Online Companion to

"Convergence of the core in assignment markets"

Yash Kanoria* Daniela Saban† Jay Sethuraman‡

A Example supporting Section 3.2

We present an example showing that the bounds in Theorem 3 are tight.

Example 1. *Suppose that there is one type of worker k and one type of employer q, i.e., $K = Q = 1$ and $u(k,q) = 0$. Fix arbitrary F that has a first moment and with density f being positive and continuous everywhere. Suppose that $\rho_k = 1/2$, implying $\rho_q = 1 - \rho_k = 1/2$. Then $\alpha_* = 0$ by symmetry, and the limiting core outcome is one in which workers with $\eta_i > 0$ match (worker i earns utility η_i), with employers with $\epsilon_j > 0$ (employer j earns utility ϵ_j). The expected number of workers with $\eta_i > 0$ is $(1 - F(0))n/2$ and this is also the expected number of employers with $\epsilon_j > 0$. However, in the n-th market, the number of workers with $\eta_i > 0$ deviates from its expectation by order $\sqrt{n}$, and similarly for the number of firms with $\epsilon_j > 0$ and, furthermore, the difference between these two numbers is $\Delta \sim \sqrt{n}$. Core α's are those that lead to exactly the same number of workers with $\eta_i - \alpha > 0$ as the number of employers with $\epsilon_i + \alpha > 0$. Consequently, core α's differ from $\alpha^* = 0$ by about $\Delta/(nf(0)) \sim 1/\sqrt{n}$ (in particular, $E[|\alpha|] = 1/\sqrt{n}$ for any core α), and the number of matched pairs scaled by n deviates by order $1/\sqrt{n}$ from its limiting value of $1/4$, i.e., $E[|N(1,1)/n - 1/4|] = \Theta(1/\sqrt{n})$. The (expected) size of the core is[1] of order $1/n$ since all core α's lie between the same pair of order statistics of η_j's*

*Graduate School of Business, Columbia University. Email: ykanoria@columbia.edu
†Graduate School of Business, Stanford University. Email: dsaban@stanford.edu
‡IEOR Department, Columbia University. Email: jay@ieor.columbia.edu

[1]Showing this rigorously requires a small (straightforward) argument to get around the dependence between the realization of productivities, and the relevant order statistics.

and $-\epsilon_i$'s. (We use that F has a first moment to obtain that $\mathrm{weight}(M) = \Theta(n)$ *w.h.p.)*

B Additional definitions and results on point processes in the unit hypercube

B.1 Additional definitions

Analogously to the definitions of the regions in Section 4.1.1, we define the regions $\mathcal{R}^k(t)$ (for appropriate k) and $\mathcal{R}^{k_1,k_2}(t,\delta)$ (for appropriate k_1, k_2), which allow us to apply the conditions (IR) and (ST) respectively. For consistency, the type t is suppressed in the definition of the associated regions, sets, and random variables (e.g. $\mathcal{R}^k := \mathcal{R}^k(t)$). Let

$$\mathcal{R}^k = \{x \in [0,1]^K : x^k \geq x^{k'} \quad \forall k' \neq k, k' \in \mathcal{K}\}. \tag{1}$$

For $k_1, k_2 \in \mathcal{K}, k_1 \neq k_2$ and for $\delta \in [0, 1/2]$, define the region

$$\mathcal{R}^{k_1,k_2}(\delta) = \left\{x \in [0,1]^K : x^{k_1} \geq x^k \quad \forall k \notin \{k_1,k_2\}, k \in \mathcal{K}, x^{k_1} \geq \delta\right\}. \tag{2}$$

The relationship between these random regions and the α's is more involved, so the explanation is delayed to the proofs. However, and analogously to the case explained above, we still care about the difference between (appropriately defined) consecutive order statistics in these regions. To that end, we define

$$\mathcal{V}^k = \{x : x = \epsilon_j^k \text{ for } \{j : \epsilon_j \in \mathcal{R}^k\}\} \tag{3}$$

$$\text{and} \qquad V^k = \max\left(\text{Difference between consecutive values in } \mathcal{V}^k \cup \{0,1\}\right) \tag{4}$$

as well as

$$\mathcal{V}^{k_1,k_2}(\delta) = \{x : x = \epsilon_j^{k_1} - \epsilon_j^{k_2} \text{ for } \{j : \epsilon_j \in \mathcal{R}^{k_1,k_2}(\delta)\}\} \tag{5}$$

and $$V^{k_1,k_2}(\delta) = \max\left(\text{Difference between consecutive values in } \mathcal{V}^{k_1,k_2}(\delta) \cup \{-1+\delta, 1\}\right). \tag{6}$$

Thus, $\mathcal{V}^k \subset [0,1]$ is the set of values of the k-th coordinate of the points lying in $\mathcal{R}^k$, and $V^k \in \mathbb{R}$ is the maximum difference between consecutive values in $\mathcal{V}^k \cup \{0,1\}$. Analogously, $\mathcal{V}^{k_1,k_2} \subset [-1+\delta, 1]$ is the set of values of the difference between the k_1-th and k_2-th coordinates of points lying in $\mathcal{R}^{k_1,k_2}$, and $V^{k_1,k_2} \in \mathbb{R}$ is the maximum difference between consecutive values in $\mathcal{V}^{k_1,k_2} \cup \{-1+\delta, 1\}$.

Using the above notation, we now define two events that will help us prove the results by providing us with a bound on the maximum difference between consecutive values in the relevant regions. Specifically, the events are defined as follows:

$$\mathcal{B}_1(t,\delta) = \left\{ \max\left(\max_{k \in \mathcal{K}} V^k(t), \max_{(k_1,k_2) \in \mathcal{K}^{(2)}} V^{k_1,k_2}(t,\delta) \right) \leq f_1(n_t, K) \right\}, \tag{7}$$

for some $f_1(n_t, K) = O^*(1/n_t^{1/K})$ defined in Lemma 1, $\delta \in [0, 1/2]$ and where $\mathcal{K}^{(2)} = \{(k_1,k_2) : k_1, k_2 \in \mathcal{K}, k_1 \neq k_2\}$. (If $K = 1$, then $\mathcal{K}^{(2)}$ is the empty set $\emptyset$ in which case we follow the convention that $\max_{\emptyset}[\,\cdot\,] = -\infty$.). In addition,

$$\mathcal{B}_2(t,\delta) = \left\{ \max_{k \in \mathcal{T}_{\mathcal{L}}} \tilde{V}^k(t,\delta) \leq f_2(n_t)/\delta^{K-1} \right\} \tag{8}$$

for some $f_2(n_t) = O^*(1/n_t)$ defined in Lemma 2, $\delta \in (0,1]$ and $\tilde{V}^k(t,\delta)$ as defined in Eq. (6).

The proof of all lemmas auxiliary to the proof of Theorem 1 assume that these events (or some subset of them) occur. As shown by the next result, that assumption does not pose a problem as these events simultaneously occur with high probability.

Theorem 1. *There exists $\hat{C} = \hat{C}(K, Q) < \infty$ such that, for any $\delta = \delta(n) \in (0, 1/2]$, the event $\bigcap_{t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}} \left(\mathcal{B}_1(t,\delta) \cap \mathcal{B}_2(t,\delta)\right)$ occurs with probability at least $1 - \hat{C}/n$.*

B.2 Results

Consider the K-dimensional unit hypercube $[0,1]^K$, and the Poisson process of uniform rate n in this hypercube, leading to N points $(\epsilon_i)_{i=1}^N$. (Note that $\mathrm{E}[N] = n$.) Here $\epsilon_i = (\epsilon_i^1, \epsilon_i^2, \ldots, \epsilon_i^K)$. Let $\mathcal{K} = \{1, 2, \ldots, K\}$ denote the set of dimension indices.

The following lemma, key to our proof of Theorem 1, says that with high probability, all the (V^k)'s and the (V^{k_1,k_2})'s are no larger than a (deterministic) function[2] of n that scales as $O^*(1/n^{1/K})$.

