Managing congestion in matching markets*

Nick Arnosti
Columbia Business School, Columbia University, New York, NY, 10027, n.arnosti@columbia.edu

Ramesh Johari
Department of Management Science and Engineering, Stanford University, Stanford, CA 94305, rjohari@stanford.edu

Yash Kanoria
Columbia Business School, Columbia University, New York, NY, 10027, ykanoria@columbia.edu

Participants in matching markets face search and screening costs when seeking a match. We study how platform design can reduce the effort required to find a suitable partner.

We study a game-theoretic model in which "applicants" and "employers" pay costs to search and screen. An important feature of our model is that both sides may waste effort: some applications are never screened, and employers screen applicants who may have already matched. We prove existence and uniqueness of equilibrium, and characterize welfare for participants on both sides of the market.

We identify that the market operates in one of two regimes: it is either *screening limited* or *application limited*. In screening-limited markets, employer welfare is low, and some employers choose not to participate. This occurs when application costs are low and there are enough employers that most applicants match, implying that many screened applicants are unavailable. In application-limited markets, applicants face a "tragedy of the commons" and send many applications that are never read. The resulting inefficiency is worst when there is a shortage of employers. We show that simple interventions – such as limiting the number of applications that an individual can send, making it more costly to apply, or setting an appropriate market-wide wage – can significantly improve the welfare of agents on one or both sides of the market.

Key words: Matching markets, decentralized, market design, operational interventions, stochastic model, mean field limit, search frictions, equilibrium, dynamics, contraction.

1. Introduction.

Driven by advances in information technology, online matching market platforms are revolutionizing how trading partners find each other, learn about each other, and ultimately consummate matches. For example, sites such as LinkedIn, Upwork, and TaskRabbit help match employers and employees; sites such as Match.com and OkCupid allow those seeking a romantic partner to browse profiles of nearby users; and travelers looking for short-term housing options can turn to Airbnb and VRBO.

* The paper **will appear in Manufacturing and Service Operations Management**, special issue on Sharing Economy and Innovative Marketplaces (accepted for publication on Nov 10, 2019).

Despite their prevalence, searching through hundreds or thousands of listings on these platforms and contacting promising options remains a costly and time-consuming process. The challenge of finding a partner is exacerbated by the fact that each user's demand for matches typically fluctuates over time. A worker on an online labor market may find themselves temporarily overwhelmed with other projects, or suddenly in need of extra work to pay the month's expenses. An OkCupid user might fail to respond to a message while on vacation, or due to a newly-formed relationship. A listing on Airbnb may be unavailable for a particular date because of a visit from the host's in-laws.

Typically, these changes in a user's status are not immediately reflected on their profile, or even known to the platform. As a result, a great deal of effort may be wasted screening profiles of users who are currently unavailable. Recent empirical work suggests that lack of information about user availability is prevalent, and that this fact has significant welfare consequences.[1]

Motivated by this observation, we develop a stochastic game theoretic model of an asynchronous dynamic two-sided matching market, and use this model to investigate operational interventions that can relieve congestion. We refer to the two sides as "applicants" and "employers". Our model exhibits several key features: (i) it is *costly to apply* for positions; (ii) it is *costly to screen* applicants; (iii) employers are *uncertain* about which applicants are *available*; and (iv) some employers make *multiple offers* to attain a successful match.

We note that our model does not include the process of endogenous price or wage formation.[2] Thus the results of our work are most directly well suited to markets where such price formation processes do not play a pivotal role, e.g., online heterosexual dating markets (where men are typically expected to "apply"). Despite the lack of endogeneous price formation in our formal model, we believe the qualitative insights of our paper are relevant to a wide range of online markets, as we discuss further below.

Our results reveal that equilibrium outcomes in our model are inefficient for both applicants and employers. Our paper is principally concerned with (1) identifying regimes where

[1] Fradkin [12] shows that on Airbnb, 49% of inquiries are rejected or ignored by the host, and only 15% of inquiries eventually lead to a transaction. This work concludes that one of the most common reasons for rejection is a "stale" (or unavailable) listing, and that an initial rejection decreases the probability that the guest eventually books *any* listing by 50%. Horton [17] finds a similar phenomenon on an online labor market.

[2] As noted above, prior academic work establishes that uncertainty about availability is a significant challenge in markets with endogenous prices, and we believe it to be at least as much of a concern in platforms where prices cannot be adjusted to signal availability.

this inefficiency is especially acute, and (2) quantifying the benefits of simple operational interventions. Below, we summarize our main contributions.

Analysis of equilibrium outcomes. In Section 5 we give a complete characterization of the equilibrium of our model. We demonstrate that the market falls in one of two regimes: it is either *screening-limited* or *application-limited*, as described below.

Screening-limited markets feature low application costs (so applicants send many applications) and enough employers that most applicants can eventually match. However, employers find that many offers are rejected by applicants who have matched elsewhere; thus employers must screen many candidates in order to hire successfully, and these screening costs offset the benefits of matching. When employers are ex-ante homogeneous, the costs are high enough to drop employer welfare to zero, and some employers choose to exit the market rather than continuing their search. If employers are heterogeneous, then the effects on their welfare are heterogeneous as well, with the most selective employers obtaining zero welfare.[3]

Application-limited markets feature low screening costs. Although employers often do well in such markets, applicants face a "tragedy of the commons," as they send many applications competing for positions. This wasteful competition has severe welfare consequences when there is a shortage of employers (implying that not all applicants can match).

In short, equilibrium welfare is low for one or both sides of the market unless the market satisfies two conditions: screening costs must be low (relative to application costs), and there must be enough employers for nearly all applicants to match.

Intervening to improve welfare. We show that simple platform interventions can significantly improve upon equilibrium outcomes. First, we show that *restricting the number of applications* that participants are allowed to send can provide substantial welfare benefits to employers in screening-limited markets, and to applicants in application-limited markets with a shortage of employers. Furthermore, gains by one side typically come at little to no cost to the other side. In fact, in screening-limited markets, an application limit can simultaneously benefit both sides of the market.

[3] This may help to explain why participants in many online marketplaces leave the platform before finding a match. For example, Fradkin [12] finds that on Airbnb, hosts rejecting proposals from guests can cause "searchers to leave the market although there are potentially good matches remaining", and Horton [17] finds that on the online labor market oDesk, when a recruited potential worker turns down the employer, "that [employer] is substantially less likely to form a match, despite having access to many substitute workers."

We also consider two alternatives to explicitly limiting applications: (i) raising application costs, or (ii) setting a market-wide price (i.e., a wage). First, the platform could make it more costly to apply, for example by charging a fee for each application or requiring the applicant to complete an extensive application form for each job. Naively, making applications more costly would seem to increase the burden on applicants, and indeed, we find that (unless payments are redistributed or application forms provide information that reduces screening costs) higher application costs lower applicant welfare. However, this approach can often improve employer welfare – in fact, Theorem 6 states that from an employer's perspective, raising application costs is equivalent to imposing an application limit. Therefore this may be a desirable intervention if limiting applications is infeasible, or as a way for the platform to monetize its services.

Alternatively, the platform could set prices in the market. Some platforms recommend prices to their sellers (e.g., Airbnb provides price recommendations to hosts [35]); as these services gain adoption, the platform is effectively setting prices through such recommendations. In our formal analysis, we suppose that on any successful match, the employer is required to pay this price to the matched applicant. We study the potential congestion-alleviating effects of such an intervention. We find that a platform-controlled wage is in some ways at least as powerful as an application limit: for any application limit, an appropriately set wage can recover the same aggregate welfare in equilibrium. This suggests that price-setting by the platform may be a viable alternative to setting application limits, although the latter may lead to better outcomes for applicants (a point that we discuss further in the conclusion).

Mean field analysis. A model capturing the features described above can quickly become intractable, due to complex agent strategies and stochastic fluctuations. A common technique in the literature (see Section 2) is to study such models in a large market ("mean field") regime, where agents respond to others' *average* behavior, which is itself predictable due to the law of large numbers. Whereas typically this approximation is made without rigorous mathematical justification, we formally prove that our mean field assumptions hold asymptotically in large markets, via a novel stochastic contraction argument.

We conclude by returning to the broader managerial consequences of our work. As noted above, our model does not include endogenous price negotiation by market participants, an important feature of a wide variety of online markets. Nevertheless, we feel the main qualitative insights of our work are relevant more broadly: the presence of congestion inefficiency for one or both sides of the market in equilibrium induced by a lack of availability information, and the ability to address this inefficiency through operational interventions such as application limits, application costs, and centrally set market-wide prices. Some evidence of the importance of this issue is that online labor markets have implemented features to overcome precisely the type of congestion effect described here.[4] A formal model including endogeneous price formation leads to substantial technical challenges beyond the scope of this work; understanding the operational interventions studied here within such a model remains an important direction for future research.

The remainder of the paper is organized as follows. In Section 2, we discuss related work. In Section 3, we present our model, and in Section 4, we discuss a formal mean field model that represents a large market limit of our original model. All our equilibrium analysis proceeds within this mean field model, and is presented in Section 5. We give analytical expressions for applicant and employer welfare in equilibrium, and demonstrate that simple operational interventions may significantly improve welfare. In Section 6, we present formal justification for using our mean field model to study markets of finite size. We conclude in Section 7.

2. Related work.

Our work relates to several strands of literature. We discuss each in turn.

2.1. Frictions in labor markets.

Our model closely resembles those used to study labor markets with costly search. This line of work was initiated by Diamond [9, 10], Mortensen [27, 28] and Pissarides [29, 30], and has since received a great deal of attention. See Rogerson et al. [31] and Wright et al. [34] for helpful surveys. Much of this literature focuses on the question of whether equilibrium outcomes are (constrained) efficient: that is, whether agents make socially optimal decisions. Broadly speaking, the conclusion is that outcomes are inefficient unless employers compete by publicly posting wages (this is called "directed search") *and* there are no congestion externalities among employers.

In contrast to the search literature, our focus is not on whether equilibrium outcomes are precisely optimal: the fact that equilibrium is inefficient in our model is unsurprising, given the absence of wages and presence of congestion externalities in our model. Instead, we take

[4] As an example, the online labor markets Elance and oDesk, which merged to form Upwork, used to have limits on the number of applications freelancers could send in a fixed time period. More recently, Upwork has moved to a system with paid Connects [11], to address congestion inefficiency.

a more operational perspective, by identifying conditions under which inefficiency is large, and studying simple interventions that can improve welfare.

One of our key findings – that imposing an application limit on workers may dramatically increase employer welfare – does not, to our knowledge, have any analogue in the search literature; the search literature has traditionally not focused on such operational questions. Search models in prior work would yield very different conclusions about the effects of such an intervention, as we now detail. Our finding is driven by three of the key modeling features mentioned in our introduction: screening is costly, employers are uncertain about availability, and employers may make multiple offers if their first offer is rejected.[5] Existing models in the search literature each lack one or more of these features. For example, Moen [26], Julien et al. [21], Burdett et al. [6] assume that applicants contact a single employer, and thus are certain to be available; Albrecht et al. [1] and Galenianos and Kircher [13] allow employers to make only one offer, so that the competition among employers is never too severe; [6] and Kircher [22] assumes that it is costless for employers to evaluate and make offers. As a result, the models in these papers would predict that an application limit offers little to no benefit to employers. We elaborate on this point, and especially the comparison to Albrecht et al. [1] and Galenianos and Kircher [13], in Appendix B.

2.2. Signaling in matching markets.

The phenomenon of congestion—i.e., that it is impossible or costly for employers to make offers to many applicants and that employers compete for the same applicants—is discussed in Roth and Xing [32]. In such cases, a reasonable design choice is to enable applicants to signal their interest to employers; see, e.g., Lee et al. [25] and Coles et al. [7] for applications in online data markets and the academic job market, respectively, and Coles et al. [8] and Halaburda et al. [16] for examples of how signaling can improve equilibrium outcomes.

Within the context of these papers, our applications can be seen as costly signals of interest. When application costs are small, applicants "signal" (apply to) many employers,

[5] Costly screening and uncertain availability imply that there is a "congestion externality" among employers. As we argue in Appendix B, when employers make multiple offers this externality is much more severe than if employers made at most one offer. Correspondingly, the benefit of an application limit (which reduces this congestion externality) is small or nonexistent in models where employers make only a single offer.

[6] Although Albrecht et al. [1] numerically study an extension in which employers can make a second offer if their first is declined, they assume that second-round offers come after first-round offers. This implies that even an employer who is constrained to make a single offer is never harmed by the fact that others make more: the presence of multiple rounds of offers *alleviates*, rather than *exacerbates*, competition among employers. Thus, their model (with or without the extension) would predict that an application limit does not significantly benefit employers.

and employers' offers are often rejected. Like Coles et al. [8] and Halaburda et al. [16], we find that offer acceptance rates rise if the platform limits applicants' ability to signal. However, whereas those papers find that higher acceptance rates result in more matches, we find that limiting applications results in (slightly) *fewer* matches. The reason for the difference is that we assume that employers can make multiple offers, each of which is costly. Thus, low acceptance rates do not reduce the number of matches formed (as they do in prior work), but do result in high search costs (which can be reduced by limiting applications).[7]

Kushnir [23] finds that signaling may have a detrimental effect including a reduction in the number of matches, and possibly in welfare. However, the source of welfare losses in that work is distinct from that in our paper: there the inefficiency is driven by information asymmetry regarding preferences of workers, and signaling can be miscoordinated, reducing the number of matches formed. In our work, agents do not have preference information beforehand, the only way to match is through the (costly) application and screening process, and increasing the number of applications sent weakly increases the number of matches that form.

We remark that the literature on search frictions in labor economics (Section 2.1) generally relies on large market assumptions that are not formally stated or justified. One exception is the recent work of Galenianos and Kircher [14], which does provide a convergence proof. However, they consider a static setting in which each applicant contacts a single firm, and acknowledge that their analysis does not easily generalize to the case where applicants contact multiple firms.

2.3. Operation of two-sided markets.

There is an emergent literature on the design of two-sided platform markets in the operations literature, that includes the current paper. The overall goal may be stated as that of optimally managing the inventory of a marketplace, where the market participants themselves constitute the inventory, and are also the ones whose needs the market must meet. The work of Allon et al. [2] is similar in spirit to our own, showing that in service marketplaces, operational efficiency may hurt market efficiency due to the involvement of strategic agents. Iyer et al. [19] and Balseiro et al. [5] also apply mean field analysis to answer market design

[7] An additional consequence of modeling screening as costly and endogenous is our finding that firm welfare may fall severely, e.g., all the way to the value of the outside option when firms are homogeneous – a conclusion that does not hold when screening is costless.

questions. (For an introduction to mean field games, see [24, 33].) In other contexts, operational optimization of dynamic matching markets has been carried out without a strategic model; see, e.g., Hu and Zhou [18], Ashlagi et al. [4], Anderson et al. [3], Gurvich et al. [15].

2.4. Empirical evidence from platform markets.

Recent empirical work by Fradkin [12] and Horton [17] concludes that on both Airbnb and oDesk, it is common for users to contact unavailable partners, and that this fact has significant welfare consequences. On AirBnB, guests are uncertain about host availability because transactions take time to complete, and hosts may not reliably update the calendars. Fradkin [12] notes that only 15% of inquiries eventually lead to a transaction, and that an initial rejection decreases the probability that the guest eventually books *any* listing by 50%. On oDesk, potential employers can browse profiles and invite workers to apply, but in practice, the best workers are often too busy to accept the invitation. Horton [17] concludes that employers are often unable to discern which workers are busy, that this fact causes many invitations to be declined, and that spurned employers are much less likely to eventually match.

3. The Model.

In the market we consider, employers and applicants arrive, interact with each other, and eventually depart. Informally, we aim to capture the following behavior:

1. Employers arrive to the market, and each posts an *opening*.
2. When applicants arrive, each *applies* to a subset of the employers currently in the market.
3. Upon exit from the system, employers may screen candidates who have applied to learn whether they are *compatible* with the job. They may make offers to compatible applicants.
4. Whenever an applicant receives an offer, he chooses whether to accept it.

We model a dynamic market, in which participants are posting and applying for jobs, screening, and making offers asynchronously. The timing of this market is described in more detail below.

3.1. Arrivals.

Our market starts empty at $t = 0$. Our dynamic markets are parameterized by $n > 0$, which describes the market size. Individual employers arrive at intervals of $1/n$, and applicants arrive at intervals of $1/(rn)$. Here $r > 0$ is a parameter that controls the relative magnitude of the two sides of the market. Employers remain in the system for a unit lifetime. Applicants depart the system according to a process that we describe below. We endow employers and applicants with unique IDs, which convey no other information about the agent.

Upon arrival, employers post an opening. They do not make any decisions at this time. Upon arrival, each applicant selects a "search intensity" $m_a \in [0, n]$, and applies to each employer currently in the system with probability m_a/n. Note that for all $t \geq 1$, there are exactly n employers in the system, and thus the expected number of applications sent by an applicant who arrives after time $t \geq 1$ is[8] m_a.

3.2. Applicant departure.

Applicants remain in the system for a maximum of one time unit. Whenever the applicant receives an offer from any employer (as described below), the applicant must choose immediately whether to accept it. If they accept, they immediately and automatically depart from the market at that time. Excluding application costs, an applicant earns a payoff of 1 as long as she is matched and zero otherwise. Thus, *it is a dominant strategy for applicants to accept the first offer they receive.* We assume henceforth that applicants follow this strategy.

3.3. Employer departure.

Each employer stays in the system for one time unit, and then departs. At the time of departure, the employer sees the set of applicants to her opening. Initially, she does not know which of these applicants are compatible for her job, nor does she know which would accept the job if offered it (i.e. which applicants remain in the system). The employer takes a sequence of "screening" and "offer" actions, instantaneously learning the result of each, until an offer she makes is accepted (causing her to exit) or she chooses to exit the market.

At each stage of the employer's sequential decision process, she may screen any unscreened applicant, thereby learning whether this applicant is compatible for the job. If the employer has yet to match, she may also make an offer to any applicant whom she has found to be compatible (the applicant responds immediately to the offer; as mentioned above, we assume that they accept it if and only if it is the first offer that they receive). Employers are also allowed to skip some or all applications and exit the market at any point.

We require that before making an offer to any applicant employers must screen this applicant and find him compatible (this assumption is justified if, for example, the cost of hiring

[8] We could instead allow applicants to directly select the number of employers to whom they apply. We chose a probabilistic specification primarily for technical convenience; we further discuss this point in Section 4.

an unqualified worker is sufficiently high). We assume that that each employer-applicant pair is compatible with probability β (independently across all such pairs), and that this is common knowledge.[9]

3.4. Utility.

If a compatible pair matches to each other, the employer earns v and the applicant earns w. We normalize $v = w = 1$, which is without loss of generality because we do not compare absolute welfare of agents on opposite sides of the market throughout most of the paper.[10] Applicants pay a cost c_a for each application that they send, and employers pay a cost c_s for each applicant that they screen. The net utility to an agent will be the difference between value obtained from any match, and costs incurred, and agents act so as to maximize their expected utility.

Assumption 1. $\max(c_a, c_s) < \beta$.

This assumption rules out the uninteresting case where costs are so high that no activity occurs in the marketplace. If $c_s \geq \beta$, then regardless of applicant behavior, it would be optimal for employers to exit the market rather than screening. Similarly, if $c_a \geq \beta$, then because employers hire only compatible applicants, no applicant strategy can earn positive surplus.

For later reference, it will be useful to consider *normalized* versions of the screening and application costs, given by

$$c'_s = c_s/\beta, \qquad c'_a = c_a/\beta. \tag{1}$$

4. The large market: A stationary mean field model.

In principle, the strategic choices facing an agent in the model described above may be quite complicated. Consider the case of an employer who knows that he has only one competitor. If he finds that one applicant has already accepted another offer, he learns that every other applicant is still looking for a job. Similar logic suggests that in thin markets, information

[9] In fact, we can relax this assumption: all of our results hold if we only assume that *conditional on applicant a having applied to an employer e*, e finds a acceptable with probability β. In particular, this allows the possibility that the parameter β is not necessarily exogenous to the platform: instead, it can be interpreted as the quality of the platform's recommendation and/or search algorithms, i.e., improved search and recommendation algorithms simply increase the value of β.

[10] The exception is Section 5.4, where we allow the platform to set the wage w. There we define $v = 2 - w$ (meaning that the employer utility is the match surplus minus the wage), and we *do* consider the total welfare across the two sides of the platform.

revealed during screening may induce significant shifts in the employer's beliefs. This could conceivably cause optimal employer behavior to be quite complex.

As the market thickens, however, one might expect that the correlations between agents on the same side of the market become weak. In particular, employers screening applicants from among a large number of workers might reasonably assume that learning that one applicant has already accepted an offer does not inform them about the availability of other applicants. Further, if the employer cannot distinguish individual applicants, each one should appear to be available with equal probability. Similarly, applicants who know nothing about individual employers may be justified in assuming that each of their applications convert to offers independently and with equal probability p.

In this section we develop a formal stationary *mean field* model for our dynamic matching market, and introduce a notion of game-theoretic equilibrium for this model; in particular, we study a model that arises from a limiting regime where the market thickens.

In our formal model, agents make the following assumptions.

Mean Field Assumption 1 (Employer Mean Field Assumption) *Each applicant in an employer's applicant set is available with probability q, and availability in the applicant set is independent across applicants.*

Mean Field Assumption 2 (Applicant Mean Field Assumption) *Each application yields an offer with probability p, independently across applications to different openings.*

Mean Field Assumption 3 (Large market assumption) *The number of applications sent by an applicant who chooses $m_a = m$ is Poisson distributed with mean m. If all applicants select $m_a = m$, the number of applications received by each employer is Poisson distributed with mean rm.*[11]

[11] Recall that our applicants apply to each employer independently with probability m/n. Allowing applicants to directly choose how many applications to send complicates our analysis of the finite system. However, in the mean field, this alternative model remains quite tractable: this simply changes the probability of receiving an offer from $1 - e^{-mp}$ to $1 - (1 - p)^m$. As we focus on settings in which application costs are small (and thus m is large and p is small), this does not substantively change our results.

We make the additional simplifying assumption that applicants cannot observe the number of other applicants to each job. If this assumption were relaxed, then in the mean field analysis each employer would receive rm applications. The probability that an employer matches in this model would become $1 - (1 - \beta q)^{rm}$ (as opposed to $1 - e^{-rm\beta q}$ in our model). Again, this does not substantively change our results.

Under these assumptions, optimal agent behavior simplifies greatly. We describe optimal employer and applicant responses in Section 4.1. For applicants, we show that there exists a unique optimal choice of m, given p; and for employers, we show that given q, their optimal response is either to employ a simple sequential screening strategy or to exit immediately (they may also randomize when indifferent between these options).

Of course p and q are not given exogenously, but rather arise endogenously from the choices made by agents. In Section 4.2, we derive "consistency checks" that p and q should satisfy, if they indeed arise from the conjectured employer and applicant strategies.

The work in Sections 4.1 and 4.2 allows us to define a *mean field equilibrium* (MFE) in Section 4.3. Informally, a mean field equilibrium consists of strategies for employers and applicants that are best responses to the stationary market dynamics that they induce. We prove that there exists a *unique* MFE. We conclude with Section 4.4, which discusses a simple intervention available to the market operator: placing a limit ℓ on the value m_a chosen by each applicant. We show that in this setting too, there exists a unique MFE.

Importantly, we prove that our mean field model is in fact the correct limit of our dynamic market as the thickness n grows. In particular, Theorems 8 and 9 in Section 6 justify our study of MFE: they state that the mean field assumptions hold as n approaches infinity, and that as a consequence any MFE is an approximate equilibrium in the game with finite but sufficiently large n.

4.1. Optimal decision rules.

We first study how agents respond when confronted with a world where the mean-field assumptions hold.

4.1.1. Applicants. As discussed in Section 3, it is a dominant strategy for applicants to accept the first offer (if any) that they receive, and we assume applicants follow this rule. Therefore the only decision an applicant a needs to make on arrival is her choice of m_a, the expected number of applications sent.

An applicant a who chooses $m_a = m$ incurs an expected cost of $c_a \cdot m$. If the applicant applies to Poisson(m) employers, and each application independently yields an offer with probability p, then at least one offer is received—i.e., the applicant matches to an employer—with probability $1 - e^{-mp}$. Thus, the expected payout of an applicant in the mean-field environment who selects $m_a = m$ is

$$W_a(m,p) = 1 - e^{-mp} - c_a m. \tag{2}$$

Applicants choose $m \geq 0$ to maximize this payoff. Because their objective is strictly concave and decays to $-\infty$ as $m \to \infty$, this problem possesses a unique optimal solution identified by first-order conditions. If $p \leq c_a$, the optimal choice is $m = 0$. Otherwise, applicants select $m = \frac{1}{p} \log\left(\frac{p}{c_a}\right)$.We define $\mathcal{M}$ to be the function that maps p to the unique optimal value of m:

$$\mathcal{M}(p) = \begin{cases} 0, & \text{if } p \leq c_a; \\ \frac{1}{p} \log\left(\frac{p}{c_a}\right), & \text{if } p > c_a. \end{cases} \tag{3}$$

4.1.2. Employers. Next, we consider the optimal strategy for employers, when Mean Field Assumption 1 holds. We consider a simple strategy, which we denote ϕ^1. An employer playing ϕ^1 sequentially screens candidates in her applicant list. When she finds a compatible applicant, she makes an offer to this candidate; otherwise, she considers the next candidate. This process repeats until one applicant accepts or no more applicants remain.

The optimal strategy for the employers is straightforward to characterize. First suppose an employer has exactly one applicant. The employer will prefer to screen the applicant if $\beta q - c_s > 0$, i.e., if $q > c_s'$; exit if $q < c_s'$; and is indifferent if $q = c_s'$. Now it is clear that if an employer has more than one applicant in her list, since all applicants are *ex ante* homogeneous from the perspective of the employer, the same reasoning holds: the employer will screen or exit immediately according to whether q is larger or smaller than c_s', respectively. (Note the essential use of Mean Field Assumption 1: if there is correlation in the availability of successive applicants in the employer's list, the preceding reasoning no longer holds.) The following proposition (proved in Appendix C) summarizes the preceding discussion.

Proposition 1. *Let ϕ^1 be the strategy of sequentially screening applicants, offering them the job if and only if they are qualified, until either an applicant is hired or no more applicants remain. Then ϕ^1 is uniquely optimal if and only if $q > c_s'$, exiting immediately is uniquely optimal if and only if $q < c_s'$, and any mixture of these strategies is optimal if $q = c_s'$.*

Motivated by this proposition, we define ϕ^α to be the strategy that plays ϕ^1 with probability α and exits immediately otherwise. Define the correspondence $\mathcal{A}(q)$ by:

$$\mathcal{A}(q) = \begin{cases} \{0\} & \text{if } q < c_s' \\ [0,1] & \text{if } q = c_s' \\ \{1\} & \text{if } q > c_s'. \end{cases} \tag{4}$$

This correspondence captures the optimal employer response, as described in Proposition 1, so that $\mathcal{A}(q) = \{\alpha \in [0,1] : \phi^\alpha$ is optimal for the employer, given $q\}$.

We define

$$W_e(\alpha, m, q) = \alpha(1 - e^{-rm\beta q})(1 - c_s'/q) \tag{5}$$

to be the expected welfare of an employer who screens with probability α, receives a number of applications that is Poisson with mean rm, and finds each applicant qualified with probability β and (independently) available with probability q.

4.2. Consistency.

In the previous section, we discussed the best responses available to employers and applicants when the mean field assumptions hold; that is, given p and q, we found the strategies that agents would adopt. However, p and q are clearly *determined* by agent strategies. In this section we identify consistency conditions that p and q must satisfy, given specified agent strategies.

We focus on strategies that could conceivably be optimal, as identified in the preceding section. We assume that all applicants choose the same $m \geq 0$, and that all employers play ϕ^α, i.e., they play ϕ^1 with probability α and exit immediately otherwise. From any m and α, we derive a unique prediction for the pair (p, q).

We emphasize at the outset that our analysis aims only to derive the correct consistency conditions under the mean field assumptions. We provide rigorous justification for these assumptions via the propositions in Section 6. As a consequence, those propositions also justify the consistency conditions described below.

We start by deriving a consistency condition for q, given p and the strategy adopted by applicants. Intuitively, q should be equal to the long-run fraction of offers that are accepted. Fix the value of m chosen by applicants, and let X be the number of offers received by a single applicant. This applicant will accept an offer if and only if $X > 0$, so the expected fraction of offers that are accepted is $P(X > 0)/E[X]$. If Mean Field Assumptions 2 and 3 hold, then X is Poisson with mean mp, so we should have

$$q = \frac{1 - e^{-mp}}{mp}. \tag{6}$$

To derive a consistency condition for p, note that when employers follow ϕ^α, only compatible applicants receive offers. Thus, p should equal β times the long-run fraction of qualified applications that result in offers. Because an applicant's availability does not influence whether they receive an offer (as it is unobserved by prospective employers), this should be equal to β times the fraction of applications by qualified *available* applicants that result in offers. Fix an employer playing ϕ^α, and let Y be the number of qualified, available applicants received by this employer. This employer successfully hires if and only if $Y > 0$ and she decides to screen. Thus, the fraction of qualified available applicants who receive offers should be $\alpha P(Y > 0)/E[Y]$. Because each applicant is qualified with probability β, and because available and unavailable applicants receive offers with the same probability (as employers do not observe availability), we conclude that the ex-ante probability that an applicant receives an offer from a given employer is $\alpha\beta P(Y > 0)/E[Y]$. Mean Field Assumptions 1 and 3 jointly imply that Y is distributed as a Poisson random variable with mean $rm\beta q$, from which we conclude that

$$p = \alpha\beta\frac{P(Y > 0)}{E[Y]} = \alpha\beta\frac{1 - e^{-rm\beta q}}{rm\beta q}, \tag{7}$$

The equations (6) and (7) are a system for p and q, given the values of m and α (as well as the parameters r and β). The following proposition states that the pair of consistency equations (6) and (7) have a unique solution (the proof is in Appendix E.1).

Proposition 2. *For fixed m, α, r, and β, there exists a unique solution (p, q) to* (6) *and* (7).

