches provides more surplus than the lower quality worker's match would. If the higher quality worker is not perfectly informed, she may believe her outside options are equivalent to those of the lower quality worker. Then, the higher quality worker wishes to continue searching only if the lower

if workers learn upon applying.

quality worker also wishes to continue searching. As such, the higher quality worker spends weakly less time searching relative to the lower quality worker, and in turn the higher quality worker has higher average tenure.

I investigate the impact of commonly suggested policy interventions targeting markets with learning. One such policy is unemployment benefits, as in principle increasing these benefits decreases the cost of search, which allows workers to be more selective. However, when workers have the option of returning to a firm that previously hired them, unemployment benefits also reduce the cost of taking a risk by applying to a competitive firm. When a firm is highly ranked by all workers, increasing unemployment benefits tempers workers' incentives to investigate other firms. In equilibrium, this can lead to underutilization of these firms and decreased surplus.

Next, I characterize the impact of improving the information about a firm's match values, through holding an open house or hiring headhunters. This has two direct effects. Workers with low match values at that firm avoid the firm, and workers with high match values at that firm can target the firm. The workers with lower firm-specific match values strictly benefit from the revelation, while the effect is ambiguous for workers with higher firm-specific match values. When the market is sufficiently congested, this revelation harms higher quality workers. This occurs because competition at other firms increases due to the increase in lower quality workers.

Last, my model provides insight regarding the influential empirical literature on search (Chade, Eeckhout, and Smith 2017). Models in this literature typically assume firms can add or subtract positions at no cost: if a worker could productively match with a firm, that firm can always add an extra job.[3] In the long run, markets can be expected to adjust, so this is a reasonable assumption. However, on the shorter timescale at which workers make strategic decisions, firms may be unable to freely add and remove positions. In many scenarios, such as union jobs or during a recession, firms cannot easily reaction to increased market demand. In markets with transient matchings, like the described examples, my model shows that calibrated search models which omit the effect of competition can yield upwards biased estimates of worker match quality. To see why, note that the incentive to search is decreasing in the quality of the current match and increasing in the expected value of the outside option. When outside options become less competitive, workers must value their current positions more greatly in order to not switch. Thus, when firms can freely add positions, workers must value their current positions more highly in order to not transfer. Understanding the quality of matches present in equilibrium is critical when evaluating the efficacy of potential counterfactual policy changes. As such, systematically biased estimates of match quality can generate incorrect conclusions about the impact of new policies.

[3]For examples of search models where firms have flexible positions, see Christensen et al. (2005), Menzio and Shi (2011), and Postel-Vinay and Turon (2010).

The model focuses on markets with non-transferable utility. Despite this, the model applies more broadly to markets where wages are determined independently of match quality. When wages are fixed across realized match-values, my qualitative results hold. Many labor markets feature fixed wages. Hall and Krueger (2012) find that less than 35% of job-seekers bargain over wage; most job-seekers accept posted wages. As a result of laws or bargaining, government and union jobs have fixed wages. In France, public school teachers are allowed to apply yearly for placement in any region (Combe, Tercieux, and Terrier 2018). Wages are fixed across regions, conditional on total experience, and so the value of a match depends primarily on the teacher's fit.

I relax the assumption of non-transferable utility in Section 6. To do so, I extend the model to a competitive equilibrium setting, where equilibrium wages are strategically chosen by firms. My qualitative results extend to the transferable utility setting. I also show that resumes—the ability of a worker to prove she has been previously hired elsewhere—play an important role in information transmission. Without them, competitive equilibria may fail to exist.

My results have several implications for designing centralized mechanisms. If a mechanism designer does not control the aftermarket, agents may be free to rematch after their initial assignment. Agents that are aware of this may alter their initial applications accordingly, skewing the initial outcome of the market. In particular, when agents have incomplete information about their match values, and that incomplete information is correlated with their potential for success on the aftermarket, standard algorithms such as deferred acceptance may no longer be strategy proof. Through understanding how agents with heterogeneous incomplete information match in decentralized markets, we can better understand the impact aftermarkets have on centralized settings.

1.2 Related Literature

This paper relates to several distinct literatures: matching with incomplete information, dynamic matching, directed search, and bandits with collisions.

There is a burgeoning literature on matching with incomplete information. Immorlica et al. (2020) consider school choice where students have incomplete preference information and dynamically learn through costly inspection. They solve the mechanism design problem of generating "regret-free stable" outcomes, wherein agents never regret their search decisions. There are two key differences between their setup and mine. First, they study a centralized one-shot school choice market, with fixed participants. Second, the cost of inspecting a school in their model is fixed and exogenous. Doval (2022), Liu et al. (2014), and Liu (2020) study stable outcomes in markets with incomplete information. Chen and Hu 2020 provide a dynamic justification for stability with incomplete information, in which firms evaluate potential employees according to their "worst" possible match values. These works focus on one-shot matching markets, in which no new participants enter the market after the game begins. In many dynamic environments, including the examples from the previous section, this fails to be the case.[4]

Similarly, recent work in the matching literature has begun to incorporate dynamics.[5] Akbarpour, Li, and Gharan (2020) consider dynamic markets with networked agents, and solve the designer's problem of choosing which agents to match. Anderson and Smith (2010) examine matchings where agents form reputations regarding their quality over time and show that positive assortative matching emerges over time. Ferdowsian, Niederle, and Yariv (2022) are at the intersection of decentralized dynamic matching and matching with incomplete information, and study the hurdles to stability that arise, even in one-shot markets. The paper shows that stringent assumptions are required to ensure stable matchings are equilibrium outcomes in markets with incomplete information. Kadam and Kotowski (2018) also consider markets with transience, and treat the problem from a more classical view of stability. They find conditions under which dynamic stability can be generated in a setting where a centralized authority may be necessary for finding the stable matching. I place more structure on agent's preferences, which enables me to study the related problem in a decentralized environment.

The directed search literature has studied labor markets where workers intentionally target firms.[6] Within the search literature, this paper connects two strains, search with marriage matching and search with frictions. Dagsvik, Jovanovic, and Shepard (1985), Jovanovic (1979), and Miller (1984) consider variations of a single-agent version of my model.

Technically, this setting is reminiscent of the multi-armed bandit setting. The solution to the standard multi-armed bandit problem was found by Whittle (1980). Weitzman (1979) studies the mathematical problem where a decision-maker chooses when to stop testing alternatives. Several papers on multi-armed bandits with collisions have recently emerged in the computer science literature, (see Liu, Mania, and Jordan 2020 and Liu et al. 2021). These papers assume that agents know how they are ranked by the other side of the market. In many practical situations, agents face uncertainty regarding their acceptance prospects not only because they are unaware of their competition, but also because they do not know how they will be ranked. Another key difference between this paper and the literature on bandits with collisions, is that extant algorithms fail to satisfy incentive compatibility when firms disagree on the rankings of workers.

[4] There are also several examples of centralized markets with incomplete information. For instance, Fernandez, Rudov, and Yariv (2022) show that standard predictions of centralized markets are not robust to perturbations of information, while Li, Wang, and Zhong (2016) experimentally tests predictions of truth telling.

[5] For a brief survey of the literature on dynamic matching, see Baccara and Yariv (2021).

[6] Chade, Eeckhout, and Smith (2017) provide a useful survey of the extensive literature studying directed search.

2 The Model

I begin by describing the agents involved. The market matches a continuum of workers with a finite set of firms, $|\mathcal{F}| = F$. Workers are categorized into *classes*, $c \in \mathcal{C}$. A worker knows her class. For instance, a worker knows her grades, or the school she graduated from. If a worker had complete information, she would also know her "type," $\theta \in c$, where $\theta = ((\theta_w^1, \theta_f^1), \ldots, (\theta_w^F, \theta_f^F)) \in [0, K]^{2F}$ is her bounded match value with each firm. Instead, workers know the distribution for each class over types. When there are a finite number of types, $m_c(\theta)$ denotes the *total* mass of class-c type-θ workers, otherwise, $m_c(\theta)$ denotes class c's density of θ. The total mass of class c, $m_c \equiv \int_\theta m_c(\theta) d\theta$, is the summation of mass across all types of class c. I normalize the total mass of workers to 1: $\int_c m_c dc = 1$. Each firm j has a hiring capacity of $m(j) > 0$, limiting the mass of workers it can hire each period. A *market* is a triplet $\mathcal{M} = (\mathcal{F}, \mathcal{C}, m)$. I assume that a worker's class already encapsulates any correlation in match values across firms. That is, learning i's match value at firm j, $\theta^j = (\theta_w^j, \theta_f^j)$, is uninformative regarding her match value at firm j', conditional on worker i's class. This can be thought of as workers learning about their fit at individual firms, as opposed to learning about their preferences for certain fields. Formally:

$$m_c(\theta) = \prod_{j \in \mathcal{F}} m_c(\theta_w^j, \theta_f^j)$$

The matching process is straightforward. Each period, every worker chooses a single firm to apply to. The choice of worker i to apply to firm j in period t is denoted by $a_i(t) = j$. Every firm observes his list of applicants, $a_j(t) = \{i | a_i(t) = j\}$, and learns his match value with each applicant, θ_f^j.[7] $A_i(t) \in \mathcal{F} \cup \varnothing$ denotes worker i's match in period t, if $A_i(t) = \varnothing$, i is *rejected*. Should a firm j hire a worker, that worker is *accepted*. A type-θ worker accepted by firm j receives θ_w^j, while the firm receives θ_f^j.[8] A rejected worker receives $\theta_w^\varnothing = 0$.[9] The set of workers hired by firm j in period t, is $A_j(t)$. Firm j's period-t profit is:

$$\pi_j(t) = \int_{\theta \in A_j(t)} \theta_f^j d\theta$$

Firms discount profits at a rate of δ. Firm j aim to maximize the sum of discounted profits,

[7] The application process can be thought of as containing an interview stage that informs firms. Interview frictions have been previously discussed in the matching literature (Lee and Schwarz 2017). To refine the focus on transient matchings in this paper, I abstract from these frictions in the interview process.

[8] The main body of the paper focuses on non-transferable utility. As discussed in the introduction, several recent empirical results support this assumption. For instance, Becker (2011) shows that as much as 40% of benefits from employment are non-wage based, implying that proper matching is critical. In Section 6, I extend the model to a transferable utility setting where firms choose match-value dependent wages. I show that the key results of the previous sections carry over to the transferable utility case.

[9] In Section 5.2, I consider the impact of adding unemployment benefits to the market, modelled by $\theta_w^\varnothing > 0$.

π_j, subject to their capacity constraints in each period:

$$\pi_j = \sum_{t=0}^{\infty} \delta_j^t \pi_j(t).$$

Firms cannot distinguish between workers with equal firm match values. For instance, if type-θ workers and type-θ' workers both apply to firm j in a given period, and $\theta_f^j = \theta_f'^j$, j must hire workers of each type with equal probability.

I depart from the standard setting of exponentially discounted utility. Instead, workers retire with probability $1-\delta$ at the end of each period, exiting the game. When a type-θ worker retires, a new type-θ worker enters the market. Critically, the new worker will no longer have the information that her predecessor acquired through play. In total, a mass of $(1-\delta)m(\theta)$ type-θ workers are born each period. This alteration to the standard interpretation of discounting has two implications. First, a worker's expected utility remains equivalent to the standard interpretation. Second, the distribution of worker-type masses remains unchanged over time, generating a stationary environment.

When a worker enters the market, she is unaware of her exact type, but knows her class. Upon acceptance by firm j, a type-θ worker learns θ_w^j.[10] Rejected workers learn only of their rejection. There is no public history, rather workers can only learn through applying to firms. Workers aim to maximize their lifetime expected utility. That is, worker i of class c, chooses how to apply to firms, to maximize:

$$U_i^c = \mathbb{E}_{m_c}\left[\sum_{t=0}^{\infty} \delta^t \theta_w^{A_i(t)}\right] \tag{1}$$

To eliminate equilibria that depend upon coordination, or time-specific applications, I refine the set of equilibria to those supported by Markovian strategies. The state variable for a worker is the tuple of her payoff relevant variables: her class, and her posterior over her type. For a given strategy profile, firms face no payoff relevant dynamic uncertainty, so their payoff relevant variable is the types of applicants.

Definition (Markovian Strategies)**.** *A worker's strategy is Markovian, if in each period, her application is only a function of her class and the posterior over her type. A firm's strategy is Markovian, if in each period, the set of workers hired is a function of the firm's mass distribution over applicant match values.*

[10] Other models of learning on the job, such as Jovanovic (1979), use more gradual learning processes, where a worker receives a noisy signal of the true match value. Since the implications generated by gradual learning have already been discussed in the search literature, I instead assume learning is immediate, to focus on the importance of competition. In Appendix section B.1, I show that the core results of this paper still hold when learning is gradual.