Lemma 1. *Let $\mathcal{R}^k$ be the region defined by Eq.* (1)*, and let $\mathcal{V}^k$ and V^k be as defined by Eqs.* (3) *and* (4) *respectively. Similarly, let $\mathcal{R}^{k_1,k_2}(\delta)$ be the region defined by Eq.* (2)*, and let $\mathcal{V}^{k_1,k_2}(\delta)$ and $V^{k_1,k_2}(\delta)$ be as defined by Eqs.* (5) *and* (6) *respectively. Fix $K \geq 1$. Then there exists $f(n,K) = O^*(1/n^{1/K})$ such that for any $\delta = \delta(n) \in [0, 1/2]$ the following holds. Let*

$$\mathcal{B}_1 = \left\{ \max\left(\max_{k\in\mathcal{K}} V^k, \max_{(k_1,k_2)\in\mathcal{K}^{(2)}} V^{k_1,k_2}(\delta) \right) \leq f(n,K) \right\}, \tag{9}$$

where $\mathcal{K}^{(2)} = \{(k_1,k_2) : k_1, k_2 \in \mathcal{K}, k_1 \neq k_2\}$. (If $K = 1$, then $\mathcal{K}^{(2)}$ is the empty set $\emptyset$ in which case we follow the convention that $\max_\emptyset[\,\cdot\,] = -\infty$.) We have

$$\Pr(\mathcal{B}_1) \geq 1 - 1/n\,.$$

Proof. Let $m = \lfloor 1/(C\log n/n)^{1/K} \rfloor$ for some $C < \infty$ that we shall choose later, and let $\Delta = 1/m$. Note that

$$\Delta \geq (C \log n/n)^{1/K}\,. \tag{10}$$

In our analysis of V^k (resp. V^{k_1,k_2}), we shall divide the interval $[0,1]$ (resp. $[-1+\delta, 1]$) into subintervals of size Δ each, and show that with high probability, each subinterval contains at least one value of $\epsilon_i \in \mathcal{R}^k$ (resp. $\epsilon_i^{k_1} - \epsilon_i^{k_2}$ for $\{i : \epsilon_i \in \mathcal{R}^{k_1,k_2}\}$). We shall find that the density of points in $\mathcal{V}^k$ (resp. $\mathcal{V}^{k_1,k_2}$) is smallest near 0 (resp. $-1+\delta$), but even for the interval $[0,\Delta]$

[2] In fact, our proof of Lemma 1 identifies a bound of $(C\log n/n)^{1/K}$ where $C = 6K(K-1)$, for sufficiently large n.

(resp. $[-1+\delta, -1+\delta+\Delta]$), the number of points is Poisson with parameter $\Theta(n\Delta^K) = \Theta(\log n)$, allowing us to obtain the desired result for appropriately chosen C.

We first present our formal argument leading to a bound on V^k, followed by a similar argument leading to a bound on V^{k_1,k_2}. Let

$$\mathcal{B}^k \equiv \bigcap_{i=0}^{m-1} \{\, [i\Delta, (i+1)\Delta] \cap \mathcal{V}^k \neq \emptyset \,\}\,, \tag{11}$$

where $\emptyset$ is the empty set. Clearly, $\mathcal{B}^k \neq \emptyset \Rightarrow V^k \le 2\Delta$. We now show that for any $k \in \mathcal{K}$, we have $\Pr(\overline{\mathcal{B}^k}) \le 1/n^{K+2}$, for appropriately chosen C. Define

$$h^j(x,\theta) = \begin{cases} x^j & \text{for } x \in [\theta, 1] \\ 0 & \text{otherwise}\,. \end{cases} \tag{12}$$

It is easy to see that $\mathcal{V}^k$ follows a Poisson process with density $nh^{K-1}(\,\cdot\,, 0)$. The number of points in interval $[i\Delta, (i+1)\Delta]$ is hence Poisson with parameter

$$n\int_{i\Delta}^{(i+1)\Delta} h^{K-1}(x)\,\mathrm{d}x = ((i+1)^K - i^K)n\Delta^K/K \ge n\Delta^K/K \ge C\log n/K\,,$$

where we used the lower bound on Δ in (10). It follows that

$$\Pr([i\Delta, (i+1)\Delta] \cap \mathcal{V}^k = \emptyset) \le \exp(-C\log n/K) = 1/n^{C/K} \le 1/n^3\,,$$

for $C \ge 3K$. We deduce by union bound over $i = 0, 1, \ldots, m-1$ and De Morgan's law on (11) that

$$\Pr(\overline{\mathcal{B}^k}) \le m/n^3 \le n^{1/K}/n^3 \le 1/n^2\,.$$

Using union bound over k we deduce that

$$\Pr\Big(\cup_k \overline{\mathcal{B}^k}\Big) \le K/n^2 \tag{13}$$

We now present a similar argument to control V^{k_1,k_2} when $K \geq 2$. Let $m' = (1-\delta)/\Delta$. (To simplify notation we assume that m' is an integer. The case where it is not an integer can be easily handled as well.) Let

$$\mathcal{B}^{k_1,k_2} \equiv \bigcap_{i=-m'}^{m-1} \{ [i\Delta, (i+1)\Delta] \cap \mathcal{V}^{k_1,k_2} \neq \emptyset \}, \tag{14}$$

where $\emptyset$ is the empty set. Clearly, $\mathcal{B}^{k_1,k_2} \neq \emptyset \Rightarrow V^{k_1,k_2} \leq 2\Delta$. We now show that for any $k_1 \neq k_2$, we have $\Pr\big(\overline{\mathcal{B}^{k_1,k_2}}\big) \leq K(K-1)/n^2$, for appropriately chosen C. It is easy to see that the two-dimensional projection $(x,y) = (\epsilon_i^{k_1}, \epsilon_i^{k_2})$ of points in $\mathcal{R}^{k_1,k_2}$ follows a two-dimensional Poisson process with density $h^{K-2}(x)\mathbb{I}(y \in [0,1])$, cf. (12). We deduce that values in $\mathcal{V}^k$ follow a one-dimensional Poisson process with density ng for $g = h^{K-2}(\,\cdot\,,\delta) * \mathbb{I}(\in [-1,0])$, where $*$ is the convolution operator. A short calculation yields

$$g(x) = \begin{cases} \big[(x+1)^{K-1} - \delta^{K-1}\big]/(K-1) & \text{for } x \in [-1+\delta, 0) \\ \big[1 - \delta^{K-1}\big]/(K-1) & \text{for } x \in [0, \delta) \\ (1 - x^{K-1})/(K-1) & \text{for } x \in [\delta, 1] \\ 0 & \text{otherwise.} \end{cases}$$

The number of points in interval $[i\Delta, (i+1)\Delta]$ is Poisson with parameter

$$n \int_{i\Delta}^{(i+1)\Delta} g(x)\, \mathrm{d}x\,.$$

Below we bound the value of this parameter for different cases on i, obtaining a bound of $(K+3)\log n$ in each case, for large enough C.

For $-m' \leq i < 0$, the smallest parameter occurs for $i = -m'$, since $g(x)$ is monotone increasing in $[-1+\delta, 0]$. Thus, the Poisson parameter is lower bounded by its value for $i = -m'$, which is

$$\begin{aligned} & n\big[\big((\delta+\Delta)^K - \delta^K\big)/K - \delta^{K-1}\Delta\big]/(K-1) \\ & \geq n\Delta^K/(K(K-1)) \geq C\log n/(K(K-1)) \geq 3\log n\,, \end{aligned}$$

for $C \geq 3K(K-1)$, using (10), and $(\delta+\Delta)^K \geq \Delta^K + K\Delta\delta^{K-1} + \delta^K$.

For $0 \leq i < m - m'$, the Poisson parameter is

$$n\big[1-\delta^{K-1}\big]\Delta/(K-1) \geq n\Delta/(2(K-1)) \geq n\Delta^K/(K(K-1)) \geq 3\log n\,,$$

using $\delta \geq 1/2$ and $K \geq 2$.

For $(m-m') \leq i < m$, the Poisson parameter is

$$n(\Delta - \Delta^K((1+i)^K - i^K)/K)/(K-1)\,.$$

A short calculation allows us to again bound this below by $(K+3)\log n$ (the bound is slack for $K > 2$). Note that

$$\begin{aligned}\Delta^K((1+i)^K - i^K) &\leq \Delta^K(m^K - (m-1)^K) = 1 - (1-\Delta)^K \\ &\leq K\Delta - K(K-1)\Delta^2/2 + K(K-1)(K-2)\Delta^3/6\,,\end{aligned}$$

where we used that $(1+i)^K - i^K$ is monotone increasing in i for $i \geq 0$. Substituting back, we obtain that the Poisson parameter is bounded by

$$n(1-(K-2)\Delta/3)\Delta^2/2 \geq n\Delta^2/4$$

for $(K-2)\Delta/3 \leq 1/2$, which occurs for sufficiently large n. Finally, $\Delta^2 \geq \Delta^K$, and hence $n\Delta^2/4 \geq n\Delta^K/4 \geq 3\log n$ for $C \geq 12$.