We refer to the unique pair (p, q) that solve (6) and (7) as a *mean field steady state* (MFSS). This pair provides a prediction of how a large market should behave, given specific strategic choices of the agents. For later reference, given strategies m and α (and parameters r and β), let $\mathcal{P}(m, \alpha; r, \beta)$ and $\mathcal{Q}(m, \alpha; r, \beta)$ denote the unique values of p and q guaranteed by Proposition 2, respectively. Because our analysis is conducted with r and β fixed, we will omit the dependence on r and β in favor of the more concise $\mathcal{P}(m, \alpha)$ and $\mathcal{Q}(m, \alpha)$.

4.3. Mean field equilibrium.

In this section we define *mean field equilibrium* (MFE), a notion of game theoretic equilibrium for our stationary mean field model. Informally, a MFE should be a pair of strategies such that (1) agents play optimally given their beliefs about the marketplace, i.e., the values p and q in the mean field assumptions; and (2) agent beliefs are consistent with the strategies being played, i.e., (p, q) is an MFSS corresponding to the agents' strategies. Section 4.1 addressed the first point; and Section 4.2 addressed the second. We define a mean field equilibrium by composing the maps defined in those sections.

Definition 1. A *mean field equilibrium* (MFE) is a pair (m^*, α^*) such that $m^* = \mathcal{M}(\mathcal{P}(m^*, \alpha^*))$ and $\alpha^* \in \mathcal{A}(\mathcal{Q}(m^*, \alpha^*))$.

In an MFE, m^* and α^* are optimal responses (under the mean field assumptions) to the steady-state (p, q) that they induce. For future reference, we define $p^* = \mathcal{P}(m^*, \alpha^*)$ and $q^* = \mathcal{Q}(m^*, \alpha^*)$. Our main theorem in this section establishes existence and uniqueness of MFE (the proof is in Appendix E.2).

Theorem 1. *Fix any r, β, c_a, c_s such that Assumption 1 holds. Then there exists a unique mean field equilibrium (m^*, α^*).*

4.4. A market intervention: application limits.

As noted in the introduction, we are interested in comparing the outcome of the market described above to the outcome when the platform operator intervenes to try to improve the welfare of employers and/or applicants. We consider a particular type of intervention: a limit on the number of applications that can be sent by any individual.

In our model with application limits, agent payoffs are identical to before, as are the strategies available to employers. Applicants, however, are restricted to selecting $m_a \le \ell$. In the corresponding mean field model, given p, applicants choose m_a to maximize $1 - e^{-m_a p} - c_a m_a$ (their expected payoff), subject to $m_a \in [0, \ell]$. The applicant objective is concave in m_a, so this problem has a unique solution given by

$$\mathcal{M}_\ell(p) = \min(\ell, \mathcal{M}(p)). \tag{8}$$

The consistency conditions are identical to those in Section 4.2. We define a *mean field equilibrium* of the market with application limit ℓ as a pair $(m_\ell^*, \alpha_\ell^*)$ solving the following pair of equations:

$$m_\ell^* = \mathcal{M}_\ell(\mathcal{P}(m_\ell^*, \alpha_\ell^*)), \quad \alpha_\ell^* \in \mathcal{A}(\mathcal{Q}(m_\ell^*, \alpha_\ell^*)). \tag{9}$$

The following proposition is an analog of Theorem 1 for the market with an application limit (the proof is in Appendix E.3).

Proposition 3. *Fix r, β, c_a, c_s such that Assumption 1 holds, and let (m^*, α^*) be the corresponding MFE in the market with no application limit. Then for any $\ell \ge 0$ there exists a unique mean field equilibrium in the market with application limit ℓ. If $m^* \le \ell$, then $(m_\ell^*, \alpha_\ell^*) = (m^*, \alpha^*)$. Otherwise, $m_\ell^* = \ell$ and α_ℓ^* is the unique solution to $\alpha_\ell^* \in \mathcal{A}(\mathcal{Q}(\ell, \alpha_\ell^*))$.*

For future reference, we define $p_\ell^* = \mathcal{P}(m_\ell^*, \alpha_\ell^*)$, $q_\ell^* = \mathcal{Q}(m_\ell^*, \alpha_\ell^*)$.

5. Welfare analysis.

In this section, we study the expected welfare of applicants and employers, both in equilibrium as well as under operational interventions such as an application limit. We compare welfare to two simple *frictionless* benchmarks:

- Applicant welfare is at most $\min(1, 1/r)$, because if $r > 1$, then not all applicants can match.
- Employer welfare is at most $\min(1, r)$, because if $r < 1$, then not all employers can match.

These benchmarks are very optimistic, as they completely ignore application and screening costs. However, when costs are small, one might hope to obtain welfare for both sides that is close to these benchmarks. We make Assumption 1 throughout this section.

In Section 5.1, we explicitly characterize employer and applicant welfare in equilibrium, without any operational intervention by the platform. We show that there are two regimes for welfare: one where the market is *application-limited* (welfare is determined by applicants' willingness to apply), and one where the market is *screening-limited* (welfare is determined by employers' willingness to screen). Equilibrium welfare of applicants and employers can fall well short of the benchmarks above, even if costs are small: when the market is application-limited, applicants' welfare can be an arbitrarily small fraction of the benchmark $\min(1, 1/r)$, and when the market is screening-limited, employers' welfare is identically *zero*.

We demonstrate in Section 5.2 that the platform can substantially improve the situation by imposing an application limit. In particular, we establish that when search costs are low, for any one side of the market, an appropriately chosen application limit can bring the welfare of that side close to the corresponding frictionless benchmark discussed above. We also provide evidence that the tradeoff between optimizing for applicants and employers is never too severe, and establish that when the market is screening-limited, then a suitable application limit simultaneously raises welfare for both sides.

In Sections 5.3 and 5.4 we discuss alternative interventions that can be used to mitigate congestion. Section 5.3 studies the effect of imposing additional application costs, while Section 5.4 considers a model in which the platform can set a market-wide price (i.e., a wage) that employers pay to applicants upon a successful match. Raising application costs can significantly improve employers' welfare (exactly as with an application limit), but always weakly reduces applicant welfare. We also show a suitable market-wide wage can achieve identical aggregate welfare to any desirable application limit.

Finally, Section 5.5 explores the generalizability of our findings to models with ex-ante heterogeneity. In it, we consider one extension in which some employers' jobs are harder than others' (lower probability of compatibility β), and another in which some applicants are more skilled (higher probability of compatibility β) than others. Numerical investigation of the solution to these models shows that our key insights from Sections 5.1 and 5.2 continue to apply.

We begin with the following notational preliminaries. Fix parameter values r (the ratio of applicants to employers), β (probability of compatibility), c_a (application cost), and c_s (screening cost). For given applicant strategy m and employer strategy α, we let $\Pi_a(m, \alpha)$ and $\Pi_e(m, \alpha)$ denote the mean field applicant and employer welfare, respectively. These are given by

$$\Pi_a(m, \alpha) = W_a(m, \mathcal{P}(m, \alpha)) \quad = 1 - e^{-m\mathcal{P}(m,\alpha)} - c_a m; \tag{10}$$

$$\Pi_e(m, \alpha) = W_e(\alpha, m, \mathcal{Q}(m, \alpha)) = \alpha \left(1 - e^{-rm\beta\mathcal{Q}(m,\alpha)}\right) \left(1 - \frac{c_s'}{\mathcal{Q}(m, \alpha)}\right), \tag{11}$$

using Eqs. (2) and (5), where we recall that $\mathcal{P}(m, \alpha)$ and $\mathcal{Q}(m, \alpha)$ are respectively the mean field acceptance probability and availability; see Proposition 2 and the subsequent discussion. The expression for $\Pi_a(m, \alpha)$ is the probability of at least one application being successful, less application costs. The expression for $\Pi_e(m, \alpha)$ is the probability of successfully hiring, times the net match value for an employer (match value minus expected screening cost per successful match).

In much of what follows, we make use of a reduced parameter set consisting of $r, c_a' = c_a/\beta$, and $c_s' = c_s/\beta$, as defined in (1). We do so because it turns out that welfare in equilibrium (with or without application limits) depends only on the value of these three parameters.

5.1. Quantifying equilibrium welfare.

We begin by providing analytical expressions for equilibrium welfare of both sides. Define Π_a^*, Π_e^* to be the applicant and employer welfare, respectively, in the unique mean-field equilibrium (m^*, α^*) guaranteed by Theorem 1. In other words,

$$\Pi_a^* = \Pi_a(m^*, \alpha^*), \qquad \Pi_e^* = \Pi_e(m^*, \alpha^*). \tag{12}$$

Examining (10) and (11) suggests that participant welfare Π_a^* and Π_e^* depend on equilibrium strategies m^*, α^*, as well as equilibrium outcomes p^* and q^*. However, the relationships between the values m^*, α^*, p^*, q^* allow us to express Π_a^* and Π_e^* in terms of only model primitives and the equilibrium acceptance probability p^*.[12] Theorem 2 provides these expressions, as well as equations that define p^*.

THEOREM 2. *There are unique solutions $\hat{p}, \overline{p} > c_a$ to the following equations:*

$$r\left(1-\frac{c_a}{\hat{p}}\right) = 1 - e^{-\frac{r\beta}{\hat{p}}\left(1-\frac{c_a}{\hat{p}}\right)} \tag{13}$$

$$1-\frac{c_a}{\overline{p}}+\frac{c_s}{\beta}\log\left(\frac{c_a}{\overline{p}}\right) = 0. \tag{14}$$

In the unique MFE, $p^* = \min(\hat{p}, \overline{p})$, and

$$\Pi_a^* = \quad 1-\frac{c_a}{p^*}+\frac{c_a}{p^*}\log\left(\frac{c_a}{p^*}\right) \tag{15}$$

$$\Pi_e^* = r\left(1-\frac{c_a}{p^*}+\frac{c_s}{\beta}\log\left(\frac{c_a}{p^*}\right)\right). \tag{16}$$

Theorem 2 identifies that the parameter space can be divided into two regions: in one, welfare is determined by $\hat{p}$, and in the other it is determined by $\overline{p}$. The interpretation for this is that $\hat{p}$ is the acceptance probability that arises when all employers choose to screen, and applicants behave according to a corresponding partial equilibrium.[13] If the screening cost c_s is sufficiently small, this outcome will be an equilibrium: we will have $p^* = \hat{p}$. On the other hand, if screening costs are large, then not all employers will choose to screen. In this case, the value p^* will be such that when applicants respond optimally, employers are exactly indifferent about whether to screen – this is the value $\overline{p}$. Motivated by this observation, we introduce the following definition.

DEFINITION 2. The market is *application-limited* if $\overline{p} > \hat{p}$ (in which case $p^* = \hat{p}$), and *screening-limited* if $\overline{p} < \hat{p}$ (in which case $p^* = \overline{p}$).

As we will see, the distinction between application-limited and screening-limited markets plays an important role in determining which side(s) have low welfare in equilibrium. Accordingly, we briefly discuss the parameter regimes that are characteristic of application-limited

[12] The expression for applicant welfare (15) comes from substituting the applicant best response function (3) into (10). The expression for employer welfare (16) comes from applying the consistency conditions (6) and (7) to (11), and then making use of the applicant best response function (3).

[13] We can interpret (13) as follows. Recall that by (3), an applicant facing acceptance probability $p > c_a$ chooses $m = \frac{1}{p}\log(p/c_a)$ and matches with probability $1 - c_a/p$. Thus, the left side of (13) represents the mass of matched applicants. Meanwhile, if employers always screen, then they match whenever they have a qualified available applicant, which occurs with probability of $1 - e^{-rm\beta q^*}$, which equals the right side of (13) using Eq. (6).

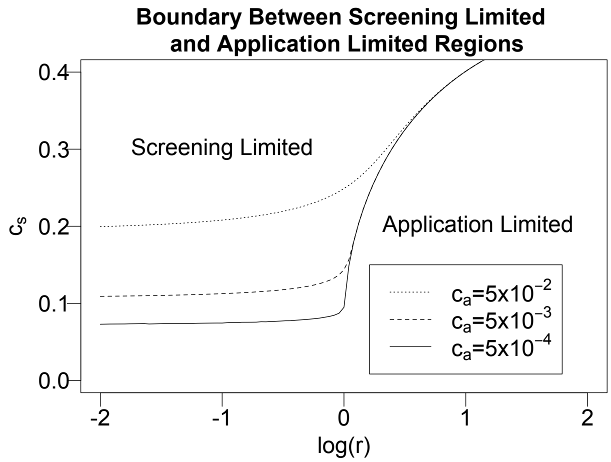

Figure 1 **The boundary between screening-limited (top) and application limited (bottom) regimes, for $\beta = 0.5$ and varying values of c_a. The y-axis gives the screening cost c_s, and the x-axis gives the log of the ratio of applicants to employers. Smaller application costs lead to a larger screening-limited region. For fixed c_s and $r < 1$, the market is screening limited for all sufficiently small c_a.**

and screening-limited markets. Figure 1 depicts the boundary between the application-limited and screening-limited regions, as a function of the model primitives r, c_a', c_s'. In order for the market to be screening limited, applicant availability must be a concern to employers. This occurs when there are more employers than applicants and application costs are small. More precisely, if $r < 1$, then for any screening cost $c_s' > 0$, the market is screening limited for all sufficiently small application costs c_a'. Conversely, if either of these conditions fail – that is, there are more applicants than employers or application costs are moderate – then the market is application limited unless screening costs are very high.[14]

While Theorem 2 gives analytical expressions for the welfare of both sides of the market, these expressions can be difficult to interpret, as they depend on the equilibrium quantity p^*. As we now demonstrate, employer and applicant welfare can fall far short of the frictionless benchmarks, even when search and application costs are small.

Theorem 3.

1. *If the market is screening-limited, then equilibrium employer welfare* $\Pi_e^* = 0$.
2. *If the market is application-limited and* $r > 1$, *then equilibrium applicant welfare*

$$\Pi_a^* \leq \frac{1}{r} \left(1 - (r-1) \log\left(\frac{r}{r-1}\right)\right).$$

[14] In particular, when $r > 1$ a necessary condition for the market to be screening limited is that $c_s' \geq \left(r \log(\frac{r}{r-1})\right)^{-1}$. Even for $r = 1.05$, this requires c_s' to be above 0.31; for $r = 1.25$, c_s' must be approximately 0.5. A second necessary condition for the market to be screening limited is that $c_s' \geq (1 - c_a')/\log(1/c_a')$. For $c_a' = 0.05$, this implies that c_s' must be above 0.31.

Theorem 3 states that when the market is screening-limited, employer welfare is zero (indeed, this a direct consequence of the definition of screening-limited: if some employers choose not to participate, they must anticipate zero welfare from the market). Meanwhile, when $r > 1$ and the market is application-limited, applicant welfare in equilibrium is substantially below the frictionless benchmark level $1/r$: the bound in Theorem 3 implies that applicant welfare is at most half of the upper bound of $1/r$ when $r = 1.4$, at most one third of $1/r$ when $r = 1.9$, and becomes a vanishingly small fraction of $1/r$ as r increases – regardless of application costs! In other words, at least one side has low welfare in equilibrium unless the market is application limited and there is a shortage of applicants ($r < 1$).[15]

5.2. Improving welfare by limiting applications.

As we now show, an application limit is a powerful operational tool for improving participants' welfare.

We let Π_a^ℓ and Π_e^ℓ be expected applicant and employer payoffs in the unique equilibrium of the market with application limit ℓ. In other words,

$$\Pi_a^\ell = \Pi_a(m_\ell^*, \alpha_\ell^*), \qquad \Pi_e^\ell = \Pi_e(m_\ell^*, \alpha_\ell^*). \tag{17}$$

The intuition for the benefits of an application limit is as follows. When the market is screening-limited, employers are choosing not to screen because the expected cost of screening completely offsets any benefits from matching. By imposing an application limit, the platform ensures that applicants will be more likely to accept offers. This reduces the screening cost required to find an available applicant, and increases employer welfare. Applicants, in turn, always suffer from a "tragedy of the commons" in the no intervention equilibrium: they do not internalize the externality that their applications incur upon others. This has especially severe consequences when the market is application-limited and $r > 1$, implying that a significant fraction of applicants remain unmatched. An application limit can reduce this costly competition. Theorem 4 formalizes these observations.

Theorem 4.

1. *Suppose the market is screening-limited. Then:*

$$\sup_\ell \Pi_e^\ell = r(1 - c_s' + c_s' \log c_s').$$

[15] In this case, both sides do fairly well, and an application limit cannot significantly help, as shown in Figure 2.

2. *Suppose the market is application limited and* $r > 1$. *Then*

$$\sup_{\ell} \Pi_a^{\ell} \geq \frac{1}{r} - \frac{c_a'}{r} + c_a' \log(c_a') \log\left(\frac{r}{r-1}\right).$$

Note that if $r \leq 1$, c_s' is small and c_a' is small enough that the market is screening limited, then the first part of Theorem 4 implies $\sup_{\ell} \Pi_e^{\ell} \approx r$, the frictionless benchmark for employers. Similarly, if $r > 1$ and c_a' is small, then the market is application limited and the second part of Theorem 4 implies that $\sup_{\ell} \Pi_a^{\ell} \approx 1/r$, the frictionless benchmark for applicants.

Figure 2 complements the insights from Theorems 3 and 4. In it, we plot the ratio of equilibrium welfare to the welfare attainable with an application limit, for both sides of the market (and for varying parameter values). The figure demonstrates that an application limit can substantially help one or both sides of the market unless the market is application limited and $r < 1$. The figure displays the screening-limited region where employers get zero welfare, and also shows that an application limit significantly benefits applicants when r is large.

The figure also reveals that an application limit substantially improves applicant welfare when c_s is high (so that the market is screening-limited). The intuition is that in this case, many employers choose not to screen, leaving applicants competing for a small number of positions: the effective market imbalance exceeds one. As a result, the tragedy of the commons is severe. Although an application limit cannot increase the number of matches, it can dramatically reduce the wasteful competition among applicants.

While there may exist choices of ℓ that result in high employer welfare and choices of ℓ that result in high applicant welfare, these choices of ℓ might not coincide. We now address this concern, both theoretically and numerically. Theorem 5 states that whenever the market is screening-limited, it is possible to choose a single application limit such that both employers and applicants are better off than they would be without any intervention.

THEOREM 5. *If the market is screening-limited, there exists* ℓ *such that* $\Pi_e^{\ell} > \Pi_e^*$ *and* $\Pi_a^{\ell} > \Pi_a^*$.

One might wonder whether Pareto improvements are possible when the market is application limited. The answer turns out to be, "not always." Indeed, if either r or c_a' is large enough, then availability never becomes a pressing concern to employers, and *any* binding application limit lowers employer welfare (this holds, for example, if $c_s' < \max\{1 - 1/r, c_a'\}$.).

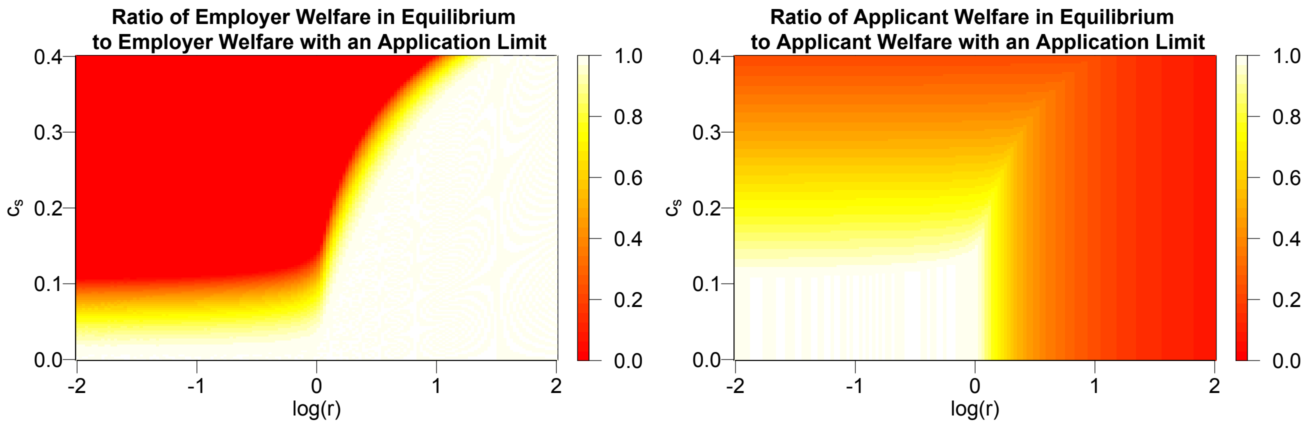

Figure 2 **Ratio of equilibrium welfare to welfare attainable with an application limit. Probability of compatibility $\beta = 0.5$, application cost $c_a = 0.005$, and we vary the screening cost c_s (y-axis) and ratio of applicants to employers r (x-axis). An application limit can substantially help one or both sides of the market unless screening costs are low and there are fewer applicants than employers. See Theorems 3 and 4.**

Even when Pareto improvements are not possible, however, an application limit may substantially improve applicant welfare at little cost to employers. Figure 3 shows that the tension between optimizing for applicants and employers is not too severe. It displays the largest fraction δ such that there exists a single limit ℓ' where employers earn an expected surplus of $\Pi_e^{\ell'} \geq \delta \sup_\ell \Pi_e^\ell$ and applicants (simultaneously) earn an expected surplus of $\Pi_a^{\ell'} \geq \delta \sup_\ell \Pi_a^\ell$. Observe that δ is never below 3/4, and is often much higher. We conjecture that it is always possible to choose a limit such that each side attains at least 3/4 of the welfare that would be obtained by optimizing solely for that side.[16]

5.3. Increasing application costs.

In this section, we study a second intervention that reduces the number of applications sent: raising the application cost. From the perspective of the employer, limiting applications and raising application costs have similar effects. To applicants, however, they look different. Although both interventions reduce the amount of competition for each opening, raising the application cost also directly harms applicants. A priori, it is not obvious which effect dominates. Theorem 6 shows that within our model, raising application costs can only hurt applicants.[17] Additionally, it formalizes the idea that the two interventions are equivalent from the perspective of employers.

[16] Both numerical results and intuition suggest that the wishes of applicants and employers are most at odds when c'_a is small and c'_s is large, or vice versa. We can prove that as $c'_a \to 0$ and $c'_s \to 1$, or as $c'_a \to 1$ with any fixed c'_s, we have $\delta \geq 3/4$.

[17] We assume that the costs paid by applicants cannot be redistributed. This is reasonable if the increased cost is not a monetary transfer, but instead an additional barrier to application (such as additional questions that each

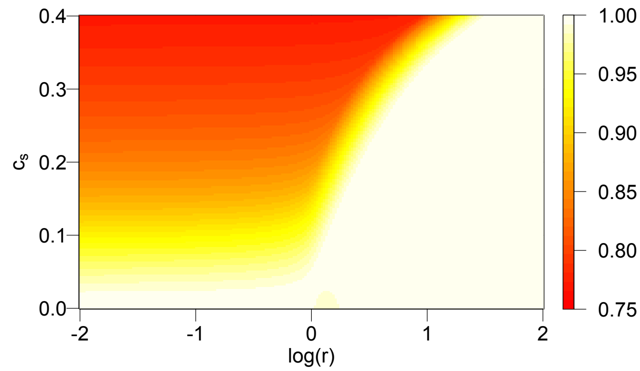

Figure 3 **For $\beta = 0.5$, $c_a = 0.005$ and varying values of c_s (y-axis) and $\log(r)$ (x-axis), this figure shows the largest fraction δ such that there exists a single limit ℓ' for which employers earn an expected surplus of $\delta \sup_\ell \Pi_e^\ell$ and applicants (simultaneously) earn an expected surplus of $\delta \sup_\ell \Pi_a^\ell$. Note that the color scale is different from that in Figure 2, and that $\delta \geq 3/4$ for all depicted parameter values. This suggests that the tension between optimizing for applicants and employers is not too severe.**

Theorem 6. *Fix values of r, c_s, β such that $c_s < \beta$, and fix the application cost $c_a < \beta$ under no intervention. Let $m^*(c)$ denote the equilibrium value of m selected by applicants when application costs are raised to $c \in [c_a, \beta]$, and let $\Pi_a^*(c)$ and $\Pi_e^*(c)$ denote the applicant welfare and employer welfare, respectively. Then (i) $\Pi_a^*(\cdot)$ is weakly decreasing; and (ii) for each $c_a \in (0, \beta)$ and each $\ell \in [0, m^*(c_a)]$, there exists a unique $c > 0$ such that $m^*(c) = \ell$. For this value of c it holds that $c \in [c_a, \beta]$ and $\Pi_e^\ell = \Pi_e^*(c)$. (Recall that Π_e^ℓ is the employer welfare with an application limit ℓ, and no change to the application cost c_a.)*

5.4. Price setting by the platform.

The model we have studied so far does not include transfers. In this section, we study an extension where the platform can set a marketwide wage w that is paid on each job performed, from the employer to the worker. The wage offers another "lever" for the platform: as wages fall, the value of an application decreases. We investigate how this can be used to manage congestion.

As before, each successful match generates a total surplus of 2 units. We allow the platform to set any wage $w \in (0, 2)$, which causes a match utility of w to accrue to the applicant and a match utility of $v = 2 - w$ to accrue to the employer. The model discussed previously corresponds to the special case where $w = 1$ is fixed. For any wage w, the expected utility

applicant must answer). Even when the additional cost is monetary, redistribution may be logistically challenging, create incentives for individuals to create multiple accounts, and may be undesirable from the point of view of the platform operator.

of an applicant in the mean-field environment with success probability p who applies with intensity m is given by the following generalization of (2):

$$W_a(m,p) = w(1-e^{-mp}) - c_a m. \tag{18}$$

The applicant best response is given by the following generalization of (3):

$$\mathcal{M}_w(p) = \begin{cases} 0, & \text{if } p \le c_a/w; \\ \frac{1}{p}\log\left(\frac{pw}{c_a}\right), & \text{if } p > c_a/w. \end{cases} \tag{19}$$

Let $c_s' = c_s/\beta$ as before. Proposition 1 remains true with the modification that the threshold on q is now $c_s'/(2-w)$. Hence (generalizing the definition of $\mathcal{A}(q)$ in (4)) we define the correspondence

$$\mathcal{A}_w(q) = \begin{cases} \{0\} & \text{if } q < c_s'/(2-w) \\ [0,1] & \text{if } q = c_s'/(2-w) \\ \{1\} & \text{if } q > c_s'/(2-w). \end{cases} \tag{20}$$

As before, the strategy ϕ^α that plays ϕ^1 with probability α and exits immediately otherwise is an employer best response for $\alpha \in \mathcal{A}_w(q)$. The expected welfare of an employer who screens with probability α in the mean-field environment is given by the following generalization of (5):

$$W_e(\alpha,m,q) = \alpha(1-e^{-rm\beta q})(2-w-c_s'/q)\,. \tag{21}$$

The definition of MFE is as before, except that for any fixed wage, the applicant and employer best responses are given by (19) and (20). In contrast to the remainder of the paper, we assume that employer and applicant welfare are quasilinear in the wage and thus measured in monetary units. We define the aggregate welfare as the sum of employer and applicant welfare, given by

$$\Pi_{\text{tot}} = r\Pi_a + \Pi_e\,. \tag{22}$$

Our main result in this section is the following.

THEOREM 7. *Consider any primitives r, β, c_a, c_s and application limit $\ell > 0$ (with $w = 1$ fixed) such that the resulting equilibrium $(m_\ell^*, \alpha_\ell^*, p_\ell^*, q_\ell^*)$ satisfies $m_\ell^* = \ell$ and $\alpha_\ell^* = 1$. If the platform sets a fixed wage $w_\ell = (c_a/p_\ell^*)\exp(\ell p_\ell^*) \in (c_a/p_\ell^*, 1]$ and no application limit, the resulting equilibrium is again $(m_\ell^* = \ell, \alpha_\ell^* = 1, p_\ell^*, q_\ell^*)$. In particular, aggregate welfare is identical in the two settings.*

The intuition for Theorem 7 is that in order to achieve the desired application intensity ℓ via a wage, we simply set w_ℓ in a manner that $\mathcal{M}_{w_\ell}(p)$ for $p=p^*_\ell$ is equal to ℓ. We show that $w_\ell \leq 1$ (as one might expect). Since the employers found it worthwhile to screen (that is, $\alpha^*_\ell = 1$) at q^*_ℓ with a wage $w=1$, they continue to screen when the wage is $w_\ell \leq 1$. The total welfare Π_{tot} remains the same as if there were an application limit of ℓ because the wage is merely a transfer and does not affect the total welfare.

We close with two remarks. First, the *distribution* of welfare is not identical in the two equilibria: applicant welfare is lower (and employer welfare higher) when wages are w_ℓ than they are with an application limit of ℓ and a wage of $w=1$. We return to this point in our closing discussion. Second, by increasing the wage, the platform can increase application intensity: an effect that an application limit cannot achieve.[18]

5.5. Heterogeneity

Our model to this point has consisted of *homogeneous* employers and applicants. This section relaxes this assumption through two extensions: one with heterogeneous employers, and another with heterogeneous applicants. In our first extension, employers have a "type" β drawn according to F_e. Any applicant to a job posted by an employer of type β is compatible with probability β. In our second extension, applicants vary in their "skill" β, which is drawn according to F_a. An applicant with skill β is compatible with probability β for each job they apply to. In both extensions, as in the base model, each match partner receives utility 1 from a compatible match. We treat an employer's type or an applicant's skill as private information: known to the individual, but not directly observable to others. Appendix A provides more details, including the definition of a mean field equilibrium for each setting, as well as expressions for the welfare of each side.

We numerically investigate the effect of imposing an application limit that holds uniformly across applicants. We find that, compared to our baseline model, the effects of this limit can be complex, as we now discuss.

With *employer heterogeneity*, an application limit may make the market attractive for employers with difficult jobs (low β), who would otherwise choose not to participate. While

[18] Numerical results suggest that within our model, this is unlikely to be beneficial. The benefit of raising wages is highest in a setting with small screening cost, and large application cost. Setting (small) $c'_s = c_s/\beta = 0.01$ and $c_a = 0.01$, we found numerically that across all values of r, the total welfare improvement from raising w is no more than about 0.3%. Even with an implausibly large $c_a = 0.1$ and the same $c'_s = 0.01$, across all values of r, the total welfare improvement from raising w is no more than about 6%.