The focus on equilibria in Markovian strategies rules out equilibria that depend on the calendar period. A strategy profile is Markovian if in that profile all agents use Markovian strategies. Similarly, an equilibrium is Markovian if the associated strategy profile is Markovian. It should be emphasized that this does not restrict the set of possible deviations when in the steady state. A Markovian strategy profile is an equilibrium only if no agent has a profitable deviation in the steady state. I assume that firms cannot commit to a hiring policy. A firm's strategy is part of an equilibrium strategy profile, only if in the steady state of that equilibrium, the firm does not wish to deviate in any period. Where appropriate, I drop time indices. Last, to deal with a trivial source of non-uniqueness, this paper will consider the matching outcomes that result in the "long run" for a Markovian strategy profile that constitutes a Perfect Bayesian Equilibrium.

Definition (Outcomes)**.** *An outcome for market $\mathcal{M}$, of a Markovian strategy profile σ, is a distribution over worker types, information sets, and applications to firms in some steady state of $\mathcal{M}$ where agents use strategy profile σ.*[11]

An outcome is the payoff relevant information for a snapshot of the market. Namely, outcomes include the proportion of information each worker has learned, as well as their choice of firms to apply to. Restricting attention to the outcomes of Markovian strategy profiles removes another trivial source of non-uniqueness. Should a worker have two strategies which differ conditional on information regarding a firm j that she never applies to under either strategy, the outcomes of the two strategies are equivalent.

3 Incomplete Information Generates Transience

3.1 Complete Information Benchmark

I begin by developing a benchmark for the case where workers are fully informed and firms and workers agree on the value of each match, i.e., preferences are *aligned*: $\forall c, \theta \in c, j \ \theta_w^j = \theta_f^j = \theta^j$. The complete information setting exemplifies the interplay between transience and information. I show that even when matches are not permanent, an outside observer may find a complete information market with transient matchings indistinguishable from a decentralized market with permanent matches. I detail the construction of equilibria, as the complete information case proves similar but distinct from the benchmark of long-lived workers with incomplete preference information.

[11] I show in Section 4.1 that every Markovian strategy profile generates a steady state, and therefore this notation is well-defined.

Definition (Complete Information Market)**.** *A market $\mathcal{M}_I$ is a complete information market if every worker class c contains a single type.*

The unique equilibrium for a complete information market, $\mathcal{M}_I$, can be computed in a simple recursive manner. First, select the firm-worker pair, j^*, c^*, with maximal match value. In any equilibrium, there must be a positive mass of class-c^* workers applying to j^*. Otherwise, class-c^* workers would benefit by deviating and applying to j^* where they are guaranteed to be hired and receive higher match values. Indeed, at least $\min\{m(c^*), m(j^*)\}$ class-c^* workers must apply to j^*. Then, either all class-c^* workers are matched, or firm j^*'s capacity is exhausted. In each step, the set of agents in the market is strictly reduced. When there are a finite number of worker classes, this procedure terminates in finite time. The "Top-Down" algorithm, found in the appendix, formalizes this procedure.

Lemma 1 (Complete Information Equilibrium)**.** *The Top-Down algorithm characterizes the unique equilibrium outcome of $\mathcal{M}_I$ when preferences are aligned.*

Decentralization inherently generates two types of inefficiencies, even in the complete information case. Workers that are qualified for top-tier positions over compete for those positions, generating significant opportunity costs, a negative externality on their type. Furthermore, workers do not consider their impact on other types of workers. For instance, a worker with two high match values would take the better of the two, even if it caused a large loss in utility for another worker type with a strong preference for one firm.

3.2 Long-Lived Workers with Incomplete Information

When agents have incomplete preference information, and have lower probabilities of exit, the equilibrium match converges to that of the complete information environment. The inefficiencies present in the complete preference information context carry over to the case where agents are long-lived, but have incomplete preference information. A new source of inefficiency also emerges, acquisition of information. When workers do not know their type, they must also invest time to determine their top possible match. In the limit, as δ goes to 1, this source of inefficiency collapses to 0. However, the inefficiency always remains strictly positive in certain markets. Relatedly, even as δ goes to 1 in the limit, certain worker types would strictly benefit from learning their match values, i.e., there exists path dependence despite vanishing time frictions. This arises due to the possibility that an unlucky hiring realization can leave a worker uninformed regarding her type. Example 1 illustrates this form of path dependence.

Example 1 (Path-Dependence)**.** *Rows represent worker types, while columns represent firms. The mass of each agent is given in the first element of the corresponding row or column. The match values for a worker and firm are given at the intersection of their row and column.*

Even when workers are long-lived, path dependence emerges in equilibrium.

Firms / *Workers*	$m(f_r) = 1/4$	$m(f_s) = 1$
$m(\theta_h) = 1/2$	$(3, 3)$	$(2, 2)$
$m(\theta_l) = 1/2$	$(1/2, 1/2)$	$(2, 2)$

For now, without justification, I focus on strategy profiles where f_r hires as many θ_h types as possible in each period. In Section 4.1, I show that in any equilibrium this will be the case. That is, despite the potential for long-term considerations, in any market and any equilibrium, firms always hire their most qualified applicants in each period.

I proceed by restricting attention to workers' strategies. The set of possible equilibrium strategy profiles can be categorized by the (possibly stochastic) number of failed attempts after which workers cease applying to f_r and instead apply to f_s forever. Naturally, the equilibrium strategy profile depends upon δ. When δ is low, workers are unwilling to invest into learning, and all apply to f_s. However, even when δ is high, no equilibrium involves all θ_h's learning their type through repeated application to f_r. To see why, consider the strategy profile where all workers apply to f_r repeatedly. Then, if all informed type-θ_h workers continued applying to f_r, each would be hired with probability 1/2. Equilibrium forces would require informed type-θ_hs to randomize between f_r and f_s such that informed type-θ_h's are indifferent between the two firms. However, this would imply that uninformed workers expect a match value less than 2 from applying to f_r, and so uninformed workers would gain nothing from learning their type. Uninformed workers would then deviate and apply to f_s. Therefore, all workers applying to f_r indefinitely could not have been an equilibrium.

When δ is sufficiently close to one, all workers will initially apply to f_r. Furthermore, since workers are not informed of their type in the event that their application fails, a portion of the workers that fail their initial application, reapply to f_r. The more workers that reapply, the lower an uninformed worker's expected payoff from f_r will be, due to the competition θ_h types face. Then, the reapplication rate will be such that workers with a single failed application are indifferent between reapplying and applying to f_s.

In any equilibrium for $\delta < 1$, uninformed workers with two failed applications to f_r never apply to f_r. To see why, suppose it was optimal for such a worker to do so. Since an uninformed worker with two failed applications places a lower weight on their probability of being type θ_h than an uninformed worker with one failed application, it must have been strictly optimal for uninformed workers with one failed application to reapply to f_r for all workers. As such, for δ

sufficiently close to 1, almost all workers will have applied at least twice to f_r. Then, because each w_h worker is hired with a minimum probability of 1/2, a minimum of 3/4 of type-θ_h workers will have been hired at least once and therefore know their type.

Informed type-θ_h workers must continue applying to f_r, otherwise taking the risk of applying to f_r to learn their type was not profitable. Then, the average rate of hire for type-θ_h workers at f_r will be below 2/3, implying they would be better off applying to f_s. This generates a contradiction, therefore all uninformed workers with two failed applications apply to f_s, completing the characterization of equilibrium for Example 1.

Notably, as $\delta \to 1$, the fact that the proportion of uninformed agents applying to f_r does not increase. However, as $\delta \to 1$, less uninformed agents are present each period, and so the mass of agents applying to f_s must increase in turn. Then, as $\delta \to 1$, the equilibrium outcome converges to that of the market with complete information.

The implications of Example 1 generalize, the equilibrium outcomes of aligned markets converge to the corresponding complete information outcomes as $\delta \to 1$. As δ increases, the opportunity cost of applying to firms with uncertain match values decrease. However, when competition is present, the gain from learning one's type also decreases as more workers attempt to learn their type. As $\delta \to 1$, workers who know they are in the class with the maximal match value must attempt to take advantage of that match value. Because any worker requires at most "finite" time to learn her type, it can be shown that on a type by type basis, as $\delta \to 1$, outcomes converge to the corresponding outcomes in the complete information case.

Lemma 2 (Long-Lived Outcomes)**.** *Let there be a finite number of worker types, and let preferences be aligned.*

For each type, as $\delta \to 1$, the equilibrium probability they apply to any given firm converges in distribution to the probability they apply to that firm in the corresponding complete information market.[12]

Despite the fact that the outcomes converge, there is a marked difference in the information present in the two economies. As shown in Example 1, for any $\delta < 1$, a minimum of 1/4 type-θ_h workers will never learn their type in equilibrium. Furthermore, in equilibrium, had a worker been informed of her type upon arriving, she would receive strictly higher expected utility than otherwise. It can be shown that the case of incomplete information is directly related to competition in the complete information setting. Had the previously described Top-Down algorithm ever entered a case of "over-subscription" then there must be incomplete information for all $\delta < 1$. To see why, note that if all workers knew their type, then all would apply to

[12]Formally, for any $\epsilon > 0$, there exists $\underline{\delta} < 1$ such that for any $\delta > \underline{\delta}$, any type θ, and any firm $j \in \mathcal{F}$, the equilibrium probability a type-θ worker applies to j in any period of the equilibrium with δ is within ϵ of the equilibrium probability that a type-θ worker applies to j in the corresponding complete information market.

their top choice until the utility from that firm was decreased to that of the next alternative. However, in such an equilibrium where there was no benefit to realizing a specific type, incomers would have no incentive to learn their type, because trying and failing yields a direct loss (not being hired in that given period) which is non-zero for any $\delta < 1$. By comparison, succeeding fails to change their expected utility, due to the indifference condition. As such, there must be path dependence in equilibrium for any market with over-subscription.

Lemma 3 (Path-Dependence)**.** *In any market with over-subscription, there exists path dependence in equilibrium for any* $\delta < 1$*.*

Then, despite the fact that market outcomes will look similar in complete information markets and markets with long-lived workers, the search patterns of workers will vary drastically.

4 Equilibrium Characterization

When workers have incomplete information and are not long-lived, they face a non-trivial trade-off. Each period, workers must decide between exploiting their information about their type to apply to firms with high expected match values and exploring new firms to learn more about their type. When determining whether to explore, workers must take into account other workers' application decisions, which generate competition over firms. This brings up a related point. Previously, it was taken for granted that a firm would always hire the most qualified applicants. In the following section, I show that firms have a profitable deviation if they do not hire their most qualified applicants.

First, Lemma 4 shows that any Markovian strategy profile generates a unique outcome.

Lemma 4 (Unique Steady State)**.** *Any Markovian strategy profile has a unique steady state.*

4.1 Hiring Thresholds

In Example 1, it was assumed that a firm would always hire its most qualified applicants in every period, without regard for how doing so would impact the future distribution over worker information. In this section, I show that firms' strategies can always be described as threshold strategies.

The optimality of threshold strategies is a direct consequence of the assumption that firms cannot commit to hiring policies at the outset. If a firm could commit to a hiring policy, then that firm could encourage certain classes of workers to apply through affirmative action policies. This would enable the firm to attract highly qualified workers of types within those classes through the promise of guaranteed jobs. However, when the firm cannot commit, he is incentivized to "cheat" and accept more highly qualified workers. Doing so allows the firm to reap the benefit from hiring the qualified worker immediately, instead of delaying. The key insight is that a firm which rejects a worker with a high match value can only do so to incentivize that worker to apply again in a future period. However, doing so simply shifts back the expected gains from matching with that worker, and so the firm would prefer to accept her in the current period.

Proposition 1 (Firms Use Hiring Thresholds)**.** *Suppose preferences are aligned,* $\forall\ \theta, j:\ \theta_w^j = \theta_f^j$*. In the steady state of any equilibrium, if firm* j *hires a worker with a match value of* θ_f^j*, then firm* j *also hires all applicants with match values above* θ_f^j*. Additionally, if firm* j *hires below its capacity,* $m(A_j) < m(j)$*, then* j *hires all applicants.*

A worker's expected utility from a firm inherently depends upon that firm's hiring decision, which in turn is a function of the application strategies of other workers. There is a natural ordering on worker types at j, θ_f^j. Proposition 1 implies that the firms' hiring decision can be summarized by the minimal hired θ_f^j and the probability with which type-θ workers are hired. The firm's hiring process is important to a worker only in how it affects that worker's hiring probabilities. This motivates a natural method of summarizing a firm's strategy, the threshold in match values at which it begins firing workers.