Choosing $C = 6K(K-1)$, in all cases the Poisson parameter is bounded below by $3\log n$. It follows that

$$\Pr([i\Delta,(i+1)\Delta] \cap \mathcal{V}^{k_1,k_2} = \emptyset) \leq \exp(-3\log n) = 1/n^3\,.$$

We deduce by union bound over i and De Morgan's law on (14) that

$$\Pr\big(\overline{\mathcal{B}^{k_1,k_2}}\big) \le 2m/n^3 \le n^{1/K}/n^3 \le 1/n^2\,, \tag{15}$$

for large enough n. Using union bound over (k_1, k_2) we deduce that

$$\Pr\Big(\cup_{(k_1,k_2)}\, \overline{\mathcal{B}^{k_1,k_2}} \Big) \le K(K-1)/n^2. \tag{16}$$

Combining (15) and (16) by union bound and using De Morgan's law, we deduce that

$$\Pr\Big[\Big(\cap_k \mathcal{B}^k \Big) \cap \Big(\cap_{(k_1,k_2)} \mathcal{B}^{k_1,k_2} \Big)\Big] \ge 1 - K^2/n^2$$

for large enough n. This implies that for large enough n, with probability at least $1 - K^2/n^2$ we have

$$\max\left(\max_k V^k, \max_{k_1,k_2} V^{k_1,k_2}\right) \le 2\Delta \le 3(C\log n/n)^{1/K} = O^*(1/n^{1/K})\,,$$

implying the main result for large enough n (note that $K^2/n^2 < 1/n$ for large enough n). For small values of n, we can simply choose $f(n,k)$ large enough to ensure that the bound holds with sufficient probability. □

Lemma 2. *For $k \in \mathcal{K}$, let $\tilde{\mathcal{R}}^k(\delta)$, $\tilde{\mathcal{V}}^k(\delta)$, and $\tilde{V}^k(\delta)$ be as defined by Eqs. (4), (5), and (6) respectively. Fix $K \ge 1$. There exists $f(n) = O^*(1/n)$ such that for any $\delta \in (0,1]$, the following occurs. Let*

$$\mathcal{B}_2 \equiv \left\{ \max_{k\in\mathcal{K}} \tilde{V}^k(\delta) \le f(n)/\delta^{K-1} \right\}\,. \tag{17}$$

Then

$$\Pr(\mathcal{B}_2) \ge 1 - 1/n\,.$$

Proof. The values in the set $\tilde{\mathcal{V}}^k \subset [0,1]$ follow a one-dimensional Poisson process with rate

$n\delta^{K-1}$. Choose $f(n) = 6\log n/n$. If $6\log n/(n\delta^{K-1}) \geq 1$ there is nothing to prove, since $\max_{k\in\mathcal{K}} \tilde{V}^k(\delta) \leq 1$ by definition. Hence assume $6\log n/(n\delta^{K-1}) < 1$. Divide $[0,1]$ into intervals of length $\Delta = f(n)/(3\delta^{K-1}) = 3\log n/(n\delta^{K-1})$ (to simplify notation, we assume that $1/\Delta \geq 2$ is an integer. The argument can easily be adapted to handle the case where $n\delta^{K-1}/(3\log n)$ is not an integer). The probability that any particular interval of length Δ does not contain a point is no more than $\exp(-3\log n) = 1/n^3$. The number of intervals of length Δ is $1/\Delta = n\delta^{K-1}/(3\log n) \leq n$ for large enough n. By union bound, with probability at least $1-1/n^2$, each Δ-interval contains at least one point, implying that $\tilde{V}^k(\delta) \leq 2\Delta = f(n)/\delta^{K-1}$ with probability at least $1-1/n^2$, as required. □

So far in this section we have considered the rate n Poisson process in $[0,1]^K$ for convenience. However, the results we have proved can easily be transported to the closely related model of n points distributed i.i.d. uniformly in $[0,1]^K$.

Lemma 3. *Consider n points distributed i.i.d. uniformly in $[0,1]^K$. Lemmas 1 and 2 hold for this model as well.*

Proof. We use a standard coupling argument along with monotonicity of the considered random variables with respect to additional points. Let $\mathcal{P}$ be a rate $n/2$ Poisson process in $[0,1]^K$. The N points are distributed i.i.d. uniformly in $[0,1]^K$ conditioned on the value of N. Let $\mathcal{B}$ be the event $N \leq n$. Clearly, $\mathcal{B}$ occurs with probability at least $1-1/n^2$. Let $\mathcal{U}$ be the process consisting of n points distributed i.i.d. in $[0,1]^K$. Conditioned on $\mathcal{B}$, we can couple the process $\mathcal{P}$ with the process $\mathcal{U}$ such that for every point in the Poisson process, there is an identically located point in $\mathcal{U}$.

We now show how to establish Lemma 2 for process $\mathcal{U}$ using such a coupling. Note that $\max_{k\in\mathcal{K}} \tilde{V}^k(\delta)$ is monotone nonincreasing as we add more points. As such, an upper bound on this quantity continues to hold if more points are added. For instance, consider $\max_{k\in\mathcal{K}} \tilde{V}^k(\delta)$. Let $\mathcal{B}'$ be the event that

$$\max_{k\in\mathcal{K}} \tilde{V}^k(\delta) \leq f(n/2)/\delta^{K-1}$$

under $\mathcal{P}$. The proof of Lemma 2 shows that $\Pr(\mathcal{B}') \geq 1 - (2/n)^2$. By union bound on $\overline{\mathcal{B}}$ and $\overline{\mathcal{B}'}$, we deduce that $\Pr(\mathcal{B} \cap \mathcal{B}') \geq 1 - 5/n^2 \geq 1 - 1/n$, for large enough n. We deduce, using a coupling as described above, that with probability at least $1 - 1/n$, for process $\mathcal{U}$ we have

$$\max_{k \in \mathcal{K}} \tilde{V}^k(\delta) \leq \tilde{f}(n)/\delta^{K-1}\,,$$

where $\tilde{f}(n) = f(n/2)$, for large enough n. (For small values of n, we can simply choose $\tilde{f}(n)$ large enough to ensure that the bound holds with sufficient probability.) Thus we have shown that Lemma 2 holds for process $\mathcal{U}$.

Lemma 1 can similarly be established for process $\mathcal{U}$ using that

$$\max\left(\max_{k \in \mathcal{K}} V^k,\ \max_{(k_1,k_2)\in\mathcal{K}^{(2)}} V^{k_1,k_2}(\delta)\right)$$

is monotone nonincreasing as we add more points. □

We now establish another result about n points $(\epsilon_j)_{j=1}^n$ distributed i.i.d. uniformly in $[0,1]^K$. This result is key to the proof of the tightness of Theorem 1 (Proposition 3).

For $\delta \in [0,1]$ let

$$\widehat{\mathcal{R}}^{k_1,k_2}(\delta) = \{x \in [0,1]^K : x^{k_1} \geq x^{k_2} - \delta\,;\ x^{k_1} \geq x^k\ \forall k \notin \{k_1,k_2\}, k \in \mathcal{K}\}. \tag{18}$$

Let $n^{k_1,k_2}(\delta)$ be the number of points in $\widehat{\mathcal{R}}^{k_1,k_2}(\delta)$.

Lemma 4. *Let $\mathcal{B}_3$ be the event that there for all $k_1, k_2 \in \mathcal{K}$ we have $n^{k_1,k_2} \geq 1 + n/K$. For $\delta = \delta(n) \geq 1/n^{0.49}$, we have that $\mathcal{B}_3$ occurs with high probability.*

Proof. A short calculation shows that the volume of $\widehat{\mathcal{R}}^{k_1,k_2}(\delta)$ is

$$v = \frac{1}{K-1}\left(1 - \frac{(1-\delta)^K}{K}\right) \tag{19}$$
$$\geq \frac{1}{K} + \frac{\delta}{K-1} - \frac{\delta^2}{2} \tag{20}$$
$$\geq \frac{1+\delta}{K} \tag{21}$$

for $\delta \leq 2/(K(K-1))$. Now, the probability of $\epsilon_j \in \widehat{\mathcal{R}}^{k_1,k_2}(\delta)$ is exactly v. It follows that n^{k_1,k_2} is distributed as Binomial(n, v). Notice $\mathrm{E}[n^{k_1,k_2}] = nv \geq n(1+\delta)/K$. We obtain

$$\Pr(n^{k_1,k_2} < 1 + n/K) \leq \exp\{-\Omega(n\delta^2)\} = \exp\{-\Omega(n^{0.02})\} = o(1) \tag{22}$$

using a standard Chernoff bound (see, e.g., Durrett 2010). Using union bound over pairs k_1, k_2 we deduce that $\overline{\mathcal{B}_3}$ occurs with probability $o(1)$; i.e., event $\mathcal{B}_3$ occurs with high probability. □

We now prove Theorem 1.

Proof of Theorem 1. By invoking Lemma 1, Lemma 2, and Lemma 3, for each t we have that w.p. at least $1 - \frac{2}{n_t}$ the event $\big(\mathcal{B}_1(t,\delta) \cap \mathcal{B}_2(t,\delta)\big)$ occurs. As the total number of types is upper-bounded by $K+Q$, we apply a union bound to conclude that w.p. at least $1 - \frac{2(K+Q)}{n^*}$, the event $\bigcap_{t \in \mathcal{T}_\mathcal{L} \cup \mathcal{T}_\mathcal{E}} \big(\mathcal{B}_1(t,\delta) \cap \mathcal{B}_2(t,\delta)\big)$ occurs, where $n^* = \min_{t \in \mathcal{T}_\mathcal{L} \cup \mathcal{T}_\mathcal{E}} n_t$. Note that, by Assumption 2, we have $n_t \geq Cn$ for every type $t \in \mathcal{T}_\mathcal{L} \cup \mathcal{T}_\mathcal{E}$ and some constant C, which completes the proof. □

C Proof of Theorem 1 (upper bound)

Proof of Proposition 1. The proposition follows immediately from stability. □

Throughout the rest of the appendix, we shall use the following characterization of $\mathcal{C}$.