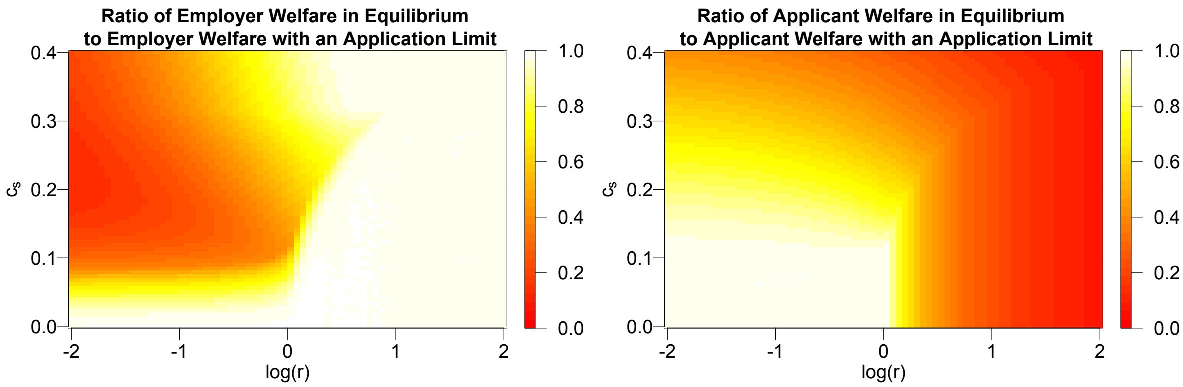

Figure 4 **A variant of Figure 2 when β is employer-specific and uniformly distributed on the interval $[0.4, 0.6]$, and $c_a = 0.005$. The screening cost c_s (y-axis) and ratio of applicants to employers r (x-axis) vary.**

a limit unambiguously benefits these employers, it may reduce the match rate for employers with easier jobs, who were already choosing to screen.

With *applicant heterogeneity*, an application limit may be binding for some applicants and not others. In some cases, the limit binds only for low-skill applicants, thereby unambiguously helping high-skill applicants and increasing the average skill perceived by employers. In others, the limit binds only for high-skill applicants. The resulting reduction in competition may attract low-skill applicants who had previously abstained.

A complete analysis of these effects is beyond the scope of this work.[19] We focus on setting an application limit to maximize the *aggregate* welfare on one side. Figures 4 and 5 show the potential benefit of an application limit for each side. Comparing to Figure 2, there are some differences: for example, with heterogeneous employers, employer welfare is never identically zero, and an application limit has little benefit when $r > 1$ (even when screening costs are high). However, the bottom line is that *our key insights remain valid*: we observe that an application limit significantly benefits employers when there is a shortage of applicants ($r < 1$) and c_s is not too small, and significantly benefits applicants when either r or c_s is large.

6. Mean field approximation.

We show that our mean field model is asymptotically valid for our finite system when the market grows large. First, we show in Theorems 8 and 9 that the mean field assumptions hold

[19] In principle, the mean field models described in Appendix A could accommodate heterogeneity in both the compatibility parameter β and the application or screening cost. Heterogeneous costs would cause participants to make different choices and prefer different application limits. Because these features are present in the model with heterogeneity in β alone, our numerical results focus on this case.

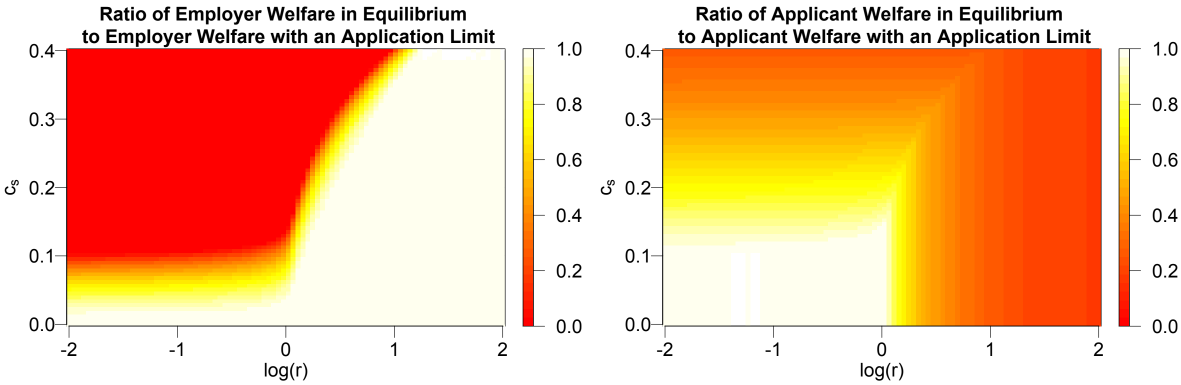

Figure 5 **A variant of Figure 2 when β is applicant-specific and uniformly distributed on the binary set $\{0.4, 0.6\}$, and $c_a = 0.005$. The screening cost c_s (y-axis) and ratio of applicants to employers r (x-axis) vary.**

as $n \to \infty$, as long as all applicants a choose $m_a = m$, and all employers e choose $\alpha_e = \alpha$.[20] Second, we use the preceding results to show Theorem 10, which states that any MFE is an approximate equilibrium in sufficiently large finite markets. Technical details of our approach are in Appendix D. Proofs are in Appendix G.

We let Binomial(n, p) denote the binomial distribution with n trials and probability of success p, and let Poisson(a) denote the Poisson distribution with mean a.

The following two theorems, we show that Mean Field Assumptions 1 and 2 hold as $n \to \infty$. In the process we also show Mean Field Assumption 3, that the number of applications received by an employer and sent by an applicant are each Poisson distributed.

THEOREM 8. *Fix r, β, m, and α. Suppose that the n-th system is initialized in its steady state distribution. Consider any employer e that arrives at $t_e \geq 0$. Let $R_e^{(n)}$ denote the number of applications received by employer e in the n-th system, and let $A_e^{(n)}$ be the number of these applicants that are still available when the employer screens.*

Then as $n \to \infty$, the pair $(R_e^{(n)}, A_e^{(n)})$ converges in total variation distance to (R, A), where $R \sim \text{Poisson}(rm)$, and conditional on R, we let $A \sim \text{Binomial}(R, q)$.

THEOREM 9. *Fix r, β, m, and α and any $m_0 < \infty$. Suppose that the n-th system is initialized in its steady state distribution. Consider any applicant a arriving at time $t_a \geq 0$, denote the value chosen by applicant a by m_a (all other applicants are assumed to choose m, and all employers are assumed to follow ϕ^α). Let $T_a^{(n)}$ denote the number of applications*

[20] Note that once we fix m and α, we have removed any strategic element from the evolution of the n-th system; and so our results are limit theorems about a certain sequence of stochastic processes.

sent by applicant a in the n-th system, and let $Q_a^{(n)}$ be the number of these applications that generate offers. Let $T \sim \text{Poisson}(m_a)$, and conditional on T, we let $Q \sim \text{Binomial}(T,p)$. Then

$$\lim_{n\to\infty}\left\{\max_{m_a\in[0,m_0]} d_{\mathrm{TV}}\Big((T_a^{(n)},Q_a^{(n)}),(T,Q)\Big)\right\}=0\,, \tag{23}$$

where $d_{\mathrm{TV}}(X,Y)$ denotes the total variation distance between the distributions of random variables X and Y that take values in the same countable set.

We use the preceding results on mean field approximation to establish that any MFE is an approximate equilibrium in sufficiently large finite markets.

THEOREM 10. *Fix r, β, c_s and c_a. Let the MFE be (m^*,α^*). For any $\varepsilon>0$, and any nonnegative integer R_0, there exists n_0 such that for all $n\geq n_0$ the following hold:*

1. *For any applicant a, if all other agents follow their prescribed mean field strategies, then applicant a can increase her expected payoff by no more than ε by deviating to any $m\neq m^*$.*
2. *For any employer e, if e receives no more than R_0 applications, and if all other agents follow their prescribed mean field strategies, then in any state in the corresponding dynamic optimization problem solved by employer e (cf. Appendix C), employer e can increase her expected payoff by no more than ε by deviating to any strategy other than ϕ^{α^*}.*

The first statement establishes that applicants cannot appreciably gain by changing the number of applications they send. The second statement makes an analogous claim for employers. In particular, if employers follow ϕ^{α^*}, they will either exit immediately; or sequentially screen and make offers to compatible candidates until such an offer is accepted, or the applicant pool is exhausted. In doing so, they will obtain information about the compatibility and availability of the subset of applicants they have already screened. Our result states that at any stage in this dynamic optimization problem, the employer cannot appreciably increase their payoff by deviating.[21]

[21] An apparent limitation of the second statement is that it applies only to employers who receive no more than R_0 applications. The fraction of employers who receive more than R_0 applications scales as $\exp(-\Omega(R_0))$; thus by choosing R large enough, this fraction can be made as small as desired. Since we show that the n-th system satisfies an appropriate stochastic contraction condition, we expect that in fact, for sufficiently large R_0, both statements in Theorem 10 hold even under arbitrary behavior by employers who receive more than R_0 applications. In the interest of brevity we choose to omit this slightly stronger result.

7. Conclusion.

Our paper presents a benchmark dynamic matching model to study welfare in markets with costly screening and uncertain availability. We characterize welfare in equilibrium for both applicants and employers, and we show how introducing an appropriately chosen application limit can significantly improve welfare for one or both sides.

We also compare the effects of an application limit to those of other available levers: either raising application costs, or lowering the wage paid to applicants. Although these interventions can lead to the same aggregate welfare as an application limit, they differ in how they *distribute* this welfare. Charging fees and lowering wages both increase aggregate welfare at the expense of applicants. While these may be appropriate for a platform looking to monetize its services or attract more employers, an application limit can yield Pareto improvements in welfare, may be more suitable if the platform is primarily concerned with applicant welfare. These considerations might explain why the tutoring platform TutorZ charges tutors for each potential client that they contact, whereas the dating platforms Coffee Meets Bagel and Tinder limit the number of likes/right swipes permitted in a certain period.

Of course, the choice between these interventions is also influenced by technological or cultural constraints not modeled in this paper. For example, an application limit may be unenforceable if applicants can easily create multiple accounts. In a setting with heterogeneity, raising the application cost might be more suitable than a one-size-fits all application limit. Conversely, charging fees to apply or intervening in the wage-setting process may be less palatable than a simple application limit. We do not assert that any one approach is universally preferable to the others: each can improve efficiency by reducing the effort required to find a suitable match. This work helps to show when (and for whom) the potential improvements are greatest.

References

[1] Albrecht J, Gautier PA, Vroman S (2006) Equilibrium directed search with multiple applications. *Review of Economic Studies* 73(4):869–891.

[2] Allon G, Bassamboo A, Cil E (2012) Large-scale service marketplaces: The role of the moderating firm. *Management Science* 58:1854–1872.

[3] Anderson R, Ashlagi I, Gamarnik D, Kanoria Y (2017) Efficient dynamic barter exchange. *Operations Research* 65(6):1446–1459.

[4] Ashlagi I, Jaillet P, Manshadi VH (2013) Kidney exchange in dynamic sparse heterogenous pools. *arXiv preprint arXiv:1301.3509* .

[5] Balseiro SR, Besbes O, Weintraub GY (2015) Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Science* 61(4):864–884.

[6] Burdett K, Shi S, Wright R (2001) Pricing and matching with frictions. *Journal of Political Economy* 109(5):1060–1085, URL `http://ideas.repec.org/a/ucp/jpolec/v109y2001i5p1060-1085.html`.

[7] Coles P, Cawley J, Levine PB, Niederle M, Roth AE, Siegfried JJ (2010) The job market for new economists: A market design perspective. *Journal of Economic Perspectives* 24(4):187–206.

[8] Coles P, Kushnir A, Niederle M (2013) Preference signaling in matching markets. *American Economic Journal: Microeconomics* 5(2):99–134.

[9] Diamond PA (1982) Aggregate demand management in search equilibrium. *Journal of Political Economy* 90(5):pp. 881–894, ISSN 00223808, URL `http://www.jstor.org/stable/1837124`.

[10] Diamond PA (1982) Wage determination and efficiency in search equilibrium. *The Review of Economic Studies* 49(2):217–227, URL `http://dx.doi.org/10.2307/2297271`.

[11] Field H (2019) Upwork to charge freelancers to bid on jobs, further squeezing those who rely on the 'gig economy'. *Entrepreneur* URL `https://www.entrepreneur.com/article/331700`.

[12] Fradkin A (2017) Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb .

[13] Galenianos M, Kircher P (2009) Directed search with multiple job applications. *Journal of Economic Theory* 144(2):445 – 471, ISSN 0022-0531.

[14] Galenianos M, Kircher P (2012) On the game-theoretic foundations of competitive search equilibrium*. *International Economic Review* 53(1):1–21, ISSN 1468-2354.

[15] Gurvich I, Ward A, et al. (2014) On the dynamic control of matching queues. *Stochastic Systems* 4(2):479–523.

[16] Halaburda H, Jan Piskorski M, Yıldırım P (2017) Competing by restricting choice: The case of matching platforms. *Management Science* .

[17] Horton JJ (2018) Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution Forthcoming in Management Science; an earlier version appeared at ACM EC 2014.

[18] Hu M, Zhou Y (2015) Dynamic matching in a two-sided market. *Available at SSRN* .

[19] Iyer K, Johari R, Sundararajan M (2014) Mean field equilibria of dynamic auctions with learning. *Management Science* 60(12):2949–2970.

[20] Janson S, Luczak T, Rucinski A (2011) *Random graphs*, volume 45 (John Wiley & Sons).

[21] Julien B, Kennes J, King IP (2000) Bidding for labor. *Review of Economic Dynamics* 3(4):619 – 649.

[22] Kircher P (2009) Efficiency of simultaneous search. *Journal of Political Economy* 117(5):861–913.

[23] Kushnir A (2013) Harmful signaling in matching markets. *Games and Economic Behavior* 80:209–218.

[24] Lasry JM, Lions PL (2007) Mean field games. *Japanese journal of mathematics* 2(1):229–260.

[25] Lee S, Niederle M, Kim HR, Kim WK (2011) Propose with a rose? signaling in internet dating markets. Working Paper 17340, National Bureau of Economic Research.

[26] Moen ER (1997) Competitive search equilibrium. *Journal of Political Economy* 105(2):pp. 385–411, ISSN 00223808, URL `http://www.jstor.org/stable/10.1086/262077`.

[27] Mortensen DT (1982) The matching process as a noncooperative bargaining game. *The Economics of Information and Uncertainty*, 233–258, NBER Chapters (National Bureau of Economic Research, Inc).

[28] Mortensen DT (1982) Property rights and efficiency in mating, racing, and related games. *The American Economic Review* 72(5):pp. 968–979, ISSN 00028282, URL `http://www.jstor.org/stable/1812016`.

[29] Pissarides CA (1984) Efficient job rejection. *The Economic Journal* 94:pp. 97–108, ISSN 00130133.

[30] Pissarides CA (1984) Search intensity, job advertising, and efficiency. *Journal of Labor Economics* 2(1):pp. 128–143, ISSN 0734306X, URL `http://www.jstor.org/stable/2535020`.

[31] Rogerson R, Shimer R, Wright R (2005) Search-theoretic models of the labor market: A survey. *Journal of Economic Literature* 43(4):959–988, URL `http://dx.doi.org/10.1257/002205105775362014`.

[32] Roth AE, Xing X (1997) Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists. *Journal of political Economy* 105(2):284–329.

[33] Weintraub GY, Benkard CL, Van Roy B (2008) Markov perfect industry dynamics with many firms. *Econometrica* 76(6):1375–1411.

[34] Wright R, Kircher P, Julîen B, Guerrieri V (2017) Directed search: A guided tour.

[35] Ye P, Qian J, Chen J, Wu Ch, Zhou Y, De Mars S, Yang F, Zhang L (2018) Customized regression model for airbnb dynamic pricing. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 932–940 (ACM).

Appendix A: Models with Agent Heterogeneity

Our main model assumes homogeneous applicants and employers. We now present mean field models that relax this assumption. Section A.1 considers a model in which some employers are less discriminating (higher β) than others, but applicants cannot distinguish employers ex ante. Section A.2 considers a model in which some applicants are more skilled (higher β) than others, but employers cannot identify an applicant's skill before screening.

In principle, one could study a model in which participants on both sides have types that are unobservable before application and screening, with the compatibility parameter a function of the types of both sides. Additionally, as noted in the text, the mean field models described below could be modified to incorporate heterogeneity in application or screening costs. These additional forms of heterogeneity would cause participants on the same side to make different choices and prefer different application limits. Because these features are present in the model with one-sided heterogeneity in β alone, our analysis focuses on this case.[22]

A.1. Employer Heterogeneity

The model is as before, except that the probability of compatibility β is employer-specific, with F_e being the cdf of β across employers. For convenience, we assume that F_e is continuous. Applicants cannot observe the employer's type β, so their strategy is summarized by their search intensity m, as before. Employers know their type and choose whether to screen applicants. If an employer with type β chooses to screen, then so will all employers with larger β. Therefore, employer strategies can be described by a threshold $\underline{\beta}$: employers with β above this threshold will choose to screen, and those below this threshold will not. For fixed m and $\underline{\beta}$, we now give the equations defining the steady-state probability of an offer p and availability q.

The equation defining q from p remains $q = g(mp)$, as given in (6). We must modify equation (7), which determines p from q. If employers perceive an availability of q, then the number of qualified available applicants to an employer with type β is Poisson$(rm\beta q)$. Assuming that this employer chooses to screen, an applicant's probability of being screened is $g(rm\beta q)$, and probability of receiving an offer is $\beta g(rm\beta q)$. Thus, we arrive at

$$p = \int_{\underline{\beta}}^{1} \beta g(rm\beta q) dF_e(\beta). \tag{24}$$

Together (6) and (24) form a consistency condition that allows us to find the mean-field steady state (p, q) associated with any applicant search intensity m and employer screening threshold $\underline{\beta}$. Let $\mathcal{P}(m, \underline{\beta})$ and $\mathcal{Q}(m, \underline{\beta})$ be the resulting steady-state quantities.

Applicant welfare is still given by (2), and the applicant best response, given application limit ℓ, remains (8) (the best response remains (3) when there is no application limit). Welfare and the best response for an employer with compatibility parameter β are given in (5) and (4) respectively, where c_s' now represents the ratio of the screening cost c_s to the employer-specific value of β. Because the best response of employers is to screen if and only if $\beta > c_s/q$, we say that a pair $(m, \underline{\beta})$ is a mean field equilibrium corresponding to application limit ℓ if $\underline{\beta} = c_s/\mathcal{Q}(m, \underline{\beta})$ and $m = \mathcal{M}_\ell(\mathcal{P}(m, \underline{\beta}))$.

A.2. Applicant Heterogeneity

In this model, applicants are differentiated by their probability β of being compatible for each job (their "skill"), which is distributed according to cdf F_a and known to each applicant. Employers cannot directly observe applicants' skill, so they screen in a uniformly random order and make offers to qualified applicants. As in the original model, we let $\alpha \in [0, 1]$ denote the fraction of employers who

[22] In particular, the equations below would remain the same if applicants differed both in their skill β and the application cost c_a, except that F_a would be a joint distribution over β and c_a and the application function M would take both parameters as arguments.

choose to screen. Because applicants with different skill levels may choose different search intensities, we summarize applicant strategies using the function $M:[0,1]\to\mathbb{R}_+$, where $M(\beta)$ gives the search intensity of applicants with skill β.

Given strategies summarized by α and M, we seek to define a consistent outcome. In the original model, we use p to denote the fraction of applications that lead to offers, and q to denote the fraction of offers that are accepted. In our present model, because skill varies across applicants, the offer probability p will as well. Therefore, our welfare expressions will be in terms of slightly different quantities: $\tilde{p}$, interpreted as the fraction of applications that are *screened*, and $\tilde{q}$, interpreted as the fraction of screenings that result in a successful hire. We now give equations defining these mean-field steady-state quantities, for any α and M.

We start by taking $\tilde{p}$ as given. Consider an applicant who searches with intensity m, has each application screened with probability $\tilde{p}$, and has skill β. Then the number of offers received by this applicant is Poisson with mean $m\beta\tilde{p}$. Because the applicant matches so long as he or she receives at least one offer, the probability of matching is $1-e^{-m\beta\tilde{p}}$, which can also be interpreted as the expected number of accepted offers. Dividing this by the expected number of total offers received reveals that the the acceptance rate for offers made to applicants with skill β is $g(m\beta\tilde{p})$. Therefore, an employer who screens an applicant with skill β will find them qualified and available with probability

$$\tilde{Q}(m,\tilde{p},\beta)=\beta g(m\beta\tilde{p}). \tag{25}$$

Employers' perceived success probability $\tilde{q}$ is then given by

$$\tilde{q}=\frac{\mathbb{E}_{\beta\sim F_a}[M(\beta)\tilde{Q}(M(\beta),\tilde{p},\beta)]}{\mathbb{E}_{\beta\sim F_a}[M(\beta)]}. \tag{26}$$

Jointly, (25) and (26) are the analog to (6).

We now determine $\tilde{p}$. For each application sent, the number of competing applications from qualified, available applicants follows a Poisson distribution with mean

$$\lambda=r\mathbb{E}_{\beta\sim F_a}[M(\beta)\tilde{Q}(M(\beta),\beta,\tilde{p})]. \tag{27}$$

Because an applicant with k qualified available competitors is screened with probability $\alpha/(k+1)$, we have the following analog to (7):

$$\tilde{p}=\alpha\sum_{k=0}^{\infty}e^{-\lambda}\lambda^k/(k+1)!=\alpha g(\lambda). \tag{28}$$

The value of $\tilde{p}$ that results from strategies α and M simultaneously solves (25), (27) and (28), and the corresponding $\tilde{q}$ is given by (26).

Given the mean field screening probability $\tilde{p}$, applicants with skill β who choose application intensity m receive utility

$$W_a(m,\beta,\tilde{p})=1-e^{-m\beta\tilde{p}}-c_a m.$$

If there is an application limit ℓ, the optimal choice for an applicant of skill β is

$$M(\beta)=\arg\max_{m\le\ell}W_a(m,\beta,\tilde{p})=\mathcal{M}_\ell(\beta\tilde{p}). \tag{29}$$

Given the success probability $\tilde{q}$, employers who choose to screen with probability α receive utility

$$W_e(\alpha,\tilde{q})=\alpha(1-e^{-\lambda})(1-c_s/\tilde{q}).$$

Employers will choose

$$\alpha\in\tilde{\mathcal{A}}(\tilde{q})=\begin{cases}\{1\} & : c_s<\tilde{q}\\ [0,1] & : c_s=\tilde{q}\\ \{0\} & : c_s>\tilde{q}.\end{cases} \tag{30}$$

A mean field equilibrium with application limit ℓ consists of an application function M and a value α that are best responses (according to (29), (30)) to the steady state quantities $\tilde{p}$ and $\tilde{q}$ that they induce (as given by (25), (26), (27), (28)).

Appendix B: Comparison with a static model.

In this appendix, we consider static versions of our model, in which employers make at most one offer. These models are very similar to those in Albrecht et al. [1] and Galenianos and Kircher [13]. Our main purpose is to show that competition among employers is much less severe when employers can make only one offer.[23] As a result, *these models predict little or no benefit to employers from an application limit*, as shown in Figure 6.

The timing of our static models is as follows:

1. Each applicant a applies to m_a employers, selected uniformly at random.
2. Each employer screens until she finds a qualified applicant or exhausts her applicant pool. Employers simultaneously make offers to the first qualified applicant that they find (if any).
3. Each applicant with an offer accepts one, choosing uniformly at random if she received multiple offers.

We consider static model versions that differ only in how the quantity m_a is determined. We begin by analyzing the case where m_a follows a Poisson distribution, as in the body of our paper, and show that the predictions of the static and dynamic models are quite different. For ease of comparison with existing work, we then introduce a model where m_a is deterministic and identical for all applicants.

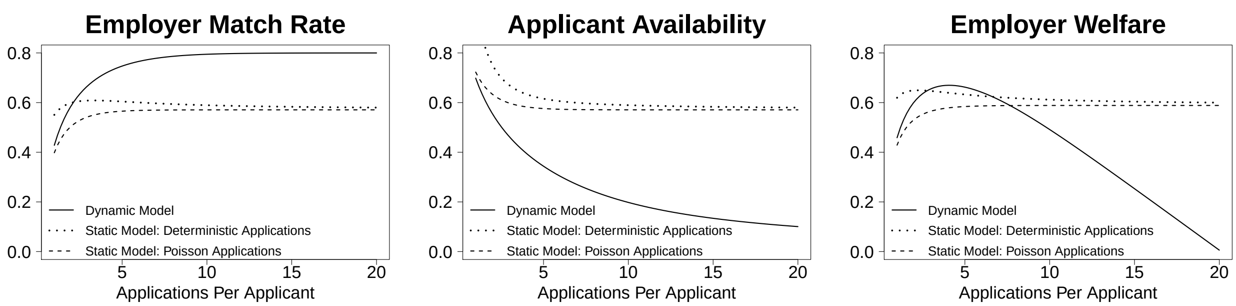

Figure 6 **Comparing outcomes and in dynamic and static models, when there are $r = 0.8$ applicants per employer and screening costs are $c_s' = 0.1$. In our dynamic model, when applicants send many applications, nearly all applicants match, resulting in low availability and low welfare for employers. When enough applications are sent ($\gtrsim 20$ per worker), employer welfare drops to zero. As a result, an application limit can significantly improve employer welfare. In both static models, employers can make at most one offer. Even with many applications, many applicants do not match, and availability remains high. In one static model (with Poisson applications) employer welfare is increasing in the number of applications sent. In the other (with deterministic applications), employer welfare hardly varies as the number of applications changes. In either case, an application limit cannot significantly benefit employers.**

B.1. Static model with Poisson m_a.

We begin by analyzing the case where m_a follows a Poisson distribution.

In our static model, employers do not screen for availability, so they make an offer if and only if they choose to screen and have a qualified applicant. As before, when applicants send an average of m applications, the number of qualified applicants received by any given employer is distributed as a Poisson random variable. Analogous to our main analysis, it follows that the probability that an applicant receives an offer from a given application is

$$\mathcal{P}(m, \alpha) = \alpha \beta g(rm\beta), \tag{31}$$

[23] As an aside, we remark that the analysis in this section can be interpreted as a demonstration that suitably restricting screening by employers can increase employer welfare — analogously to the result that an application limit can increase applicant welfare.

where $g(x) = (1 - e^{-x})/x$. Our previous consistency condition was $p = \alpha\beta g(rm\beta q)$, so the only change is that q no longer appears in the argument of $g(\cdot)$ (since employers do not screen for availability). As we shall see, this small change makes a big difference to employer welfare.

Because the number of applications sent follows a Poisson distribution, and each application triggers an offer with probability $\mathcal{P}(m, \alpha)$, the probability that a given applicant receives at least one offer is

$$M(m, \alpha) = 1 - e^{-m\mathcal{P}(m,\alpha)} \tag{32}$$

The consistency condition for mean-field availability q remains

$$\mathcal{Q}(m, \alpha) = M(m, \alpha)/(m\mathcal{P}(m, \alpha)). \tag{33}$$

We then have

$$\Pi_e(m, \alpha) = r(M(m, \alpha) - c_s' m\mathcal{P}(m, \alpha)). \tag{34}$$

Under the relatively mild condition $c_s' < e^{-1/r}$, one can show that employer welfare as given in (34) is monotone increasing with respect to m, for $\alpha = 1$. In this case, *an application limit can never help employers*. This contrasts sharply with the conclusion of our dynamic model, where for any c_s' and any $r \leq 1$, employer welfare is zero if the number of applications is sufficiently large (and an application limit vastly increases employer welfare).

The disagreement arises because when employers send at most one offer, unless there are far more employers than applicants, a significant fraction of applicants will go unmatched. These applicants would accept any offer that they receive, so from the perspective of each employer, the probability that their offer is accepted never falls too low. By contrast, in our dynamic model employers continue to screen until they match or exhaust their applicant pool. As a result, whenever there are more employers than applicants, if applicants send enough applications (and employers are willing to screen them), most applicants will find a job, and availability, as perceived by employers, will plummet. In other words, the ability of employers to continue to make multiple offers intensifies the competition between them.

B.2. Static model with fixed m_a.

To clarify the relationship between the model sketched above and that in Albrecht et al. [1], we now consider a version of the static model in which each applicant sends exactly m applications. The number of qualified applicants who apply to a given employer remains Poisson, so (31) remains valid. Now, however, each applicant matches with probability

$$M^D(m, \alpha) = 1 - (1 - \mathcal{P}(m, \alpha))^m \tag{35}$$

In fact, (35) with $\alpha = \beta = 1$ is exactly the match function defined in Proposition 1 of Albrecht et al. [1]. The expressions (33) and (34) for availability and employer welfare remain valid, after substituting M^D for M.

The conclusions from this model are qualitatively similar to those from our previous static model in which the number of applications sent is Poisson. One curiosity is that the number of matches formed is no longer increasing in m, the number of applications sent. However, in most cases employer welfare is very insensitive to m, for the same reason as before: so long as employers can make only one offer, many applicants remain unemployed and available. As a result, *an application limit provides little, if any, benefit to employers*. We show this in Figure 6 for the case where $r = 0.8$ and $c_s' = 0.1$.

Appendix C: Employer Best Response (Section 4.1.2)

Employer actions, strategy space, and optimization. Recall that an employer posts their job and returns one time unit later to take a sequence of "screening" and "offer" actions, instantaneously learning the result of each, before leaving. The employer is only permitted to make an offer to an applicant she has previously found to be compatible. For clarity, we now provide a more complete specification of the employer actions and strategy space before proceeding to prove Proposition 1.

When the employer returns to process applications, the following events occur in sequence, but in an instant of real time:

- She learns the *number* of applications received.
- The employer now has two options: *screen* (a uniformly random applicant) or *exit*. If she screens, she immediately learns if the applicant is compatible.
- If the employer has not left, she must choose between *offer* (to a specific applicant who was found to be compatible but not yet offered), *screen* (a uniformly random unscreened applicant) or *exit*. If she makes an offer, then she immediately learns if it was accepted or not. If the offer is accepted, she immediately leaves. If she screens, she immediately learns if the applicant was compatible.
- The previous step repeats until the employer leaves, either because one of her offers is accepted, or because there are no more applicants, or because the employer chose to exit.

Note that an employer who receives k applications is able to dynamically choose a sequence of up to $2k$ actions, i.e., a finite number of actions, before exiting, since each applicant can be screened at most once and made an offer at most once. The employer strategy space consists of all adapted dynamic policies that specify how to choose a permitted action at each stage. The employer acts to maximize expected net utility, i.e., match surplus less the total cost of screening (cf. Section 3.4).