Definition (Hiring Threshold)**.** *A hiring threshold for firm* j *is a tuple* (v_j, p_j)*, which consists of a match value and a hiring probability.*

Crucially, in the steady state of an equilibrium, thresholds are time invariant. Because firm j uses a threshold strategy, there is a hired worker with minimal match value. To determine firm j's threshold, begin by finding the set of workers that apply to j within a single period. Let θ be that worker's type. Set v_j to that worker's θ_f^j, and p_j to type-θ's probability of hire. If the mass of applicants received by j is below its capacity, $m(a_j) \leq m(j)$, define its threshold as $(0, 1)$, all applicants to j are hired.

Thresholds suffice to describe firm behavior in equilibrium. Workers can use v_j and p_j to compute their expected utility from applying to firm j. If, for a worker type θ, θ^j is above v_j, then type-θ workers are always hired by j. If θ^j is below v_j, type-θ workers are never hired by j. Last, if $\theta^j = v_j$, then type-θ workers are hired by j with probability p_j.

It will be convenient to rank thresholds in the lexicographic order. A threshold, (v_j, p_j), is greater than another threshold, (v'_j, p'_j), if $v_j > v'_j$ or $v_j = v'_j$ and $p_j \geq p'_j$. $(v, p) = \{(v_j, p_j)\}_{j \in \mathcal{F}}$ will refer to the vector of thresholds.

4.2 Equilibrium Worker Strategies

Next, I discuss the characterization of a worker's equilibrium strategy. As previously mentioned, workers tradeoff exploring firms where their match value is still unknown, and exploiting firms that are known to be profitable matches. I discuss a technique from the Multi-Armed Bandit literature, which provides an alternate solution method that holds for any market where all firms hire all applicants. This section then extends the technique to settings where the hiring probabilities are endogenous.

Let h_t^i denote the private match value history for a worker i of age t, $h_t^i = (\theta_w^{a_i(1)}, \theta_w^{a_i(2)}, \ldots, \theta_w^{a_i(t-1)})$.

Definition (Gittins Index)**.** *The Gittins index for worker i at firm j given history h_t^i, $GI_i(j, h_t)$, is the solution to the following optimal stopping problem, where τ can depend upon the realization of match values:*

$$GI_i(j, h_t) = \max_{\tau} \frac{\mathbb{E}_{i,h_t}\left[\sum_{t=1}^{\tau} \delta^t \theta_w^j\right]}{\mathbb{E}_{i,h_t}\left[\sum_{t=1}^{\tau} \delta^t\right]}$$

Gittins indices provide a simple characterization of the benefits from learning. Intuitively, the Gittins index of firm j for a class-c worker is her expected match value at j, weighted by her ability to strategically reapply to j or cease applying to j conditional on the realized match value. That is, if the worker discovers a high match value at j she can take advantage of the positive realization and reapply to j instead of leaving immediately as she would after discovering a low match value at j. The ability to strategically reapply biases the Gittins index upward from the standard expected value, and implicitly incorporates a benefit for learning. When describing the Gittins index of a newly arrived worker, I omit the trivial match history to conserve space.

Determining the optimal worker policy can be done tractably, thanks to a result from the operations literature. Rather than worrying about the timing of learning and equilibrium effects, simply compare the equilibrium Gittins indices associated with each firm. Naturally, these Gittins indices are endogenous objects, dependent on the proportion of workers applying to each firm.

As discussed previously, declaring firm j's threshold allows for the computation of each worker's initial Gittins index at j. What stopping strategy τ determines j's Gittins index for a worker who has not yet visited j? It can be shown that the following heuristic solves for $GI(j)$. Choose an arbitrary value x and apply to j. If the resulting match value is above x, stay with j forever, $\tau = \infty$. If the match value is below x, leave j immediately, $\tau = 1$. This procedure determines a value for the stopping problem if x is equal to the stopping problem. The maximal value of x determines the Gittins index. Lemma 5 formalizes this heuristic computation.

Let $\psi_\theta(v, p)$ be a type-θ worker's match values given thresholds (v, p). That is:

$$\psi_\theta^j(v,p) \equiv \begin{cases} \theta_w^j & \theta_f^j > v_j \\ p_j \theta_w^j & \theta_f^j = v_j \\ 0 & \theta_f^j < v_j \end{cases}$$

Lemma 5 (Gittins Index Value)**.** *$GI_c(j)$ for a class-c worker that has not yet applied to j is characterized by the fixed point solution to the following functional equation:*

$$GI_c(j) = \frac{\mathbb{P}\left[\psi_w^j < GI_c(j)\right]\mathbb{E}\left[\psi_w^j|\psi_w^j < GI_c(j)\right] + \frac{1}{1-\delta}\mathbb{P}\left[\psi_w^j \geq GI_c(j)\right]\mathbb{E}\left[\psi_w^j|\psi_w^j > GI_c(j)\right]}{\mathbb{P}\left[\psi_w^j < GI_c(j)\right] + \frac{1}{1-\delta}\mathbb{P}\left[\psi_w^j > GI_c(j)\right]}$$

In Example 1, determining the value of $GI_i(f_s)$ is trivial, because each worker knows her match value at f_s, for any choice of τ, $GI_i(f_s) = 2$. Computing $GI_i(f_r)$ requires slightly more work. First, consider the case where $m(f_r) = 1$, that is, f_r is not capacity constrained. If a worker applies to f_r and receives a match value of $1/2$, she immediately knows her type is θ_l. Then, applying to f_r a second time, increases both the numerator of $GI_i(f_r)$ by $\delta 1/2$ and the denominator by δ. If $GI_i(f_r) > 1/2$, then it can be shown that the solution of the stopping time problem induced by $GI_i(f_r)$ never involves applying a second time to f_r after observing $\theta_w^{f_r} = 1/2$. If the worker had applied to f_r and received a match value of 3, she would know that her type was θ_h. An extension of the previous logic shows that if $GI_i(f_r) < 3$, the solution to the stopping time problem involves applying to f_r a second time upon observing $\theta_w^{f_r} = 3$. Putting everything together, the solution to $GI_i(f_r)$ is found by infinitely applying to f_r if type θ_h and applying only once if type θ_l, yielding:

$$GI_i(f_r) = \frac{1/2\left(\frac{3}{1-\delta} + 1/2\right)}{1/2\left(\frac{1}{1-\delta} + 1\right)}$$

The intuitive solution to the Gittins index stopping problem is later generalized in Lemma 5, where I characterize the solution for any arbitrary distribution of match values at a firm. Notably, when computing $GI_i(f_r)$, the distribution over possible match values at f_s was not utilized. Even when the number of firms, F, increases, the number of computations to determine all $GI_i(j)$ also increases linearly in F, one additional step for each additional firm.

Standard results in the operations literature show under certain conditions, determining the optimal policy takes two steps: in each period compute the Gittins index for each firm, then apply to the firm with the highest Gittins index (Whittle 1980). Previous results involving directed search settings or Gittins indices have required that only one decision maker is present or that the decisions made by each decision maker are independent of one another.[13]

4.3 Alignment Implies Uniqueness

In this section, I show that when firms and workers agree on the quality of each match, i.e., preferences are *aligned*, there exists a unique equilibrium outcome under a mild condition on payoffs. To do so, I first develop the following example, which illustrates the specificity in payoffs required to generate non-unique equilibrium outcomes.

Example 2 (Non-Unique Outcomes)**.** *Preferences are aligned. There are two worker classes, $\mathcal{C} = \{1, 2\}$, $m_1 = 1/2$, $m_2 = 1/2$, $F = 2$, and $\delta = 0$. When workers are guaranteed to be hired, and indifferent over firms, there exists a multiplicity of equilibria.*

Workers \ *Firms*	$m(f_r) = 1/2$	$m(f_s) = 1/2$
$m_1(\theta_h) = 1/4$	$(4, 4)$	$(3, 3)$
$m_1(\theta_l) = 1/4$	$(2, 2)$	$(3, 3)$

Workers \ *Firms*	$m(f_r) = 1/2$	$m(f_s) = 1/2$
$m_2(\theta) = 1/2$	$(1, 1)$	$(1, 1)$

Class-1 workers are indifferent between the two firms, and furthermore are never fired in equilibrium. Therefore, incoming class-1 workers are indifferent between the two firms in any equilibrium. There exists a continuum of equilibrium outcomes characterized by the proportion of class 1 that applies to f_r.

The non-uniqueness of equilibrium outcomes cannot be trivially fixed by requiring each firm's Gittins index to be unique in the absence of firm hiring capacities. To see why, return to Example 2 and suppose class 1 contained several additional types, such that match values of $1/2$ at f_r and $1/4$ at f_s were possible. In the absence of firm hiring constraints, class 1's Gittins index at f_r, $GI_1(f_r)$, would be higher than $GI_1(f_s)$. In practice, this difference doesn't factor into class-1

[13]For examples in economics, see Papageorgiou 2018, Rothschild 1974, Urgun 2021, and Weitzman 1979.

workers' decision-making, because in any equilibrium, workers with match values below 1 are never hired.

Furthermore, this construction did not hinge on the absence of learning. Consider the same market, with some $\delta > 0$. A simple computation can be used to evaluate $GI_1(f_r) = \frac{1/2\frac{4}{1-\delta}+1/2\cdot 2}{1/2\frac{1}{1-\delta}+1/2\cdot 1}$ while $GI_1(f_s) = 3$. For $\delta > 0$, $GI_1(f_r) > GI_1(f_s)$ due to the option value of switching to j_2 if a low utility is realized. However, a small adjustment to θ_h^r can set $GI_1(f_r) = GI_2(f_s)$. Replacing $\theta_h^r = 4$ with $\theta_h^r = 4 - \delta$ ensures that the two Gittins indices are equal once more. In the resulting market, there again exists a continuum of equilibrium outcomes. As such, a more stringent condition is needed to rule out this trivial form of non-uniqueness.

Strict dynamic preferences requires workers to not face indifference between two firms when each firm hires deterministically. This assumption depends on δ, but only rules out a finite number of parameters from an infinite set. A non-generic set of knife-edge cases are eliminated through the strict dynamic preferences assumption. For instance, the parameter set which generated multiple equilibrium outcomes in Example 2 is eliminated by strict dynamic preferences, preventing "higher-ranked" workers from mixing at the top. Phrased intuitively, the equilibrium Gittins indices at any two non-competitive firms cannot be equal. It should be emphasized that this assumption is exogenous, the assumption only depends on utilities and masses.

Assumption 1 (Strict Dynamic Preferences)**.** *Let every type of worker class c have probability of hire 0 or 1 at firm j and j'. The strict dynamic preferences assumption holds if for every class c there does not exist a type $\theta \in i$ such that the following hold:*

- *θ has probability 1 of hire at j*
- *$GI_c(j) = GI_c(j')$, $\theta^j = GI_c(j')$, or $\theta^j = \theta^{j'}$*

First, no strict dynamic preferences implies that two strategy profiles, which induce different equilibrium outcomes, must also generate different hiring thresholds. The claim is formally proved in the appendix. Given this, Proposition 2 shows that equilibrium must necessarily be unique in this environment.

Proposition 2 (Uniqueness)**.** *When preferences are aligned and satisfy strict dynamic preferences, if σ and σ' are equilibrium strategy profiles, then equilibrium outcomes are the same under σ and σ'.*

To provide intuition, suppose σ and σ' are distinct equilibrium outcomes. Then, the thresholds induced by σ and σ' must be distinct. Without loss of generality, let firm j have a higher threshold under σ than σ', $(v_f(\sigma), p_f(\sigma)) > (v_f(\sigma'), p_f(\sigma'))$. Then, f must have more high quality applicants under σ relative to σ', otherwise it could not have a higher threshold. Those applicants must come from another firm f' under σ'. For those applicants to switch to f, f' must be less attractive under σ. However, when preferences are aligned, thresholds imply that the quality of workers hired is monotonic, $(v_{f'}(\sigma), p_{f'}(\sigma)) > (v_{f'}(\sigma'), p_{f'}(\sigma'))$. Iterating this logic implies that there exists a set of firms with higher thresholds. However, since all involved firms have binding capacity constraints, the same total mass of workers must be hired. Furthermore, all hiring thresholds are higher implies that total worker utility is also higher. At least one class, c, must be strictly better off under σ. Since hiring thresholds were lower under σ', c workers can profitably deviate to their strategy under σ. Then, σ' could not have been an equilibrium outcome.

The uniqueness of equilibrium transient matchings is useful for the prospective econometrician. When estimating structural parameters, the statistician does not need to worry about the problem of equilibrium selection. Instead, the unique equilibrium can be determined and exploited.

4.4 Gross Substitutes in Equilibrium

In this section, I proceed by describing equilibrium strategies, utilizing firm thresholds to characterize equilibria. Importantly, because workers evaluate firms as if they were bandits, their demand for firms satisfies

First, Lemma 9, can be extended from settings where all firms have unbounded capacities to equilibrium settings where firms can only hire a limited subset of workers.