Lemma 1 (Size of the core)**.** *Let M be the unique maximum-weight type-matching. Assume Assumption 2, fixed type-type compatibilities $u(\cdot,\cdot)$, idiosyncratic productivities being i.i.d. $U(0,1)$, and that at least one $u(\cdot,\cdot)$ is greater than -2 (otherwise there will no matched agent pairs). For each pair of types $(k,q) \in \mathcal{T}$, let $N(k,q)$ denote the number of matches between agents of type k*

and agents of type q. Then, with probability $1-1/n$, the size of the core $\mathcal{C}$ (cf. Definition 1) is bounded as

$$\mathcal{C} = \Theta\left(\frac{\sum_k \sum_q N(k,q)|\alpha_{kq}^{\max} - \alpha_{kq}^{\min}|}{\sum_k \sum_q N(k,q)}\right).$$

Proof. To prove this lemma, we need to show that the total number of matches $\sum_k \sum_q N(k,q)$ is within a constant factor of the social welfare with probability $1-1/n$. The total social welfare is bounded above by $n(2+\max u(\cdot,\cdot)) = O(n)$ and the total number of matches is bounded above by n. We shall show that both these quantities are also $\Omega(n)$ with probability $1-1/n$, which will then imply the lemma. Suppose that $u(k,q) > -2$. By Assumption 2, we know that there are at least Cn agents each of type k and type q. Let $\delta = (u(k,q)+2)/3$. Let

$$\mathcal{M}_k = \mathcal{M}_k(\delta) = \{i \in k : \eta_i^q \geq 1-\delta\}, \tag{23}$$

and similarly,

$$\mathcal{M}_q = \mathcal{M}_q(\delta) = \{j \in q : \epsilon_j^k \geq 1-\delta\}. \tag{24}$$

Clearly, one can find a matching (not necessarily a stable outcome) consisting of $\min(|\mathcal{M}_k|, |\mathcal{M}_q|)$ pairs, each with one agent from $\mathcal{M}_k$ and the other agent from $\mathcal{M}_q$. The weight of every edge/pair in this matching is at least $u(k,q)+2(1-\delta) = \delta$, and hence the total weight of the matching is at least $\delta \min(|\mathcal{M}_k|, |\mathcal{M}_q|)$. A standard Chernoff bound yields that $|\mathcal{M}_k| \geq Cn \min(\delta/2, 1)$ with probability $1-1/n^2$, and similarly for $|\mathcal{M}_q|$. It follows that the total weight of the matching is at least $Cn\delta \min(\delta/2,1) = \Omega(n)$ with probability $1-1/n$, as needed. Since stable outcomes maximize social welfare it follows that the total social welfare is $\Omega(n)$ in stable outcomes.

It remains to bound the number of matches in stable outcomes. We claim that either all agents in $\mathcal{M}_k$ are matched or all agents in $\mathcal{M}_q$ are matched. Otherwise, there is a blocking pair consisting of an unmatched agent from each set. It follows that there are at least $\min(|\mathcal{M}_k|, |\mathcal{M}_q|)$ matched agents in maximum-weight matching (the matching in all stable outcomes) and, in particular, at least $Cn\delta \min(\delta/2,1) = \Omega(n)$ matched pairs under the lower bound obtained above on $|\mathcal{M}_k|$ and $|\mathcal{M}_q|$. □

The above characterization differs from the original definition (Definition 1) only in that the size of the core is scaled by the number of matched agents instead of being scaled by the social welfare. Lemma 1 allows us to bound $\mathcal{C}$ by controlling $|\alpha_{kq}^{\max} - \alpha_{kq}^{\min}|$.

We are now ready to present the complete proof of Theorem 1. Before moving on to the key lemmas, we introduce some definitions. For every type $t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}$, we define $\vartheta(t)$ as $\vartheta(t) = \{k \in \mathcal{T}_{\mathcal{L}} : N(k,t) > 0\}$ when $t \in \mathcal{T}_{\mathcal{E}}$ and $\vartheta(t) = \{q \in \mathcal{T}_{\mathcal{E}} : N(t,q) > 0\}$ when $t \in \mathcal{T}_{\mathcal{L}}$. That is, $\vartheta(t)$ is the set of neighbors of t in the graph $G(M)$. Recall that, given a type $t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}$, we denote by $D(t)$ the *dimension* of the idiosyncratic productivity vector associated with agents of type t. That is, $D(t) = K$ if $t \in \mathcal{T}_{\mathcal{E}}$ and $D(t) = K$ if $t \in \mathcal{T}_{\mathcal{L}}$.

Given a type $t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}$ we denote by $\nu(t)$ (or simply ν) the points in t. That is, for each agent j of type t, we define ν_j as follows:

$$\nu_j = \begin{cases} \epsilon_j & \text{if } t \in \mathcal{T}_{\mathcal{E}} \\ \eta_j & \text{if } t \in \mathcal{T}_{\mathcal{L}}. \end{cases}$$

For a fixed $t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}$ and $t' \in \vartheta(t)$, let $\beta_{tt'}$ be defined as:

$$\beta_{tt'} = \begin{cases} -\alpha_{tt'} & \text{if } t \in \mathcal{T}_{\mathcal{E}} \\ \alpha_{tt'} - u(t,t') & \text{if } t \in \mathcal{T}_{\mathcal{L}}. \end{cases} \tag{25}$$

Note that $\beta_{tt'}$ can be interpreted as the price that agents of type t pay to match to type t'.

Using the above notation, we can rewrite the conditions in Proposition 2 associated with a fixed type $t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}$ as follows

(ST) For every $k, k' \in \vartheta(t)$:

$$\min_{j \in t \cap M(k)} \nu_j^k - \nu_j^{k'} \geq \beta_{kt} - \beta_{k't} \geq \max_{j \in t \cap M(k')} \nu_j^k - \nu_j^{k'}.$$

(IM) For every $k \in \vartheta(t)$:

$$\min_{j \in t \cap M(k)} \nu_j^k \geq \beta_{kt} \geq \max_{j \in q \cap U} \nu_j^k.$$

As all the ν variables are in $[0,1]$, then the above conditions can be interpreted as geometric conditions in the $[0,1]^{D(t)}$-hypercube.

Lemma 2. *Consider the unique maximum-weight type-matching M and a type $t \in \mathcal{T}_{\mathcal{L}} \cup \mathcal{T}_{\mathcal{E}}$. Let $\mathcal{F}_1(t)$ be the event*

$$\mathcal{F}_1(t) = \{t \text{ is marked in } G(M) \text{ and at least one agent in } t \text{ is matched}\}; \tag{26}$$

that is, t has at least one unmatched and one matched agent. Let the events $\mathcal{B}_1(t,\delta)$ and $\mathcal{B}_2(t,\delta)$ be as defined by Eqs. (7) *and* (8) *respectively. Under $\mathcal{F}_1(t) \cap \mathcal{B}_1(t,\delta) \cap \mathcal{B}_2(t,\delta)$, we have*

$$\max_{t' \in \vartheta(t)} \left(\alpha_{t,t'}^{\max} - \alpha_{t,t'}^{\min}\right) \leq \max\left(2f_1(n_t, D(t)) + \delta, f_2(n_t)/\delta^{D(t)-1}\right),$$

where f_1 and f_2 agree with those in the definitions of events $\mathcal{B}_1(t,\delta)$ and $\mathcal{B}_2(t,\delta)$ respectively.

The proof of Lemma 2 is partitioned into two lemmas. Given a core solution (M,α), let the event $\mathcal{D}(t,\delta)$ be defined as

$$\mathcal{D}(t,\delta) = \{\beta_{tz} \geq \delta \quad \forall z \in \vartheta(t)\}. \tag{27}$$

We interpret $\mathcal{D}(t,\delta)$ as the event that *all* prices seen by type t agents (weakly) exceed δ. Lemma 3 below deals with $\mathcal{D}(t,\delta)$ whereas Lemma 4 deals with the complement $\overline{\mathcal{D}(t,\delta)}$. Together they imply Lemma 2.

Lemma 3. *Consider a core solution (M,α) and a type t. Let the events $\mathcal{F}_1(t)$, $\mathcal{D}(t,\delta)$, and $\mathcal{B}_2(t,\delta)$ be as defined by Eqs.* (26), (27), *and* (8) *respectively. Under $\mathcal{F}_1(t) \cap \mathcal{D}(t,\delta) \cap \mathcal{B}_2(t,\delta)$, we have $\max_{t' \in \vartheta(t)} \left(\alpha_{t,t'}^{\max} - \alpha_{t,t'}^{\min}\right) \leq f_2(n_t)/\delta^{D(t)-1}$, where f_2 is as defined in the statement of Lemma 2.*

Proof. The intuition is that when all prices seen by type t agents (weakly) exceed δ, we can individually bound the variation within the core of each price: for each type $k \in \vartheta(t)$ on the other side, we consider agents of type t who are not interested in types other than k and bound the gap between consecutive order statistics of ν_j^k for such agents.

