Online appendix for "Managing congestion in matching markets"

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.[1]

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

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

Together (6) and (1) 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 $m = \mathcal{M}_\ell(\mathcal{P}(m, \underline{\beta}))$.

[1] 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.

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 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

$$\tilde{Q}(m, \tilde{p}, \beta) = \beta g(m\beta\tilde{p}). \tag{2}$$

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

$$\tilde{q} = \frac{\mathbb{E}_{\beta \sim F_a}[M(\beta)\tilde{Q}(M(\beta), \tilde{p}, \beta)]}{\mathbb{E}_{\beta \sim F_a}[M(\beta)]}. \tag{3}$$

Jointly, (2) and (3) 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

$$\lambda = r\mathbb{E}_{\beta \sim F_a}[M(\beta)\tilde{Q}(M(\beta), \beta, \tilde{p})]. \tag{4}$$

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

$$\tilde{p} = \alpha \sum_{k=0}^{\infty} e^{-\lambda}\lambda^k/(k+1)! = \alpha g(\lambda). \tag{5}$$

The value of $\tilde{p}$ that results from strategies α and M simultaneously solves (2), (4) and (5), and the corresponding $\tilde{q}$ is given by (3).

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

$$M(\beta) = \arg\max_{m \le \ell} W_a(m, \beta, \tilde{p}) = \mathcal{M}_\ell(\beta\tilde{p}). \tag{6}$$

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

$$\alpha \in \tilde{\mathcal{A}}(\tilde{q}) = \begin{cases} \{1\} & : c_s < \tilde{q} \\ [0,1] & : c_s = \tilde{q} \\ \{0\} & : c_s > \tilde{q}. \end{cases} \tag{7}$$

A mean field equilibrium with application limit ℓ consists of an application function M and a value α that are best responses (according to (6), (7)) to the steady state quantities $\tilde{p}$ and $\tilde{q}$ that they induce (as given by (2), (3), (4), (5)).

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 [3]. Our main purpose is to show that competition among employers is much less severe when employers can make only one offer.[2] As a result, *these models predict little or no benefit to employers from an application limit*, as shown in Figure 1.

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.

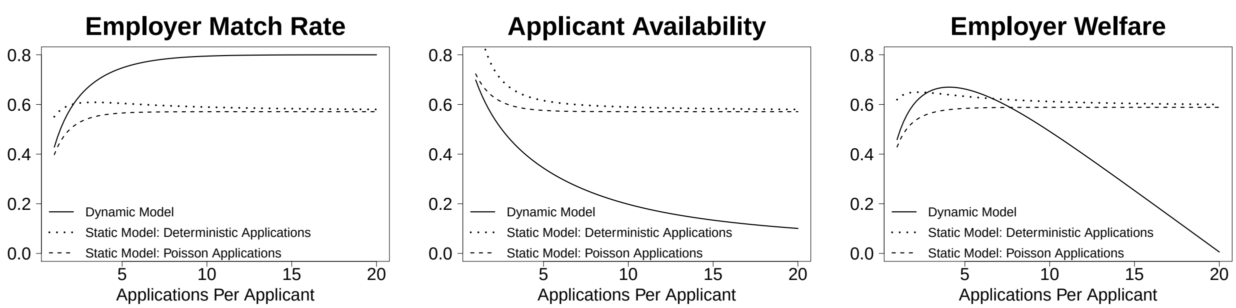

Figure 1 **Comparing outcomes and in dynamic and static models, when there are $r = 0.8$ applicants per employer and screening costs are $c_s' = 0.1$. In our dynamic model, when applicants send many applications, nearly all applicants match, resulting in low availability and low welfare for employers. 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.**

[2] 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.

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

$$\mathcal{P}(m,\alpha) = \alpha\beta g(rm\beta), \tag{8}$$

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

$$M(m,\alpha) = 1 - e^{-m\mathcal{P}(m,\alpha)} \tag{9}$$

The consistency condition for mean-field availability q remains

$$\mathcal{Q}(m,\alpha) = M(m,\alpha)/(m\mathcal{P}(m,\alpha)). \tag{10}$$

We then have

$$\Pi_e(m,\alpha) = r(M(m,\alpha) - c_s' m\mathcal{P}(m,\alpha)). \tag{11}$$

Under the relatively mild condition $c_s' < e^{-1/r}$, one can show that employer welfare as given in (11) 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 $r \le 1$, 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 (8) remains valid. Now, however, each applicant matches with probability

$$M^D(m,\alpha) = 1 - (1 - \mathcal{P}(m,\alpha))^m \tag{12}$$

In fact, (12) with $\alpha = \beta = 1$ is exactly the match function defined in Proposition 1 of Albrecht et al. [1]. The expressions (10) and (11) 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 1 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. □

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.[3] 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:

$$\Sigma(t) = S(t)/(rn); \tag{13}$$

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:

$$\alpha\left(1 - \left(1 - \frac{\beta m}{n}\right)^{S(t)}\right) = \alpha(1 - \rho^{\Sigma(t)}), \tag{14}$$

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 (14) 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 1 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:

$$\Sigma(\gamma) = \frac{1}{k}\sum_{j=0}^{k-1} \gamma^j = \frac{1 - \gamma^k}{k(1-\gamma)}. \tag{15}$$

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

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

On the other hand, as in our mean field analysis, we can develop a "consistency check" that γ must satisfy using (14). Assume that $k \geq \eta/r$ and define $\gamma(\Sigma)$ as follows:

$$\gamma(\Sigma) = 1 - \alpha\left(\frac{1 - e^{-\eta\Sigma}}{rk\Sigma}\right), \tag{16}$$

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 (14): $1 - \gamma(\Sigma)$ 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 (16); 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 (15)-(16), 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 the full technical report [2].

Lemma 2. *Suppose that $k \geq \eta/r$. There is a unique pair of real numbers (γ^*, Σ^*) that simultaneously solve* (15) *and* (16).

Lemma 3. *Suppose that $k \geq \eta/r$. Let (γ_k^*, Σ_k^*) denote the unique solution to* (15) *and* (16) *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. (15). 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 1. *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 the full technical report [2]). We can then use those results to prove that MFE is an approximate equilibrium in large finite markets (Theorem 10, proven in the full technical report [2]).

References

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

[2] Arnosti N, Johari R, Kanoria Y (2019) Managing congestion in matching markets. *Available at SSRN 2427960* .

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