Lemma 6 (Worker Strategy)**.** *For any equilibrium strategy profile σ, for any private history h_t^i, each worker i applies to $j \in \arg\max_j GI_i^\sigma(j, h_t^i)$.*

Lemma 6 shows that the optimal worker strategy for a given set of thresholds can be simply determined. For each worker i, determine each firm's Gittins index, potentially updating the probability of each match value based on h_t^i. Each worker i applies to a firm with maximal Gittins index, $j \in \arg\max_j GI_i(j)$. This application strategy prescribes a distribution over learning paths for each worker type, pinning down the mass of informed worker types and therefore the proportion of each type applying to each firm. Type θ's demand, $D^\theta(v, p)$, is defined as the vector of θ's hiring rates across firms, when the vector of thresholds is (v, p). Similarly, *aggregate demand*, $D(v, p)$ is defined as the integral over all types' demands.

Definition (Aggregate Demand)**.**

$$D(v,p) = \left(\int_\theta D^\theta_{j_1}(v,p) d\theta, \int_\theta D^\theta_{j_2}(v,p) d\theta, \dots, \int_\theta D^\theta_{j_F}(v,p) d\theta \right)$$

The characterization of optimal worker strategies as index based has an additional feature. When a firm increases its threshold, demand at other firms can only increase. To see why, note that workers apply to the firm with the maximal Gittins index. Increasing a firm's threshold can only decrease that firm's Gittins index for all types. The increase in threshold will not change other firms' demands unless it causes the Gittins index of one type at a firm to exceed their Gittins index at another firm. Doing so, will shift that type's demand to the other firm. It can be shown that the effect on that type's demand in future periods does not exceed the immediate effect. More generally, increasing the thresholds of a subset of firm cannot decrease the demand of firms outside that subset. Namely, aggregate demand satisfies the *gross substitutes* condition, used in Kelso and Crawford (1982) and Gul and Stacchetti (2000), to prove existence of equilibria.

Definition (Gross Substitutes)**.** *Let $\boldsymbol{F}$ be a set of firms, $\boldsymbol{F} \subset \mathcal{F}$, and $(v, p), (v', p')$ be two thresholds, such that $(v_j, p_j) = (v'_j, p'_j)$ for $j \in \mathbf{F}$ and $(v_j, p_j) \geq (v'_j, p'_j)$ otherwise. Gross substitutes is satisfied if demand for $\boldsymbol{F}$ is greater under (v_j, p_j) than (v'_j, p'_j):*

$$\forall j \in \boldsymbol{F} : D_j(v, p) \geq D_j(v', p')$$

Proposition 3 (Aggregate Demand)**.** *Aggregate demand satisfies gross substitutes.*

Given Proposition 3 the equilibrium can be found through a tâtonnement procedure. The *tâtonnement procedure* follows five steps.[14] 1) Treat every firm as if it had unbounded capacity, i.e., $(v_j, p_j) = (0, 1)$. 2) For every worker class and private history, compute the Gittins index of each firm. Using these Gittins indices, determine individual demand and therefore aggregate demand for each firm. 3) Arbitrarily select a firm j such that j's aggregate demand exceeds its capacity, $D_j(v, p) > m(j)$. 4) Decrease the rate of hire for the lowest match value at firm j continuously, $p_j \to 0$. If at any point, the decrease in p_j causes any worker's Gittins index at j to equal another firm's Gittins index, repeat step 2 for that worker's class—compute their application policy and corresponding demand, updating aggregate demand accordingly. If p_j reaches zero, repeat the process with the next lowest worker type. Stop when $D_j \leq m(j)$. Importantly, Proposition 3 implies that this step never causes a decrease in demand for other firms.[15] 5) Return to step 3 if there exists another firm j' such that $D_{j'} > m(j')$.[16] The proof

[14] The algorithm is formally defined in the appendix.

[15] This step requires a technical adjustment when worker types have positive mass to avoid cycles, see the appendix for the formal treatment.

[16] This characterization of equilibrium behavior is similar in principle to Azevedo and Leshno (2016). There are two key differences. First, my formulation allows for settings where firms are indifferent over a positive mass of workers, for instance, when worker types are discrete. Second, aggregate demand is not only a function of workers' initial applications, but also must account for applications conditional on learning.

of Theorem 1 shows that the tâtonnement procedure terminates in a finite number of iterations, one for each firm.

Theorem 1 (Equilibrium Characterization)**.** *The tâtonnement procedure culminates in an equilibrium strategy profile in at most F steps.*

Standard bandit techniques such as the Gittins index are normally only valid in decision problems. Despite this, the characterization through hiring thresholds provides the necessary step for the proof of Theorem 1. These thresholds reduce the best response problem into worker specific decision problems, each of which can then be solved through computation of Gittins indices. Furthermore, this procedure only requires a finite number of iterations, making implementability simple.

The tâtonnement procedure shares several features with the worker proposing Deferred Acceptance algorithm. Workers begin by proposing to their ideal positions, then reconsider as they are rejected. In settings where preferences are not aligned, there may be multiple equilibria. However, the equilibria can be ranked through a weak order in terms of worker and firm preferences. Consider two equilibrium strategy profiles, σ and σ', where the thresholds induced by σ are greater than those induced by σ', $(v,p)(\sigma) \geq (v,p)(\sigma')$. Workers must prefer the equilibrium under σ' to σ, while firms prefer σ to σ'.

Proposition 4 (Threshold Ranking)**.** *If σ and σ' are equilibrium strategy profiles, and $(v,p)(\sigma) \geq (v,p)(\sigma')$, then for all $c \in \mathcal{C}, U^c(\sigma) \leq U^c(\sigma')$, and for all $j \in \mathcal{F}$, $\pi_j(\sigma) \geq \pi_j(\sigma')$.*

Importantly, if the expected utility under σ were above that under σ', a class c worker could deviate to her strategy under σ, be hired weakly more often, and receive greater utility. Conversely, firms receive higher utility under σ because their thresholds are higher, implying a higher quality for each worker hired.

4.5 Tenure

On the individual level, the effects of incomplete preference information also naturally capture several of the puzzles mentioned in the introduction. One consistent trend in the data is that higher earners or those who attended better schools are less likely to state that given the chance they would have changed their major or the school they attend (Network 2017). For instance, while 50% of US adults reported that they would change an important education decision, that number drops to 40% when considering adults at top schools and further drops to 23% when considering adults with incomes above $250,000. Lemma 7 shows that despite having potentially higher outside options, those with higher current matches are more likely to be satisfied with any given match. s_i will denote the average stopping age for worker i.

Definition (Stopping Age)**.** *Let $t_s(w)$ be the first period t such that $A_\tau(w) = A_{\tau+1}(w) \forall \tau \geq t$, then $s_i \equiv \mathbb{E}[t_s(w)]$.*

Lemma 7 (High-Quality Workers Search Less)**.** *Suppose θ_h and θ_l are both types in class c, and θ_h is a weakly higher type than θ_l; that is, $\theta_h^j \geq \theta_l^j \forall\ f$.*

Then, in equilibrium, $s_{\theta_h} \leq s_{\theta_l}$.

Lemma 7 is a natural consequence of the optimal search policy. Workers apply for the firm with the maximal Gittins index. Since types θ_h and θ_l are members of the same class, c, type-θ_h workers and type-θ_l workers would have identical Gittins indices across firms they have not yet applied to. That is, their option value is identical across firms. However, type-θ_h workers always realize a higher match value than type-θ_l workers realize. As such, if a type-θ_l worker is satisfied with a firm j, then a type-θ_h worker would be satisfied with j as well. Therefore, whenever type-θ_l workers are willing to stop searching at a given firm, type-θ_h workers would also stop searching at that firm. When matches are permanent, this would manifest as type-θ_h workers reporting higher levels of satisfaction with their current matches, in line with the Gallup report.

5 Policy Interventions

After characterizing the unique equilibrium of the market game, I evaluate the impact of two common policy interventions. The presence of incomplete preference information directly affects the outcomes of these policy interventions through altering their search patterns. While there are many instruments that could be used to increase the quality of matches, I focus on headhunters and unemployment benefits as they are commonly used tools in practice, with the specific intent of altering workers search patterns.

5.1 Headhunters

A natural question is the impact of hidden information, are all workers affected equally due to the presence of uncertainty? I show when firms are congested, incomplete preference information benefits more competitive workers. When there is incomplete preference information, workers must apply to multiple firms to find their best matches. This search comes at the cost of time and firings for low-quality workers. The ability of low-quality workers to apply to other firms is restricted, leaving more spots open for competitive workers.

Suppose firm j were to hire a headhunter. The headhunter then could send out offers to every worker, informing them of their match value at firm j, (θ_w^j, θ_f^j). As an aside, if instead the headhunter only contacted workers that were hired in the resulting equilibrium, the result would remain unchanged. Workers that were not contacted, would learn that they were poor fits at j, and avoid applying there regardless.

I define the market with revelation at firm j as the market induced by all workers learning their match value at j.

Definition (Revelation at Firm j)**.** *Let market $\mathcal{M} = (\mathcal{F}, \mathcal{C}, m)$ be given.*

Then, market $\mathcal{M}_j = (\mathcal{F}', \mathcal{C}', m')$ is induced by revelation at firm j if:

1. $\mathcal{F}' = \mathcal{F}$,
2. $\mathcal{C}' = \cup_{c \in \mathcal{C}} \left\{ \cup_{x \in \mathbb{R}^+} \{\theta | ((\theta_w^j, \theta_f^j), \theta_{-j}) = \theta \in c\} \right\}$,
3. $m'(\theta') = m(\theta')$.

The second condition of revelation at firm j involves splitting each worker class c into k_c worker classes, where each new class involves a different match value at firm j. Condition 3 simply requires that the mass function is proportionally distributed.

Example 3 illustrates how headhunters can benefit low-quality workers to the detriment of high-quality workers.

Example 3 (Revelation)**.** *Revealing information at f_r benefits type-θ_l workers through inducing them to avoid f_r. This harms type-θ_h workers through increasing competition at f_s.*

Workers \ *Firms*	$m(f_r) = 1/4$	$m(f_s) = 1/2$
$m(\theta_h) = 1/2$	$(3, 3)$	$(2, 2)$
$m(\theta_l) = 1/2$	$(1/2, 1/2)$	$(2, 2)$

When $\delta = 1/2$, equilibrium payoffs for type θ_h and θ_l respectively are $U(\theta_h) = \frac{9}{4}$ and $U(\theta_l) = \frac{5}{4}$. The equilibrium strategy profile is simple for δ this low. Workers randomize between the two firms when uninformed, and apply to f_r if type θ_h or to f_s if rejected initially from f_r. In equilibrium, when θ_l types are rejected by applying to f_r, the proportion that can apply to f_s is reduced. The low capacity of f_r requires θ_h types to apply to f_s in equilibrium, and so they benefit from the rejections suffered by θ_l types.

Now consider $\mathcal{M}_{f_r}$. θ_l types know they will never be hired by f_r. As such, they only apply to f_s. In doing so, they increase the level of competition at f_s. This weakens the outside option for θ_h types who previously benefitted from f_s, but no longer have a safety valve for over-competition at f_r. In market $\mathcal{M}_{f_r}$, equilibrium payoffs are $U(\theta_h) = U(\theta_l) = \frac{7}{4}$. θ_l types are made better off by the revelation, while θ_h types are worse off. In this stylized example, complete information harms θ_h types relative to θ_l types because the information itself provides no aggregate increase in efficiency.

In general, firm j hiring a headhunter has two effects; a sorting effect, where higher types can immediately apply to a higher quality match, and an avoidance effect, where lower types apply to safer options. When j is sufficiently congested, the former effect vanishes, leaving only the latter, benefiting lower types on the whole. If j isn't sufficiently congested, the result of this policy is ambiguous. Proposition 5 characterizes these two effects, and provides a sufficient condition for less competitive types to be better off than more competitive types.

Two definitions will be useful, first—a market is congested if each firm fully utilizes its capacity in equilibrium.

Definition (Congestion)**.** *Firm j is congested, if in equilibrium, $D^j = m(j)$. The market is congested if every firm is congested.*

Second, to economize on notation, let $U^j_{\mathcal{M}}(\theta)$ denote type θ's payoff from firm j in market $\mathcal{M}$.

$$U^j_{\mathcal{M}}(\theta) = \mathbb{E}_{\mathcal{M}}\left[\sum_{t=0}^{\infty} \delta^t \mathbf{1}_{w(t)=j}\theta^j_w\right]$$

Proposition 5 (Impact of Headhunters)**.** *Suppose $\mathcal{M}$ is congested and $|\mathcal{C}| = 1$.*

Consider two worker types θ_h and θ_l such that $\theta^j_h > \theta^j_l$.