We now prove Proposition 1. To do so, we use the following fact, which is immediate since compatibility in our model is binary:

Fact 1 *Whenever an employer screens and finds an applicant compatible, it is a dominant strategy to make an offer to this applicant immediately.*

Proof of Proposition 1. We prove the proposition by induction on the number of applications received by a employer. Any employer with applicants must start by exiting or screening. We begin by considering a employer with a single applicant. If the employer exits, she earns a payoff of zero, and if she begins by screening, then makes an offer if the applicant is compatible, she earns an expected payoff of $q\beta - c_s = \beta(q - c'_s)$. It follows that if $q > c'_s$, the employer chooses to screen, if $q < c'_s$, the employer should exit immediately, and if $q = c'_s$, the employer is indifferent between these options. Thus, the proposition holds for a employer with a single applicant.

Now consider a employer with $k+1$ applicants. Again, any strategy must either exit or begin by screening.

Suppose now that $q > c'_s$. Then exiting immediately earns a payoff of zero, which is less than that earned by playing ϕ^1, so the optimal strategy must begin by screening. Fact 1 implies that after screening the first applicant, it is optimal to offer the applicant the job if and only if they are qualified, causing the employer to match and exit if the applicant is both qualified and available, and otherwise to play the optimal strategy with k employers, which is ϕ^1 by our inductive assumption. In other words, ϕ^1 is uniquely optimal among strategies that start by screening, and therefore among all strategies.

Suppose that $q < c'_s$. Again by Fact 1, we know that among strategies that begin by screening, the best one will make an offer if and only if the candidate is qualified, causing the employer to match and exit if the candidate is both qualified and available, and otherwise will play the optimal strategy with k employers, which is to exit (by our inductive assumption). This strategy earns $\beta(q - c'_s) < 0$, so it cannot be optimal. It follows that exiting immediately must be uniquely optimal.

When $q = c'_s$, it is clear that both ϕ^1 and exiting immediately have an expected payout of zero, and therefore so does any mixture between the two. No other strategy can do strictly better by identical logic to that used above; any strategy earning positive surplus must begin by screening, and therefore earn no more than $\beta(q - c'_s)$ (which is zero) plus $1 - q\beta$ times the optimal payout with k applicants, which is zero by our inductive assumption. $\square$

Appendix D: Mean field limit: technical analysis.

In this section we develop the technology required to prove the key approximation results (Theorems 8 and 9, and ultimately Theorem 10).

We begin by formalizing the stochastic process of interest, when m and α are fixed.[24] Note that in our original model, applicants decide where to apply when they arrive to the system; however, for purposes of stochastic analysis, we obtain an equivalent system if we *realize applicant applications only when employers depart*. In particular, we consider the following stochastic system parameterized by n. Individual applicants arrive at intervals of length $1/rn$, as before. Let $S(t)$ denote the number of applicants in the system at time t. In addition, we define $\Sigma(t)$ as the *normalized* number of applicants in the system:

$$\Sigma(t) = S(t)/(rn); \tag{36}$$

note that $\Sigma(t) \leq 1$ for any $N(t)$ that can arise. At intervals of length $1/n$ (corresponding to employer departures), there are opportunities for at most a single applicant in the system to match and hence depart early. At each such employer departure time t, the probability of a departure of an applicant is:

$$\alpha \left(1 - \left(1 - \frac{\beta m}{n} \right)^{S(t)} \right) = \alpha(1 - \rho^{\Sigma(t)}), \tag{37}$$

where we define $\rho := (1 - \beta m/n)^{rn}$. Note that $\rho \to \exp(-\eta)$ as $n \to \infty$, where we define $\eta := rm\beta$.

The preceding equation (37) is derived as follows. As before, with probability α an employer screens using strategy ϕ, and exits immediately otherwise (in which case no applicant departs early). Every applicant that arrived in the last one time unit applied to the departing employer with probability m/n. Any such applicant that has already departed cannot match to the given employer. On the other hand, among the remaining applicants, if even one of them is compatible with the employer, then employer following ϕ is sure to find a match. Thus at least one departure occurs as long as there is at least one available, compatible applicant that applied to the departing employer. Note that under ϕ, each of the applicants in the system at time t is equally likely to depart, so for each applicant the probability of departure is $\alpha(1 - \rho^{\Sigma(t)})/(rn\Sigma(t))$.

To capture the state at time t, we must track the residual lifetimes of all applicants in the system. To simplify this tracking, a key instrument in our analysis is a "binned" version of the stochastic process $S(t)$, defined as follows. Fix an integer k, and let $S_j(t)$ be the number of applicants that have been in the system for a time between j/k and $(j+1)/k$ units, for $j = 0, 1, \ldots, k-1$. Let $X_j(t) = S_j(t)/(rn)$; note that $\Sigma(t) = \sum_{j=0}^{k-1} X_j(t)$. Our fundamental result (Proposition 4 below) proves a concentration result for the vector-valued stochastic process $X(t)$.

What does $X(t)$ concentrate around? To develop intuition, let's think of the matching process from the perspective of the applicants that arrive in a fixed interval of length $1/k$. In the large market limit we would expect $\Sigma(t)$ to be unchanging over time, and each of these applicants should match in successive intervals of length $1/k$ with a constant probability; or equivalently, their survival probability is a constant γ in each such interval. Looking back in time, then, in steady state we should expect that the vector $X(t)$ satisfies $X_j(t) = \gamma X_{j-1}(t)$ for $j = 1, \ldots, k-1$, with $X_0(t) = 1/k$. With this inspiration (and in an abuse of notation), we define $\Sigma(\gamma)$ as:

$$\Sigma(\gamma) = \frac{1}{k} \sum_{j=0}^{k-1} \gamma^j = \frac{1 - \gamma^k}{k(1-\gamma)}. \tag{38}$$

(Note that $\Sigma(1) = 1$.)

[24] Throughout this section we assume $m > 0$.

On the other hand, as in our mean field analysis, we can develop a "consistency check" that γ must satisfy using (37). Assume that $k \geq \eta/r$ and define $\gamma(\Sigma)$ as follows:

$$\gamma(\Sigma) = 1 - \alpha\left(\frac{1 - e^{-\eta\Sigma}}{rk\Sigma}\right), \tag{39}$$

where we take $\gamma(0) = 1 - \alpha\eta/(rk)$ (this is the limit of the preceding quantity as $\Sigma \to 0$). This equation is an approximate version of (37): $1-\gamma(\Sigma)$ represents an estimate of the probability that an individual applicant in the system is matched in the next $1/k$ time units, if the current (normalized) number of available applicants is Σ, and k and n are "large" (specifically $k = \omega(1)$ and $n = \omega(k)$).

In our analysis we will require the following basic facts regarding (39); we record them here for later reference.

Lemma 1. *Let $h(\Sigma) = (1 - e^{-\eta\Sigma})/\Sigma$, so that $\gamma = 1 - \alpha h(\Sigma)/(rk)$. Then*
1. $-\eta^2 \leq \frac{dh}{dx} \leq 0$.
2. $0 \leq \frac{d\gamma}{d\Sigma} \leq \frac{\alpha\eta^2}{rk}$.
3. $\frac{\alpha(1-e^{-\eta})}{rk} \leq 1 - \gamma \leq \frac{\alpha\eta}{rk}$
4. $\frac{d}{d\Sigma}(1-\gamma)\Sigma \geq \frac{\alpha\eta e^{-\eta}}{rk}$.

The following two results are critical to our analysis: they establish the uniqueness of a solution to the pair of equations (38)-(39), and show that in the limit $k \to \infty$ this solution is the unique MFSS (p, q) guaranteed by Proposition 2. The proofs are in Appendix G.1.

Lemma 2. *Suppose that $k \geq \eta/r$. There is a unique pair of real numbers (γ^*, Σ^*) that simultaneously solve* (38) *and* (39).

Lemma 3. *Suppose that $k \geq \eta/r$. Let (γ_k^*, Σ_k^*) denote the unique solution to* (38) *and* (39) *guaranteed by Lemma 2. Then as $k \to \infty$, $k(1-\gamma_k^*) \to mp$, and $\Sigma_k^* \to q$, where (p, q) is the unique MFSS guaranteed by Proposition 2.*

The interpretation of Lemma 3 is as follows. Note that under the mean field assumptions, each applicant sends a Poisson distributed number of applications, with mean m; and each application independently succeeds with probability p. Since the applicant's applications are independent to each employer, it follows that in the mean field model the applicant's lifetime in the system is an exponential random variable of mean $1/mp$ truncated to be less than or equal to 1 (since applicants only live for at most a unit lifetime). In other words, the rate of applicant departure is mp. On the other hand, for fixed k the rate of departure is approximated by $k(1-\gamma_k^*)$, so we should expect the latter quantity to approach mp. Similarly, observe that Σ^* is meant to be an estimate of the steady state (normalized) number of applicants in the system; we should expect that this approaches the applicant availability q in the mean field model.

Let X^* be the vector given by $X_j^* = (\gamma^*)^j/k$ for $j = 0, \ldots, k-1$, so that $\Sigma^* = \sum_{j=0}^{k-1} X_j^*$ from Eq. (38). The main result that enables our three mean field theorems is the following proposition, that shows that $X(t)$ concentrates around X^*. Note that this is a very strong result, because it precludes drift of the $X(t)$ process away from X^* as time grows. We achieve this result by using a stochastic concentration argument on the process $X(t)$.

Proposition 4. *Fix $m_0 \in [1, \infty)$ and $\alpha \in [0, 1]$. There exists $C = C(r, m_0, \beta, \alpha) < \infty$ such that for any $m \in [1/m_0, m_0]$, for any $n > C$, $k = \lfloor n^{1/3} \rfloor$ the following is true: For any $t > C \log n$, and any starting state at time* 0*, we have $\mathbb{E}[\|X(t) - X^*\|_1] \leq Cn^{-1/6}$. If the starting state at time* 0 *is drawn from the steady state distribution, we have $\mathbb{E}[\|X(t) - X^*\|_1] \leq Cn^{-1/6}$ for all $t \geq 1$.*

This proposition provides the concentration result required to establish that the mean field assumptions hold asymptotically (Theorems 8 and 9, proven in Appendix G.5). We can then use those results to prove that MFE is an approximate equilibrium in large finite markets (Theorem 10, proven in Appendix G.6).

Appendix E: Proofs: Section 4

E.1. Section 4.2: Mean field steady states.

In this appendix we study mean field steady states (MFSS), i.e., solutions (p,q) to the equations (6) and (7). We prove Proposition 2, which states that for any r,β,m,α, there exists a unique MFSS. We also show in Lemma 6 that the MFSS values p and q are monotonic in m and α. This will be useful in proving results about mean field equilibria in Section E.2.

For notational convenience, we define the function g as follows:

$$g(0)=1,\quad g(x)=\frac{1-e^{-x}}{x} \text{ for } x>0. \tag{40}$$

LEMMA 4. *The function $g:[0,\infty)\to(0,1]$ defined by (40) is continuous and strictly decreasing. Further:*

1. $g'(x)\geq -g(x)/x$ *and* $g''(x)/g'(x)>-1$ *for* $x>0$*;*
2. $\lim_{x\to 0} g'(x)=-1/2$*, and* $\lim_{x\to 0} g''(x)=1/3$*;*
3. $e^{-g^{-1}(c)}\leq c^2$ *for* $0<c<1$.

Proof of Lemma 4. Note first that g is continuous at zero. Differentiating, for $x>0$ we have

$$g'(x)=\frac{(1+x)e^{-x}-1}{x^2}=\frac{e^{-x}-g(x)}{x}. \tag{41}$$

Monotonicity of g follows from applying the inequality $1+x<e^x$ for $x>0$, which implies (on rearranging terms) that $e^{-x}<g(x)$. Having established that the expressions in (41) are negative, it follows that $g'(x)>-g(x)/x$ for $x>0$.

We now prove that $g''(x)/g'(x)>-1$, which is equivalent to $g'(x)+g''(x)<0$ (the inequality has reversed because $g'<0$). Basic algebra reveals that

$$-x(g'(x)+g''(x))=(g(x)+2g'(x)) \tag{42}$$

$$=\frac{1}{x^2}\left((2+x)e^{-x}+x-2\right) \tag{43}$$

Thus, $g'(x)+g''(x)<0$ if and only if $(2+x)e^{-x}+x-2>0$. Differentiating this expression, we see that

$$\frac{d}{dx}\left((2+x)e^{-x}+x-2\right)=1-e^{-x}(1+x)\geq 0,$$

so the expression is minimized at $x=0$, when it takes the value zero.

If we apply L'Hospital's rule to (41), we obtain

$$\lim_{x\to 0} g'(x)=\lim_{x\to 0}\frac{e^{-x}-g(x)}{x}=\lim_{x\to 0} -e^{-x}-g'(x).$$

Rearranging, we see that $2\lim_{x\to 0} g'(x)=-1$, implying that $\lim_{x\to 0} g'(x)=-1/2$. Rearranging (42) reveals that

$$\lim_{x\to 0} g''(x)=\lim_{x\to 0} -g'(x)-\frac{g(x)+2g'(x)}{x}=\frac{1}{2}-\lim_{x\to 0}(g'(x)+2g''(x)), \tag{44}$$

where we have used the fact that $\lim_{x\to 0} g'(x)=-1/2$ and applied L'Hospital's rule. Rearranging (44), we see that $3\lim_{x\to 0} g''(x)=1$.

Finally, by rearranging we observe that $e^{-g^{-1}(c)}\leq c^2$ if and only if $1-c^2+2c\log c\geq 0$. It is straightforward to check that the expression $1-c^2+2c\log c$ is decreasing in c, and is equal to zero at $c=1$.

Proof of Proposition 2. Recall from (6)-(7) that given r, m, β, α, we define a MFSS as any solution (p,q) to

$$p = \alpha\beta g(rm\beta q), \qquad q = g(mp). \tag{45}$$

Note that the preceding system trivially has a single solution when $m = 0$: namely, $p = \alpha\beta$ and $q = 1$. Therefore we focus on the case where $m > 0$. We proceed by showing that the preceding system has a unique solution.

For this purpose it is useful to rewrite (45) as providing two functions that yield p in terms of q. In particular, for $m > 0$, we note that (p,q) is an MFSS if and only if $p = p_1(q) = p_2(q)$, where $p_1, p_2 : [g(m), 1] \to [0,1]$ are defined by

$$p_1(x; m, \alpha) = \alpha\beta g(rm\beta x), \qquad p_2(x; m) = g^{-1}(x)/m. \tag{46}$$

Here the notation $f(x;y)$ indicates that the function f is parameterized by y. The chosen lower bound of $g(m)$ for the domain arises from the fact that any mean-field steady-state (p,q) corresponding to (m,α) must satisfy $q = g(mp) \geq g(m)$ (since g is strictly decreasing).

We show in Lemma 5 that for $m > 0$ and any r, β, α, there is a unique point where p_1 and p_2 intersect; this point is the unique MFSS.

The following lemma, used in the preceding proof, also establishes some useful properties of MFSS.

LEMMA 5. *Fix r, β, and α, and $m > 0$. Then there exists a unique pair $(p,q) \in [0, \alpha\beta] \times [g(m), 1)$ such that $p = p_1(q; m, \alpha) = p_2(q; m)$.*

Furthermore: (1) the functions p_1 and p_2 defined in (46) are monotonically decreasing in q; and (2) for $q' < q$, we have $p < p_1(q'; m, \alpha) < p_2(q'; m)$; and for $q' > q$, we have $p > p_1(q'; m, \alpha) > p_2(q'; m)$.

Proof. For the duration of the proof we suppress the dependence of p_1 and p_2 on α and m.

Note that on the domain $[g(m), 1)$, p_1 begins "below" p_2 and finishes "above" it, i.e. (because $g(\cdot) \leq 1$ by Lemma 4), we have $p_1(g(m)) \leq \alpha\beta < 1 = p_2(g(m))$ and $p_1(1) > 0 = p_2(1)$. Since p_1 and p_2 are continuous, this implies that a mean-field steady-state exists. This is illustrated in Figure 7.

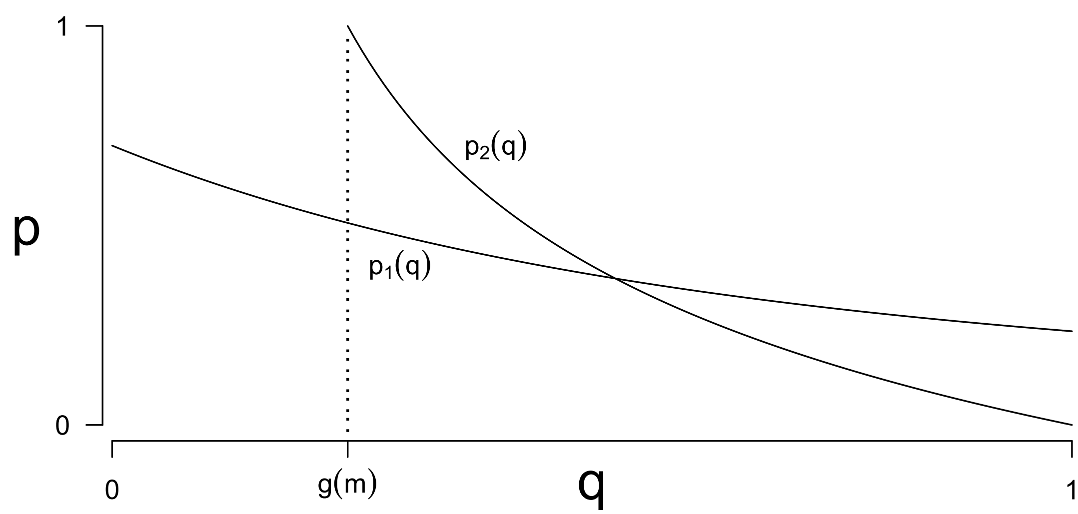

Figure 7 A depiction of $p_1(q)$ and $p_2(q)$; we prove in Lemma 5 that they have a unique intersection point.

Once we know that the intersection of p_1 and p_2 is unique, the final claim of the lemma follows immediately. Thus, all that remains to show is that p_1 and p_2 have a unique intersection point.

Note that because g is decreasing by Lemma 4, so are both p_1 and p_2. To show uniqueness, we will show that whenever p_1 and p_2 intersect, the curve p_2 is decreasing more "steeply," i.e. $\frac{\partial p_1}{\partial x} > \frac{\partial p_2}{\partial x}$.

Equivalently, we wish to show that whenever $p_1(q) = p_2(q) = p$, we have $\frac{\partial p_1}{\partial x}(q) \big/ \frac{\partial p_2}{\partial x}(q) < 1$ (the inequality reverses because both partials are negative). But

$$\frac{\partial p_1}{\partial x}(q) \Big/ \frac{\partial p_2}{\partial x}(q) = (\alpha\beta r m\beta g'(rm\beta q))\,(mg'(mp)) \tag{47}$$

$$< \left(\frac{\alpha\beta g(rm\beta q)}{q}\right)\left(\frac{g(mp)}{p}\right). \tag{48}$$

$$= \left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = 1. \tag{49}$$

The first line follows from implicit differentiation of p_2; the second from the inequality $|g'(x)| < g(x)/x$ proven in Lemma 4; and the third from the fact that (p,q) is a mean-field steady-state, i.e. $p_1(q) = p_2(q) = p$.

Having established the existence of a unique mean-field steady-state, it will be useful when proving later results to understand how $\mathcal{P}(m,\alpha)$ and $\mathcal{Q}(m,\alpha)$ vary with m and α. We require the following identity, which can be easily derived by rearranging (6)-(7):

$$\alpha(1 - e^{-rm\beta q}) = r(1 - e^{-mp}) \tag{50}$$

There is a straightforward interpretation of this identity. Each employer matches if and only if they screen (which they do with probability α), and they have a qualified available applicant (which occurs with probability $1 - e^{-rm\beta q}$, because each applicant is available with probability q and qualified with probability β). On the other hand, each applicant matches if and only if one of her applications receives an offer (which occurs with probability $1 - e^{-mp}$). Because the number of employers and applicants who match must be equal, we conclude that if p and q are the MFSS consistent with strategies m and α, then (50) must hold.

The following lemma provides some monotonicity properties of $\mathcal{P}(m,\alpha)$ and $\mathcal{Q}(m,\alpha)$.

Lemma 6. *For $(m,\alpha) \in (0,\infty) \times (0,1]$, the quantity $\mathcal{P}(m,\alpha)$ is strictly decreasing in m and strictly increasing in α, and $\mathcal{Q}(m,\alpha)$ is strictly decreasing in m and in α. The quantity $m\mathcal{P}(m,\alpha)$ is strictly increasing in m and in α. For $\alpha > 0$,*

$$\lim_{m\to\infty} \mathcal{P}(m,\alpha) = \alpha\beta f_2(\alpha/r),\ \mathcal{P}(0,\alpha) = \alpha\beta.$$
$$\lim_{m\to\infty} \mathcal{Q}(m,\alpha) = f_2(r/\alpha),\ \mathcal{Q}(0,\alpha) = 1,$$

where f_2 is defined by

$$f_2(x) = \begin{cases} 0 & : x \le 1 \\ g(\log \frac{x}{x-1}) & : x > 1. \end{cases} \tag{51}$$

Proof. Define the functions p_1 and p_2 as in (46). Recall from Lemma 5 that $\mathcal{P}(m,\alpha)$ and $\mathcal{Q}(m,\alpha)$ are the unique solution (p,q) to $p_1(q;m,\alpha) = p_2(q;m) = p$. For $x < \mathcal{Q}(m,\alpha)$, we have $p_1(x) < p_2(x)$ and for $x > \mathcal{Q}(m,\alpha)$, we have $p_1(x) > p_2(x)$.

We first prove the statements about monotonicity in α. Suppose that $\alpha < \alpha'$. Then $p_2(\mathcal{Q}(m,\alpha);m) = p_1(\mathcal{Q}(m,\alpha);m,\alpha) < p_1(\mathcal{Q}(m,\alpha);m,\alpha')$. It follows immediately from Lemma 5 that $\mathcal{Q}(m,\alpha) > \mathcal{Q}(m,\alpha')$. Furthermore, since $\mathcal{P}(m,\alpha) = p_2(\mathcal{Q}(m,\alpha);m)$ and $\mathcal{P}(m,\alpha') = p_2(\mathcal{Q}(m,\alpha');m)$ and p_2 is monotonically decreasing, this implies that $\mathcal{P}(m,\alpha) > \mathcal{P}(m,\alpha')$.

We now prove monotonicity with respect to m. It follows from (50) that for fixed r and $\alpha > 0$,

$$m'\mathcal{Q}(m',\alpha) < m\mathcal{Q}(m,\alpha) \Leftrightarrow m'\mathcal{P}(m',\alpha) < m\mathcal{P}(m,\alpha). \tag{52}$$

Let $m' > m$, and suppose that $m'\mathcal{P}(m',\alpha) < m\mathcal{P}(m,\alpha)$. It follows that

$$\mathcal{Q}(m',\alpha) = g(m'\mathcal{P}(m',\alpha)) > g(m\mathcal{P}(m,\alpha)) = \mathcal{Q}(m,\alpha),$$

which contradicts (52). Therefore, our supposition was incorrect; it must be that both $m'\mathcal{P}(m',\alpha) > m\mathcal{P}(m,\alpha)$ and $m'\mathcal{Q}(m',\alpha) > m\mathcal{Q}(m,\alpha)$. By definition (see (6) and (7)), we have $\mathcal{Q}(m,\alpha) = g(m\mathcal{P}(m,\alpha))$ and $\mathcal{P}(m,\alpha) = \alpha\beta g(r\beta m\mathcal{Q}(m,\alpha))$. Applying the fact that g is decreasing (see Lemma 4), we conclude that $\mathcal{Q}(m',\alpha) < \mathcal{Q}(m,\alpha)$ and $\mathcal{P}(m',\alpha) < \mathcal{P}(m,\alpha)$, as claimed.

Having established monotonicity of $\mathcal{P}$ and $\mathcal{Q}$, we move on to evaluating their limits as $m \to \infty$. When $r < \alpha$, (50) implies that $1 - e^{r\beta m\mathcal{Q}(m,\alpha)} \le r/\alpha$, so $m\mathcal{Q}(m,\alpha)$ is bounded above. This implies both that $\mathcal{Q}(m,\alpha) \to 0$ as $m \to \infty$ and that $\mathcal{P}(m,\alpha) = \alpha\beta g(r\beta m\mathcal{Q}(m,\alpha))$ is bounded away from zero. It follows from (50) that

$$\lim_{m\to\infty} 1 - e^{-r\beta m\mathcal{Q}(m,\alpha)} = \lim_{m\to\infty} \frac{r}{\alpha}(1 - e^{-m\mathcal{P}(m,\alpha)}) = \frac{r}{\alpha},$$

so $r\beta m\mathcal{Q}(m,\alpha) \to -\log(1 - \frac{r}{\alpha}) = \log\left(\frac{\alpha/r}{\alpha/r-1}\right)$ and $\mathcal{P}(m,\alpha) = \alpha\beta g(r\beta m\mathcal{Q}(m,\alpha)) \to \alpha\beta f_2(\alpha/r)$.

Analogously, when $r > \alpha$, (50) implies that $1 - e^{-m\mathcal{P}(m,\alpha)} \le \alpha/r$. This means that $m\mathcal{P}(m,\alpha)$ is bounded above, so $\mathcal{P}(m,\alpha) \to 0$ and $\mathcal{Q}(m,\alpha) = g(m\mathcal{P}(m,\alpha))$ is bounded away from zero. Applying (50) we get that

$$\lim_{m\to\infty} 1 - e^{-m\mathcal{P}(m,\alpha)} = \lim_{m\to\infty} \frac{\alpha}{r}(1 - e^{-rm\beta\mathcal{Q}(m,\alpha)}) = \frac{\alpha}{r},$$

so $m\mathcal{P}(m,\alpha) \to \log\left(\frac{r/\alpha}{r/\alpha-1}\right)$ and $\mathcal{Q}(m,\alpha) = g(m\mathcal{P}(m,\alpha)) \to f_2(r/\alpha)$.

Finally, when $r = \alpha$, (50) implies that $\alpha\beta\mathcal{Q}(m,\alpha) = \mathcal{P}(m,\alpha) = \alpha\beta g(r\beta m\mathcal{Q}(m,\alpha))$. Since g is strictly decreasing, $g(r\beta mq) \le g(r\beta m)$, and $g(r\beta m) \to 0$ as $m \to \infty$. We conclude that $g(r\beta mq) \to 0$ uniformly in q as $m \to \infty$; since $\mathcal{Q}(m,\alpha)$ is the solution to $q = g(r\beta mq)$, it follows that $\mathcal{Q}(m,\alpha) \to 0$ as well as $m \to \infty$. Since $\mathcal{P}(m,\alpha) = \alpha\beta\mathcal{Q}(m,\alpha)$, this implies $\mathcal{P}(m,\alpha) \to 0 = f_2(1)$, completing the proof.

E.2. Section 4.3: Mean field equilibrium.

We now prove Theorem 1, which states that mean field equilibria exist and are unique.

Proof of Theorem 1. We prove the theorem in two steps. First, we fix the employer strategy to be ϕ^α, and allow applicants to respond optimally. We show in Lemma 7 that for any $\alpha \in [0,1]$, there exists a unique strategy m such that

$$m = \mathcal{M}(\mathcal{P}(m,\alpha)). \tag{53}$$

Let m_α denote the unique choice of m satisfying (53).

Second, we endogenize the employer's choice of α; Lemma 10 shows that there is exactly one value of α such that $\alpha \in \mathcal{A}(\mathcal{Q}(m_\alpha,\alpha))$, i.e. such that (m_α,α) is a mean field equilibrium.

We start with the following lemma, regarding "partial" equilibrium among the applicants given a fixed value of α.

Lemma 7. *Given $\alpha \in [0,1]$, there exists a unique value m_α satisfying:*

$$m_\alpha = \mathcal{M}(\mathcal{P}(m_\alpha,\alpha)). \tag{54}$$

Furthermore: (1) $m_\alpha = 0$ if and only if $\alpha \le c'_a$; and (2) $\mathcal{P}(m_\alpha,\alpha)$ is strictly increasing in α.

Proof. Refer to Figure 8. Any solution to (53) is equivalent to finding a solution to the following system of equations:

$$p = \mathcal{P}(m,\alpha); \quad m = \mathcal{M}(p).$$

As discussed in Section 4.2, the value $\mathcal{P}(m,\alpha)$ is the (unique) solution to

$$\mathcal{P}(m,\alpha) = \alpha\beta g(rm\beta g(m\mathcal{P}(m,\alpha))) = \alpha\beta g(r\beta(1 - e^{-m\mathcal{P}(m,\alpha)})/\mathcal{P}(m,\alpha)). \tag{55}$$

We now incorporate the fact that m should be optimally chosen by applicants. Because $g(x) \le 1$ (see Lemma 4), we have $\mathcal{P}(m,\alpha) \le \alpha\beta$. Thus, if $\alpha \le c'_a = c_a/\beta$, then $\mathcal{P}(m,\alpha) \le c_a$ for all m, and thus $\mathcal{M}(\mathcal{P}(m,\alpha)) = 0$ for all m, so $m = 0$ is the unique solution to (53).

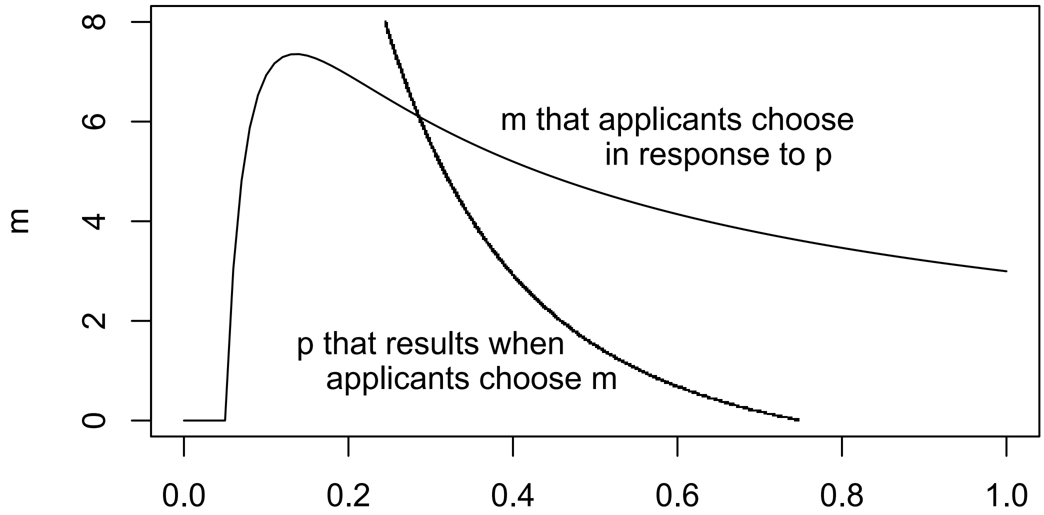

Figure 8 A visualization of Lemma 7. The value $\mathcal{P}(m,\alpha)$ that results when applicants choose m is the solution to (55). Lemma 7 states that this line intersects with the applicant best response function $\mathcal{M}(p)$ at a unique point.