Let $D = D(t)$. Fix $k \in \vartheta(t)$ and consider the orthotope $\tilde{\mathcal{R}}^k = \tilde{\mathcal{R}}^k(t, \delta)$ as defined by Eq. (4). As $\mathcal{D}(t, \delta)$ occurs, $\beta_{tz} \geq \delta$ for all $z \in \vartheta(t)$ and therefore $\tilde{\mathcal{R}}^k$ can contain only points corresponding to agents in $M(k) \cup U$. By using the notation introduced above, condition (IM) in Proposition 2 implies that $\alpha_{kt}^{\max} - \alpha_{kt}^{\min} \leq \min_{j \in t \cap M(k)} \nu_j^k - \max_{j \in q \cap U} \nu_j^k$. However, $\min_{j \in t \cap M(k)} \nu_j^k - \max_{j \in q \cap U} \nu_j^k \leq \min_{j \in \tilde{\mathcal{R}}^k \cap M(k)} \nu_j^k - \max_{j \in \tilde{\mathcal{R}}^k \cap U} \nu_j^k \leq \tilde{V}^k(t, \delta)$, where $\tilde{V}^k(t, \delta)$ is as defined by Eq. (6). Therefore, for each $k \in \vartheta(t)$ we must have $\alpha_{kt}^{\max} - \alpha_{kt}^{\min} \leq \tilde{V}^k(t, \delta)$. Finally, under $\mathcal{B}_2(t, \delta)$ we have $\max_{k \in \vartheta(t)} \tilde{V}^k(t, \delta) \leq f_2(n_t)/\delta^{D-1}$, which completes the result. □

Lemma 4. *Consider a core solution (M, α) and a type t. Let $\mathcal{F}_1(t)$ be the event defined in Eq. (26). Let the event $\mathcal{B}_1(t, \delta)$ be as defined by Eq. (7), and let the event $\overline{\mathcal{D}(t, \delta)}$ denote the complement of the event defined by Eq. (27). Under $\mathcal{F}_1(t) \cap \overline{\mathcal{D}(t, \delta)} \cap \mathcal{B}_1(t, \delta)$, we have $\max_{t' \in \vartheta(t)} \left(\alpha_{t,t'}^{\max} - \alpha_{t,t'}^{\min} \right) \leq 2 f_1(n_t, D(t)) + \delta$, where f_1 is as defined in the statement of Lemma 1.*

Proof. Here at least one price seen by type t is less than δ, since we are considering $\overline{\mathcal{D}(t, \delta)}$. The first part of our proof controls the variation in the core of the price that is the smallest at α; we call the corresponding type k^*. The second part of our proof controls relative variation in prices faced by type t, and then uses the control on the price of type k^* from the first part to control prices of other types $k \neq k^*$.

Suppose that $t \in \mathcal{T}_{\mathcal{E}}$. Consider the unit hypercube in $\mathbb{R}^K$. For each $j \in \mathcal{E}$ such that $\tau(j) = t$, let $\epsilon_j \in [0, 1]^K$ denote the vector of realizations of ϵ_j^k for every $k \in \mathcal{T}_{\mathcal{L}}$. By condition (ST) in Proposition 2, we can partition the $[0, 1]^K$ hypercube into $|\vartheta(t)|$ regions such that all the points ϵ corresponding to agents matched to $k \in \vartheta(t)$ must be contained in the corresponding region (these regions might also contain points corresponding to unmatched agents). In particular, for each $k \in \vartheta(t)$, we define $Z(k) \subseteq [0, 1]^K$ to be the region corresponding to type k, with $Z(k) = \cap_{k' \in \vartheta(t),\ k' \neq k} \{x \in [0, 1]^K : x_k - x_{k'} \geq \alpha_{k't} - \alpha_{kt}\}$. Note that the region $Z(k)$ can contain only points corresponding to agents matched to k or unmatched.

Let $k^* = \mathrm{argmax}_{k \in \mathcal{T}_{\mathcal{L}}} \{\alpha_{tk} : \ k \in \vartheta(t)\}$, and let $\mathcal{R}^{k^*} = \mathcal{R}^{k^*}(t)$ be as defined by Eq. (1). By condition (ST) in Proposition 2, we have that for all $k \in \vartheta(t)$,

$$\min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} - \epsilon_j^k \geq \alpha_{kt} - \alpha_{k^*t} \geq \max_{j \in q \cap M(k)} \epsilon_j^{k^*} - \epsilon_j^k.$$

As $\alpha_{kt} - \alpha_{k^*t} \leq 0$ for all $k \in \vartheta(t)$, we must have $\mathcal{R}^{k^*} \subseteq Z(k^*)$. Let $V^{k^*} = V^{k^*}(t)$ be as defined in Eq. (4). We claim that $\alpha_{k^*,t}^{\max} - \alpha_{k^*,t}^{\min} \leq V^{k^*}$. To see why this holds, consider two separate cases. First, suppose that there is at least one point corresponding to an unmatched agent in $\mathcal{R}^{k^*}$. By condition (IM) in Proposition 2, we must have $\min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} \geq -\alpha_{k^*t} \geq \max_{j \in t \cap U} \epsilon_j^{k^*}$. Hence, $\alpha_{k^*,t}^{\max} - \alpha_{k^*,t}^{\min} \leq \min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} - \max_{j \in t \cap U} \epsilon_j^{k^*} \leq V^{k^*}$, as desired. For the second case, suppose that all points in $\mathcal{R}^{k^*}$ correspond to matched agents. As $\max_{j \in t \cap U} \epsilon_j^{k^*} \geq 0$, we must have $\alpha_{k^*,t}^{\max} - \alpha_{k^*,t}^{\min} \leq \min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} \leq \min_{j \in \mathcal{R}^{k^*}} \epsilon_j^{k^*} \leq V^{k^*}$, as the difference between 0 and the $\min_{j \in \mathcal{R}^{k^*}} \epsilon_j^{k^*}$ is upper-bounded by V^{k^*}. Therefore, we conclude that $\alpha_{k^*,t}^{\max} - \alpha_{k^*,t}^{\min} \leq V^{k^*}$.

Next, we consider the bound for any arbitrary type $k \in \vartheta(t)$. By condition (ST) in Proposition 2, we have that for all $k \in \vartheta(t)$,

$$\alpha_{k^*t}^{\max} + \min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} - \epsilon_j^k \geq \alpha_{kt} \geq \alpha_{k^*t}^{\min} + \max_{j \in q \cap M(k)} \epsilon_j^{k^*} - \epsilon_j^k.$$

Therefore,

$$\alpha_{kt}^{\max} - \alpha_{kt}^{\min} \leq \alpha_{k^*t}^{\max} - \alpha_{k^*t}^{\min} + \min_{j \in t \cap M(k^*)} \left(\epsilon_j^{k^*} - \epsilon_j^k\right) - \max_{j \in q \cap M(k)} \left(\epsilon_j^{k^*} - \epsilon_j^k\right).$$

From our previous bound, we have that $\alpha_{k^*t}^{\max} - \alpha_{k^*t}^{\min} \leq V^{k^*}$. We now want an upper bound on $\min_{j \in t \cap M(k^*)} (\epsilon_j^{k^*} - \epsilon_j^k) - \max_{j \in q \cap M(k)} (\epsilon_j^{k^*} - \epsilon_j^k)$. Let $\mathcal{R}^{k^*,k} = \mathcal{R}^{k^*,k}(t, \delta)$ and $V^{k^*,k} = V^{k^*,k}(t)$ be as defined by Eqs. (2) and (6). We shall show that $\min_{j \in t \cap M(k^*)} (\epsilon_j^{k^*} - \epsilon_j^k) - \max_{j \in q \cap M(k)} (\epsilon_j^{k^*} - \epsilon_j^k) \leq V^{k^*,k} + \delta$. (Recall that, under $\overline{\mathcal{D}(t, \delta)}$, we have $\delta \geq -\alpha_{k^*t}$.)