1.
$$U^j_{\mathcal{M}_j}(\theta_h) - U^j_{\mathcal{M}}(\theta_h) \geq U^j_{\mathcal{M}_j}(\theta_l) - U^j_{\mathcal{M}}(\theta_l)$$

2. *If $m(\theta^j_h) > m(j)$:*
$$U_{\mathcal{M}_j}(\theta_h) - U_{\mathcal{M}}(\theta_h) \leq U_{\mathcal{M}_j}(\theta_l) - U_{\mathcal{M}}(\theta_l)$$

Part 1 of Proposition 5 shows that a higher type receives more net utility from firm j relative to a lower type. Part 2, whose condition was satisfied in Example 3, states that on net, a higher type gains less from headhunters when j is heavily congested. There are two forces behind the second result: the nature of the outside option and the independence of match values across firms. In the decentralized markets I study, the cost of a poor match is embedded in the opportunity cost. A centralized market, wherein a fired worker could immediately apply to another firm, would eliminate this force. In such a setting, this result would reverse, headhunters would solely benefit higher quality workers. Similarly, if types were correlated across firms, that is if $\theta_h^j > \theta_l^j$ were to imply $\theta_h^{j'} > \theta_l^{j'}$, then the competition effect would vanish, because θ_h would not be effected by the generated increase in θ_l's applying to f_s.

5.2 Unemployment Benefits in Congested Markets

A common intuition for economists would be that increasing unemployment benefits further encourages search, because it reduces the loss from unsuccessful matching. In turn, unemployment benefits would be believed to improve equilibrium efficiency as workers are better matched to firms. $\theta_w^{\varnothing}$ has a straightforward interpretation in this context, $\theta_w^{\varnothing}$ is the benefit for rejected workers, or the "unemployment benefits." This section considers the impact when the utility from being rejected in a given period is raised from $\theta_w^{\varnothing} = 0$ to $\theta_w^{\varnothing} > 0$. Example 4 shows that the answer is somewhat nuanced. *Negative learning*, a worker's incentive to learn about a less desirable firm, is suppressed by unemployment benefits.

Example 4 (Unemployment Benefits)**.** *As depicted in Figure S1, increasing unemployment benefits when there is an agreed upon top firm can decrease equilibrium welfare.*

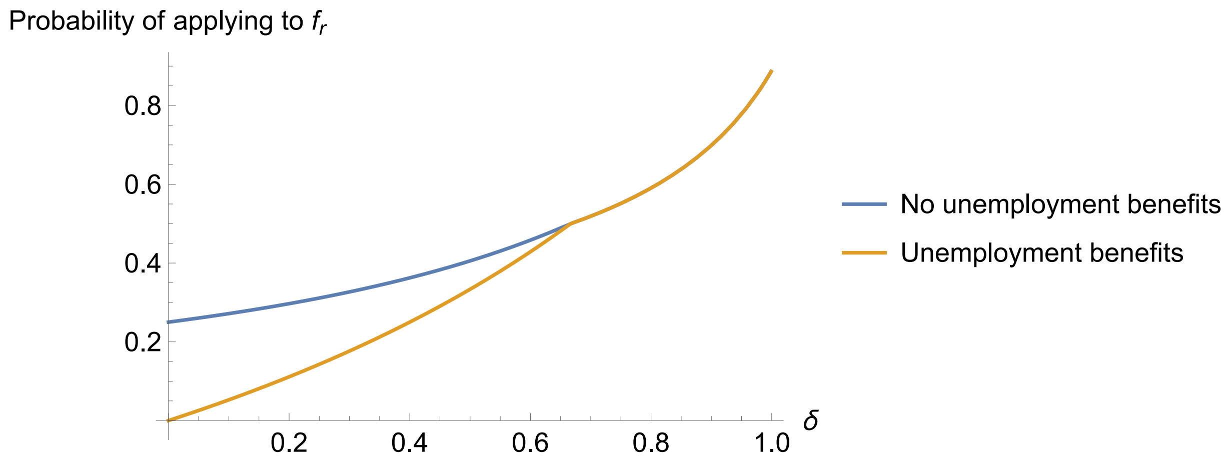
Firms / *Workers*	$m(f_r) = 1/4$	$m(f_s) = 1/4$
$m(\theta_h) = 1/2$	$(3, 3)$	$(3, 3)$
$m(\theta_l) = 1/2$	$(1, 1)$	$(3, 3)$

Safety nets can reduce the overall efficiency of equilibrium. Indeed, when the initial f_r application probability drops below 1/4 in Example 4, the total equilibrium utility, net of the safety net, is reduced. The key insight here is that f_s is more attractive than f_r. The only thing preventing more applications to f_s is the high level of competition and resulting low probability of hire. As $\theta_w^{\varnothing}$ increases, the loss from rejection decreases. This incentivizes more applicants to f_s, which in turn increases the probability of rejection.

6 Transferable Utility

The preceding sections of the paper required non-transferable utility, because the motivating examples such as residency markets or dating markets either had fixed wages or lacked transferable utility. Nonetheless, in many settings such as labor markets we might expect wages to adjust in equilibrium. The key results of the previous section only required that the gross substitutes

FIGURE S1: The x-axis represents δ, the probability does not retirement. The blue line shows the proportion of uninformed workers that apply to f_r when $\theta_w^{\varnothing} = 0$. The yellow line shows the proportion of uninformed workers that apply to f_r when $\theta_w^{\varnothing} = 1$.



condition were satisfied. I show that incorporating transferable utility into the model still allows for gross substitutes and therefore does not change the qualitative results.

In this section, firms chose wages conditional on types in a competitive equilibrium setting. Formally, I consider a separate "wage game". In the wage game, all firms simultaneously announce a *payment function* $\phi_j : \mathcal{C} \times \mathbb{R}_+ \to \mathbb{R}$, denoting the payment from firm j to a worker conditional on match value and class. Importantly, payments cannot be differentiated between workers within a class with equal match values. To maintain the connection with the previous section, I assume the prenumeration match values remain unchanged, namely if a type-θ worker matches to firm j, both agents receive θ^j in addition to the transfer.

After period 0, wages remain fixed, and workers apply to firms as before, with value from matching altered accordingly. Firms evaluate their profit through their per period expected profit, denoted by $\pi(\{\phi_j\}_{j\in\mathcal{F}})$. When a worker and firm match, they receive $\theta_w^j + \phi_j(\theta)$ and $\theta_f^j - \phi_j(\theta)$ respectively.

Contracts must satisfy limited liability, the size of any transfer cannot exceed the gain from the matching for either party:

Assumption 2 (Limited Liability)**.**

$$-\theta_w^j \leq \phi_j(\theta) \leq \theta_f^j$$

Competitive equilibrium is defined in the standard manner. A profile of payment functions constitutes a competitive equilibrium if no firm could change its payment function and profit:

Definition (Competitive Equilibrium)**.** *A profile of payment functions* $\{\phi_j\}_{j\in\mathcal{F}}$ *constitutes a* *competitive equilibrium if for all f, ϕ:*

$$\pi\left(\{\phi_j\}_{j\in\mathcal{F}}\right) \geq \pi\left(\phi, \{\phi_{j'}\}_{j'\neq f}\right)$$

This characterization of equilibrium implicitly prohibits dynamic punishments. Allowing for dynamic punishments would enable a folk-theorem style argument which could rationalize all wages when firms are sufficiently patient.

Importantly, for any profile of payment functions, the resulting market is effectively a new market with non-transferable utility. Then, workers still evaluate firms as if they were endogenous bandits, albeit with differing rewards. As such, aggregate demand satisfies gross substitutes, and therefore an equilibrium can be characterized using the algorithm in Theorem 1.

Proposition 6 (Wages Satisfies Gross Substitutes)**.** *For any profile of wages $\{\phi_j\}_{j\in\mathcal{F}}$, aggregate demand satisfies gross substitutes*

6.1 Non-Existence of Competitive Equilibrium

Even in such a simple environment, equilibria aren't guaranteed to exist. This comes from the combined facts that the demand for a firm isn't continuous in its wage, and workers can apply multiple times.

Example 5 (Competitive Equilibria Need Not Exist)**.**

Workers \ *Firms*	$m(f_r) = 1/2$	$m(f_s) = 2$
$m(\theta_h) = 1$	$(3, 3)$	$(2, 2)$
$m(\theta_l) = 1$	$(1/2, 1/2)$	$(2, 2)$

Lemma 8. *Example 5 has no competitive equilibrium.*

First, note that the minimum proportion of workers f_s can hire is given by the strategy profile where all workers initially apply to the f_r, then low type workers subsequently apply to f_s. To see why, note that if a low type worker is willing to remain at f_r after learning her type, then all high type workers must strictly prefer to remain at f_r. Furthermore, this implies that an uninformed worker must strictly prefer to work at f_r. However, then the mass of workers applying to f_r is greater than $1/2$, and low type workers are hired with probability 0. As such, they would prefer to work at f_s, unless the wage is exactly -2, in which case they are indifferent. By standard bargaining arguments, this implies that $\phi_s = -2$ and all low type workers apply to f_s. Then, a lower bound for π_s in equilibrium is given by $1/2\delta \cdot 4 = 2\delta$.

To see why Lemma 8 holds, suppose a competitive equilibrium existed in Example 5. There are three possible worker strategies. Either all workers apply to f_s forever; all workers initially apply to f_r, then type-θ_h workers remain at f_r and type-θ_l workers migrate to f_s forever; or workers randomize between the previous two options.

First, suppose in the competitive equilibrium, all workers applied to f_s forever. Then, f_r would hire no workers, and so receive 0 in profit. f_r would deviate and set payment for θ_h equal to $6-\epsilon$, since this will generate positive profit for f_r. However, f_s will never offer a wage higher than $2\delta = (2-w)2$ or $w = 2(1-\delta/2)$, as otherwise f_s could revert to a wage of -2. Since $2(1-\delta/2) < 6$ at least one of f_r or f_s must have a profitable deviation.

Next, suppose the equilibrium was in the second category. Then, f_s must set its wage to -2 as listed above. However, f_r then can deviate to a wage of $-3+\epsilon$ while still attracting high types if its wage was above -3. Last, this implies f_s could set a wage just above ϕ_r, a marginal increase in payment that earns f_s a strict increase in workers of $1-1/2\delta$.

Last, suppose the equilibrium was in the third category. Then, workers are exactly indifferent between the two firms. However, a firm could increase its wage marginally to capture all workers. This generates a strict increase in profit, unless that firm was earning zero profit from each worker. For f_r this implies $\phi_r = 3$, while for f_s this implies $\phi_s = 2$. However, then f_s could deviate to $\phi_s = -2$ as detailed above.

The driving force behind Lemma 8 is that firms cannot distinguish workers with different outside options. This reduction in the dimension of possible wages, restricts the ability to find an equilibrium. It is worth noting that while on the surface this may appear similar to issues such as the non-existence that arises from adverse selection in Rothschild and Stiglitz (1978), the issue is categorically different here. In their work, non-existence results from free firm entry, as firms currently in the market are harmed by entry. In transient markets, the firms already present in the market face incentives to change their pricing structures, in order to manipulate the learning behavior of workers in the market.

6.2 Resumes

The non-existence in the previous example is resolved by the inclusion of resumes. In this context, a resume allows a worker to prove she had been previously hired by a separate firm, in return for higher wages. Resumes allow for wages to be conditioned not only upon the direct match value, but also upon a worker's previous history. Formally, a *resume-dependent payment function* is a payment function defined on a larger domain $\phi_j : \mathcal{C} \times \mathbb{R}_+ \times \{A_i^j(\tau)\}_{\tau=0}^t \to \mathbb{R}$.

To see how the inclusion of resumes generates a competitive equilibrium in Example 5, note that now f_s could choose a wage that incorporates the worker's outside option. The following wages are an equilibrium for δ sufficiently high, type-θ_l workers receive 0 from f_r and -2 from

f_s, while type-θ_h workers receive 1 from f_r and 2 from f_s upon providing a resume and -2 otherwise. In equilibrium, all workers apply to f_r initially, and type-θ_l workers transition to working at f_s.

The non-existence issue was due to the requirement that firms treat workers of a single class with equal match values identically. Through introducing resumes, workers could prove that their outside options were distinct. Thus, resumes play an important role in information transmission.

7 Discussion

This paper develops a framework for analyzing transient matching when workers learn through experimentation. Firm capacity constraints force workers to anticipate other workers' application decision. Nonetheless, firms can be evaluated as endogenous bandits. Once firms' hiring decisions are described as thresholds, techniques from the multi-armed bandit literature allow for a simple description of the optimal worker policy. Importantly, the aggregation of demand satisfies the gross substitutes condition, given which equilibrium can be characterized. Workers search patterns match data from labor markets, high-quality workers report higher satisfaction—despite not having better information.

Utilizing the equilibrium, the subsequent results show that the nature of both learning and competition are critical to understanding the impact of policy interventions. Well-intentioned but naïve interventions, such as hiring headhunters or increasing unemployment benefits, may generate unexpected effects in congested markets.