Thus, we can substitute (53) into (3) to obtain:

$$p = \alpha\beta g(r\beta(1-\beta c_a'/p)/p). \tag{56}$$

It suffices to show that (56) has a unique solution $p \in (0,\alpha\beta)$, as the desired (unique) m_α is then $\mathcal{M}(p)$.

To see this, multiply each side by $r/(\alpha p)$ and substitute $x = \beta c_a'/p$ to get

$$\frac{r}{\alpha} = \frac{r}{c_a'} x g(rx(1-x)/c_a') = \frac{1-e^{-(r/c_a')x(1-x)}}{(1-x)}. \tag{57}$$

By Lemma 8 (see below), the right side of (57) is strictly increasing in x, takes the value zero at $x=0$, and approaches r/c_a' as $x \to 1$. Because $\alpha > c_a'$, this implies that there is a unique solution x to (57) and therefore a unique solution p to (56). Furthermore, evaluating the right side of (57) at $x = c_a'/\alpha$ indicates that the solution x to (57) is greater than c_a'/α and thus the solution $p = \beta c_a'/x$ to (56) is less than $\alpha\beta$.

All that remains to show is that $\mathcal{P}(m_\alpha,\alpha)$ is strictly increasing in α. For $\alpha \le c_a'$, we have $m_\alpha = 0$, so $\mathcal{P}(m_\alpha,\alpha) = \mathcal{P}(0,\alpha) = \alpha\beta$, which is strictly increasing. When $\alpha > c_a'$, we have $\mathcal{P}(m_\alpha,\alpha) = \beta c_a'/x(\alpha)$, where $x(\alpha)$ is the solution to (57). Lemma 8 implies that the right side is strictly increasing in x, so $x(\alpha)$ is strictly decreasing in α and thus $\mathcal{P}(m_\alpha,\alpha)$ is strictly increasing.

Our proof of Lemma 7 used the following fact.

Lemma 8. *For any $a>0$ the function $y(x) = \frac{1-e^{-ax(1-x)}}{1-x}$ for $x \in [0,1)$ is strictly increasing in x, taking the value 0 when $x=0$ and approaching a as $x \to 1$.*

Proof. Evaluating $y(0)$ and $\lim_{x\to 1} y(x)$ is straightforward, so we move on to proving that $y(x)$ is strictly increasing. Differentiating with respect to x and rearranging terms, we get

$$\frac{dy}{dx} = \frac{ae^{-ax(1-x)}}{(1-x)^2}\left(2x^2 + \frac{1}{a}(e^{ax(1-x)}-1) - 3x + 1\right). \tag{58}$$

We wish to show that this expression is positive for $x \in (0,1)$. The first term is clearly positive, so let's consider the second term. Since $e^{ax(1-x)} \ge 1 + ax(1-x)$, we have for $x \in (0,1)$:

$$\left(2x^2 + \frac{1}{a}(e^{ax(1-x)}-1) - 3x + 1\right) \ge 2x^2 + x(1-x) - 3x + 1 = (1-x)^2 > 0. \tag{59}$$

Before proceeding, we require some properties of the solution m_α to (54).

LEMMA 9. *$\mathcal{Q}(m_\alpha,\alpha)$ is nonincreasing in α. If $c'_a<1$, $\mathcal{Q}(m_\alpha,\alpha)=1$ for $\alpha\in[0,c'_a]$ and is strictly decreasing for $\alpha\in(c'_a,1]$. If $c'_a\ge 1$, $\mathcal{Q}(m_\alpha,\alpha)=1$ for all $\alpha\in[0,1]$.*

Proof. From Lemma 7, we have that $m_\alpha=0$ if and only if $\alpha\le c'_a$. When $m_\alpha=0$, we have $\mathcal{Q}(m_\alpha,\alpha)=g(0)=1$. If $c'_a<1$, for $\alpha>c'_a$, Lemma 7 implies that $m_\alpha>0$ and thus by (3) we have:

$$m_\alpha\mathcal{P}(m_\alpha,\alpha)=\log\left(\frac{\mathcal{P}(m_\alpha,\alpha)}{\beta c'_a}\right). \tag{60}$$

Lemma 7 implies that the right hand side (and therefore the left hand side) is strictly increasing in α. But

$$\mathcal{Q}(m_\alpha,\alpha)=g(m_\alpha\mathcal{P}(m_\alpha,\alpha)). \tag{61}$$

Because g is a strictly decreasing function (see Lemma 4), this completes the proof.

The following lemma completes the proof of Theorem 1, by endogenizing the employers' choice of α.

LEMMA 10. *Define m_α to be the unique solution to (53). If $c'_s<1$, there is a unique value of $\alpha\in[0,1]$ such that*

$$\alpha\in\mathcal{A}(\mathcal{Q}(m_\alpha,\alpha)), \tag{62}$$

i.e. such that (m_α,α) is a mean field equilibrium.

Proof. Our proof leverages the following intuition. Consider fixing the employers' strategy to be ϕ^1, i.e., employers always enter and screen. If in that case the resulting applicant availability is high enough (under the optimal seller response), then this will be an MFE. On the other hand, if the resulting applicant availability is too low, some employers will choose to exit the market. The key phenomenon that we exploit is that in this case, applicant availability is monotonically increasing as employers leave the market, increasing until exactly the point where employers are indifferent between entering and exiting. This indifference point is precisely where applicant availability q is equal to the (scaled) screening cost c'_s.

Formally, recall from Section 4.1.2 that

$$\mathcal{A}(q)=\begin{cases}\{0\}, & \text{if } q<c'_s;\\ [0,1], & \text{if } q=c'_s;\\ \{1\}, & \text{if } q>c'_s.\end{cases} \tag{63}$$

If $\mathcal{Q}(m_1,1)\ge c'_s$, then $\alpha=1$ solves (62). For any $\alpha<1$, since $c'_s<1$, it follows by Lemma 9 that regardless of the value of c_a, we have $\mathcal{Q}(m_\alpha,\alpha)>c'_s$. Thus, for any $\alpha<1$, $\alpha\notin\mathcal{A}(\mathcal{Q}(m_\alpha,\alpha))=\{1\}$, so $\alpha=1$ is the unique solution to (62).

Now suppose that $\mathcal{Q}(m_1,1)<c'_s$. By Lemma 9, for any $\alpha\le c'_a$, we have $\mathcal{Q}(m_\alpha,\alpha)=1>c'_s$. By continuity and monotonicity of $\mathcal{Q}$ (see Lemma 9), there exists exactly one value $\alpha'\in(c'_a,1)$ such that $\mathcal{Q}(m_{\alpha'},\alpha')=c'_s$. Clearly, $\alpha'\in\mathcal{A}(\mathcal{Q}(m_{\alpha'},\alpha'))=\mathcal{A}(c'_s)=[0,1]$, so α' solves (62). Furthermore, for any $\alpha<\alpha'$, we have $\mathcal{Q}(m_\alpha,\alpha)>c'_s$ and thus $\mathcal{A}(\mathcal{Q}(m_\alpha,\alpha))=\{1\}$. For any $\alpha>\alpha'$, we have $\mathcal{Q}(m_\alpha,\alpha)<c'_s$ and thus $\mathcal{A}(\mathcal{Q}(m_\alpha,\alpha))=\{0\}$. Therefore, α' is the unique solution to (62).

E.3. Section 4.4: the regulated market.

In this section we prove Proposition 3, which states that there is a unique equilibrium in the regulated market. It is the same as the equilibrium of the unregulated market if the limit ℓ does not bind, and otherwise involves applicants selecting $m_a=\ell$.

Our analysis requires a partial equilibrium characterization of the employers' behavior given a choice of m by applicants. In particular, for a fixed m, we show there exists a unique value α_m satisfying:

$$\alpha_m\in\mathcal{A}(\mathcal{Q}(m,\alpha_m)), \tag{64}$$

where we recall from Section 4.1.2 that:

$$\mathcal{A}(q) = \begin{cases} \{0\}, & \text{if } q < c_s'; \\ [0,1], & \text{if } q = c_s'; \\ \{1\}, & \text{if } q > c_s'. \end{cases} \tag{65}$$

We have the following lemma.

LEMMA 11. *For any fixed m, there exists a unique solution α_m to (64). Furthermore, the quantity α_m is weakly decreasing in m.*

Proof. Lemma 6 states that $\mathcal{Q}(m,1)$ is decreasing in m. Let $\overline{m} = \sup\{m : \mathcal{Q}(m,1) \geq c_s'\}$, which may be infinite.

If $m \leq \overline{m}$, then $\mathcal{Q}(m,1) \geq c_s'$ and $\alpha_m = 1$ solves (64). Furthermore, for any $\alpha < 1$, $\mathcal{Q}(m,\alpha) > \mathcal{Q}(m,1)$ by Lemma 6, and thus $\alpha \notin \mathcal{A}(\mathcal{Q}(m,\alpha)) = \{1\}$.

If $m > \overline{m}$, then $\mathcal{Q}(m,1) < c_s'$, so $1 \notin \mathcal{A}(\mathcal{Q}(m,1)) = \{0\}$, and $0 \notin \mathcal{A}(\mathcal{Q}(m,0)) = \mathcal{A}(1) = \{1\}$. It follows that any solution α_m to (64) must satisfy $0 < \alpha_m < 1$, and thus must satisfy $\mathcal{Q}(m,\alpha_m) = c_s'$. Lemma 6 states that $\mathcal{Q}(m,\alpha)$ strictly increases continuously to 1 as α decreases, implying both that there is a unique α_m such that $\mathcal{Q}(m,\alpha_m) = c_s'$, and that this value α_m is decreasing in m.

LEMMA 12. *For $m > 0$, the pair (m,α_m) is an MFE if and only if $h(m) = -\log c_a'$, where*

$$h(m) = m\mathcal{P}(m,\alpha_m) - \log\left(\mathcal{P}(m,\alpha_m)/\beta\right). \tag{66}$$

The function h is strictly increasing, and unbounded, with $h(0) = 0$.

Proof. Note that a mean-field equilibrium is exactly a pair (m,α_m) such that $m = \mathcal{M}(\mathcal{P}(m,\alpha_m))$. From the definition of $\mathcal{M}$ in (3) it follows that the pair (m,α_m) with $m > 0$ is an equilibrium if and only if $m = \frac{1}{\mathcal{P}(m,\alpha_m)} \log\left(\frac{\mathcal{P}(m,\alpha_m)}{c_a}\right)$, or equivalently

$$m\mathcal{P}(m,\alpha_m) - \log(\mathcal{P}(m,\alpha_m)/\beta) = -\log c_a'. \tag{67}$$

We already know from Theorem 1 that for any $c_a' < 1$ and any β, there is a unique equilibrium with $m > 0$, i.e. a unique solution to $h(m) = -\log c_a'$. This implies that h is invertible (if $h(m) = h(m')$ for $m \neq m'$, then for some c_a', there would be multiple mean-field equilibria). Because h is continuous, this also implies that h is monotonic. Furthermore, the existence of a solution to $h(m) = -\log c_a'$ for any $c_a' < 1$ (by Theorem 1) implies that $h(0) = 0$ and that h is unbounded.

We now prove the desired result.

Proof of Proposition 3. Let (m^*,α^*) be the mean-field equilibrium in the original game, and let (m_ℓ^*,α_ℓ^*) be any equilibrium of the game with application limit ℓ, meaning that $m_\ell^* = \mathcal{M}_\ell(\mathcal{P}(m_\ell^*,\alpha_\ell^*))$, and $\alpha_\ell^* \in \mathcal{A}(\mathcal{Q}(m_\ell^*,\alpha_\ell^*))$.

Immediately from (8) (which states that $\mathcal{M}_\ell(p) = \min(\ell,\mathcal{M}(p))$), we get that if there is an equilibrium of the regulated market with $m_\ell^* < \ell$, it must be that (m_ℓ^*,α_ℓ^*) is an equilibrium in the unregulated market and thus $(m_\ell^*,\alpha_\ell^*) = (m^*,\alpha^*)$. Therefore, the only candidate equilibria in the game with application limit ℓ are (m^*,α^*) and (ℓ,α_ℓ).

It is clear that (m^*,α^*) is an equilibrium of the game with application limit ℓ if and only if $m^* \leq \ell$. To complete the proof of Proposition 3, we claim that the pair (ℓ,α_ℓ) is an equilibrium of the regulated market if and only if $m^* \geq \ell$.

To prove this claim, first suppose that $m^* < \ell$. Then $-\log c_a' = h(m^*) < h(\ell)$ by Lemma 12. Conversely, if $m^* \geq \ell$, then $-\log c_a' = h(m^*) \geq h(\ell)$. In other words, for $c_a' \leq 1$,

$$m^* \geq \ell \Leftrightarrow h(\ell) \leq -\log c_a'. \tag{68}$$

Straightforward manipulation of the equation (66) defining h and an application of (3) reveals that for $\ell > 0$ and $c_a' \leq 1$,

$$\ell \leq \mathcal{M}(\mathcal{P}(\ell,\alpha_\ell)) \Leftrightarrow h(\ell) \leq -\log c_a'. \tag{69}$$

Combining (68) and (69), we get that $m^* \geq \ell$ if and only if $\ell \leq \mathcal{M}(\mathcal{P}(\ell, \alpha_\ell))$. Furthermore, $\ell \leq \mathcal{M}(\mathcal{P}(\ell, \alpha_\ell))$ implies $\mathcal{M}_\ell(\mathcal{P}(\ell, \alpha_\ell)) = \ell$, and conversely $\ell > \mathcal{M}(\mathcal{P}(\ell, \alpha_\ell))$ implies $\mathcal{M}_\ell(\mathcal{P}(\ell, \alpha_\ell)) = \mathcal{M}(\mathcal{P}(\ell, \alpha_\ell))$ (using $\mathcal{M}_\ell(p) = \min(\ell, \mathcal{M}(p))$ from (8)). Putting it all together, we get that

$$m^* \geq \ell \Leftrightarrow \ell \leq \mathcal{M}(\mathcal{P}(\ell, \alpha_\ell)) \Leftrightarrow \ell = \mathcal{M}_\ell(\mathcal{P}(\ell, \alpha_\ell)). \tag{70}$$

In other words, $m^* \geq \ell$ if and only if (ℓ, α_ℓ) is an equilibrium of the game with application limit ℓ.

Appendix F: Proofs: Section 5.

F.1. Preliminaries

We begin with four results used as building blocks in our proofs of Theorems 2, 3, 4 and 5.

The first of these results, Proposition 5, shows that the welfare expressions for Π_a and Π_e given in (10) and (11), respectively, can be expressed as functions of only the model parameters (r, c'_a, c'_s), the value α selected by employers, and the quantity $m\mathcal{P}(m, \alpha)$.

Our second result, Proposition 6, shows that optimizing over application limits is equivalent to optimizing over applicants' match probability.

Our third result, Proposition 7, shows that both applicant and employer welfare are unimodal in the application limit ℓ.

Proposition 5. *For any m and α, we have*

$$\Pi_a(m, \alpha) = 1 - e^{-m\mathcal{P}(m,\alpha)} - c'_a m\mathcal{P}(m, \alpha)/(\alpha g(-\log(1 - \frac{r}{\alpha}(1 - e^{-m\mathcal{P}(m,\alpha)})))) \tag{71}$$

$$\Pi_e(m, \alpha) = r(1 - e^{-m\mathcal{P}(m,\alpha)} - c'_s m\mathcal{P}(m, \alpha)), \tag{72}$$

where $g(\cdot)$ was defined in (40). *Furthermore,*

$$\Pi_a^* = 1 - (1 + m^* p^*) e^{-m^* p^*}. \tag{73}$$

$$\Pi_e^* = r(1 - e^{-m^* p^*} - c'_s m^* p^*). \tag{74}$$

In fact,

Proof. Recall that the mean field equations (6) and (7) imply that $\mathcal{P}(m, \alpha)$ and $\mathcal{Q}(m, \alpha)$ solve the following equations

$$rm\mathcal{P}(m, \alpha)\mathcal{Q}(m, \alpha) = \alpha(1 - e^{-rm\beta\mathcal{Q}(m,\alpha)}) = r(1 - e^{-m\mathcal{P}(m,\alpha)}). \tag{75}$$

Applying this to (11) yields (72), from which (74) follows immediately.

To get to (71), we note that $c_a m = (\alpha\beta) c'_a m/\alpha$. The mean-field equation (7) for $\mathcal{P}(m, \alpha)$ implies that

$$\alpha\beta = \mathcal{P}(m, \alpha)/g(rm\beta\mathcal{Q}(m, \alpha)) = \mathcal{P}(m, \alpha)/g(-\log(1 - \frac{r}{\alpha}(1 - e^{-m\mathcal{P}(m,\alpha)}))),$$

where the final equality follows from solving the second part of (75) for $rm\beta\mathcal{Q}(m, \alpha)$. Combining these facts and substituting into (10) yields (71). We obtain (73) by applying (3) (which gives the applicant's best response as a function of p) to (10).

From Proposition 5, we have the following corollary.

Lemma 13. *For any $\alpha < 1$ and any $m > 0$, there exists $m' < m$ such that $\Pi_a(m', 1) > \Pi_a(m, \alpha)$ and $\Pi_e(m', 1) = \Pi_e(m, \alpha)$.*

Proof. Note that by Lemma 6, $\mathcal{Q}(m, 1) < \mathcal{Q}(m, \alpha) \leq \mathcal{Q}(0, \alpha) = 1$, and furthermore $\mathcal{Q}(\cdot, 1)$ is continuously decreasing. It follows that for some $m' < m$, $\mathcal{Q}(m', 1) = \mathcal{Q}(m, \alpha)$. Because the mean-field consistency equation (6) states that $\mathcal{Q} = g(m\mathcal{P})$, it follows that $m'\mathcal{P}(m', 1) = m\mathcal{P}(m, \alpha)$. This fact, combined with (72) from Proposition 5, implies that $\Pi_e(m', 1) = \Pi_e(m, \alpha)$. Furthermore, applying (10) from Proposition 5, we see that

$$\Pi_a(m', 1) = 1 - e^{-m'\mathcal{P}(m',1)} - c_a m' = 1 - e^{-m\mathcal{P}(m,\alpha)} - c'_a m' > 1 - e^{-m\mathcal{P}(m,\alpha)} - c_a m = \Pi_a(m, \alpha).$$

Lemma 13 states that the Pareto frontier of (Π_e, Π_a) consists only of points where $\alpha = 1$. When $\alpha = 1$, Proposition 5 gives Π_a and Π_e as functions of only r, c'_a, c'_s and the quantity $m\mathcal{P}(m,1)$. By Lemma 6, the quantity $m\mathcal{P}(m,1)$ is strictly increasing in m, so rather than optimizing over the choice of m, one can equivalently optimize over the choice of $m\mathcal{P}(m,1)$ or over the fraction of applicants that match, given by

$$y(m) = 1 - e^{-m\mathcal{P}(m,1)}, \tag{76}$$

a monotone function of $m\mathcal{P}(m,1)$. Motivated by this, for fixed r, c'_a, c'_s we define

$$\tilde{\Pi}_a(y) = y(1 - c'_a \frac{\log(1-y)}{y} \frac{\log(1-ry)}{ry}) \tag{77}$$

$$\tilde{\Pi}_e(y) = r(y + c'_s \log(1-y)), \tag{78}$$

so that

$$\Pi_a(m,1) = \tilde{\Pi}_a(y(m)), \qquad \Pi_e(m,1) = \tilde{\Pi}_e(y(m)). \tag{79}$$

We define y^* to be the applicant match rate in equilibrium, given by

$$y^* = 1 - e^{-m^* p^*}, \tag{80}$$

We now show that the problem of choosing an application limit ℓ to optimize Π_a and Π_e is equivalent to choosing an applicant match rate y to optimize $\tilde{\Pi}_a$ and $\tilde{\Pi}_e$.

Proposition 6.

$$\sup_\ell \Pi_a^\ell = \sup_{y \le y^*} \tilde{\Pi}_a(y) \qquad \sup_\ell \Pi_e^\ell = \sup_{y \le y^*} \tilde{\Pi}_e(y)$$

Proof. Proposition 3 states that $m^*_\ell = \min(m^*, \ell)$, so without loss of generality we can restrict attention to limits that bind – that is, limits $\ell \le m^*$. Define

$$\bar{\ell} = \sup\{\ell \le m^* : \alpha^*_\ell = 1\}.$$

For $\ell > \bar{\ell}$, we have $\alpha^*_\ell < 1$ by definition of $\bar{\ell}$. In this case, Lemma 13 implies that both sides weakly prefer a lower volume of applications m', and Proposition 3 states that this lower volume of applications can be achieved by setting $\ell = m'$. In other words, there is no loss to only considering limits $\ell \le \bar{\ell}$ (i.e. limits such that all employers choose to screen):

$$\sup_\ell \Pi_a^\ell = \sup_{\ell \le \bar{\ell}} \Pi_a^\ell \qquad \sup_\ell \Pi_e^\ell = \sup_{\ell \le \bar{\ell}} \Pi_e^\ell \tag{81}$$

For such ℓ, we have

$$\Pi_a^\ell = \Pi_a(\ell, 1) = \tilde{\Pi}_a(y(\ell)), \qquad \Pi_e^\ell = \Pi_e(\ell, 1) = \tilde{\Pi}_e(y(\ell)). \tag{82}$$

It follows that

$$\sup_{\ell \le \bar{\ell}} \Pi_a^\ell = \sup_{\ell \le \bar{\ell}} \tilde{\Pi}_a(y(\ell)), \qquad \sup_{\ell \le \bar{\ell}} \Pi_e^\ell = \sup_{\ell \le \bar{\ell}} \tilde{\Pi}_e(y(\ell)). \tag{83}$$

It follows from Lemma 6 that the function $y(\cdot)$ defined in (76) is strictly increasing, with $y(0) = 0$. Furthermore, we claim that $y(\bar{\ell}) = y^*$, so that

$$\sup_{\ell \le \bar{\ell}} \tilde{\Pi}_a(y(\ell)) = \sup_{y \le y^*} \tilde{\Pi}_a(y), \qquad \sup_{\ell \le \bar{\ell}} \tilde{\Pi}_e(y(\ell)) = \sup_{y \le y^*} \tilde{\Pi}_e(y). \tag{84}$$

Combining (81), (82), (84) yields the claimed result.

All that remains is to prove that $y(\overline{\ell}) = y^*$. If $\overline{\ell} = m^*$, this is immediate: $\alpha^* = 1$ by definition of $\overline{\ell}$, so $p^* = \mathcal{P}(m^*, 1)$ and

$$y(\overline{\ell}) = y(m^*) = 1 - e^{-m^*\mathcal{P}(m^*,1)} = 1 - e^{-m^*p^*} = y^*.$$

Otherwise, $\overline{\ell} < m^*$, and $\Pi_e^{\overline{\ell}} = \Pi_e(\overline{\ell}, 1) = 0$ (it cannot be that $\Pi_e^{\overline{\ell}} > 0$, or else we would have $\Pi_e^{\overline{\ell}+\varepsilon} > 0$ for all sufficiently small ε, implying that $\alpha^*_{\overline{\ell}+\varepsilon} = 1$, contradicting the definition of $\overline{\ell}$). But if $\Pi_e(\overline{\ell}, 1) = 0$, then (11) implies that $\mathcal{Q}(\ell, 1) = c'_s = \mathcal{Q}(m^*, \alpha^*)$. Because $\mathcal{Q}(\ell, 1) = g(\ell\mathcal{P}(\ell, 1))$ and $\mathcal{Q}(m^*, \alpha^*) = g(m^*p^*)$ it follows from the fact that g is strictly monotonic that $\ell\mathcal{P}(\ell, 1) = m^*p^*$, and therefore that $y(\overline{\ell}) = 1 - e^{-m^*p^*} = y^*$, as claimed.

We conclude our final result for this section, which shows that both applicant and employer welfare are unimodal in the application limit ℓ for values of ℓ such that $\alpha^* = 1$. Before stating the result, we note that using Proposition 7 we can conclude that the welfare on each side is unimodal in ℓ for the entire range of possible values of ℓ, including if $\alpha^* < 1$: Let $\ell_0 = \sup\{\ell : \alpha^* = 1\}$. We will show that welfare on each side is non-increasing in ℓ above ℓ_0. For all $\ell > \ell_0$, as ℓ increases we know that α^* will fall so as to keep $\Pi_e^{\ell} = 0$ and $q^*_\ell = c'_s$ fixed, hence keeping $m^*_\ell p^*_\ell$ fixed as per (6). So since m^*_ℓ is non-decreasing in ℓ, we conclude that $\Pi_a^{\overline{\ell}} = 1 - e^{-m^*_\ell p^*_\ell} - c_a m^*_\ell$ is non-increasing in ℓ. This establishes unimodality of welfare on both sides for all values of ℓ.

Proposition 7. *For any r, c'_a, c'_s, the functions $\Pi_a(\cdot, 1)$, $\Pi_e(\cdot, 1)$, $\tilde{\Pi}_a(\cdot)$, and $\tilde{\Pi}_e(\cdot)$ are increasing at zero, and are unimodal. $\Pi_a(\cdot, 1)$ and $\tilde{\Pi}_a(\cdot)$ have a unique local maximum, and $\Pi_e(\cdot, 1)$ and $\tilde{\Pi}_e(\cdot)$ either have a unique local maximum or are strictly increasing.*

Proof. We define

$$\mathrm{AW}(x) = 1 - e^{-x} - c'_a x / g(-\log(1 - r(1 - e^{-x}))) \tag{85}$$

$$\mathrm{EW}(x) = r(1 - e^{-x} - c'_s x) \tag{86}$$

Proposition 5 establishes that $\Pi_e(m, 1) = \mathrm{EW}(m\mathcal{P}(m, 1)) = \tilde{\Pi}_e(1 - e^{-m\mathcal{P}(m,1)})$, and $\Pi_a(m, 1) = \mathrm{AW}(m\mathcal{P}(m, 1)) = \tilde{\Pi}_a(1 - e^{-m\mathcal{P}(m,1)})$. Lemma 6 states that $m\mathcal{P}(m, 1)$ is increasing in m. Thus, the unimodality of $\Pi_a(\cdot, 1)$, $\Pi_e(\cdot, 1)$, $\tilde{\Pi}_a$ and $\tilde{\Pi}_e$ follow from the unimodality of AW and EW.

It is straightforward to show that EW is concave, with positive first derivative at zero. Thus, all that remains is to prove that for all $r > 0$ and $c'_a \in (0, 1)$, AW is unimodal. Because AW has a continuous first derivative, $\mathrm{AW}'(0) = 1 - c'_a > 0$, and AW is negative for sufficiently large x, it suffices to show that there is a unique solution to $\mathrm{AW}'(x) = 0$.

For the purposes of this proof, for fixed r we define $a(x) = -\log(1 - r(1 - e^{-x}))$, $b(x) = g(a(x))$, and $u(x) = x/b(x)$, so that $\mathrm{AW}(x) = 1 - e^{-x} - c'_a u(x)$. Then

$$\mathrm{AW}'(x) = e^{-x} - c'_a u'(x) = 0 \Leftrightarrow u'(x)e^x = 1/c'_a. \tag{87}$$

Note that

$$u'(x) = \frac{1}{b(x)} - x\frac{b'(x)}{b(x)^2}, \tag{88}$$

so $u'(0)e^0 = 1$. It follows from (87) that $\mathrm{AW}'(x) = 0$ has a unique solution for all $c'_a \in (0, 1)$ if and only if $u'(x)e^x$ is (strictly) increasing and unbounded. We see that

$$\frac{d}{dx}u'(x)e^x = e^x(u'(x) + u''(x)).$$

To show that $u'(x)e^x$ is increasing and unbounded, we will show that $u'(x) + u''(x) > 1$.

By differentiating (88), we see that

$$u'(x) + u''(x) = \frac{1}{b(x)}\left(1 - \frac{xb'}{b} - \frac{xb''}{b} + 2x\left(\frac{b'}{b}\right)^2 - 2\frac{b'}{b}\right).$$

Note that $b'(x) = g'(a(x))a'(x) < 0$, so every term in the above sum except $-x\frac{b''}{b}$ is clearly positive. Since $b(x) \le 1$, to show that $u'(x) + u''(x) > 1$, it suffices to show that $-\frac{xb'}{b} - \frac{xb''}{b} > 0$, or equivalently, $b'(x) + b''(x) < 0$, or equivalently

$$\frac{b''(x)}{b'(x)} > -1. \tag{89}$$

We note (omitting the algebra) that

$$b'(x) = g'(a(x))a'(x) \tag{90}$$
$$b''(x) = g''(a(x))a'(x)^2 + g'(a(x))a''(x). \tag{91}$$
$$a'(x) = re^{a(x)-x} \tag{92}$$
$$a''(x) = \frac{r-1}{r}e^x a'(x)^2 = (r-1)e^{a(x)}a'(x) \tag{93}$$

We apply (90), (91), followed by (93), to conclude that

$$\frac{b''(x)}{b'(x)} = a'(x)\frac{g''(a(x))}{g'(a(x))} + \frac{a''(x)}{a'(x)} \tag{94}$$
$$= a'(x)\frac{g''(a(x))}{g'(a(x))} + (r-1)e^{a(x)}. \tag{95}$$

Note that $a'(x) > 0$, and by Lemma 4, $g''(h)/g'(h) > -1$. Hence,

$$\frac{b''(x)}{b'(x)} > -a'(x) + (r-1)e^{a(x)}.$$

Now apply (92) and rearrange to get that

$$-a'(x) + (r-1)e^{a(x)} = -e^{a(x)}(1 - r(1 - e^{-x})) = -1,$$

completing the proof of (89).

F.2. Proofs of Theorems in Section 5.