Note that all points in $\mathcal{R}^{k^*,k}$ must correspond to agents matched to k^* or matched to k, as the region $\mathcal{R}^{k^*,k}$ cannot contain unmatched agents without violating condition (IM). Furthermore, as $\mathcal{R}^{k^*} \subseteq Z(k^*)$ and $\mathcal{R}^{k^*} \cap \mathcal{R}^{k^*,k} \neq \emptyset$, at least one point in $\mathcal{R}^{k^*,k}$ corresponds to an agent matched to k^*. We now consider two separate cases, depending on whether $\mathcal{R}^{k^*,k}$ contains at least one point matched to k. First, suppose that $\mathcal{R}^{k^*,k}$ contains at least one point matched to

k. Then, the bound trivially applies as

$$\min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} - \epsilon_j^k - \max_{j \in t \cap M(k)} \epsilon_j^{k^*} - \epsilon_j^k \leq \min_{j \in \mathcal{R}^{k^*,k} \cap M(k^*)} \epsilon_j^{k^*} - \epsilon_j^k - \max_{j \in \mathcal{R}^{k^*,k} \cap M(k)} \epsilon_j^{k^*} - \epsilon_j^k \leq V^{k^*,k}.$$

Otherwise, $\mathcal{R}^{k^*,k}$ contains only points matched to k^*. In that case,

$$\min_{j \in t \cap M(k^*)} \epsilon_j^{k^*} - \epsilon_j^k - \max_{j \in q \cap M(k)} \epsilon_j^{k^*} - \epsilon_j^k \leq \min_{j \in \mathcal{R}^{k^*,k}} \epsilon_j^{k^*} - \epsilon_j^k - (1 + \alpha_{k^*t}) \leq V^{k^*,k} + \delta,$$

as desired. Overall, we have shown that:

$$\max_{k \in \vartheta(t)} \left(\alpha_{tk}^{\max} - \alpha_{tk}^{\min} \right) \leq \max \left(V^{k^*}, \max_{k \in \vartheta(t)} \left(V^{k^*} + V^{k^*,k} + \delta \right) \right).$$

Under $\mathcal{B}_1(t, \delta)$ we have $\max \left(V^{k^*}, \max_k V^{k^*,k} \right) \leq f_1(n_t, K)$, implying

$$\max \left(V^{k^*}, \max_{k \in \vartheta(t)} \left(V^{k^*} + V^{k^*,k} + \delta \right) \right) \leq 2 f_1(n_t, K) + \delta,$$

as desired.

To conclude, we briefly discuss the changes when $t \in \mathcal{T}_{\mathcal{L}}$. Consider the unit hypercube in $\mathbb{R}^Q$. For each $j \in \mathcal{L}$ such that $\tau(j) = t$, let $\eta_j \in [0,1]^Q$ denote the vector of realizations of η_j^q for every $q \in \mathcal{T}_{\mathcal{E}}$. For each $q \in \vartheta(t)$, we define $Z(q) \subseteq [0,1]^Q$ to be the region corresponding to type q. The main difference with the case in which $t \in \mathcal{T}_{\mathcal{E}}$ is that we need to define the regions $Z(q)$ in terms of $\tilde{\eta}$ instead of η. To that end, let $\beta_{kq} = \alpha_{kq} - u(k,q)$. By the (ST) condition in Proposition 2, we must have:

$$\min_{i \in t \cap M(q')} \tilde{\eta}_i^{q'} - \tilde{\eta}_i^q \geq \alpha_{tq'} - \alpha_{tq} \geq \max_{i \in t \cap M(q)} \tilde{\eta}_i^{q'} - \tilde{\eta}_i^q,$$

or equivalently,

$$\min_{i \in t \cap M(q')} \eta_i^{q'} - \eta_i^q \geq \beta_{tq'} - \beta_{tq} \geq \max_{i \in t \cap M(q)} \eta_i^{q'} - \eta_i^q.$$

By using β instead of α, the same geometric intuition as before applies. Then, we define $Z(q) = \cap_{q' \in \vartheta(t),\, q' \neq q} \{x \in [0,1]^Q : x_q - x_{q'} \geq \beta_{qt} - \beta_{q't}\}$. To select q^*, we just select the one with smallest β_{qt}. The rest of the proof remains the same. $\square$

Proof of Lemma 2. Lemma 2 immediately follows from Lemmas 3 and 4. $\square$

Lemma 5. *Consider the unique maximum-weight type-matching M and a type $t \in \mathcal{T}_{\mathcal{L}} \cap \mathcal{T}_{\mathcal{E}}$. Let $\mathcal{F}_2(t)$ be the event*

$$\mathcal{F}_2(t) = \{\textit{all agents in } t \textit{ are matched}\}.$$

Let the event $\mathcal{B}_1(t,\delta)$ be as defined by Eq. (7). Under $\mathcal{F}_2(t) \cap \mathcal{B}_1(t,\delta)$, for every $t^* \in \vartheta(t)$ we have $\max_{t' \in \vartheta(t)} \left(\alpha^{\max}_{t,t'} - \alpha^{\min}_{t,t'}\right) \leq \left(\alpha^{\max}_{t,t*} - \alpha^{\min}_{t,t*}\right) + 2f_1(n_t, D(t)) + 2\delta$, where f_1 agrees with the one in the definition of $\mathcal{B}_1(t,\delta)$.

Proof. This proof is similar to the proof of Lemma 4. Consider a core solution (M,α). Let $D = D(t)$. Fix a type $t^* \in \vartheta(t)$, and let $k^* = \text{argmax}_{k \in \vartheta(t)} \beta_{tk}$. We start by showing that, under $\mathcal{F}_2(t) \cap \mathcal{B}_1(t,\delta)$, we must have $\alpha^{\max}_{tk^*} - \alpha^{\min}_{tk^*} \leq \left(\alpha^{\max}_{t,t*} - \alpha^{\min}_{t,t*}\right) + f_1(n_t, D(t)) + 2\delta$. If $k^* = t^*$, the claim follows trivially. Otherwise, let $\mathcal{R}^{k^*,t^*} = \mathcal{R}^{k^*,t^*}(t,\delta)$ and $V^{k^*,t^*} = V^{k^*,t^*}(t,\delta)$ be as defined by Eqs. (2) and (6). We show that $\min_{j \in t \cap M(k^*)} \nu^{k^*}_j - \nu^{t^*}_j - \max_{j \in q \cap M(t^*)} \nu^{k^*}_j - \nu^{t^*}_j \leq V^{k^*,t^*} + \delta$. To that end, note that all points in $\mathcal{R}^{k^*,t^*}$ must correspond to agents matched to k^* or matched to t^*, as under $\mathcal{F}_2(t)$ all agents in t are matched. Furthermore, by the definition of k^*, $\mathcal{R}^{k^*,t^*}$ must contain a point corresponding to an agent matched to k^*. We now consider two separate cases, depending on whether $\mathcal{R}^{k^*,t^*}$ contains at least one point corresponding to an agent matched to t^*. First, suppose that $\mathcal{R}^{k^*,t^*}$ contains at least one point corresponding to an agent matched to t^*. Then,

$$\min_{j \in t \cap M(k^*)} \nu^{k^*}_j - \nu^{t^*}_j - \max_{j \in t \cap M(t^*)} \nu^{k^*}_j - \nu^{t^*}_j \leq \min_{j \in \mathcal{R}^{k^*,t^*} \cap M(k^*)} \nu^{k^*}_j - \nu^{t^*}_j - \max_{j \in \mathcal{R}^{k^*,t^*} \cap M(k)} \nu^{k^*}_j - \nu^{t^*}_j \leq V^{k^*,t^*}.$$

Otherwise, $\mathcal{R}^{k^*,t^*}$ contains only points matched to k^*. In that case,

$$\min_{j \in t \cap M(k^*)} (\nu^{k^*}_j - \nu^{t^*}_j) - \max_{j \in t \cap M(t^*)} (\nu^{k^*}_j - \nu^{t^*}_j) \leq \min_{j \in \mathcal{R}^{k^*,t^*}} (\nu^{k^*}_j - \nu^{t^*}_j) - 1 \leq V^{k^*,t^*} + \delta,$$

as desired. By condition (ST) in Proposition 2, we must have:

$$\alpha_{tk^*}^{\max}-\alpha_{tk^*}^{\min} \le \alpha_{tt^*}^{\max}-\alpha_{tt^*}^{\min}+\min_{j\in t\cap M(k^*)}(\nu_j^{k^*}-\nu_j^{t^*})-\max_{j\in t\cap M(t^*)}(\nu_j^{k^*}-\nu_j^{t^*}) \le \alpha_{tt^*}^{\max}-\alpha_{tt^*}^{\min}+V^{k^*,t^*}+\delta.$$

Next, consider an arbitrary $k \in \vartheta(t)$ with $k \neq t^*, k^*$. By condition (ST) in Proposition 2, we must have

$$\alpha_{kt}^{\max} - \alpha_{kt}^{\min} \le \alpha_{k^*t}^{\max} - \alpha_{k^*t}^{\min} + \min_{j\in t\cap M(k^*)}(\nu_j^{k^*} - \nu_j^{k}) - \max_{j\in q\cap M(k)}(\nu_j^{k^*} - \nu_j^{k}).$$

Let $\mathcal{R}^{k^*,k} = \mathcal{R}^{k^*,k}(t,\delta)$ and $V^{k^*,k} = V^{k^*,k}(t)$ be as defined by Eqs. (2) and (6). By repeating the same arguments as before, we can show that $\min_{j\in t\cap M(k^*)}(\nu_j^{k^*}-\nu_j^{k})-\max_{j\in q\cap M(k)}(\nu_j^{k^*}-\nu_j^{k}) \le V^{k^*,k} + 2\delta$. Hence,

$$\alpha_{kt}^{\max} - \alpha_{kt}^{\min} \le \alpha_{k^*t}^{\max} - \alpha_{k^*t}^{\min} + V^{k^*,k} + \delta \le \alpha_{tt^*}^{\max} - \alpha_{tt^*}^{\min} + V^{k^*,t^*} + V^{k^*,k} + 2\delta.$$

To conclude, note that

$$\max_{k\in\vartheta(t)}\left(\alpha_{kt}^{\max} - \alpha_{kt}^{\min}\right) \le \left(\alpha_{tt^*}^{\max} - \alpha_{tt^*}^{\min}\right)+2\left(\max_{k\in\vartheta(t)} V^{k^*,k}\right)+2\delta \le \left(\alpha_{tt^*}^{\max} - \alpha_{tt^*}^{\min}\right)+2f_1(n_t,D)+2\delta,$$

where the last inequality follows from the fact that $\mathcal{B}_1(t,\delta)$ occurs by hypothesis. □

We can now proceed to the proof of the main theorem.