The results also imply careful consideration should be taken before changing centralized mechanisms. Such mechanisms often feature decentralized aftermarkets, where the incoming information from the original centralized market can radically shift the final outcome. This paper provides a first step towards better understanding the effects changes of mechanism rules can have on the aftermarkets of centralized mechanisms.

References

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan (2020). "Thickness and information in dynamic matching markets". In: *Journal of Political Economy* 128.3, pp. 783–815.

Anderson, Axel and Lones Smith (2010). "Dynamic matching and evolving reputations". In: *The Review of Economic Studies* 77.1, pp. 3–29.

Azevedo, Eduardo M and Jacob D Leshno (2016). "A supply and demand framework for two-sided matching markets". In: *Journal of Political Economy* 124.5, pp. 1235–1268.

Baccara, Mariagiovanna and Leeat Yariv (2021). "Dynamic matching". In: *Online and Matching-Based Market Design.*

Becker, Dan (2011). "Non-wage job characteristics and the case of the missing margin". In: *Available at SSRN 1805761.*

Chade, Hector, Jan Eeckhout, and Lones Smith (2017). "Sorting through search and matching models in economics". In: *Journal of Economic Literature* 55.2, pp. 493–544.

Chen, Yi-Chun and Gaoji Hu (2020). "Learning by matching". In: *Theoretical Economics* 15.1, pp. 29–56.

Christensen, Bent Jesper, Rasmus Lentz, Dale T Mortensen, George R Neumann, and Axel Werwatz (2005). "On-the-job search and the wage distribution". In: *Journal of Labor Economics* 23.1, pp. 31–58.

Combe, Julien, Olivier Tercieux, and Camille Terrier (2018). "The design of teacher assignment: Theory and evidence". In: *Unpublished paper, University College London.[1310].*

Dagsvik, John, Boyan Jovanovic, and Andrea Shepard (1985). "A foundation for three popular assumptions in job-matching models". In: *Journal of Labor Economics* 3.4, pp. 403–420.

Doval, Laura (2022). "Dynamically stable matching". In: *Theoretical Economics* 17.2, pp. 687–724.

Ferdowsian, Andrew, Muriel Niederle, and Leeat Yariv (2022). *Decentralized Matching with Aligned Preferences.* Tech. rep. mimeo.

Fernandez, Marcelo Ariel, Kirill Rudov, and Leeat Yariv (2022). "Centralized matching with incomplete information". In: *American Economic Review: Insights* 4.1, pp. 18–33.

Gittins, John, Kevin Glazebrook, and Richard Weber (2011). *Multi-armed bandit allocation indices.* John Wiley & Sons.

Gul, Faruk and Ennio Stacchetti (2000). "The English auction with differentiated commodities". In: *Journal of Economic theory* 92.1, pp. 66–95.

Hall, Robert E and Alan B Krueger (2012). “Evidence on the incidence of wage posting, wage bargaining, and on-the-job search”. In: *American Economic Journal: Macroeconomics* 4.4, pp. 56–67.

Immorlica, Nicole, Jacob Leshno, Irene Lo, and Brendan Lucier (2020). “Information acquisition in matching markets: The role of price discovery”. In: *Available at SSRN 3705049*.

Jovanovic, Boyan (1979). “Job matching and the theory of turnover”. In: *Journal of political economy* 87.5, Part 1, pp. 972–990.

Kadam, Sangram V and Maciej H Kotowski (2018). “Multiperiod matching”. In: *International Economic Review* 59.4, pp. 1927–1947.

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

Lee, Robin S and Michael Schwarz (2017). “Interviewing in two-sided matching markets”. In: *The RAND Journal of Economics* 48.3, pp. 835–855.

Li, Bingbing, Stephanie Wang, and Xiaohan Zhong (2016). “Timing is everything? An experimental investigation of incomplete information in centralized matching mechanisms”. In: *Manuscript, Univ. Pittsburgh. http://www. pitt. edu/swwang/papers/Matching. pdf*.

Liu, Lydia T, Horia Mania, and Michael Jordan (2020). “Competing bandits in matching markets”. In: *International Conference on Artificial Intelligence and Statistics*. PMLR, pp. 1618–1628.

Liu, Lydia T, Feng Ruan, Horia Mania, and Michael I Jordan (2021). “Bandit learning in decentralized matching markets”. In: *Journal of Machine Learning Research* 22.211, pp. 1–34.

Liu, Qingmin (2020). “Stability and bayesian consistency in two-sided markets”. In: *American Economic Review* 110.8, pp. 2625–66.

Liu, Qingmin, George J Mailath, Andrew Postlewaite, and Larry Samuelson (2014). “Stable matching with incomplete information”. In: *Econometrica* 82.2, pp. 541–587.

Menzio, Guido and Shouyong Shi (2011). “Efficient search on the job and the business cycle”. In: *Journal of Political Economy* 119.3, pp. 468–510.

Menzio, Guido, Irina A Telyukova, and Ludo Visschers (2016). “Directed search over the life cycle”. In: *Review of Economic Dynamics* 19, pp. 38–62.

Miller, Robert A (1984). “Job matching and occupational choice”. In: *Journal of Political economy* 92.6, pp. 1086–1120.

Network, Strada Education (2017). *On second thought: US adults reflect on their education decisions*.

Papageorgiou, Theodore (2018). “Large firms and within firm occupational reallocation”. In: *Journal of Economic Theory* 174, pp. 184–223.

Postel-Vinay, Fabien and Helene Turon (2010). "On-the-job search, productivity shocks, and the individual earnings process". In: *International Economic Review* 51.3, pp. 599–629.

Rothschild, Michael (1974). "A two-armed bandit theory of market pricing". In: *Journal of Economic Theory* 9.2, pp. 185–202.

Rothschild, Michael and Joseph Stiglitz (1978). "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information". In: *Uncertainty in economics*. Elsevier, pp. 257–280.

Urgun, Can (2021). "Restless Contracting". In: *mimeo*.

Voorneveld, Mark and Henk Norde (1997). "A characterization of ordinal potential games". In: *Games and Economic Behavior* 19.2, pp. 235–242.

Weitzman, Martin L (1979). "Optimal search for the best alternative". In: *Econometrica: Journal of the Econometric Society*, pp. 641–654.

Whittle, Peter (1980). "Multi-armed bandits and the Gittins index". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 42.2, pp. 143–149.

A Proofs

I begin by showing that standard results from the operations literature apply when firms hire all applicants. The proof of Lemma 9 provides the basis for the bandit related proofs in the rest of the appendix.

Lemma 9. *Let σ be an equilibrium, such that, in any subgame, all firms hire all applicants each period. In any period, each worker i applies to some $j \in \arg\max_j GI_i^{\sigma}(j, h_t^i)$.*

Proof of Lemma 9

This proof follows directly from Theorem 2.1 of (Gittins, Glazebrook, and Weber 2011) which states:

Theorem 2.1: *A policy for a simple family of alternative bandit processes is optimal if it is an index policy with respect to the Gittins index of each bandit process.*

Then, it must be shown that the set of firms acts as a simple family of alternative bandit processes for each worker. First, since firms hire all applicants, each worker i faces a fixed decision problem for any strategy profile σ, independent of the other workers' strategies. For any possible worker strategy profile, each worker is always hired, no matter which firm she applies to. As such, the payoff for a type-θ worker from applying to firm j, is simply θ_w^j in any equilibrium.

Each firm is then a bandit process that can either be activated or frozen. A frozen firm provides no payoff, while an activated firm provides θ_w^j. Each worker must activate exactly one firm each period. The independence of firm match values—conditional on a worker's class—implies that the set of firms is a simple family of alternative bandit processes. Theorem 2.1 from Gittins, Glazebrook, and Weber (ibid.) then applies, and the result follows. ■

Proof of Lemma 1

I begin by describing the Top-Down algorithm in detail. For any complete-information market $\mathcal{M}_I = (\mathcal{F}, \mathcal{C}, m)$, the Top-Down algorithm proceeds as follows:

Top-Down Algorithm:

Let $\mathcal{E} = \mathcal{F} \cup \mathcal{C}$ be the set of agents present in the market.

1. Find the largest worker-firm match value: $(c^*, j^*) = \arg\max_{c,j \in \mathcal{E}} \theta_c^j$, the finiteness of $\mathcal{C}$ implies this is well-defined.

2. Find class-c^*'s second-highest match value: $j_2 = \arg\max_{j \neq j^*} \theta_{c^*}^j$.

3. Match c^* to j^*, which leads to one of two possible outcomes:

(a) Under Subscription—When all class-c^* workers apply to j^*, j^* is still preferred by class-c^* workers relative to j_2:

$$\frac{m(j^*)}{m(c^*)}\theta_{c^*}^{j^*} \geq \theta_{c^*}^{j_2}$$

i. All class-c^* workers apply to j^* forever.

ii. Firm j^*'s capacity is reduced by the mass of applicants: $m(j^*) = \max\{0, m(j^*) - m(c^*)$.

iii. Reduce to the submarket without class-c^*, $\mathcal{E} = \mathcal{E} \setminus \{c^*\}$. Remove j^* from the market as well if $m(j^*) = 0$.

(b) Over Subscription—When too many class-c^* workers apply to j^*, j_2 becomes preferable:

$$\frac{m(j^*)}{m(c^*)}\theta_{c^*}^{j^*} < \theta_{c^*}^{j_2}$$

i. c^* randomizes between j^* and j_2 to generate indifference: c^* applies to j^* with probability $m(j^*)\frac{\theta_{i^*}^{j^*}}{\theta_{c^*}^{j_2}}$. The remaining probability, $1-m(j^*)\frac{\theta_{i^*}^{j^*}}{\theta_{c^*}^{j_2}}$, will be accounted for in a future step.

ii. Reduce to the submarket without j^*, $\mathcal{E} = \mathcal{E} \setminus \{j^*\}$.

4. In the new submarket $\mathcal{E}$, again select the highest match value: $(c^*, j^*) = \arg\max_{c,j\in\mathcal{E}} \theta_c^j$.

(a) If j^* was not previously selected, repeat steps 2 and 3.

(b) Otherwise, if c^* was previously oversubscribed:

i. The proportion of c^* applying to the original firm must be increased to keep c^*'s payoffs from both firms equal.

ii. Redistribute workers between the two (or more) firms keeping c^*'s payoff from each firm equal, continuing until all of c^* has been allocated, or the payoffs from each firm are equal to those of the next top choice of c^*.

(c) If j^* was previously selected as j_2, increase the proportion of the previous class-c^* applying to the previous j^* to equalize the previous c^*'s utility between the previous j^* and the current j^*.

5. Repeat the above steps until either $\mathcal{E} \subset \mathcal{C}$ or $\mathcal{E} \subset \mathcal{F}$

6. Any remaining workers apply to any firm and are rejected. All remaining firms operate below capacity.

Now I prove the claim. Suppose there exists a strategy profile σ' whose outcome does not coincide with the Top-Down algorithm. The Top-Down can be used to construct a strategy profile σ. Because σ' and σ differ in their outcome, there is a minimum k such that in the kth iteration, there exists a class c, such that c's strategy profile differs between σ and σ'. By construction, this implies that some member of class c has payoff under σ' below its payoff under σ. Furthermore, all matches corresponding to higher match values are equivalent between σ and σ' by construction. Then, under σ', a class c worker could deviate to their firm under σ and receive higher expected utility every period. Therefore, σ' could not have been an equilibrium. ■

Proof of Lemma 2

To prove the claim, I first characterize the equilibrium strategy profile in the limit as $\delta \to 1$. I then show that the resulting outcome coincides with the equilibrium outcome in the market with complete information. As $\delta \to 1$, the maximal match-value must be fully utilized, either by the associated firm reaching capacity, or by the associated type fully applying to the firm. Otherwise, for sufficiently high δ, a worker of the class with the maximal match value would benefit from applying immediately to the firm, and learning their match value. This implies that an algorithm similar to the Top-Down algorithm can be used to characterize long-lived workers' strategies.

One key alteration is necessary, because workers have incomplete information, rather than specifying the application strategy of a given match value, the algorithm now specifies the application strategy for an entire class. As such, rather than agents behaving identically every period, workers of a given class follow a descending chain of applications hunting for their top match value. When doing so, the mass of each type within a class must decrease by δ every step as workers retire along the way.

To proceed, consider the previous Top-Down algorithm, where workers now follow the strategy for their entire class, stopping searching once they have found their proscribed match-value. Notably, when a firm's capacity is not exhausted in a given iteration, workers immediately learn if they would achieve the relevant match value upon applying once. Either they are accepted and learn their actual match value, or they are rejected and learn that they do not have the match value specified in that step of the procedure. If a firm's capacity were to be exhausted for a given iteration, it would be possible for a rejected worker to have access to the associated match-value but simply be unlucky. Then, equilibrium behavior may require that workers of such a class apply multiple times to the same firm. Similar to the complete information setting, one of two cases results, either over subscription or under subscription.