Proof of Theorem 2. Note that (13) can be rewritten as

$$r = \frac{1 - e^{-\frac{r}{c'_a}\frac{c_a}{\hat{p}}(1-\frac{c_a}{\hat{p}})}}{1 - \frac{c_a}{\hat{p}}}.$$

Uniqueness of $\hat{p}$ follows from Lemma 8, which states that the right side of the above expression is strictly increasing in $c_a/\hat{p}$, equal to $\frac{1-\exp(-r\beta(1-\beta))}{1-\beta} \le \frac{1-(1-r\beta(1-\beta))}{1-\beta} = r\beta < r$ for $c_a/\hat{p} = c_a$ (i.e., for $\hat{p} = 1$) and approaching $r/c'_a > r$ as $c_a/\hat{p} \to 1$ (i.e., for $\hat{p} \to c_a^+$). Happily, we can also infer $\hat{p} \in (c_a, 1)$ at the solution.

To show uniqueness of $\overline{p}$, we note that (14) is equivalent to

$$c'_s = g(\log(\overline{p}/c_a)). \tag{96}$$

Because g is invertible (see Lemma 4), it follows that $\overline{p} = c_a e^{g^{-1}(c'_s)}$. Happily, we can also infer $\overline{p} > c_a$.

We now turn to the welfare expressions. We first show that given a value p^*, the welfare expressions in Theorem 2 are correct. We then show that $p^* = \min(\hat{p}, \overline{p})$.

Given a value p^*, applicants choose m to maximize welfare. The optimal choice of m is given by the function $\mathcal{M}$ in (3). For $p^* \ge c_a$, we have

$$m^* = \mathcal{M}(p^*) = \frac{1}{p^*}\log(p^*/c_a). \tag{97}$$

Combining this with Proposition 5, which gives Π_a^* and Π_e^* as functions of m^*p^*, yields the equations in Theorem 2. We remark that these expressions for welfare on each side are derived using the applicant best response condition as well as the consistency conditions (6) and (7), but do *not* rely upon the employer best response condition. Hence they apply as long as all employers are using the same strategy ϕ^α for some fixed α, and workers are playing a partial equilibrium in response.

We complete the proof by showing that $p^* = \min(\hat{p}, \overline{p})$. Note that this claim passes the sanity check $p^* \in (c_a, 1)$ since $\hat{p} \in (c_a, 1)$ and $\overline{p} > c_a$.

We first consider the case $\hat{p} \leq \overline{p}$. As in (54), define m_1 be the solution to

$$m_1 = \mathcal{M}(\mathcal{P}(m_1, 1)). \tag{98}$$

Note that $\mathcal{P}(m_1, 1) = \hat{p}$ by definition of $\hat{p}$. We will show that $(m^*, \alpha^*) = (m_1, 1)$ is an equilibrium, from which it follows that $p^* = \mathcal{P}(m^*, \alpha^*) = \mathcal{P}(m_1, 1) = \hat{p}$. Suppose applicants are each using the application strategy m_1, and that employers are each using the strategy ϕ^1, i.e., all employers screen. By definition of m_1 as per (98), we know that m_1 is a partial equilibrium between workers. It remains to check that employers are playing a best response, i.e., that the expected welfare of employers $\Pi_e(m_1, 1)$ is non-negative (the alternative for an employer is to not screen, and obtain a welfare of zero). Since the employer welfare expression (11) applies to any partial equilibrium between workers (as argued above), it applies to $(m_1, 1)$ and hence the employer welfare is given by

$$\Pi_e(m_1, 1) = r\left(1 - \frac{c_a}{\hat{p}} + \frac{c_s}{\beta}\log\left(\frac{c_a}{\hat{p}}\right)\right) = r\log(\hat{p}/c_a)\big(g(\log(\hat{p}/c_a)) - c_s'\big)\,. \tag{99}$$

Comparing with (96) which says $g(\log(\overline{p}/c_a)) - c_s' = 0$, we deduce that $g(\log(\hat{p}/c_a)) - c_s' \geq 0$ since $\hat{p} \leq \overline{p}$ and $g(\cdot)$ is monotone decreasing by Lemma 4. Recalling that $\hat{p} > c_a$ as argued above, we infer that $\log(\hat{p}/c_a) > 0$. It follows that $\Pi_e(m_1, 1) \geq 0$, and so employers are indeed playing a best response, and $(m^*, \alpha^*) = (m_1, 1)$ is an equilibrium.

We now turn to the case $\hat{p} > \overline{p}$. In this case, it is straightforward to verify that there is no equilibrium with $\alpha^* = 1$: in such an equilibrium, the reasoning above implies that $p^* = \hat{p}$ and $\Pi_e^* < 0$, so employers would, in fact, prefer to leave without screening, a contradiction. Thus, it must be that $\alpha^* \in (0, 1)$. By (4) this implies that $c_s' = q^*$. Because (3) and (6) imply that $q^* = g(m^*p^*) = g(\log(p^*/c_a))$, comparing with (96), we must have $p^* = \overline{p}$.

Proof of Theorem 3. The fact that $\Pi_e^* = 0$ when the market is screening-limited follows immediately from Theorem 2, as one can directly evaluate Π_e^* taking $p^* = \overline{p}$.

The bound on Π_a^* uses the fact that by (73) in Proposition 5, $\Pi_a^* = 1 - (1 + m^*p^*)e^{-m^*p^*}$. This expression is increasing in m^*p^*, so to derive upper-bounds on Π_a^*, it suffices to provide upper-bounds on m^*p^*. Lemma 6 states that when $r > 1$, $m\mathcal{P}(m, \alpha) \leq \log\frac{r}{r-1}$ for all m and[25] α. Substituting this bound into (73) completes the proof.

LEMMA 14. *Let*

$$z(y) = -y/\log(1 - y). \tag{100}$$

Then z decreasing in y. For $c \in (0, 1)$, we have $z(1 - c) \geq c$.

Proof. Note that $-y/\log(1 - y) = g(-\log(1 - y))$. The monotonicity of z follows from monotonicity of g (Lemma 4) and of $-\log(1 - y)$.

The second part of the lemma follows because for $c \in (0, 1)$, we have $1 - c + c\log c \geq 0$ (to see this, note that the derivative of the left side is $\log c < 0$, and the left side approaches zero as $c \to 1$). Rearranging, we get $c \leq (1 - c)/(-\log c) = z(1 - c)$.

[25] This has a simple interpretation. Of course, when $r > 1$, applicants match with probability at most $1/r$. But the proportion of applicants who match is $1 - e^{-m\mathcal{P}(m,\alpha)}$. From this, we conclude that $r(1 - e^{-m\mathcal{P}(m,\alpha)}) \leq 1$, which can be rearranged to give $m\mathcal{P}(m, \alpha) \leq \log\frac{r}{r-1}$.

Proof of Theorem 4. Proposition 6 states that $\sup_\ell \Pi_e^\ell = \sup_{y \le y^*} \tilde{\Pi}_e(y)$. Proposition 7 (and the argument just before it) establishes that $\tilde{\Pi}_e$ is unimodal, and initially increasing. Because $\tilde{\Pi}_e(0) = 0$, it follows that if $y^* > 0$ and $\tilde{\Pi}_e(y^*) = 0$, then y^* is larger than the maximizer of $\tilde{\Pi}_e$, so

$$\sup_{y \le y^*} \tilde{\Pi}_e(y) = \sup_{y \in [0,1]} \tilde{\Pi}_e(y) = r(1 - c'_s + c'_s \log(c'_s)),$$

with the final equality holding because the unconstrained maximizer of $\tilde{\Pi}_e$ is $y = 1 - c'_s$.

Proposition 6 also states that $\sup_\ell \Pi_a^\ell = \sup_{y \le y^*} \tilde{\Pi}_a(y)$. Proposition 7 states that $\tilde{\Pi}_a$ is unimodal (initially increasing, then decreasing). We will show that when the market is application-limited, $\tilde{\Pi}'(y^*) < 0$, from which it follows that

$$\sup_\ell \Pi_a^\ell = \sup_{y \le y^*} \tilde{\Pi}_a(y) = \sup_{y \le [0,1]} \tilde{\Pi}_a(y). \tag{101}$$

In particular, for any $r > 1$ define $y' = (1 - c'_a)/r$. Then (101) implies that

$$\begin{aligned}
\sup_\ell \Pi_a^\ell &\ge \tilde{\Pi}_a(y') \\
&= y'(1 - \frac{c'_a}{z(y')z(ry')}) \\
&\ge y'(1 - \frac{c'_a}{z(1/r)z(ry')}) \\
&= \frac{1}{r} - \frac{c'_a}{r} + c'_a \log(c'_a) \log\left(\frac{r}{r-1}\right),
\end{aligned}$$

where the first equality follows by definition of $\tilde{\Pi}_a$ and the function z (see Lemma 14), and the second inequality follows because $y' \le 1/r$ and z is positive and decreasing (Lemma 14).

It remains to prove that $\tilde{\Pi}'(y^*) < 0$ when the market is application-limited.

Recall that $\tilde{\Pi}_a(y(m)) = \Pi_a(m,1)$ for all m, where y is as defined in (76). It follows that for any m,

$$y'(m)\tilde{\Pi}'(y(m)) = \frac{d}{dm}\Pi_a(m,1) = (m\frac{d}{dm}\mathcal{P}(m,1) + \mathcal{P}(m,1))e^{-m\mathcal{P}(m,1)} - c_a$$

When the market is application limited, $\alpha^* = 1$, $y^* = y(m^*)$, and $p^* = \mathcal{P}(m^*,1)$, so the above simplifies to

$$\begin{aligned}
y'(m^*)\tilde{\Pi}'(y^*) &= (m^* \frac{d}{dm}\mathcal{P}(m^*,1) + p^*)e^{-m^* p^*} - c_a \\
&= m^* e^{-m^* p^*} \frac{d}{dm}\mathcal{P}(m^*,1) \\
&< 0,
\end{aligned}$$

where the second equality follows because the applicant best response function given in (3) implies that $p^* e^{-m^* p^*} - c_a = 0$, and the inequality follows because $\mathcal{P}(m,1)$ is decreasing by Lemma 6. Lemma 6 also states that $m\mathcal{P}(m,1)$ is strictly increasing in m, from which it follows that $y(m) = 1 - e^{-m\mathcal{P}(m,1)}$ is strictly increasing in m, so $y'(m^*) > 0$ and therefore $\tilde{\Pi}'(y^*) < 0$.

Proof of Theorem 5. By Theorem 2, if the market is screening-limited, then $\alpha^* < 1$ and $\Pi_e^* = 0$. It follows from Lemma 13 that there exists $m' < m^*$ such that $\Pi_a(m',1) > \Pi_a(m^*,\alpha^*)$ and $\Pi_e(m',1) = 0 = \Pi_e(m^*,\alpha^*)$.

Choose $\varepsilon > 0$ small enough that $\Pi_a(m' - \varepsilon, 1) > \Pi_a(m^*, \alpha^*)$. (This is possible by continuity.) We claim that the outcome when participants choose strategies $(m' - \varepsilon, 1)$ Pareto dominates the equilibrium outcome. To see this, note that by (11), $\Pi_e(m',1) = 0$ implies that $\mathcal{Q}(m',1) = c'_s$; by Lemma 6, it follows that $\mathcal{Q}(m' - \varepsilon, 1) > c'_s$, so $\Pi_e(m' - \varepsilon, 1) > 0$.

Furthermore, it follows from Proposition 3 that if the operator sets an application limit $\ell = m' - \varepsilon$, then the strategies chosen by applicants and employers are $(m' - \varepsilon, 1)$. (The fact that employers select $\alpha = 1$ in this case follows from the fact that $\mathcal{A}(q) = \{1\}$ when $q > c'_s$, and $\mathcal{Q}(m' - \varepsilon, 1) > c'_s$.)

Proof of Theorem 6. We first establish the second bullet by showing that for each $\ell \in [0, m^*(c_a)]$, there exists a unique $c > 0$ such that $m^*(c) = \ell$, and that this $c \in [c_a, \beta]$. We will show that $m^*(c)$ is strictly decreasing in c, with $m^*(\beta) = 0$. From there, using Lemma 11, we can infer that the employer equilibrium strategy is identical in the two cases: when there is an application limit ℓ and when the application cost is raised to the unique corresponding value c, since the applicant equilibrium strategies are identical in the two cases (and β and c_s are also identical in the two cases). It will follow that $\Pi_e^\ell = \Pi_e^*(c)$.

The aforementioned claim follows because Lemma 12 states that $m^*(c)$ is the unique solution to $h(m^*(c)) = -\log(c/\beta)$, where the function h is defined in (66), and depends only on r, c_s, β. Lemma 12 also states that $h(0) = 0$, and h is strictly increasing, from which it follows that m^* is strictly decreasing and $m^*(\beta) = 0$.

As for applicant welfare (the first bullet), by Proposition 5, $\Pi_a^* = 1 - (1 + m^*p^*)e^{-m^*p^*}$, which is increasing in m^*p^*. Thus, it suffices to prove that the quantity m^*p^* is weakly decreasing in the application cost c_a. By the applicant best response function (3), $m^*p^* = \log(p^*/c_a)$, so it suffices to prove that c_a/p^* is weakly increasing in c_a.

We know from Theorem 2 that for any c_a such that the market is screening-limited, we have that $\Pi_e^* = 1 - c_a/p^* + c_s' \log(c_a/p^*) = 0$, or equivalently,

$$c_s' = (1 - c_a/p^*)/\log(p^*/c_a) = g(\log(p^*/c_a)).$$

This implies that c_a/p^* is constant (equal to $e^{-g^{-1}(c_s')}$) on the set

$$\{c_a : \text{ the market defined by } (c_a, c_s, r, \beta) \text{ is screening-limited}\}.$$

If c_a is such that the market is application-limited, then rearranging (13) in Theorem 2, reveals that c_a/p^* is the solution x to

$$r = (1 - e^{-\frac{r\beta}{c_a}x(1-x)})/(1-x)\,. \tag{102}$$

By Lemma 8, the expression on the right is increasing in x, and it is clearly decreasing in c_a for fixed x, implying that the solution x to (102) is increasing in c_a, as claimed.

Remark 1. The proof of Theorem 6 establishes an equivalence between the equilibrium that results from an application limit of $\ell \in [0, m^*(c_a)]$, and from raising the application cost to the unique c such that $m^*(c) = \ell$. At this c, the equilibrium satisfies $p^*(c) = p_\ell^*$, $q^*(c) = q_\ell^*$ and $\alpha^*(c) = \alpha_\ell^*$. Here $p^*(c)$, $q^*(c)$ and $\alpha^*(c)$ denote the equilibrium values when the application cost is raised to c.

Proof of Theorem 7. Fix r, β, c_a, c_s and application limit $\ell > 0$ such that the resulting equilibrium with $w = 1$ fixed is $(m_\ell^*, \alpha_\ell^*, p_\ell^*, q_\ell^*)$, and satisfies $m_\ell^* = \ell$ and $\alpha_\ell^* = 1$. Let the equilibrium under no intervention be $(m^*, \alpha^*, p^*, q^*)$. Note further that $p_\ell^* \in (0,1)$ and $q_\ell^* \in (0,1)$, and $m^* \geq \ell$ using Proposition 3.

Recall the definition $w_\ell = (c_a/p_\ell^*)\exp(\ell p_\ell^*)$. Note that $w_\ell > c_a/p_\ell^*$ since $p_\ell^* > 0$. We can further show that $w_\ell \leq 1$ as follows: The second bullet of Theorem 6 (and Remark 1 after the proof of Theorem 6) tells us that the equilibrium $(m_\ell^*, \alpha_\ell^*, p_\ell^*, q_\ell^*)$ can be recreated by the single intervention of changing the application cost, in particular, there is a unique application cost which results in the same equilibrium strategies, and this application cost is *higher* than c_a. Now, observe from Eq. (19) that the applicant best response depends on the ratio (w/c_a) so increasing c_a with fixed w is equivalent (in terms of resulting equilibrium strategies) to holding c_a fixed while decreasing w (and holding $v = 1$ fixed). Thus, there is a unique *lower* value $w \leq 1$ such that the same equilibrium strategies occur. Now, under $w = w_\ell$ and $p = p_\ell^*$, the applicant best response is exactly an application intensity of ℓ as per Eq. (19). It follows that w_ℓ is that unique wage level and hence $w_\ell \leq 1$.

Now consider the alternate intervention where there is no application limit, but the wage w is set as $w = w_\ell$, so that employers get a value $v = 2 - w_\ell$ from a match. We can verify that $(m_\ell^* = \ell, \alpha_\ell^* = 1, p_\ell^*, q_\ell^*)$ constitutes a mean field equilibrium:

- Since $(\ell, \alpha_\ell^*, p_\ell^*, q_\ell^*)$ is the MFE under application limit ℓ, we know that if applicants use application intensity m_ℓ^* and employers use strategy α_ℓ^* then the resulting MFSS is (p_ℓ^*, q_ℓ^*). It remains to verify that the applicants and employers are each playing a best response.
- **Applicant best response:** The applicant best response under no application limit and wage w_ℓ is given by Eq. (19) with $w = w_\ell$. Plugging in the value of w_ℓ and $p = p_\ell^*$, the second case arises since $w_\ell > c_a/p_\ell^*$, and we see that the applicant best response is an application intensity of $\frac{1}{p_\ell^*}\log(\exp(\ell p_\ell^*)) = \ell$. Thus, we have verified that applicants are playing a best response in the claimed MFE.
- **Employer best response:** The employer best response under no application limit and wage w_ℓ is given by Eq. (20). We know that the employer best response was $\alpha_\ell^* = 1$ even for a wage $w = 1$. So with a lower wage $w_\ell \le 1$ to be paid, we have $c_s'/(2 - w_\ell) \le c_s'$ and so from Eq. (20), it follows that $1 \in \mathcal{A}_{w_\ell}(q_\ell^*)$, i.e., $\alpha_\ell^* = 1$ remains a best response.

Since the equilibrium strategies on both sides of the market are the same in the two settings, it follows that under the natural coupling between the two systems, the expenditures of application and screening costs are the same, and the matches formed are also the same. Only the wage paid differs across the two settings. As a result, the total welfare is the two settings is identical (the distribution of welfare across the two sides of the market is different due to the difference in wages).

Appendix G: Proofs: Appendix D and Theorems in Section 6.

This section culminates with proofs of the key mean field limit theorems of our paper: Theorems 8 and 9, which we then use to prove Theorem 10. The section is organized as follows. First, we establish some basic results on our binned stochastic system; these are in Section G.1.

We then outline the proof of Proposition 4 in Section G.2. This proof relies on two lemmas that we prove in Sections G.3 and G.4.

We then use Proposition 4 to prove that our mean field assumptions are valid in the limit; this is the content of Theorems 8 and 9, proven in Section G.5. Finally, we use validity of the mean field assumptions to establish that MFE is an approximate equilibrium in large finite markets, cf. Theorem 10, proven in Section G.6.

Throughout, we let Binomial(n, p) denote the binomial distribution with n trials and probability of success p, and let Binomial$(n,p)_k = \mathbb{P}(\text{Binomial}(n,p) = k)$. We let Poisson$(a)$ denote the Poisson distribution with mean a, and let Poisson$(a)_k = \mathbb{P}(\text{Poisson}(a) = k)$.

G.1. Proofs of Lemmas 1, 2, and 3.

Proof of Lemma 1. Note that $\frac{dh}{dx} = \frac{(1+\eta x)e^{-\eta x} - 1}{x^2}$. The first inequality in item 1 comes from substituting $e^{-\eta x} \ge 1 - \eta x$, the second from substituting $1 + \eta x \le e^{\eta x}$. Item 2 follows from the fact that $\gamma = 1 - \frac{\alpha}{rk} h(\Sigma)$. The third item comes from the fact that $(1-\gamma)$ is monotone in Σ (from item 2), and $\Sigma \in [0,1]$. The final item comes from the fact that $\frac{d}{d\Sigma}(1-\gamma)\Sigma = \frac{d}{d\Sigma}\frac{\alpha(1-e^{-\eta\Sigma})}{rk} = \frac{\alpha\eta e^{-\eta\Sigma}}{rk}$, which is decreasing in Σ (and $\Sigma \le 1$).

Proof of Lemma 2. For $\alpha = 0$ we immediately find that $\gamma^* = 1$ and $\Sigma^* = 1$ is the unique solution.

Assume $\alpha > 0$. Rearranging (39), we get

$$r = \frac{\alpha(1-e^{-\eta\Sigma})}{k(1-\gamma)\Sigma} = \frac{\alpha(1-e^{-\eta\Sigma})}{1-\gamma^k}, \tag{103}$$

where the final expression follows from substituting (38). Differentiate the expression on the right with respect to γ, thinking of Σ as the function of γ specified by Eq. (38): the result is

$$\frac{\alpha}{(1-\gamma^k)^2}\left(k\gamma^{k-1}(1-e^{-\eta\Sigma}) + \eta e^{-\eta\Sigma}(1-\gamma^k)\frac{d\Sigma}{d\gamma}\right). \tag{104}$$

Since $\frac{d\Sigma}{d\gamma} > 0$ for Σ given by Eq. (38), the expression above is positive, so the right side of (103) is increasing in γ. Further, since

$$\lim_{\gamma\to 0}\frac{\alpha(1-e^{-\eta\Sigma})}{1-\gamma^k} = \alpha(1-e^{-\eta/k}) < \alpha\eta/k \le r < \infty = \lim_{\gamma\to 1}\frac{\alpha(1-e^{-\eta\Sigma})}{1-\gamma^k}, \tag{105}$$

there is a unique solution γ^* to (103), leading to a unique value Σ^* from Eq. (38).

Proof of Lemma 3. Note that $0 \le \Sigma_k^* \le 1$. Further:

$$k(1-\gamma_k^*) = \alpha \left(\frac{1-e^{-\eta \Sigma_k^*}}{r\Sigma_k^*} \right). \tag{106}$$

This remains bounded for all $\Sigma_k^* \in [0,1]$. Taking subsequences if necessary, therefore, we assume without loss of generality that $k(1-\gamma_k^*) \to m\hat{p}$ and $\Sigma_k^* \to \hat{q}$ as $k \to \infty$.

To establish the result, we show that $(\hat{p}, \hat{q})$ must satisfy (6) and (7). Rewriting and taking limits in (38), we have:

$$\Sigma_k^* = \frac{1-\left(1-\frac{k(1-\gamma_k^*)}{k}\right)^k}{k(1-\gamma_k^*)} \qquad \to \frac{1-e^{-m\hat{p}}}{m\hat{p}}.$$

And similarly, taking limits in (106) and substituting for η we have:

$$m\hat{p} = \alpha \left(\frac{1-e^{-rm\beta\hat{q}}}{r\hat{q}} \right).$$

Thus $(\hat{p}, \hat{q})$ satisfy (6) and (7), as required.

G.2. Proof of Proposition 4.

The proposition follows from a stochastic contraction result that shows that for any starting state that is significantly far from X^*, the distance between the state and X^* decays exponentially (a *contraction*) until the state becomes close to X^*. We prove this contraction with two key steps. First, Lemma 15 says that given a starting state, in the large system limit the state contracts towards X^* in time. Second, Lemma 16 involves a routine argument that the expected distance of the true state in a (stochastic) finite market from the state in the large system limit is small for time increments that are not large.

Throughout our proof, we think of the state of the system at t as being the subset of applicants who are still available at t (with their arrival times) from among the applicants who arrived during the last one time unit. The applications to a specific employer are “revealed” just before the employer exits. If at time 0, the set of applications by applicants in the system has been revealed, we simply need to move forward 1 time unit before we can use our above description of the state. Thus, our analysis holds for times $t \ge 1$.

In order to prove Proposition 4, we first show that for fixed $m_0 \in [1,\infty)$ and $\alpha \in [0,1]$, there exists $\kappa' = \kappa'(r, m_0, \beta, \alpha) > 0$ and $C' = C'(r, m_0, \beta) < \infty$ such that for any $n > C'$, for any k, any $m \in [1/m_0, m_0]$, and any starting state, we have

$$\mathbb{E}\|X(t+1/k) - X^*\|_1 \le (1-\kappa'/k)\|X(t) - X^*\|_1 + C' \left(\frac{1}{\sqrt{n}} + \frac{1}{k^{3/2}} \right). \tag{107}$$

so that by iterating we have:

$$\mathbb{E}\|X(t+i/k) - X^*\|_1 \le (1-\kappa'/k)^i\|X(t) - X^*\|_1 + C' \left(\frac{1}{\sqrt{n}} + \frac{1}{k^{3/2}} \right) k/\kappa' \quad \text{for all } i, \tag{108}$$

using $\sum_{j=0}^{i} (1-\kappa'/k)^j < \sum_{j=0}^{\infty} (1-\kappa'/k)^j = k/\kappa'$ to bound the second term.

To establish the existence of κ' that yields (107), define $\hat{X}(t+1/k)$ by $\hat{X}_0(t+1/k) = 1/k$, and for $0 \le j \le k-1$, $\hat{X}_{j+1}(t+1/k) = \gamma X_j(t)$, where γ is given by (39), i.e.,

$$\gamma = 1 - \frac{\alpha(1-e^{-\eta\|X(t)\|_1})}{rk\|X(t)\|_1}. \tag{109}$$

We use the following two lemmas to establish (107).

Lemma 15. *Fix $m_0 \in [1,\infty)$. There exists $\kappa = \kappa(r, m_0, \beta) > 0$ such that for any k, any $\alpha \in [0,1]$, any $m \in [1/m_0, m_0]$, and any $X(t)$ we have*

$$\|\hat{X}(t+1/k) - X^*\|_1 \le (1-\alpha\kappa/k)\|X(t) - X^*\|. \tag{110}$$

The preceding lemma is essential to our argument: it establishes that the deterministic analog of our stochastic system obeys a contraction, a fact that we can use to control errors over arbitrarily large time horizons. (Lemma 15 is proven below.)

Lemma 16. *Fix $m_0 < \infty$. There exists $C' = C'(m_0, r, \beta) < \infty$ such that for $n > C'$, any $m \le m_0$, any k, any $\alpha \in [0,1]$, and any $X(t)$, we have*

$$\mathbb{E}\|X(t+1/k) - \hat{X}(t+1/k)\|_1 \le C'(n^{-1/2} + k^{-3/2}). \tag{111}$$

The preceding result is natural: it establishes that the stochastic system $X(t)$ concentrates around the deterministic variant $\hat{X}(t)$, as we expect. (The proof of this lemma is deferred to Appendix G.4.)

Using the triangle inequality, we have

$$\mathbb{E}\|X(t+1/k) - X^*\|_1 \le \|\hat{X}(t+1/k) - X^*\|_1 + \mathbb{E}\|X(t+1/k) - \hat{X}(t+1/k)\|_1\,. \tag{112}$$

The relation (107) follows immediately from Lemmas 15 and 16 with $\kappa' = \alpha\kappa$.

To complete the proof of Proposition 4, note that $\|X(t_0) - X^*\|_1 \le \|X(t_0)\|_1 + \|X^*\|_1 = \Sigma(X(t_0)) + \Sigma(X^*) \le 2$ for any t_0. Use (108) to track the evolution from $t_0 = t - i/k$ to t where we choose $i = kC\log n$. Choosing $C = (1+2C')/\kappa'$ and using $(1-\kappa'/k)^i\|X(t_0) - X^*\|_1 \le \exp(-i\kappa'/k)\cdot 2\exp(-C\kappa'\log n)\cdot 2 \le 1/n$, we have

$$\mathbb{E}[\|X(t) - X^*\|_1] \le Cn^{-1/6}\,, \tag{113}$$

yielding the first part of the proposition.

Now consider the second part of the proposition. For any starting state, the system reaches steady state as $t\to\infty$. Let $X_{\rm ss}$ be the steady state distribution of X. Then we have $\lim_{t\to\infty} X(t) = X_{\rm ss}$. It follows from the dominated convergence theorem that $\mathbb{E}[\|X_{\rm ss} - X^*\|_1] = \lim_{t\to\infty}\mathbb{E}[\|X(t) - X^*\|_1] \le Cn^{-1/6}$ using (113). The second part of the proposition follows immediately since if the starting state is distributed as per the steady state distribution, the state at time t is also distributed as per the steady state distribution, so the same bound holds for all $t \ge 0$. This completes the proof.

G.3. Proof of Lemma 15.

As noted above, the technical novelty in our argument comes in Lemma 15, which establishes a contraction for the deterministic analog of our binned stochastic system. In the limit, a fraction $1-\gamma$ of sellers match and leave in an interval of length $1/k$, and the remaining ones move into the next bin of length $1/k$ (except for the last bin, from which all sellers leave). Crucially, we show that the measure of sellers that leave by virtue of matching, $(1-\gamma)\Sigma$, is increasing in Σ, at a rate that is bounded below, cf. Lemma 1 part 4. Thus if the number of available sellers exceeds the steady state value, then sellers match faster than usual, whereas if the number of available sellers is less than the steady state value, then sellers match slower than usual; thus making the number trend towards the steady state value. We believe this contraction phenomenon and our analytical approach extends to matching setting beyond the one considered here. The proof of this lemma follows.

Proof of Lemma 15. For simplicity, we use X to denote $X(t)$, $\hat{X}$ for $X(t+1/k)$, Σ for $\Sigma(X)$, and γ for $\gamma(\Sigma)$.