Proof of Theorem 1 (upper bound). Let $n^* = \min_{t\in\mathcal{T}_\mathcal{L}\cup\mathcal{T}_\mathcal{E}} n_t$. Under Assumption 2, we have that $n^* = \Theta(n)$. Let $\delta = 1/(n^*)^{1/\max(K,Q)}$. For each $t \in \mathcal{T}_\mathcal{L}\cup\mathcal{T}_\mathcal{E}$, let the events $\mathcal{B}_1(t,\delta)$ and $\mathcal{B}_2(t,\delta)$ be as defined by Eqs. (7) and (8) respectively. We start by showing that, under $\bigcap_{t\in\mathcal{T}_\mathcal{L}\cup\mathcal{T}_\mathcal{E}}\left(\mathcal{B}_1(t,\delta)\cap\mathcal{B}_2(t,\delta)\right)$, we must have $\mathcal{C} \le O^*\left(1/n^{1/\max(K,Q)}\right)$. To that end, construct the type-adjacency graph $G(M)$ as defined in Section 4. For each vertex v, we denote by $d(v)$ the minimum distance between v and any marked vertex (that is, $d(v) = 0$ if v is marked, $d(v) = 1$ if v is unmarked and has a marked neighbour, and so on). By Lemma 1, we know that w.p. 1, each connected component of $G(M)$ must contain at least one marked vertex, and so $d(v)$ is well defined for all v. Let $C_d = \{v \in C : d(v) = d\}$; that is, C_d is the set of vertices that are at distance d from a marked vertex. We now show the result by induction in d. In particular, we show that, under $\bigcap_{t\in\mathcal{T}_\mathcal{L}\cup\mathcal{T}_\mathcal{E}} (\mathcal{B}_1(t,\delta)\cap\mathcal{B}_2(t,\delta))$, for each $t \in C_d$ we have that $\max_{k\in\vartheta(t)} \left(\alpha_{tk}^{\max} - \alpha_{tk}^{\min}\right) \leq g_d(n^*, \max(K,Q))$ for some $g_d(n^*, \max(K,Q)) = O^*(1/n^{1/\max(K,Q)})$.

We start by showing that the claim holds for the base case $d = 0$. For each $t \in C_0$, either all agents in t are unmatched or at least one agent is matched. In the former case, we can just ignore type t as it will not contribute to the size of the core. In the latter, we note that w.p. 1 the event $\mathcal{F}_1(t)$ as defined in the statement of Lemma 2 must hold. Therefore, we can apply Lemma 2 to obtain $\max_{t'\in\vartheta(t)} \left(\alpha_{t,t'}^{\max} - \alpha_{t,t'}^{\min}\right) \leq \max\left(2f_1(n_t, D(t)) + \delta, f_2(n_t)/\delta^{D(t)-1}\right)$, where f_1 and f_2 are as defined in the statement of the lemma. To conclude the proof of the base case, let

$$g_0(n^*, \max(K,Q)) = \max\left(2f_1(n^*, \max(K,Q)) + \delta, f_2(n^*)/\delta^{\max(K,Q)-1}\right).$$

By the definition of f_1, f_2, and δ, together with Assumption 2, we have $g_0(n^*, \max(K,Q)) = O^*(1/n^{1/\max(K,Q)})$. Therefore, we have shown that, for every $t \in C_0$, we have

$$\max_{k\in\vartheta(t)} \left(\alpha_{tk}^{\max} - \alpha_{tk}^{\min}\right) \leq g_0(n^*, \max(K,Q)).$$

Now suppose that the result holds for all $d' \leq d$; we want to show that it holds for $d+1$. Fix $t \in C_{d+1}$. By the definition of C_{d+1}, we have that all agents in t must be matched and therefore, w.p. 1, the event $\mathcal{F}_2(t)$ as defined in the statement of Lemma 5 occurs. Moreover, there must exist a t^* such that the vertex corresponding to t^* is C_d and $t^* \in \vartheta(t)$. By induction, we have that $\left(\alpha_{tt^*}^{\max} - \alpha_{tt^*}^{\min}\right) \leq g_d(n^*, \max(K,Q))$ for $g_d(n^*, \max(K,Q)) = O^*(1/n^{1/\max(K,Q)})$. Furthermore, by Lemma 5, we know that under $\mathcal{F}_2(t)\cap\mathcal{B}_1(t,\delta)$, we have $\max_{t'\in\vartheta(t)} \left(\alpha_{t,t'}^{\max} - \alpha_{t,t'}^{\min}\right) \leq \left(\alpha_{t,t*}^{\max} - \alpha_{t,t*}^{\min}\right) + 2f_1(n_t, D(t)) + 2\delta$, where $\mathcal{B}_1(t,\delta)$ is as defined by Eq. (9) and f is as defined in the statement of Lemma 1. Therefore, by letting $g_{d+1}(n^*, \max(K,Q)) = g_d(n^*, \max(K,Q)) + 2f_1(n^*, \max(K,Q)) + 2\delta$, we have shown that with probability at least $1 - \frac{d+1}{n^*}$, we have

$$\max_{k\in\vartheta(t)} \left(\alpha_{tk}^{\max} - \alpha_{tk}^{\min}\right) \leq g_{d+1}(n^*, \max(K,Q)),$$

with $g_{d+1}(n^*, \max(K,Q)) = O^*(1/n^{1/\max(K,Q)})$.

Next, we note that $\max_v d(v)$ is upper-bounded by $K+Q$. Hence, for every $t \in \mathcal{T}_\mathcal{L} \cup \mathcal{T}_\mathcal{E}$, we have $\max_{k\in\vartheta(t)} \left(\alpha_{tk}^{\max} - \alpha_{tk}^{\min}\right) \leq g_{K+Q}(n^*, \max(K,Q))$ for $g_{K+Q}(n^*, \max(K,Q)) = O^*(1/n^{1/\max(K,Q)})$ and, therefore,

$$\max_{t\in\mathcal{T}_\mathcal{L}\cup\mathcal{T}_\mathcal{E}} \max_{k\in\vartheta(t)} \left(\alpha_{tk}^{\max} - \alpha_{tk}^{\min}\right) \leq g_{K+Q}(n^*, \max(K,Q)).$$

To conclude, by Theorem 1 we have that with probability at least $1 - \frac{2(K+Q)}{n^*}$, the event $\bigcap_{t\in\mathcal{T}_\mathcal{L}\cup\mathcal{T}_\mathcal{E}} \left(\mathcal{B}_1(t,\delta) \cap \mathcal{B}_2(t,\delta)\right)$ occurs. In all other cases, we just use the fact that the size of the core is upper-bounded by a constant $C < \infty$. Hence,

$$\begin{aligned}
\mathrm{E}[\mathcal{C}] &= \frac{\sum_{(k,q)\in\mathcal{T}_\mathcal{L}\times\mathcal{T}_\mathcal{E}} N(k,q)\left(\alpha_{kq}^{\max} - \alpha_{kq}^{\min}\right)}{\sum_{(k,q)\in\mathcal{T}_\mathcal{L}\times\mathcal{T}_\mathcal{E}} N(k,q)} \\
&\leq (K+Q) g_{K+Q}(n^*, \max(K,Q)) + C\frac{2(K+Q)}{n^*} \\
&= O^*\left(\frac{1}{\sqrt[\max(K,Q)]{n}}\right)
\end{aligned}$$

implying the main result for large enough n (note that $\frac{2(K+Q)}{n^*} = \Theta^*(1/n)$). □

D Theorem 1 (lower bound): Proof of Proposition 3

Again, we make use of the characterization of $\mathcal{C}$ in Lemma 1.

Proof of Proposition 3. We consider the sequence of markets described in Section 4.2. We remark that the gaps in values of n in the described sequence can easily be filled in, leaving the core and its size essentially unaffected.[3]

Claim 1. *For this market, all labor agents of types different from k_* will be matched in the core.*

Proof. We know that there is some employer j who is either unmatched or matched to a labor agent i' of type k_*. Consider any matching where a labor agent i of type $k \neq k_*$ is unmatched.

[3] In the described construction, we have $n = (2K-1)\tilde{n} + 1$ for $\tilde{n} = 1, 2, \ldots$, but intermediate values of n can be handled by having slightly fewer workers of type k_*, which leaves our analysis essentially unaffected.