Altered Top-Down algorithm:

1. Find the largest match value: $(i^*, j^*) = \arg\max_{w,f\in\mathcal{E}} \theta_i^j$
2. Find i^*'s next best match value: $j_2 = \arg\max_{f\neq j^*} \theta_{i^*}^j$

3. There are two possible cases:

   - Under Subscription - even if all i^* apply to j^*, j^* is still preferable to j_2: $\frac{m(j^*)}{m(i^*)}\theta_{i^*}^{j^*} \geq \theta_{i^*}^{j_2}$.
     - All i^* apply to j^* forever.
     - Therefore, i^* can be removed from consideration
     - Firm j^*'s capacity is reduced by $\min\{m(j^*), m(i^*)\}$
     - The mass of all class-c non-type-i^* workers is reduced by $(1-\delta)$
   - Over Subscription - all i^* applying to j^* makes j_2 preferable: $\frac{m(j)}{m(i)}v(m, f) < v(w, j_2)$.
     - This implies it isn't immediately a dominant strategy for all i to apply to j, since the probability of rejection is sufficiently high to disincentivize them.
     - However, at least enough to equalize the expected payoffs must apply, namely $m(j)\frac{v(w,f)}{v(w,j_2)}$.
     - Remove firm j from all other worker's consideration, as no other workers will be able to successfully apply to j

4. From the reduced matrix, select the next highest match value with i', j'

5. If $i' \neq w$ for any previous i, repeat the above

6. Similarly, if i' was in an under subscription case, repeat the above

7. If i' faced over subscription in its previous appearance, then the values need to be adjusted correctly

   - Increase the number of applicants to the original firm j and the new firm j' to keep both payoffs equal, while decreasing them until either these workers run out, or the payoffs are equal to those of the next firm, j_2'

8. Since this algorithm iterates over match values, removes at least one from consideration each time, and there are a finite number of match values, the algorithm eventually terminates yielding the unique match.

Type-convergence holds inductively. Note that there are a finite number of steps, and each step differs by a continuous function of δ. First, the top type converges as $\delta \to 1$. Then, the distance between the outcomes of $\mathcal{M}$ and $\mathcal{M}^I$ is a continuous function of δ. When types are discrete, the rest converge immediately as well. ■

Proof of Lemma 3

At the first step in which over subscription occurs in the algorithm of Lemma 1, select the associated type-i^* and firms j^* and j_2. Members of type-i^* then either they apply to j^* or j_2. Those that fail to apply to j^* will never return, while those who apply to j^* may be hired with positive probability. This generates path dependence in equilibrium. ■

Proof of Proposition 1

Consider a firm j that is hiring below capacity, namely j rejects applicants while $m(A_j) < m(j)$. Importantly, rejecting some such applicant, i, has two impacts that could potentially benefit j: the rejection could cause i to return to j in a later period, and the rejection could trigger a rejection cycle wherein a more preferable worker i' is rejected by another firm causing i' to apply to j.

Hiring i immediately front-loads the match-value from i, avoiding the loss from i possibly retiring before returning to j. Note that it cannot be the case that i eventually returns to j and applies more times to j than i would if j had accepted immediately. Eventually i's information sets under the original strategy and the deviation must coincide, at which point the Markovian nature of i's strategy forces i to apply to the same firm under both strategy profiles. Then, it follows that j's payoff from that point forward is unchanged. Since i was accepted by j at most once before i's information converged under the original strategy profile, j strictly benefits from the deviation.

By assumption, preferences are aligned. It is known that aligned markets have no rejection cycles (Voorneveld and Norde 1997). Namely, it cannot be the case that $\theta_i^{j'} > \theta_{i'}^{j'}, \theta_i^j > \theta_i^{j'}$, and $\theta_{i'}^{j'} > \theta_{i'}^j$. Then, j can also not benefit from rejecting i to attempt to attract other workers.

Last, if j is congested, then it must have excess applications with match values equal to its lowest match value. In order to accept a high ranked applicant, j would then need to reject one of its lowest ranked applicant. However, even if it dissuades one such applicant through rejection, another with equivalent match value will be available to replace it. Then, the previous arguments imply that j has a profitable deviation.

■

Proof of Lemma 4 A Markovian strategy for firm j, is a mapping from the types of workers, a_j, to acceptances. Notably, because worker's strategies are Markovian, worker's payoff relevant information is summarized by her belief regarding her type. For a given worker j, and private history h_t^j, let j's posterior regarding her type be given by $p \in \Delta\Theta$. Then, the payoff-relevant state space is a distribution over p, Ψ.

Consider the Markov chain over Ψ. Workers' applications follow the same initial distribution over applications. Furthermore, there are a finite number of firms, and as such workers must eventually converge in belief, independent of the order in which they apply to firms, implying that their final applications are a fixed quantity. Then, there exists a subset $V \subset \Psi$, such that V is irreducible. Furthermore, because new workers enter the market every period, the Markov chain is aperiodic. Then, by the Steady State Theorem, there exists a steady state. ■

Proof of Lemma 5

Suppose for worker i, under an arbitrary strategy profile σ, for some firm j, $GI_i^\sigma(j, h_t^i) = x$. I show that this value of $GI_i^\sigma(j, h_t^i)$ is achieved by the strategy described in the statement of the lemma: if the realization of ϕ_i^j is above x set $\tau = \infty$, otherwise set $\tau = 1$. If exactly x was realized, any stopping time yields an equivalent outcome. Assume throughout that the realization is given by $y > x$. It will be helpful to refer to the value of the stopping time problem that characterizes $GI_i^\sigma(j, h_t^i)$. I let $gi(f, h_t^i, \tau)$ be the value of the stopping time problem that characterizes $GI_i^\sigma(j, h_t^i)$, given a possibly suboptimal stopping time, τ. Formally:

$$gi(f, h_t^i, \tau) \equiv \frac{\mathbb{E}_i\left[\sum_{t=1}^{\tau} \delta^t \phi_i^j\right]}{\mathbb{E}\left[\sum_{t=1}^{\tau} \delta^t\right]}$$

I use induction to show that, for any stopping time $\tau < \infty$, $gi(f, h_t^i, \tau + 1) > gi(f, h_t^i, \tau)$.

To begin, I consider the case where $\tau = 1$:

Let $x = p/q$ where $q > 0$, p and q are not necessarily integers, no rationality assumption is made.

$$\begin{aligned} \frac{p + \delta y}{q + \delta} &> p/q \\ p + \delta y &> \frac{p(q + \delta)}{q} \\ \delta q y &> \delta p \\ q y &> p \\ y &> p/q = x \end{aligned}$$

Next, suppose that the claim holds for $\tau \in \{1, 2, \dots, k\}$. A similar computation implies the inductive step also holds for $\tau + 1$.

$$\begin{aligned} \frac{p + \delta^{k+1} y}{q + \delta^{k+1}} &> p/q \\ \implies y &> p/q = x \end{aligned}$$

Then, for any k, $\frac{p+\delta^k y}{q+\delta^k} > p/q$, therefore $gi^\sigma(f, h_t^i, \tau) = GI_i^\sigma(j, h_t^i)$ only if $\tau \in \{1, \infty\}$.

Conversely, the argument also shows that when $y < x$, setting $\tau = 1$ is optimal.

Then, the claim is proven, $GI_i^\sigma(j, h_t^i)$ is characterized by the strategy described in the statement of the lemma. ■

Before proving Proposition 2, I prove a useful lemma. Strict dynamic preferences implies that two strategy profiles that induce different equilibrium outcomes must also generate different hiring thresholds.

Lemma 10. *Under strict dynamic preferences, if σ and σ' generate distinct equilibrium outcomes, then there must exist some firm j with different hiring thresholds under σ and σ'.*

Proof of Lemma 10

Suppose not. That is, σ and σ' induce different outcomes, but every firm has identical thresholds under σ and σ'. Then, every worker faces identical Gittins indices under σ and σ', at every firm, under any informational partition. In particular, since σ and σ' have distinct outcomes, there exists a type θ and information set I such that type θ with information I makes different choices under σ and σ' and is hired with positive probability. That is, type θ applies for firm j with greater probability under σ than under σ'. Similarly, type θ applies for another firm j' with greater probability under σ' than σ. Since forward induction policies are optimal, this implies that $GI_c^\sigma(j) = GI_c^{\sigma'}(j')$. Then, strict dynamic preferences implies that type θ must be hired with intermediate probability for either firm j or firm j'. This immediately implies that type θ has positive mass.

First, suppose type θ was hired with intermediate probability at j. As the non-zero mass of type θs leave j, if no other workers start applying to j, the probability of being hired increases for remaining type θs due to the reduced competition, thereby increasing $GI_\theta^{\sigma'}(j)$ contradicting the claim. To prevent this, there must be another type θ' which has begun applying to firm j in increased numbers. However, type θ's willingness to do so implies that there exists another firm j'' such that $GI_{\theta'}^\sigma(j) = GI_{\theta'}^\sigma(j'')$. Again, by assumption 1 this type must be hired with intermediate probability at one of these firms. Since type θ was hired with intermediate probability at firm j, type θ' must be hired with intermediate probability at firm j''. Repeating this line of logic implies that there exists a cycle of workers, each facing equal Gittins indices at least two firms and hired with intermediate probability at one such firm. However, inherently the total mass of workers is fixed, for every mass of workers that leave a firm, an equal mass must take their place. But marginal workers have a probability less than 1 of being hired, while those taking their place do not. Then, such a cycle cannot keep the total amount of workers hired at involved firms equal. Therefore, the thresholds at those firms must either increase or decrease, contradicting the original assumption. ■

Proof of Proposition 2

Suppose σ and σ' are both equilibria, with distinct outcomes. Lemma 10 then implies that there exists some firm j such that firm j's threshold under σ is distinct from its threshold under σ'.

Without loss of generality, suppose firm j's threshold is higher under σ than under σ'. Then, consider worker i, where i previously applied to a firm j' but now applies to j. A positive mass of such workers must exist, otherwise j's threshold could not have increased. One of two cases must have occurred to generate this change in application, either worker i's original firm's GI decreased, $GI_c^{\sigma'}(j') < GI_c^{\sigma}(j')$, or worker i's GI at their new firm increased, $GI_c^{\sigma'}(j) > GI_c^{\sigma}(j)$. However, $GI_c^{\sigma'}(j)$ cannot be larger than $GI_c^{\sigma}(j)$ because firm j's threshold has increased, therefore every worker has a weakly lower GI than before at firm j. Then, $GI_c^{\sigma'}(j') < GI_c^{\sigma}(j')$, implying that the threshold at firm j' has increased.

Then, the same line of logic as before implies that there exists another non-zero mass of workers applying to firm j', and their original firm must have increased its hiring thresholds. Since there are a finite number of firms, eventually this chain must return to a firm it has previously visited, generating a set of firms with higher thresholds under σ' relative to σ.

Consider the implications of such a set, it cannot be the case that a smaller total of workers are being hired at these firms under σ, otherwise thresholds would fall accordingly. Furthermore, since thresholds are higher, the total utility of hired workers must be higher under σ relative to σ'. Then, there must be at least one worker with strictly greater utility than before. Then, such a worker could have profitably deviated in σ' to its strategy under σ. However, this contradicts the fact that σ' was an equilibrium. ■

Proof of Lemma 6 Lemma 9 showed that in equilibrium workers must apply to a firm with maximal Gittins index when capacities were unconstrained. Repeating the proof of Lemma 9, updating the rewards from firm j, replacing θ_w^j with $\psi_\theta^j(v, p)$ immediately implies that the set of capacity constrained firms is still a simple family of multi-armed bandits. The result follows. ■

Proof of Proposition 3

I begin by proving a stronger claim for individual demand. I show that when whenever any type θ faces a Multi-Armed Bandit problem, and θ's Gittins index for any firm j, $GI_\theta(j)$, is decreasing in j's threshold, (v_j, p_j), type-θ's demand will be decreasing in j's threshold as well. Gross substitutes is satisfied in multi-armed bandit problems because the Gittins index for each individual firm is only a function of the rewards from that firm. Then, increasing the threshold of firm j_2 cannot affect the Gittins index of firm j_1. Furthermore, the realized demand over each firm is weakly increasing in that firm's Gittins index. As such, raising the thresholds of a set of firms, decreases the Gittins indices of those firms, but fails to change the Gittins indices of other firms. Then, applications to the original set of firms must weakly decrease.

Theorem 2. *(Simple bandits imply gross substitutes) Suppose for all θ, type-θ workers face a simple multi-armed bandit problem, and for any firm j, $GI_\theta(j)$ is decreasing in (v_j, p_j). Then, type-θ's demand satisfies gross substitutes.*

Proof of Theorem 2 Lemma 9 implies that, when firms hire all workers, workers apply to a firm with maximal Gittins index each period. However, for any firm j and threshold (v_j, p_j), an auxiliary firm j' can be defined, where type-θ's match values are given by $\psi_\theta(v_j, p_j)$. Then, j' hires all workers, and so Lemma 9 applies.