We first prove Lemma 15 for the case $\Sigma \ge \Sigma^*$. By Lemma 1 part 2, this implies $\gamma \ge \gamma^*$. Note that

$$|a-b| = (b-a) + 2[a-b]_+\,, \tag{114}$$

which we use to get that

$$\|X - X^*\|_1 = \sum_{j=0}^{k-1} (X_j - X_j^*) + 2[X_j^* - X_j]_+ = \Sigma - \Sigma^* + 2\sum_{j=0}^{k-1} [X_j^* - X_j]_+ \,. \tag{115}$$

Recall that $\hat{X}_0 = X_0^* = 1/k$. It follows that

$$\begin{aligned}
\|\hat{X} - X^*\|_1 &= \sum_{j=1}^{k-1} \left| X_j^* - \hat{X}_j \right| = \sum_{j=0}^{k-2} \left| \gamma^* X_j^* - \gamma X_j \right| \\
&= \sum_{j=0}^{k-2} \gamma X_j - \gamma^* X_j^* + 2[\gamma^* X_j^* - \gamma X_j]_+ \\
&= \gamma\Sigma - \gamma X_{k-1} - \gamma^*\Sigma^* + \gamma^* X_{k-1}^* + 2\sum_{j=0}^{k-2} [\gamma^* X_j^* - \gamma X_j]_+ \\
&= \gamma\Sigma - \gamma^*\Sigma^* - \left|\gamma^* X_{k-1}^* - \gamma X_{k-1}\right| + 2\sum_{j=0}^{k-1} [\gamma^* X_j^* - \gamma X_j]_+ \,,
\end{aligned} \tag{116}$$

where both the first and last lines use (114). Now, by Lemma 1 part 4,

$$\begin{aligned}
\gamma\Sigma - \gamma^*\Sigma^* &= \Sigma - \Sigma_* + [(1-\gamma^*)\Sigma^* - (1-\gamma)\Sigma] \\
&\le (\Sigma - \Sigma^*) \left(1 - \frac{\alpha\eta e^{-\eta}}{rk}\right).
\end{aligned} \tag{117}$$

Also, using $\gamma \ge \gamma^*$ and the upper bound on γ^* from Lemma 1, item 3, we have

$$[\gamma^* X_j^* - \gamma X_j]_+ \le [\gamma^* X_j^* - \gamma^* X_j]_+ = \gamma^*[X_j^* - X_j]_+ \le \left(1 - \frac{\alpha(1-e^{-\eta})}{rk}\right) [X_j^* - X_j]_+ \,. \tag{118}$$

If we define $\kappa = \min_{m \in [1/m_0, m_0]} \frac{1}{r} \min(\eta e^{-\eta}, 1 - e^{-\eta}) > 0$, then substituting (117) and (118) into (116) yields

$$\begin{aligned}
\|\hat{X} - X^*\|_1 &\le (1 - \alpha\kappa/k)(\Sigma - \Sigma^*) - |\gamma^* X_{k-1}^* - \gamma X_{k-1}| + 2(1 - \alpha\kappa/k) \sum_{j=0}^{k-1} [X_j^* - X_j]_+ \\
&\le (1 - \alpha\kappa/k)\|X - X^*\|_1,
\end{aligned} \tag{119}$$

where the second line follows from dropping the absolute value term and applying (115).

The proof for the complementary case $\Sigma \le \Sigma^*$ is analogous.

G.4. Proof of Lemma 16.

Note that

$$\mathbb{E}\left|X_j - \hat{X}_j\right| \le \left|\mathbb{E}[X_j] - \hat{X}_j\right| + \mathbb{E}\left|X_j - \mathbb{E}[X_j]\right| \le \left|\mathbb{E}[X_j] - \hat{X}_j\right| + \sqrt{\mathrm{Var}(X_j)}, \tag{120}$$

where we have applied the triangle inequality and Jensen's inequality in turn. Sum over j to get

$$\mathbb{E}\|X(t+1/k) - \hat{X}(t+1/k)\|_1 \le \|\mathbb{E}[X(t+1/k)] - \hat{X}(t+1/k)\|_1 + \sum_{j=0}^{k-1} \sqrt{\mathrm{Var}(X_j(t+1/k))}. \tag{121}$$

Fix an applicant a who is in the system at time t and has arrived within the last $1 - 1/k$ time units. Let I_a be the indicator that a is matched in the next $1/k$ time units, and let $\tilde{\gamma} = \mathbb{E}[1 - I_a]$ be the probability that this applicant is still in the system at time $t + 1/k$. To prove Lemma 16 we use the following two lemmas; Lemma 17 is used to prove Lemma 18. The proofs of both lemmas are deferred below.

Lemma 17. *Fix $m_0 < \infty$. There exists $C = C(m_0, r, \beta) < \infty$ such that for $n > C$, any $m \le m_0$, any k, any $\alpha \in [0,1]$, and any starting state, the following hold:*

$$|\tilde{\gamma} - \gamma| \le C/k^2 \tag{122}$$

For applicants a and a' who arrived between $t-1+1/k$ and $t+1/k$ we have

$$\mathrm{Var}(I_a) \le C/k \tag{123}$$

$$|\mathrm{Cov}(I_a, I_{a'})| \le 1/k^3 \text{ for } a \ne a' \tag{124}$$

Lemma 18. *Fix $m_0 < \infty$. There exists $C = C(m_0, r, \beta) < \infty$ such that for $n > C$, any $m \le m_0$, any k, any $\alpha \in [0,1]$, and any starting state, the following hold*

$$\|\mathbb{E}[X(t+1/k)] - \hat{X}(t+1/k)\|_1 \le (C + \eta/r)/k^2 \le (C + \eta/r)k^{-3/2}, \tag{125}$$

and

$$\sum_{j=0}^{k-1} \sqrt{\mathrm{Var}(X_j(t+1/k))} \le \sqrt{C/r}\; n^{-1/2} + k^{-3/2}. \tag{126}$$

Here C is the same as in Lemma 17.

Using these two lemmas we can complete the proof of Lemma 16.

Proof of Lemma 16. Note that Lemma 18 implies that

$$\|\mathbb{E}[X(t+1/k)] - \hat{X}(t+1/k)\|_1 + \sum_{j=0}^{k-1} \sqrt{\mathrm{Var}(X_j(t+1/k))} \le \sqrt{C/r}\; n^{-1/2} + (\eta/r + C + 1)k^{-3/2}, \tag{127}$$

using Lemma 17. If we take $C' = \max(\sqrt{C/r}, \eta/r + C + 1)$, then (121) and (127) imply Lemma 16.

The remainder of this section is devoted to proving Lemmas 18 and 17. We require the following preliminary result.

Lemma 19. *Recall $\rho = (1 - m\beta/n)^{rn}$. Let $\tilde{h}(\Sigma) = (1 - \rho^\Sigma)/\Sigma$. Then we have $(\log \rho)^2 \le \frac{d\tilde{h}}{dx} \le 0$.*

Proof. Analogous to the proof of Lemma 1 item 1.

Proof of Lemma 17. Note that for all $t' \in (t, t+1/k)$, we have

$$\Sigma(t) - 1/(rk) \le \Sigma(t') \le \Sigma(t) + 1/k\,, \tag{128}$$

since the number of new arrivals in the system is rn/k, the number of "match" departures is at most n/k, and $\Sigma(t') = N(t')/(rn)$.

Fix a, and suppose that there are $rn\Sigma$ available applicants when an employer possibly screens and exits. The probability that a exits at this opportunity is $\tilde{h}(\Sigma)/(rn) = \frac{\alpha(1-\rho^\Sigma)}{rn\Sigma}$and this is a monotone decreasing function of Σ by Lemma 19 item 1. For convenience we define $\tilde{h}(x) = \alpha \cdot (-\log \rho)/rn$ for $x \le 0$. Since there are n/k opportunities to depart, Eq. (128) implies

$$\left(1 - \frac{\alpha}{rn}\tilde{h}(\Sigma(t) + 1/k)\right)^{n/k} \le \mathbb{E}[1 - I_a] = \tilde{\gamma} \le \left(1 - \frac{\alpha}{rn}\tilde{h}(\Sigma(t) - 1/(rk))\right)^{n/k} \tag{129}$$

Recall that $\lim_{n\to\infty} \rho = \exp(-\eta)$. By Lemma 19 item 1 we have that

$$\tilde{h}(\Sigma(t) - 1/rk) \le \tilde{h}(\Sigma(t)) + \frac{2\eta^2}{rk}\,, \quad \text{and} \tag{130}$$

$$\tilde{h}(\Sigma(t) + 1/k) \ge \tilde{h}(\Sigma(t)) - \frac{2\eta^2}{k}. \tag{131}$$

for large enough n. It follows by substitution into (129) that

$$\left(1-\frac{1}{rn}\tilde{h}(\Sigma(t))-\frac{2\eta^2}{rnk}\right)^{n/k}\leq\tilde{\gamma}\leq\left(1-\frac{1}{rn}\tilde{h}(\Sigma(t))+\frac{2\eta^2}{r^2nk}\right)^{n/k}. \tag{132}$$

Using the inequality

$$1-m\varepsilon\leq(1-\varepsilon)^m\leq 1-m\varepsilon+\frac{1}{2}m^2\varepsilon^2, \tag{133}$$

we get

$$1-\frac{\alpha}{rk}\tilde{h}(\Sigma(t))-\frac{2\alpha\eta^2}{rk^2}\leq\tilde{\gamma}\leq 1-\frac{\alpha}{rk}\tilde{h}(\Sigma(t))+\frac{2\alpha\eta^2}{r^2k^2}+\frac{1}{2}\left(\frac{\alpha}{rk}\right)^2\tilde{h}(\Sigma(t))^2. \tag{134}$$

for large enough k. Now $\rho=\exp(-\eta)+O(1/n)$. Also, $|\partial\tilde{h}/\partial\rho|=\rho^{\Sigma-1}\leq 1/\rho\leq 2\exp(\eta)$ for all $\Sigma\in[0,1]$, for large enough n. It follows that

$$\tilde{h}(\Sigma)=h(\Sigma)+O(1/n)\,. \tag{135}$$

Since we have $\gamma=1-\frac{1-e^{-\eta\Sigma}}{rk\Sigma}=1-\frac{1}{rk}h(\Sigma)$, combining Eqs. (134) and (135), we obtain $|\tilde{\gamma}-\gamma|$ is $\mathcal{O}(\alpha/k^2)$, establishing (122) in Lemma 17.

Note that

$$\mathrm{Var}(I_a)=\tilde{\gamma}(1-\tilde{\gamma})\leq 1-\tilde{\gamma}. \tag{136}$$

Using Eq. (134) and $\tilde{h}(\Sigma(t))\leq\tilde{h}(0)=-\log(\rho)=\eta+O(1/n)$ from Lemma 19, we have

$$1-\tilde{\gamma}\leq\frac{\alpha\eta}{rk}+O(\alpha/k^2)\,,$$

which proves (123) from Lemma 17 for applicants that arrived between $t-1+1/k$ and t, and were available at t. For any applicant a that arrived after t, the probability of still being available at $t+1/k$ is even larger than $\tilde{\gamma}$, i.e., this probability is in $[\tilde{\gamma},1]$, leading to $\mathrm{Var}(I_a)\leq\tilde{\gamma}(1-\tilde{\gamma})$ since $\tilde{\gamma}\geq 1/2$ leading to (123) for a.

Finally, we bound $\mathrm{Cov}(I_a,I_{a'})$. We define a bipartite 'interaction' graph $G=(V_S,V_B,E)$ whose vertex sets are $V_S=\mathcal{S}(t)\cup\mathcal{S}(t+1/k)$, i.e., the applicants who arrive between $t-1$ and $t+1/k$, and $V_B=\mathcal{B}(t+1/k)\backslash\mathcal{B}(t)$, i.e., the set of employers who arrive between t and $t+1/k$. For $s\in V_S$ and $b\in V_B$, we decide on the presence of edge (s,b) as follows: Let τ denote the time of arrival of a. Then if $b\in\mathcal{B}(\tau)$, i.e., e is in the system when a arrives, we set $\mathbb{I}((s,b)\in E)=\mathbb{I}(s\in M_b)$, i.e., we include edge (s,b) if applicant a applies to employer e. If $b\notin\mathcal{B}(\tau)$, i.e., e is not in the system when a arrives, we draw $\mathbb{I}((s,b)\in E)\sim\mathrm{Bernoulli}(m/n)$, independent of everything else.

Note that the interaction graph G as defined above is a bipartite Erdos-Renyi graph with $rn(1+1/k)$ vertices on one side, n/k vertices on the other, and edge probability m/n independently between vertices on the two sides. The following fact is immediate (see, e.g., Janson et al. [20]) :

Fact 2 *Fix k such that $r(1+1/k)(1/k)m<1$. There exists $C<\infty$ such that the following occurs. For any $\varepsilon>0$ there exists $n_0<\infty$ such that for all $n>n_0$, with probability at least $1-\varepsilon$, no connected component in G has more than $C\log n$ vertices.*

Here the threshold for existence of a giant component is $r(1+1/k)(1/k)m=1$. Thus it is sufficient to ensure that $r(1+1/k)(1/k)m\leq 2rm/k<1$. In fact, fixing k, Fact 2 holds uniformly for all $m<m_0=k/(2r)$, since the size of the largest connected component is monotone in the edge probability.

Clearly, I_a depends only on the connected component $\mathcal{C}_a$ containing a and similarly for $I_{a'}$. Choose arbitrary $\varepsilon>0$. Let $\mathcal{E}_1$ be the event that no connected component has size more than $C\log n$. Fact

2 tells us that $\mathbb{P}(\mathcal{E}_1^c) \le \varepsilon$. Let $\mathcal{E}_2$ be the event that $a' \notin \mathcal{C}_a$. Clearly, $\mathbb{P}(\mathcal{E}_2^c|\mathcal{E}_1) \le C\log n/(rn) \le \varepsilon$ for large enough n. We deduce that

$$\mathbb{P}(\mathcal{E}_1^c \cup \mathcal{E}_2^c) = \mathbb{P}(\mathcal{E}_1^c) + \mathbb{P}(\mathcal{E}_2^c \cap \mathcal{E}_1) \le \mathbb{P}(\mathcal{E}_1^c) + \mathbb{P}(\mathcal{E}_2^c|\mathcal{E}_1) \le 2\varepsilon \tag{137}$$

Reveal $\mathcal{C}_a$. If $|\mathcal{C}_a| \ge C\log n$ or $a' \in \mathcal{C}_a$, declare 'failure'. Here $|\mathcal{C}_a|$ denotes the number of vertices in $\mathcal{C}_a$. Suppose failure does not occur.

Let $\mathcal{C}_s = \mathcal{C}$ be the revealed connected component. Since failure has not occurred we know that $\mathcal{C}$ contains no more than $C\log n$ nodes and does not contain a'. Now consider the conditional distribution of $\mathcal{C}_{a'}$ given $\mathcal{C}_s = \mathcal{C}$.

Claim 1. *Consider any candidate connected component $\mathcal{C}'$, containing a' and not overlapping with $\mathcal{C}$. We have*

$$\mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}') = \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}'|\mathcal{C}_s = \mathcal{C})(1-m/n)^{|\mathcal{C}'|_a|\mathcal{C}|_b + |\mathcal{C}'|_b|\mathcal{C}|_a},$$

where $|\mathcal{C}|_a$ denotes the number of applicants in component $\mathcal{C}$, and $|\mathcal{C}|_b$ denotes the number of employers in component $\mathcal{C}$.

Proof of claim. The distribution of the rest of G conditioned on $\mathcal{C}_s = \mathcal{C}$ has edge (s,b) present iid with probability m/n if both a and e are not in $\mathcal{C}$, and not present if one of a or e is present in $\mathcal{C}$. The result follows from a standard revelation argument on $\mathcal{C}_{a'}$.

We have

$$\mathbb{E}(I_{a'}) = \sum_{\mathcal{C}' \ni a'} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}')\mathbb{E}(I_{a'}|\mathcal{C}_{a'} = \mathcal{C}'), \tag{138}$$

and

$$\mathbb{E}[I_{a'}|I_a = 1] = \sum_{\mathcal{C} \ni s, \mathcal{C}' \ni a'} \mathbb{P}(\mathcal{C}_a = \mathcal{C}|I_a = 1)\mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}'|\mathcal{C}_a = \mathcal{C})\mathbb{E}[I_{a'}|\mathcal{C}_{a'} = \mathcal{C}'] \tag{139}$$

using the fact that I_a—$\mathcal{C}_a$—$\mathcal{C}_{a'}$—$I_{a'}$ form a Markov chain. Below we argue that the sum in Eq. (139) is very close to the sum in Eq. (138).

Let $\pi \equiv \mathbb{E}[I_a] = \mathbb{E}[I_{a'}]$. If $\pi = 0$, which might occur for instance if a arrives just before $t + 1/k$, we immediately have $\mathrm{Cov}(I_a, I_{a'}) = 0$. As such, we assume $\pi > 0$ in what follows. We deduce from Eq. (137) that

$$\mathbb{P}(\mathcal{E}_1^c \cup \mathcal{E}_2^c | I_s = 1) \le 2\varepsilon/\pi \tag{140}$$

It follows from Eqs. (139) and (140) that

$$\mathbb{E}[I_{a'}|I_a = 1] = \sum_{\substack{\mathcal{C} \ni s, \mathcal{C}' \ni a' \text{ s.t. } \mathcal{C} \cap \mathcal{C}' = \Phi, \\ |\mathcal{C}| \le C\log n, |\mathcal{C}'| \le C\log n}} \mathbb{P}(\mathcal{C}_a = \mathcal{C}|I_a = 1)\mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}'|\mathcal{C}_a = \mathcal{C})\mathbb{E}[I_{a'}|\mathcal{C}_{a'} = \mathcal{C}'] \\ + \ \delta_1, \tag{141}$$

where $0 \le \delta_1 \le \mathbb{P}(\mathcal{E}_1^c \cup \mathcal{E}_2^c | I_a) \le 2\varepsilon/\pi$. Now using Claim 1, we know that for such $(\mathcal{C}, \mathcal{C}')$ we have

$$\mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}'|\mathcal{C}_a = \mathcal{C}) = \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}')(1 + \delta_{\mathcal{C},\mathcal{C}'}),$$

where $0 \le \delta_{\mathcal{C},\mathcal{C}'} \le \varepsilon$ for large enough n. It follows that

$$\sum_{\substack{\mathcal{C}' \ni a' \text{ s.t. } \mathcal{C} \cap \mathcal{C}' = \Phi, \\ |\mathcal{C}'| \le C\log n}} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}'|\mathcal{C}_a = \mathcal{C})\mathbb{E}[I_{a'}|\mathcal{C}_{a'} = \mathcal{C}'] \\ = (1 + \delta_2) \sum_{\substack{\mathcal{C}' \ni a' \text{ s.t. } \mathcal{C} \cap \mathcal{C}' = \Phi, \\ |\mathcal{C}'| \le C\log n}} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}')\mathbb{E}[I_{a'}|\mathcal{C}_{a'} = \mathcal{C}'] \tag{142}$$

for some $0 \le \delta_2 \le \varepsilon$. Now,

$$\mathbb{P}(|\mathcal{C}_{a'}| \ge C \log n) \le \varepsilon \tag{143}$$

and

$$\mathbb{P}(v \in \mathcal{C}_{a'} \,|\, |\mathcal{C}_{a'}| \le C \log n) \le C \log n/(kn)$$

for any agent v, whether employer or applicant, hence for any $\mathcal{C}$ s.t. $a' \notin \mathcal{C}$ and $|\mathcal{C}| \le C \log n$ we have

$$\mathbb{P}(\mathcal{C} \cap \mathcal{C}_{a'} \neq \Phi \,|\, |\mathcal{C}_{a'}| \le C \log n) \le (C \log n)^2/(kn) \le \varepsilon \tag{144}$$

for large enough n. Combining Eqs. (143) and (144), we obtain

$$\mathbb{P}\big((|\mathcal{C}_{a'}| \ge C \log n) \cup (\mathcal{C} \cap \mathcal{C}_{a'} \neq \Phi)\big) \le 2\varepsilon\,. \tag{145}$$

Plugging in to Eq. (142) we obtain

$$\begin{aligned}
&\sum_{\substack{\mathcal{C}' \ni a' \text{ s.t. } \mathcal{C} \cap \mathcal{C}' = \Phi,\\ |\mathcal{C}'| \le C \log n}} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}' | \mathcal{C}_a = \mathcal{C}) \mathbb{E}[I_{a'} | \mathcal{C}_{a'} = \mathcal{C}'] \\
&= (1+\delta_2)\left(-\delta_3 + \sum_{\mathcal{C}' \ni a'} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}') \mathbb{E}[I_{a'} | \mathcal{C}_{a'} = \mathcal{C}']\right) \\
&= \delta_4 + \sum_{\mathcal{C}' \ni a'} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}') \mathbb{E}[I_{a'} | \mathcal{C}_{a'} = \mathcal{C}']
\end{aligned} \tag{146}$$

for some $0 \le \delta_3 \le 2\varepsilon$, leading to $3\varepsilon \le -2\varepsilon(1+\varepsilon) \le \delta_4 \le \varepsilon$. Plugging Eq. (146) back into Eq. (139), we obtain

$$\begin{aligned}
\mathbb{E}[I_{a'} | I_a = 1] &= \sum_{\substack{\mathcal{C} \ni s \text{ s.t.}\\ |\mathcal{C}| \le C \log n}} \mathbb{P}(\mathcal{C}_a = \mathcal{C} | I_a = 1)\left(\delta_4 + \sum_{\mathcal{C}' \ni a'} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}') \mathbb{E}[I_{a'} | \mathcal{C}_{a'} = \mathcal{C}']\right) \\
&= \delta_5 + \left(\sum_{\mathcal{C}' \ni a'} \mathbb{P}(\mathcal{C}_{a'} = \mathcal{C}') \mathbb{E}[I_{a'} | \mathcal{C}_{a'} = \mathcal{C}']\right)\left(\sum_{\substack{\mathcal{C} \ni s \text{ s.t.}\\ |\mathcal{C}| \le C \log n}} \mathbb{P}(\mathcal{C}_a = \mathcal{C} | I_a = 1)\right)
\end{aligned} \tag{147}$$

where $|\delta_5| \le |\delta_4| \le 3\varepsilon$. The first term in the product is simply $\pi' \equiv \mathbb{P}(I_{a'} = 1) = \mathbb{E}[I_{a'}]$, as noted in Eq. (138). The second term in the product is $\mathbb{P}(|\mathcal{C}_a| \le C \log n | I_a = 1) \ge 1 - \mathbb{P}(\mathcal{E}_1 | I_a = 1) \ge 1 - \varepsilon/\pi$, where again $\pi = \mathbb{P}(I_a = 1) \ge 1/(C_1 k)$. We deduce that

$$\mathbb{E}[I_{a'} | I_a = 1] = \delta_6 + \pi' \tag{148}$$

where $|\delta_6| \le \varepsilon(3 + 1/\pi)$ for large enough n. We have

$$\begin{aligned}
\mathrm{Cov}(I_a, I_{a'}) &= \mathbb{E}[I_a I_{a'}] - \mathbb{E}[I_a]\mathbb{E}[I_{a'}] \\
&= \mathbb{E}[I_a]\mathbb{E}[I_{a'} | I_a = 1] - \pi\pi' \\
&= \pi(\pi' + \delta_6) - \pi\pi' = \pi\delta_6
\end{aligned}$$

Hence

$$|\mathrm{Cov}(I_a, I_{a'})| \le \varepsilon(3\pi + 1) \le 4\varepsilon$$

Choosing $\varepsilon = 1/(4k^3)$ yields the desired result.

Proof of Lemma 18. To establish (125), note that for $j = 0, 1, \ldots, k-1$

$$\left|\mathbb{E}[X_j(t+1/k)] - \hat{X}_j(t+1/k)\right| = |\tilde{\gamma} - \gamma|\, X_{j-1}(t) \le |\tilde{\gamma} - \gamma|\,/k \le C/k^3, \tag{149}$$

with the final inequality coming from Lemma 17 Eq. (122). Meanwhile, using the fact that newly arrived applicants are matched with probability no greater than $\alpha\eta/(rk)$ in time $(1/k)$, we have that

$$\left|\mathbb{E}[X_0(t+1/k)] - \hat{X}_0(t+1/k)\right| \le (1/k)\max(\alpha\eta/(rk), 1-\gamma) = \alpha\eta/(rk^2), \tag{150}$$

where we used Lemma 1 item 2. Summing (149) over j and combining it with (150) yields (125).

We now turn our attention to the variance term, i.e. (126). For $j \ge 1$,

$$\begin{aligned}
\mathrm{Var}(X_j(t+1/k)) &= \frac{1}{(rn)^2}\mathrm{Var}\left(\sum_{s \in N_{j-1}(t)} I_a\right) \\
&= \frac{1}{(rn)^2}\left(|N_{j-1}(t)|\,\mathrm{Var}(I_a) + 2\binom{|N_{j-1}(t)|}{2}\mathrm{Cov}(I_a, I_{a'})\right) \\
&\le X_{j-1}(t)\mathrm{Var}(I_a)/rn + X_{j-1}(t)^2\mathrm{Cov}(I_a, I_{a'}). \\
&\le \frac{C}{rnk^2} + \frac{1}{k^5},
\end{aligned} \tag{151}$$

where the final line comes from the fact that $X_{j-1}(t) \le 1/k$ and Lemma 17 Eqs. (123) and (124). By the concavity of the square root function, this implies

$$\sqrt{\mathrm{Var}(X_j(t+1/k))} \le \sqrt{C}/(k\sqrt{rn}) + 1/k^{5/2} \tag{152}$$

Summing (152) over j yields (126).

G.5. Proofs of Theorems 8 and 9.

We require the following lemma.

Lemma 20 **(Le Cam's inequality)**. *Let $X_1, \ldots, X_n$ be Bernoulli random variables with $\mathbb{P}(X_i = 1) = p_i$. Let $S_n = \sum_j X_j$, and $\lambda_n = \sum_i p_i$. Then:*

$$\sum_{j \ge 0}\left|\mathbb{P}(S_n = j) - \frac{e^{-\lambda_n}\lambda_n^j}{j!}\right| \le 2\sum_{i=1}^{n} p_i^2. \tag{153}$$

The preceding inequality is a concentration result for the Poisson approximation to the binomial distribution; in particular it implies the following result.

Corollary 1. *Fix $\zeta_0 < \infty$. For any $\zeta < \infty$, as $n \to \infty$, we have that* Binomial$(n, \zeta/n)$ *converges in total variation distance to* Poisson(ζ). *Further, this convergence is uniform over $\zeta \in [0, \zeta_0]$.*

Proof. Let $p_i = \zeta/n$ for all i in Lemma 20; then (153) bounds the total variation distance between Binomial$(n, \zeta/n)$ and Poisson(ζ) by $\zeta^2/n \le \zeta_0^2/n \to 0$ as $n \to \infty$.

Proof of Theorem 8. We proceed as follows:

$$\begin{aligned}
&\sum_{\ell, a}\left|\mathbb{P}(R_e^{(n)} = \ell, A_e^{(n)} = a) - \mathbb{P}(R = \ell, A = a)\right| \\
&\quad = \sum_{\ell, a}\left|\mathbb{P}(A_e^{(n)} = a | R_e^{(n)} = \ell)\mathbb{P}(R_e^{(n)} = \ell) - \mathbb{P}(A = a | R = \ell)\mathbb{P}(R = \ell)\right| \\
&\quad \le \sum_{\ell, a}\mathbb{P}(A_e^{(n)} = a | R_e^{(n)} = \ell)\left|\mathbb{P}(R_e^{(n)} = \ell) - \mathbb{P}(R = \ell)\right| \\
&\qquad + \sum_{\ell, a}\mathbb{P}(R = \ell)\left|\mathbb{P}(A_e^{(n)} = a | R_e^{(n)} = \ell) - \mathbb{P}(A = a | R = \ell)\right|.
\end{aligned} \tag{154}$$

Since the system begins in steady state, any employer arriving to the system is visible to rn applicants that each apply to the given employer with probability m/n; i.e., the number of applications $R_e^{(n)}$ received by such an employer e follows a Binomial$(rn, m/n)$ distribution. Thus $R_e^{(n)}$ converges in total variation distance to R by Corollary 1, so the first summation in (154) approaches zero as $n\to\infty$ (since for every ℓ, $\sum_a \mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell)=1$).

We thus focus on the second summation in (154). Note that for all a such that $0\le a\le \ell$, $\mathbb{P}(A=a|R=\ell)=\binom{\ell}{a}q^a(1-q)^{\ell-a}$. To simplify notation, let $t=t_e+1$ denote the exit time of employer e.

Given $\varepsilon>0$, let $C_n(\varepsilon)$ be the event that $\|X(t)-X^*\|_1\le\varepsilon$ in the n-th system. Recall that we realize the randomness as follows: at the time of exit, we independently determine whether each applicant that arrived in the last time unit applied to this employer. Thus in particular the number of applications that employer e receives is independent of the state $X(t)$ at her exit time t, and so we conclude that $R_e^{(n)}$ is independent of $C_n(\varepsilon)$. Letting $k=\lfloor n^{1/3}\rfloor$, from Proposition 4, for n sufficiently large it follows that:

$$\mathbb{P}(C_n(\varepsilon)|R_e^{(n)}=\ell)\ge 1-\frac{C}{\varepsilon n^{1/6}}$$

for an appropriate constant C, by Markov's inequality. Note that on $C_n(\varepsilon)$ it follows that $|\Sigma(t)-\Sigma^*|\le\varepsilon$ as well, since $\Sigma(t)=\|X(t)\|_1$ and $\Sigma^*=\|X^*\|_1$.

Now again, since we realize applications at the time of exit of the employer, note that conditional on the value of $S(t)$ as well as $R_e^{(n)}=s$, the employer receives exactly a applications from available applicants with the following probability:

$$\mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell,S(t))=\frac{\binom{S(t)}{a}\binom{rn-S(t)}{\ell-a}}{\binom{rn}{\ell}}. \tag{155}$$

For the moment, assume that $0<q<1$, and choose n large enough and ε small enough so that $0<\Sigma^*-\varepsilon<\Sigma^*+\varepsilon<1$ (cf. Lemma 3). On the event $C_n(\varepsilon)$, we know that $\Sigma^*-\varepsilon<\Sigma(t)<\Sigma^*+\varepsilon$; since $S(t)=rn\Sigma(t)$, we conclude that on $C_n(\varepsilon)$ both $S(t)$ and $rn-S(t)$ are $\Theta(n)$. Thus for fixed ℓ and a, we can approximate the preceding probability with the equivalent calculation as if the available and unavailable applicants were sampled with replacement. It follows that:

$$\left|\mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell,C_n(\varepsilon))-\binom{\ell}{a}(\Sigma^*)^a(1-\Sigma^*)^{\ell-a}\right|\le f(\varepsilon),$$

where $f(\varepsilon)\to 0$ as $\varepsilon\to 0$. Finally, taking n large enough, we can assume that $|\Sigma^*-q|\le\varepsilon$, from which it follows that:

$$\left|\mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell,C_n(\varepsilon))-\binom{\ell}{a}q^a(1-q)^{\ell-a}\right|\le\hat f(\varepsilon),$$

where $\hat f(\varepsilon)\to 0$ as $\varepsilon\to 0$. Putting the steps together, we find that:

$$\begin{aligned}
&|\mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell)-\mathbb{P}(A=a|R=\ell)|\\
&\qquad\le\left|\mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell,C_n(\varepsilon))-\binom{\ell}{a}q^a(1-q)^{\ell-a}\right|+|\mathbb{P}(C_n(\varepsilon)^c|\\
&\qquad\le\hat f(\varepsilon)+\frac{C}{\varepsilon n^{1/6}}.
\end{aligned}$$

Take $n\to\infty$, then $\varepsilon\to 0$ to conclude that:

$$\lim_{n\to\infty}\left|\mathbb{P}(A_e^{(n)}=a|R_e^{(n)}=\ell)-\mathbb{P}(A=a|R=\ell)\right|=0.$$

In the case where $q=0$ or $q=1$, a direct analysis of (155) can be used to establish the preceding result; we omit the details.