Then, let $\mathbf{F}$ be a set of firms, $\mathbf{F} \subset \mathcal{F}$, and $(v, p), (v', p')$ be two vectors of thresholds, such that $(v_j, p_j) = (v'_j, p'_j)$ for $j \in \mathbf{F}$ and $(v_j, p_j) \geq (v'_j, p'_j)$ otherwise. Because the thresholds for firms in $\mathbf{F}$ remain unchanged, their Gittins indices are also equal under (v, p) and (v', p'). However, the Gittins indices for firms outside of $\mathbf{F}$, must be weakly lower than before under (v, p). Then, the total demand for firms in $\mathbf{F}$ must weakly increase, as the corresponding Gittins indices are always higher in relative terms.

■

Returning to the proof of Proposition 3, it must be shown that for a given set of thresholds, aggregating across workers preserves changes in demand. This immediately follows from monotonicity of integration. By the definition of aggregate demand, for a given firm $j \in \mathbf{F}$, $D_j(v, p) = \int_\theta D^\theta_j(v, p)d\theta$. Then, if for all θ, $D^\theta_j(v, p) \geq D^\theta_j(v', p')$, it must be that $D_j(v, p) \geq D_j(v', p')$. As such, when every individual type's demand satisfies gross substitutes, so must aggregate demand. The result follows. ■

Before addressing Theorem 1, I prove the following helpful lemma:

Lemma 11. *Suppose after some private history, h^θ_t, which occurs with positive probability for type θ; type-θ applies to j with positive probability, $v_j = \theta^j$, and $GI_\theta(j, h^\theta_t) = GI_\theta(j', h^\theta_t)$. Then, a marginal decrease in p_j yields a discontinuous decrease in D^j.*

Proof. Forward induction implies a form of independence of irrelevant alternatives, reducing $GI_\theta(j)$ below $GI_\theta(j')$ only affects type θ's choice regarding whether to apply to j or j'. In particular, there is exactly one change, type θs that would have applied to j apply to j' instead. Afterwards, type θs will either remain at j' or apply next to j.

Even if type θs later return to j, a period by period comparison shows that the type θ demand for j has decreased. In the first period, the number applying to j is zero, since all have applied to j'. Then, type θs either remain at j' forever, in which case the claim follows, or they proceed by applying to j. However, even should type θs apply to j, upon having applied to both firms, the impact of decreasing $GI_\theta(j)$ has washed out. Type θs now have applied to both firms. Subsequently, type θs will be weakly more likely to apply to firm j'. Last, since $\delta < 1$ front loading applications to firm j' decreases the total number of type θs applying to firm j. $\square$

Proof of Theorem 1

I begin by describing the tâtonnement procedure more detail:

Tâtonnement procedure:

1. Begin by setting all thresholds to $(v_j, p_j) = (1, 0)$.

2. Determine the aggregate demand vector, D.

3. If no firms are hiring beyond their capacities; $\forall j, D_j < m(j)$, terminate.

4. Otherwise, take an arbitrary firm j that is over capacity, and impose its capacity continuously:

   (a) Select the worst match quality θ^j for firm j that is hired with positive probability. Let $\underline{v}_j = \min_{\theta|\theta^j \geq v_j} \theta^j$ and let c be the class of $\theta \in \arg\min_{\theta|\theta \geq v_j} \theta^j$.

   (b) Continuously decrease θ's hire probability at j, $p_j \to 0$.

   (c) Three events could result:

      i. Lowering p_j equates demand and capacity, $D^j = m(j)$, continue to step 5.

      ii. Type $\underline{v}_j$ workers are never hired, $p_j = 0$, return to step 4a and repeat.

      iii. Lowering p_j equalizes c's Gittins indices at j and at least one another firm j', $GI_c(j) = GI_c(j')$. Increase type θs demand for j' at this information set while lowering their demand for j until the demand for j is 0 or $D^j = m(j)$. Proceed to step 4a or 5 accordingly.

5. Again select firm j such that $D^j > m(j)$, if no such j exists, terminate. Otherwise, repeat the process in step 4 with the following change. In step 4(c)iii, if j' is a firm that was not selected in a previous step, then repeat step 4(c)iii without alteration. However, if that firm was selected in a previous step, the above method leads to cycles. To see why, note that Lemma 11 implies an infinitesimal decrease in p_j can cause demand in j and j' to oscillate. Instead, select all firms j' that were previously selected, and simultaneously lower all corresponding $p_{j'}$s along with p_j in proportions such that $GI_c(j) = GI_c(j') \forall j'$ while also reallocating demand between the set of firms accordingly such that no firm in the set is left with excess capacity.

This process must end either through all selected firms equating demand and capacity or with some $p_{\tilde{f}} = 0$. In the first event, return to the beginning of step 5, in the latter event, step 4(c)iii can be resumed as normal.

By design, thresholds never decrease under the tâtonnement procedure. Furthermore, step 5 ensures that previously selected firms never face overdemand. Firm j's step only concludes when $D^j = m(j)$. The only unselected firms are those such that $D^j(0,1) \leq m(j)$. Then, the procedure must conclude within a finite number of iterations, at most one per firm. By construction, worker incentives are incorporated through the Gittins index and threshold characterization, therefore the procedure finds an equilibrium, concluding the argument. ■

It is worth emphasizing that the tâtonnement procedure converges due to the optimality of forward induction. In particular, had type values been correlated between firms, increasing the threshold of one firm could *decrease* demand at another firm. Alternatively phrased, independence of types implies gross substitutes. Gross substitutes requires that the demand for a firm is weakly increasing in the prices (thresholds) of other firms. Gross substitutes proves to be a weaker sufficient condition in order for firm thresholds to characterize equilibrium. Notably, as shown in Appendix section B.1, a model of gradual learning—where instead of receiving θ_w^j from a match, a worker received a noisy signal of θ_w^j—would also converge to equilibrium under the tâtonnement procedure as the gross substitutes condition would still be satisfied.

Proof of Proposition 5

To begin, note that congestion implies that the total number of workers hired does not change under information revelation. In particular, market congestion implies that several workers who applied to firm j under $\mathcal{M}$ were fired. Furthermore, the fired workers must have weakly lower match values than the workers who are hired. If θ_h was maximally hired at firm j, then the total payoff for θ_h at firm j cannot increase. However, since θ_l must decrease its level of applications to firm j relative to θ_h, θ_l must apply in larger quantities to firms other than j. $|\mathcal{C}| = 1$ requires θ_h and θ_l to share a class. In expectation, their match values at other firms are equivalent. As such, the net effect is a reduction in the proportion of θ_l fired at j, increasing the congestion at the remaining set of firms. This causes the average payoff of workers at those firms to decrease. This argument proves parts 1 and 2 of the claim. When $m(j) < m(\theta_h)$ this can be the only effect on θ_h type workers, and so their payoff decreases, however low type workers see a commensurate increase in payoff since they are no longer being fired from firm j. Therefore, the change in payoffs must be greater for θ_l relative to θ_h. ■

Proof of Lemma 7

Since types θ_h and θ_l are in class c, workers of either type have Gittins indices that are equal across every firm. As a reminder, the optimal strategy every period for a worker was to apply to the firm with the highest Gittins index. In particular, a worker stops searching if her match value at a given firm is above the Gittins index of any other firm. That is, whenever θ_ls stop searching at some firm j, $GI_{\theta_l}(j) \geq GI_{\theta_l}(j') \forall j'$. Since both worker types are in class c, this implies that both types follow the same probability distribution over initial applications. Furthermore, $\theta_h^j \geq \theta_l^j$ by assumption, and so $\theta_h^j \geq GI_{\theta_h}(j')$. Therefore, $s_{\theta_h} \leq s_{\theta_l}$. ■

Proof of Proposition 6 For any market, $\mathcal{M}$, and profile of payment functions $\{\phi_j\}_{j\in\mathcal{F}}$, observe that a new market, $\mathcal{M}_\phi$, can be defined to incorporate payment functions into match values.

Definition. *Let market $\mathcal{M} = (\mathcal{F}, \mathcal{C}, m)$ be given.*

Then, market $\mathcal{M}_\phi = (\mathcal{F}', \mathcal{C}', m')$, where:

1. $\mathcal{F}' = \mathcal{F}$,
2. $\mathcal{C}' = \cup_{c\in\mathcal{C}}\{\theta' = (\theta_w^1 + \phi_1(\theta), \theta_f^1 - \phi_1(\theta), \dots) | \theta \in c\}$,
3. $m'(\theta') = m(\theta)$.

Then, observe that $\mathcal{M}_\phi$ satisfies all the assumptions of the model. Limited liability implies that match values are always above zero. Furthermore, because wages could only be conditioned on match value, not type, the types in $\mathcal{M}_\phi$ satisfy independence conditional on a worker's class. Then, Proposition 3 shows that aggregate demand in $\mathcal{M}_\phi$ satisfies gross substitutes. ■

B Extensions Satisfy Gross Substitutes

This section shows that the bandit structure ensures that gross substitutes is satisfied in more general settings. To do so, I build on Lemma 6 to extend the model in a natural manner. In each instance, the set of firms remains a simple family of multi-armed bandits, and furthermore, worker's Gittins indices are decreasing in each firms' thresholds. Then, Lemma 6 implies that gross substitutes is satisfied, and the tâtonnement algorithm of Theorem 1 can be used to characterize an equilibrium.

B.1 Gradual Learning

Throughout the paper so far, it has been assumed that workers learn immediately—a type-θ worker that is hired by firm j learns θ_w^j. In many scenarios, learning requires time, or is noisy.

Here, I show that in markets where workers receive a noisy signal of their match value, aggregate demand still satisfies gross substitutes.

Noise is parametrized through a normal distribution. When a type-θ worker, i, is hired by firm j in period t, i observes her utility from the match, which is given by $\theta_w^j + \epsilon_i(t)$. $\epsilon_i(t)$ is drawn iid each period from a normal distribution, $\epsilon_i(t) \sim N(0, \xi)$, where $\xi > 0$.

Lemma 12. *Aggregate demand satisfies gross substitutes when learning is gradual.*

Proof of Lemma 12

To begin, I note that workers face a simple family of multi-armed bandit processes. To see why, note that for any firm j, associated thresholds, (v_j, p_j), and type-θ; the expected payoff for type-θ from applying to j is equivalent to type-θ's expected payoff from applying to j when learning was instant.

Then, it remains to be shown that $GI_\theta(j, h_t^\theta)$ is decreasing in (v_j, p_j). A simple interchange argument proves the point. Suppose the optimal stopping solution to $GI_\theta(j, h_t^\theta)$ yielded a larger value for some $(v'_j, p'_j) > (v_j, p_j)$. Then, utilize the stopping solution for (v'_j, p'_j) in place of the original stopping solution for (v_j, p_j), with one key difference. Observe that the value from the new solution, under (v_j, p_j) can be decomposed into two terms, one corresponding to matches would have been received under (v'_j, p'_j), and a second corresponding to the value for the additional matches due to the difference between (v_j, p_j) and (v'_j, p'_j). Then, possibly through garbling the original stopping solution, a new stopping solution can be characterized that mimics the original stopping solution for matches above (v'_j, p'_j) on average, while treating matches between (v'_j, p'_j) and (v_j, p_j) as if they had led to rejection. This new stopping solution's value can then be decomposed into two terms, one of which is equal to the original stopping solution's value, and the second which includes the benefit from matching for a single period, and is therefore positive.

Last, Lemma 6 implies that individual demand satisfies gross substitutes, and therefore aggregate demand satisfies gross substitutes. ■

B.2 Learning Through Interviews

Similar to the previous result, when workers learn at the interview stage, the gross substitutes condition holds. Formally, when a type-θ worker applies to firm j, the worker learns θ_w^j, regardless of whether she is hired.

Lemma 13. *Aggregate demand satisfies gross substitutes when workers learn through interviewing.*

Proof of Lemma 13

Learning through interviewing simplifies the previous constructions. Previously, when a type-θ worker applied to a firm with threshold (v_j, p_j), if $\theta_w^j > v_j$ she learned as much. If instead $\theta_w^j \leq v_j$, then she may have been unable to tell whether $theta_w^j = v_j$ and she was unlucky or whether $\theta_w^j < v_j$. Now, she faces a simple family of multi-armed bandit processes where the reward from each bandit j is the realization of Ψ_θ^j in expectation. Then, by construction, $GI_\theta(j, h_t^\theta)$ is decreasing in (v_j, p_j) and therefore Lemma 6 implies that individual demand satisfies gross substitutes. It follows that aggregate demand satisfies gross substitutes as well. ■