Finally, to conclude the proof, note that for each ℓ:

$$\sum_{a=0}^{\ell} |\mathbb{P}(A_e^{(n)} = a | R_e^{(n)} = \ell) - \mathbb{P}(A = a | R = \ell)|$$

converges to zero as $n \to \infty$. Further, the preceding quantity is bounded above by $\ell + 1$, which is integrable against the Poisson(rm) distribution; so by the dominated convergence theorem we conclude that the second term in (154) converges to zero as $n \to \infty$, as required.

Proof of Theorem 9. The result is trivial for $m_a = 0$ so we assume $m_a > 0$ henceforth.

We start with a similar approach to the proof of Theorem 8, as follows:

$$\begin{aligned}\sum_{\ell,a} & \left|\mathbb{P}(T_a^{(n)} = \ell, Q_a^{(n)} = a) - \mathbb{P}(T = \ell, Q = a)\right| \\ & \leq \sum_{\ell,a} \mathbb{P}(Q_a^{(n)} = a | T_a^{(n)} = \ell) \left|\mathbb{P}(T_a^{(n)} = \ell) - \mathbb{P}(T = \ell)\right| \\ & \quad + \sum_{\ell,a} \mathbb{P}(T = \ell) \left|\mathbb{P}(Q_a^{(n)} = a | T_a^{(n)} = \ell) - \mathbb{P}(Q = a | T = \ell)\right|. \end{aligned} \tag{156}$$

Since the system begins in steady state, any arriving applicant will find n employers in the system upon arrival, and applies to each of these employers independently with probability m_a/n; i.e., the number of applications $T_a^{(n)}$ sent by such an applicant a follows a Binomial$(n, m_a/n)$ distribution. Thus $T_a^{(n)}$ converges in total variation distance to T by Corollary 1, so the first summation in (156) approaches zero as $n \to \infty$, uniformly for all $m_a \in (0, m_0]$.

As before, therefore, we focus our attention on the second summation in (156). Note that for all a such that $0 \leq a \leq \ell$, we have $\mathbb{P}(Q = a | T = \ell) = \binom{\ell}{a} p^a (1-p)^{\ell - a}$.

First, we fix an upper bound L on the number of applications that we consider by applicant a. In particular, note that we have:

$$\sum_{\ell > L, a} \mathbb{P}(T = \ell) \left|\mathbb{P}(Q_a^{(n)} = a | T_a^{(n)} = \ell) - \mathbb{P}(Q = a | T = \ell)\right| \leq \sum_{\ell \geq L} (\ell + 1) \mathbb{P}(T = \ell), \tag{157}$$

and the last expression does not depend on n and goes to zero as $L \to \infty$, uniformly for all $m_a \leq m_0$. Thus it suffices to show that for fixed L, the sum:

$$\sum_{\ell \leq L, a} \mathbb{P}(T = \ell) \left|\mathbb{P}(Q_a^{(n)} = a | T_a^{(n)} = \ell) - \mathbb{P}(Q = a | T = \ell)\right| \tag{158}$$

goes to zero as $n \to \infty$, uniformly for all $m_a \leq m_0$.

Let $\mathcal{I}(\ell)$ denote the set of all 2^ℓ possible outcome vectors $\mathbf{I}$ that are possible when the applicant sends ℓ applications. Note that (158) is less than or equal to:

$$\sum_{\ell \leq L, \mathbf{I} \in \mathcal{I}(\ell)} \mathbb{P}(T = \ell) \left|\mathbb{P}(\mathbf{I} | T_a^{(n)} = \ell) - p^{\sum_i I_i} (1-p)^{\ell - \sum_i I_i}\right|, \tag{159}$$

where we take advantage of the fact that Q is distributed as Binomial(ℓ, p) when $T = \ell$. Since $\ell \leq L$, it suffices to show that for any ℓ and each $\mathbf{I} \in \mathcal{I}(\ell)$, the quantity

$$\left|\mathbb{P}(\mathbf{I} | T_a^{(n)} = \ell) - p^{\sum_i I_i} (1-p)^{\ell - \sum_i I_i}\right| \tag{160}$$

goes to zero as $n \to \infty$. Now note that the preceding quantity is less than or equal to:

$$\int \left|\mathbb{P}(\mathbf{I} | T_a^{(n)} = \ell, \mathbf{t}) - p^{\sum_i I_i} (1-p)^{\ell - \sum_i I_i}\right| \, d\mathbb{P}_n(\mathbf{t} | T_a^{(n)} = \ell). \tag{161}$$

Here the integral is over all possible vectors of departure times for employers to whom the applicant submitted an application. Note that $\mathbf{t}$ has an atomic distribution that varies with n, so the integral reduces to a sum over feasible $\mathbf{t}$. Note also that this quantity does not depend on m_a at all (since we are conditioning on applicant a making ℓ applications). In fact, it turns out that we do not need to consider m_a for the rest of the proof.

We argue as follows. Fix $k = \lfloor n^{1/3} \rfloor$. When $T_a^{(n)} = \ell$, we use $b_1, \ldots, b_\ell$ to denote the employers that a applied to, and without loss of generality we let $t_1 \le t_2 \le \cdots \le t_\ell$ denote the departure times of these employers, and let $\mathbf{t}$ denote the vector of departure times. We define "rounded" departure times by rounding the true departure times up to the nearest multiple of $1/k$; denote these as $t'_i = \lceil k t_i \rceil / k$ for $i = 1, 2, \ldots, \ell$, and let $\mathbf{t'}$ denote the vector of rounded departure times. We use the rounded departure times so that we can apply the analysis that yields (108); that result applies to the "binned" process of available applicants, binned on time increments of length $1/k$.

In our analysis, we make use of a coupled process of available applicants, but with a particular employer b_i and applicant a removed during $(t'_{i-1}, t'_i]$ (with $t'_0 = 0$). Let $S^{-i}(t)$ denote the state of this process at time t for $t \le t'_i$. We couple this process to the original process so that each employer (besides b_i) receives the same set of applications, and screens them in the same order (or does not screen at all in both systems). So $S^{-i}(t)$ is identical to $S(t)$ for $t \le t'_{i-1}$, and further for $t \in (t'_{i-1}, t_i)$ except for the removal of a. The states are further nearly identical until t'_i if none of the applicants who applied to b_i applied to another employer who departed in (t_i, t'_i) (the event $E_n(i+1)$ defined below): in this case the systems can differ only on applicant a and any applicant who received an offer from b_i.

We also use I_i to denote the outcome on the i'th application; i.e., $I_i = 1$ if applicant a receives an offer from employer b_i, and $I_i = 0$ otherwise. Let $\mathbf{I}$ denote the vector of application outcomes. Fix $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots \varepsilon_\ell > 0$. Let $C_n(0)$ be the event that the $\mathbf{t'}$ are all distinct. Let $D_n(J)$ be the event that no more than $2mrn/k$ distinct applicants apply to employers who depart during (t_J, t'_J); this will hold with high probability for large n. Define $C_n(1)$ be the event that $C_n(0)$ occurs; and in the $S^{-1}(\cdot)$ process, $D_n(1)$ occurs and $\|X(t'_1) - X^*\|_1 \le \varepsilon_1$. Let $C_n(J)$ for $J > 1$ be iteratively defined as the event that $C_n(J-1)$ occurs; and in the $S^{-J}(\cdot)$ process, $D_n(J)$ occurs, $\|X(t'_J) - X^*\|_1 \le \varepsilon_J$, and no applicant applies to both b_{J-1} and some other employer who departs in the interval $(t_{J-1}, t'_{J-1}]$. We remark here that we consider the modified process $S^{-i}(\cdot)$, and employ event $D_n(J)$ in the definition of $C_n(J)$ for convenience; these can be viewed as minor technical details.

In what follows, to economize on notation for a vector $\mathbf{x}$, we let $\mathbf{x}_{i:j} = (x_i, x_{i+1}, \ldots, x_j)$.

We proceed as follows. First, we condition on $\mathbf{I}_{1:i-1}$ and $\Sigma(t_i)$, and consider the probability employer b_i makes applicant a an offer. Recall that the employer only learns compatibility but not availability by screening; a receives an offer from b_i if and only if she is screened by employer b_i and is also compatible.

Given that the employer follows ϕ^α, $\Sigma(t_i)$ and $\mathbf{I}_{1:i-1}$ form a sufficient statistic to determine whether employer i screens a, as follows. Let d denote the number of competing applicants that employer i receives. Applicant a receives an offer from employer b_i if she is compatible with b_i, and no compatible, available applicant is screened by employer b_i before a. The scaled number of available applicants besides a at time t_i is $\Sigma(t_i)$ (possibly with a $O(1/n)$ adjustment if applicant a is also available; we skip this minor detail). It follows that the probability a receives an offer is:

$$\beta \sum_{d=0}^{rn\Sigma(t_i)} \binom{rn\Sigma(t_i)}{d} \left(\frac{m}{n}\right)^d \left(1 - \frac{m}{n}\right)^{rn\Sigma(t_i) - d} \left\{ \frac{\alpha}{d+1} \sum_{j=0}^{d} (1-\beta)^j \right\}.$$

The expression in brackets simplifies to:

$$\frac{\alpha}{d+1} \cdot \frac{1 - (1-\beta)^{d+1}}{\beta},$$

so the probability that a receives an offer becomes:

$$\sum_{d=0}^{rn\Sigma(t_i)} \text{Binomial}\left(rn\Sigma(t_i), \frac{m}{n}\right)_d \left\{\frac{\alpha(1-(1-\beta)^{d+1})}{d+1}\right\}. \tag{162}$$

At this point we recall the derivation of (7); from that calculation it follows that:

$$\sum_{d=0}^{\infty} \text{Poisson}(rm\Sigma)_d \cdot \frac{\alpha(1-(1-\beta)^{d+1})}{d+1} = \frac{\alpha(1-e^{-rm\beta\Sigma})}{rm\Sigma} = \alpha\beta g(rm\beta\Sigma), \tag{163}$$

where $g(x)=(1-e^{-x})/x$. Further, observe that if $\Sigma=q$ in the preceding expression, then (7) implies the right hand side of the preceding expression is equal to p.

We proceed by showing that on $C_n(i)$, (162) is well approximated by (163) with $\Sigma=q$. This requires three steps: first, showing that the binomial distribution in (162) is well approximated by a Poisson distribution; second, exploiting the fact that $\Sigma(t_i)$ is close to Σ^*; and third, taking n large enough so that Σ^* is close to q.

We proceed as follows. Let $h(d)=\alpha(1-(1-\beta)^{d+1})/(d+1)$. Note that $0\le h(d)\le 1$ for all d. Further, note that we have:

$$\begin{aligned}
&\left|\left(\sum_{d=0}^{rn\Sigma(t_i)} h(d)\text{Binomial}\left(rn\Sigma(t_i), \frac{m}{n}\right)_d\right) - p\right| \\
&\quad\le \sum_{d=0}^{\infty} h(d)\left|\text{Binomial}\left(rn\Sigma(t_i), \frac{m}{n}\right)_d - \text{Poisson}(rm\Sigma(t_i))_d\right| \\
&\qquad + \left|\sum_{d=0}^{\infty} h(d)\left(\text{Poisson}(rm\Sigma(t_i))_d - \text{Poisson}(rm\Sigma^*)_d\right)\right| \\
&\qquad + \left|\sum_{d=0}^{\infty} h(d)\left(\text{Poisson}(rm\Sigma^*)_d - \text{Poisson}(rmq)_d\right)\right|.
\end{aligned}$$

By Lemma 20, the first summation on the right is bounded above by a K/n for some constant K. We now use (163) to simplify the second and third summations. The second summation reduces to $|\alpha\beta(g(rm\beta\Sigma(t_i))-g(rm\beta\Sigma^*))|$, which for large enough n on $C_n(i)$, using that $|\Sigma(t_i)-\Sigma(t_i')|=O(t_i'-t_i)=O(1/k)$, is bounded above by $f(\varepsilon_i+1/k)$ for some f such that $f(x)\to 0$ as $x\to 0$. Finally, the third summation above reduces to $|\alpha\beta(g(rm\beta\Sigma^*)-g(rm\beta q))|$, which can be made less than or equal to ε_i for n large enough. Summarizing then, by taking n large enough, we can ensure that on $C_n(i)$ we have:

$$\left|\left(\sum_{d=0}^{rn\Sigma(t_i)} h(d)\text{Binomial}\left(rn\Sigma(t_i), \frac{m}{n}\right)_d\right) - p\right| \le \hat{f}(\varepsilon_i+1/k),$$

for some $\hat{f}$ such that $\hat{f}(\varepsilon_i)\to 0$ as $\varepsilon_i\to 0$.

Analogously, in the case where employer i does not make an offer to applicant a, it immediately follows that $\mathbb{P}(I_i=0|T_a^{(n)}=\ell,\mathbf{t},\mathbf{I}_{1:i-1},C_n(i))$ is well approximated by $1-p$.

To simplify notation, let $A_{\ell,\mathbf{t}}^{(n)}$ be the event that $T_a^{(n)}=\ell$ and the vector of employer departure times this applicant applied to is $\mathbf{t}$. For now, consider only those $A_{\ell,\mathbf{t}}^{(n)}$ such that $C_n(0)$ holds (this occurs with high probability, cf. Eq. (171) below). To summarize then, we conclude that by taking n large enough, we have:

$$\left|\mathbb{P}(I_i|A_{\ell,\mathbf{t}}^{(n)},\mathbf{I}_{1:i-1},C_n(i)) - p^{I_i}(1-p)^{1-I_i}\right| \le \hat{f}(\varepsilon_i+1/k). \tag{164}$$

Next, we consider the conditional probability $\mathbb{P}(C_n(i)|A_{\ell,\mathbf{t}}^{(n)},\mathbf{I}_{1:i-1},C_n(i-1))$. As a preliminary step, note that $\mathbb{P}(D_n(i)|A_{\ell,\mathbf{t}}^{(n)},\mathbf{I}_{1:i-1},C_n(i-1))\le 1-\exp(-\Theta(n/k))$ by the Chernoff bound, since each of

$(1+1/k)rn$ applicants has, independently, applied to some employer who departs in (t_i, t_i') with probability no more than $(m/n)(n/k) = m/k$, so the likelihood that over $2mrn/k$ distinct applicants have applied to this set of employers is $\exp(-\Theta(n/k))$. Also, let $E_n(i)$ be the event that none of the applicants who applied to an employer who departs in (t_{i-1}, t_{i-1}') also applied to employer b_i. By the union bound, we have $\mathbb{P}(E_n(i)|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1), D_n(i)) \le (m/n)(2mrn/k) = 2m^2/k$ under $D_n(i-1)$. It is straightforward to check that this bound remains $O(1/k)$ even if we further condition on I_{i-1}, because the conditional probability $\mathbb{P}(I_{i-1}|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1), D_n(i))$ is bounded away from zero and one for sufficiently large n.

We now rely on (108) to control $\mathbb{P}(C_n(i)|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1))$. Now observe that on $C_n(i-1)$, we know that $\|X(t_{i-1}') - X^*\|_1 \le \varepsilon_{i-1}$ for the $S^{-(i-1)}(\cdot)$ process. To iterate our argument to the departure of the i'th employer, we need to control the state $X(t_i')$ in the process $S^{-i}(\cdot)$, i.e., for the system where employer b_i and applicant a are removed during (t_{i-1}', t_i') (but employer b_{i-1} is included). For this, we only need to make an adjustment to $X(t_{i-1}')$ if employer b_{i-1} matched to an applicant a' other than a in the original system. If this indeed happened, then, on $E_n(i)$, we only need to adjust one coordinate of $X(t_{i-1}')$ downward by $1/n$ corresponding to a' being unavailable in $S^{-i}(t_{i-1}')$, and this small adjustment will not affect our analysis (we omit this detail below). For n large, it follows that an analogous result to (108) holds for the evolution of $X(t)$ between t_i' and t_{i-1}', from which we can conclude that for large enough n there holds:

$$
\begin{aligned}
&\mathbb{E}[\|X(t_i') - X^*\|_1 | A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1), E_n(i)] \\
&\qquad \le \varepsilon_{i-1} + \frac{C'}{\kappa'} \frac{1}{n^{1/6}} + O(1/n).
\end{aligned}
$$

The first term follows from the contraction term in (108); the second term follows because $k = \lfloor n^{1/3} \rfloor$. By Markov's inequality we obtain:

$$
\mathbb{P}(\|X(t_i') - X^*\|_1 < \varepsilon_i | A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1), E_n(i)) \ge 1 - \frac{\varepsilon_{i-1}}{\varepsilon_i} - \frac{2C'}{\kappa'\varepsilon_i} \frac{1}{n^{1/6}}. \tag{165}
$$

Combining, we obtain

$$
\begin{aligned}
&\mathbb{P}(C_n(i)|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1)) \\
&\qquad \ge \mathbb{P}(C_n(i)|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1), E_n(i))(1 - \mathbb{P}(E_n(i)^c)) \\
&\qquad \ge \mathbb{P}(C_n(i)|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1), E_n(i)) - \mathbb{P}(E_n(i)^c) \\
&\qquad \ge 1 - \frac{\varepsilon_{i-1}}{\varepsilon_i} - \frac{3C'}{\kappa'\varepsilon_i} \frac{1}{n^{1/6}}
\end{aligned} \tag{166}
$$

for large enough n. This bound applies for any $i > 1$. For $i = 1$—the first employer that applicant a applies to—we can directly apply Proposition 4 together with Markov's inequality to conclude that for sufficiently large n, we have:

$$
\mathbb{P}(C_n(1)|A_{\ell,\mathbf{t}}^{(n)}, C_n(0)) \ge 1 - \frac{2}{\varepsilon_1 n^{1/6}}. \tag{167}
$$

Observe that by the definition of conditional probability, together with the fact that $C_n(i) \subset C_n(i-1)$, we have:

$$
\begin{aligned}
&\mathbb{P}(\mathbf{I}_{i:\ell}|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1)) \\
&\quad = \mathbb{P}(\mathbf{I}_{i:\ell}|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i)) \cdot \mathbb{P}(C_n(i)|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1)) \\
&\qquad + \mathbb{P}(\mathbf{I}_{i:\ell}|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i)^c, C_n(i-1)) \cdot \mathbb{P}(C_n(i)^c|A_{\ell,\mathbf{t}}^{(n)}, \mathbf{I}_{1:i-1}, C_n(i-1)).
\end{aligned} \tag{168}
$$

If we consider the first term in the expansion above, we have again by the definition of conditional probability that:

$$\begin{aligned}\mathbb{P}(\mathbf{I}_{i:\ell}|A^{(n)}_{\ell,\mathbf{t}},\mathbf{I}_{1:i-1},C_n(i))\\ =\mathbb{P}(\mathbf{I}_{i+1:\ell}|A^{(n)}_{\ell,\mathbf{t}},\mathbf{I}_{1:i},C_n(i))\cdot\mathbb{P}(I_i|A^{(n)}_{\ell,\mathbf{t}},\mathbf{I}_{1:i-1},C_n(i)).\end{aligned} \tag{169}$$

The last piece we need is the following simple bound on the difference of two products, where $0\le a_i,b_i\le 1$:

$$|a_1b_1-a_2b_2|\le|a_1-a_2|+|b_1-b_2|. \tag{170}$$

We now return to our overall goal: namely, we wish to show that the absolute value

$$\left|\mathbb{P}(\mathbf{I}|A^{(n)}_{\ell,\mathbf{t}})-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}\right|$$

becomes small for $A^{(n)}_{\ell,\mathbf{t}}$ such that $C_n(0)$ holds. To show this, we iterate and combine (164), (166), (167), (168), (169), and (170). For example, after the first step of this iteration, we obtain that:

$$\begin{aligned}\left|\mathbb{P}(\mathbf{I}|A^{(n)}_{\ell,\mathbf{t}})-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}\right| &= \Big|\mathbb{P}(\mathbf{I}|A^{(n)}_{\ell,\mathbf{t}},C_n(1))\mathbb{P}(C_n(1)|A^{(n)}_{\ell,\mathbf{t}})\\ &\quad+\mathbb{P}(\mathbf{I}|A^{(n)}_{\ell,\mathbf{t}},C_n(1))\mathbb{P}(C_n(1)^c|A^{(n)}_{\ell,\mathbf{t}})\\ &\quad-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}(\mathbb{P}(C_n(1)|A^{(n)}_{\ell,\mathbf{t}})+\mathbb{P}(C_n(1)^c|A^{(n)}_{\ell,\mathbf{t}}))\Big|\\ &\le\Big|\mathbb{P}(\mathbf{I}_{2:\ell}|A^{(n)}_{\ell,\mathbf{t}},I_1,C_n(1))\mathbb{P}(I_1|A^{(n)}_{\ell,\mathbf{t}},C_n(1))\\ &\quad-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}\Big|+\mathbb{P}(C_n(1)^c|A^{(n)}_{\ell,\mathbf{t}})\\ &\le\left|\mathbb{P}(\mathbf{I}_{2:\ell}|A^{(n)}_{\ell,\mathbf{t}},I_1,C_n(1))-p^{\sum_{i>1} I_i}(1-p)^{\ell-\sum_{i>1} I_i}\right|\\ &\quad+\hat{f}(\varepsilon_1+1/n^{1/3})+O(n^{-1/6}).\end{aligned}$$

The first step follows by conditioning, i.e., (168); the second step follows by (169), together with the fact that the absolute value of the difference of two probabilities is bounded above by one; and the third step follows by applying (170) to the absolute value, and then using (164) and (167). Continuing in this manner, we obtain:

$$\begin{aligned}&\left|\mathbb{P}(\mathbf{I}|T^{(n)}_a=\ell,\mathbf{t}\text{ s.t. }C_n(0))-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}\right|\\ &\qquad\le\sum_i\hat{f}(\varepsilon_i+1/n^{1/3})+\sum_{i=2}^{\ell}\frac{\varepsilon_{i-1}}{\varepsilon_i}+O(n^{-1/6}).\end{aligned}$$

We know that

$$\mathbb{P}(C_n(0)^c|T^{(n)}_a=\ell)\le L^2/k=O(n^{-1/3}) \tag{171}$$

using an elementary union bound over events that two rounded departure times are identical. To complete the proof, returning to (161), and recalling the integral is in fact a sum over feasible $\mathbf{t}$, we deduce that:

$$\begin{aligned}&\int\left|\mathbb{P}(\mathbf{I}|T^{(n)}_a=\ell,\mathbf{t})-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}\right|\,d\mathbb{P}_n(\mathbf{t}|T^{(n)}_a=ell)\\ &\le\sum_i\hat{f}(\varepsilon_i+1/n^{1/3})+\sum_{i=2}^{\ell}\frac{\varepsilon_{i-1}}{\varepsilon_i}+O(n^{-1/6})+\mathbb{P}(C_n(0)^c|T^{(n)}_a=\ell)\\ &=\sum_i\hat{f}(\varepsilon_i+1/n^{1/3})+\sum_{i=2}^{\ell}\frac{\varepsilon_{i-1}}{\varepsilon_i}+O(n^{-1/6}).\end{aligned}$$

Note that $\varepsilon_1,\ldots,\varepsilon_\ell$ were arbitrary; thus if we first take $n\to\infty$, and then take $\varepsilon_i\to 0$ in such a way that every ratio $\varepsilon_{i-1}/\varepsilon_i\to 0$ as well, then we conclude that the right hand side of the preceding expression approaches zero as $n\to\infty$. Returning to (160), we conclude that for each ℓ and $\mathbf{I}\in\mathcal{I}(\ell)$ we have:

$$\left|\mathbb{P}(\mathbf{I}|T_a^{(n)}=\ell)-p^{\sum_i I_i}(1-p)^{\ell-\sum_i I_i}\right|\to 0$$

as $n\to\infty$, as required.

G.6. Proof of Theorem 10.

We now give a proof of Theorem 10, using Theorems 8 and 9.

Proof of Theorem 10. We first prove the result for applicants. Consider an applicant a who can choose the value of m_a, while other agents play their prescribed mean field strategies. Clearly, choosing $m_a=0$ leads to a utility of 0, whereas choosing $m_a>1/c_a$ leads to a negative expected utility since the expected application cost exceeds 1. (In particular, note that $m^*<1/c_a$.) Hence, it suffices to show that:

CLAIM 2. *For large enough n, playing $m_a=m^*$ is additively ε-optimal for applicant a among $m_a\in[0,1/c_a]$.*

(Since $m_a=0$ has greater utility than any $m_a>1/c_a$, this will imply that $m_a=m^*$ is additively ε-optimal among $m_a\in[0,\infty)$.)

We prove the claim using Theorem 9 with $m_0=1/c_a$. Now, the expected utility of applicant a who chooses m_a is

$$\Pr(\text{Applicant } a \text{ gets at least one offer under } m_a)-c_a m_a\,.$$

Since m^* is a best response under mean field assumptions, it follows that

$$1\cdot\Pr(Q(m_a)>0)-c_a m_a\;\le\;\Pr(Q(m^*)>0)-c_a m^* \tag{172}$$

for any m_a. Theorem 9 implies that there exists n_0 such that for any $n>n_0$, we have

$$\max_{m_a\in[0,m_0]} d_{\mathrm{TV}}\big((T_a^{(n)}(m_a),Q_a^{(n)}(m_a)),(T(m_a),Q(m_a))\big)\;\le\;\varepsilon/2\,.$$

where we have made the dependence on m_a explicit for convenience. It follows that

$$|\Pr(Q(m_a)>0)-\Pr(Q_a^{(n)}(m_a)>0)|\;\le\;\varepsilon/2\qquad\forall m_a\in[0,m_0]\,.$$

Combining with Eq. (172), we obtain that

$$\Pr(Q_a^{(n)}(m_a)>0)-c_a m_a\;\le\;\Pr(Q_a^{(n)}(m^*)>0)-c_a m^*+\varepsilon\qquad\forall m_a\in[0,m_0]\,,$$

implying the claim.

We now prove the result for employers who receive no more than R_0 applications. Think of employer e as first selecting a uniformly random order among her applicants, and then screening them (or not) in that order. With this order, denote the applicants by $s_1,s_2,\ldots,s_\ell$ (here ℓ denotes the number of applications) and let $\mathcal{S}$ denote the set of all applicants. Let $\mathbf{I}_b^{(n)}$ denote the applicant availability vector at the time that e departs, with this order. (So $\mathbf{I}_b^{(n)}\in\{0,1\}^\ell$, where a 1 represents that the corresponding applicant is available.) Let $\mathbf{I}$ denote an analogously defined vector for a hypothetical employer who receives $R\sim\mathrm{Poisson}(rm)$ applications of which $A\sim\mathrm{Binomial}(R,q)$ are from applicants who are still available. It follows from Theorem 8, and the fact that applicants are considered in a uniformly random order, that:

CLAIM 3. $(R_b^{(n)},A_b^{(n)},\mathbf{I}_b^{(n)})$ *converges in total variation distance to* $(R,A,\mathbf{I})$.

Let $\mathcal{S}_i=\{s_j:j\le i\}$. Let $\mathcal{S}_i'$ denote the subset of $\mathcal{S}_i$ that e makes offers to. (So, for $i<\ell$, employer e does not screen s_{i+1} if one of $\mathcal{S}_{i'}$ accepts an offer.) The next claim follows from Claim 3 above.

CLAIM 4. *Fix $\varepsilon' > 0$ and $R_0 < \infty$. There exists $n_1 < \infty$ such that the following holds. For any non-negative integer $\ell \le R_0$, any $i < \ell$ and any subset $S_i' \subseteq \mathcal{S}_i$ with fixed indices in $\{1, 2, \ldots, i\}$ we have*

$$\left|\Pr\big(s_{i+1} \text{ is available} \,\big|\, |\mathcal{S}| = \ell\,, \mathcal{S}_i' = S_i'\,, \mathcal{S}_i' \text{ are not available}\big) - q\right| \le \varepsilon'\,. \tag{173}$$

The prescribed mean field strategy for employer e is ϕ^{α^*}. For $\alpha^* < 1$, we argue that not screening at all is additively ε optimal (for large enough n): First, note that $c_s' \ge \beta q$, since not screening is a best response under the employer mean field assumption. Then we know, from Claim 4, that each time the employer chooses to screen an applicant, the net expected benefit is no more than $\beta\varepsilon' \le \varepsilon'$. Using an elementary martingale stopping argument, it follows that the overall net utility of any other strategy (relative to the utility of 0 obtained by not screening) is no more than $\mathbb{E}[\bar{R}\varepsilon'] \le R_0\varepsilon'$, where $\bar{R}$ is the number of applicants screened under ϕ^1, using $\bar{R} \le \ell \le R_0$. Choosing $\varepsilon' \le \varepsilon/R_0$ we obtain the desired result.

Similarly for $\alpha^* > 0$, we argue that ϕ^1 (keep screening until you hire or run out of applicants) is additively ε-optimal (for large enough n): First, note that $c_s' \le \beta q$, since screening is a best response under the employer mean field assumption. Then we know, from Claim 4, that each time the employer chooses to screen the next applicant, the net expected benefit is at least $-\beta\varepsilon' \ge -\varepsilon'$, irrespective of the realization of $\mathcal{S}_i'$ (all these applicants rejected their offers). Any other strategy can be better than ϕ^1 by at most ε' in expectation for each applicant screened under ϕ^1 but not under the alternative strategy. Using an elementary martingale stopping argument, it follows that ϕ^1 is additively $\mathbb{E}[\underline{R}]\varepsilon' \le R_0\varepsilon'$ optimal, where $\underline{R}$ is the number of applicants screened under ϕ^1, but not under the alternative strategy. Choosing $\varepsilon' \le \varepsilon/R_0$ we obtain the desired result.

Finally, since ϕ^0 is ε-optimal for $\alpha^* < 1$ and ϕ^1 is ε-optimal for $\alpha^* > 0$, we conclude that ϕ^{α^*} is ε-optimal, irrespective of the value of α^*, for large enough n.