Now $\Phi(i',j) = \epsilon_{i'} + \eta_j^{k_*} \le 1+1 = 2$, whereas $\Phi(i,j) \ge u(k,1) = 3$, and hence the weight of such a matching can be increased by instead matching j to i. It follows that in any maximum-weight matching, all labor agents of a type different from k_* are matched. Finally, recall that every core outcome lives on a maximum-weight matching; cf. Proposition 2. □

Among agents $i \in k_*$, exactly one agent will be matched, specifically agent $i_* = \arg\max_{i \in k_*} \eta_i$. Let j_* be the agent matched to i_* (break ties arbitrarily). Recall that core solutions always live on a maximum-weight matching, and in case of multiple maximum-weight matchings, the set of vectors α such that (M, α) is a core solution is the same for any maximum-weight matching M. This allows us to suppress the matching, and talk about a vector α being in the core; cf. Proposition 2. The (IR) conditions in Proposition 2 for the pair of types $(k_*, 1)$ are

$$\eta_{i_*} \ge \alpha_{k_*} \ge \max_{i \in k_* \setminus i_*} \eta_i \,, \tag{28}$$

and the slack condition $\alpha_{k_*} \ge -\epsilon_{j_*}^{k_*}$. The (IM) conditions for types $(k,1)$ for $k \ne k_*$ are

$$3 + \min_{i \in k} \eta_i \ge \alpha_k \ge - \min_{j \in M(k)} \epsilon_j^k \,. \tag{29}$$

The stability conditions are

$$\min_{j \in M(k)} \epsilon_j^k - \epsilon_j^{k'} \ge \alpha_{k'} - \alpha_k \ge \max_{j \in M(k')} \epsilon_j^k - \epsilon_j^{k'} \,, \tag{30}$$

for all $k \ne k'$. It is easy to see that Eq. (30) with $k' = k_*$ implies that $\alpha_k \le 2$ for all $k \ne k_*$. Hence, the upper bound in Eq. (29) is slack. Consider the left stability inequality with $k' = k_*$. As Eq. (28) implies that $\alpha_{k_*} \ge 0$, we must have

$$\alpha_k \ge - \min_{j \in M(k)} \epsilon_j^k - \epsilon_j^{k_*} \ge - \min_{j \in M(k)} \epsilon_j^k$$

implying that the lower bound in (29) is also slack. Thus a vector α is in the core if and only if conditions (28) and (30) are satisfied.

For simplicity, we start with the special case of $K = 2$, with the two types of labor being k

and k_*. To obtain intuition, notice that from Eq. (28) we have $\alpha_{k_*} \xrightarrow{\tilde{n}\to\infty} 1$ in probability, and when we use this together with Eq. (30) we obtain $\alpha_k \xrightarrow{\tilde{n}\to\infty} 2$ in probability. (We do not use these limits in our formal analysis below.) Hence, we focus on Eq. (28) together with

$$\min_{j\neq j_*} \epsilon_j^k - \epsilon_j^{k_*} \geq \alpha_{k_*} - \alpha_k \geq \epsilon_{j_*}^k - \epsilon_{j_*}^{k_*}\,, \tag{31}$$

where $j_* = \arg\min_j \epsilon_j^k - \epsilon_j^{k_*}$. Now, $X_j = \epsilon_j^k - \epsilon_j^{k_*}$ are distributed i.i.d. with density $U[0,1] * U[-1,0]$, which is

$$f(x) = \begin{cases} 1 - |x| & \text{for } |x| \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

(Note that if we draw $\tilde{n}+1$ samples from this distribution, it is not hard to see that $\mathrm{E}[(\min_{j\neq j_*} X_j) - X_{j_*}] = \Theta(1/\sqrt{\tilde{n}})$.) We lower bound the expected core size as follows. Let $X_j = \epsilon_j^k - \epsilon_j^{k_*}$. Let $\mathcal{B}$ be the event that exactly one of the X_j's is in $[-1, -1 + 1/\sqrt{\tilde{n}}]$, and no X_j is in $[-1 + 1/\sqrt{\tilde{n}}, -1 + 2/\sqrt{\tilde{n}}]$. Under f the probability of being in $[-1, -1 + 1/\sqrt{\tilde{n}}]$ is $1/(2\tilde{n})$ and the probability of being in $[-1 + 1/\sqrt{\tilde{n}}, -1 + 2/\sqrt{\tilde{n}}]$ is $3/(2\tilde{n})$. It follows that

$$\Pr(\mathcal{B}) = \binom{\tilde{n}+1}{1, 0, \tilde{n}} \frac{1}{2\tilde{n}} \big(1 - 2/\tilde{n}\big)^{\tilde{n}} = \Omega(1)\,. \tag{33}$$

Claim 2. *Consider the case of $K = 2$. Under event $\mathcal{B}$, for any core vector (α_{k_*}, α_k), for any value $\alpha'_k \in [\alpha_{k_*} + 1 - 2/\sqrt{\tilde{n}}, \alpha_{k_*} + 1 - 1/\sqrt{\tilde{n}}]$, we have that vector $(\alpha_{k_*}, \alpha'_k)$ is in the core. In particular, $\mathcal{C} = \Omega(1/\sqrt{\tilde{n}})$.*

Proof. Eq. (31) is satisfied since event $\mathcal{B}$ holds. Since, $\alpha_{k'}$ can take any value in an interval of length $1/\sqrt{\tilde{n}}$, it follows that $\mathcal{C} = \Omega(1/\sqrt{\tilde{n}})$ under $\mathcal{B}$. □

Combining Claim 2 with Eq. (33), we obtain that $\mathrm{E}[\mathcal{C}] = \Omega(1/\sqrt{\tilde{n}})$, as desired.

We now construct a similar argument for $K > 2$, with $\mathcal{K} = \mathcal{T}_{\mathcal{L}} \backslash \{k_*\}$ being the other labor types, all of whose agents are matched. It again turns out that $\alpha_{k_*} \xrightarrow{\tilde{n}\to\infty} 1$ in probability, and when we use this together with Eq. (30) we obtain $\alpha_k \xrightarrow{\tilde{n}\to\infty} 2\ \forall k \in \mathcal{K}$ in probability (but we do not prove or use these limits).

Considering only the dimensions in $\mathcal{K}$ (recall $|\mathcal{K}| = K - 1$ here) of each ϵ_j, let $\mathcal{B}_3$ be the event as defined in Lemma 4 with $\delta = 1/n^{0.51}$.

Claim 3. *Let $\underline{k} = \arg\min_{k\in\mathcal{K}} \alpha_k$ and let $\bar{k} = \arg\max_{k\in\mathcal{K}} \alpha_k$. Under event $\mathcal{B}_3$, we claim that*

$$\alpha_{\bar{k}} - \alpha_{\underline{k}} \leq \delta. \tag{34}$$

Proof. From Proposition 2, we know that the set of core α's is a linear polytope, and hence it is immediate to see that the set of θ's is an interval. Let $\underline{k} = \arg\min_{k\in\mathcal{K}} \alpha_k$ and let $\bar{k} = \arg\max_{k\in\mathcal{K}} \alpha_k$. Under event $\mathcal{B}_3$, we claim that $\alpha_{\bar{k}} - \alpha_{\underline{k}} \leq \delta$. We can argue this by contradiction. Suppose that $\alpha_{\bar{k}} - \alpha_{\underline{k}} > \delta$. One can see that all j's such that $\epsilon_j^{\mathcal{K}} \in \widehat{\mathcal{R}}^{\bar{k},\underline{k}}(\delta)$, cf. (18), will be matched to type $\bar{k}$, with the possible exception of j_*. Thus, under $\mathcal{B}_3$, the number of employers matched to type $\bar{k}$ is bounded below by

$$n^{\bar{k},\underline{k}} - 1 \geq ((K-1)\tilde{n} + 1)/(K-1) > \tilde{n}\,,$$

which is a contradiction, implying (34). □

The above claim bounds the maximum difference between α's corresponding to any pair of types in $\mathcal{K}$. Intuitively, note that all types in $\mathcal{K}$ have the same u and therefore the same distribution for the θ variables of the agents in such a type. Moreover, all types in $\mathcal{K}$ have the same number of agents. Hence, one would expect the α's to be equal. While true in the limit, for each finite n we need to account for the stochastic fluctuations in a given realization. Therefore, we can show that no pair of α's in $\mathcal{K}$ can differ by more than δ. The next claim follows immediately from Claim 3.

Claim 4. *Let $k \in \mathcal{K}$ be an arbitrary type. Under event $\mathcal{B}_3$, we claim that*

$$\max_{k'\in\mathcal{K}} \left(\epsilon_j^{k'} - \epsilon_j^{k}\right) \leq \delta \quad \forall j \in M(k). \tag{35}$$

Proof. By Claim 3, we have that under $\mathcal{B}_3$, $|\alpha_k - \alpha_{k'}| \leq \delta$ for all $k' \in cK$. By the stability condition in Eq. (30), we have

$$\delta \geq \alpha_k - \alpha_{k'} \geq \epsilon_j^{k'} - \epsilon_j^k \quad \forall j \in M(k),\ \forall k' \in \mathcal{K}.$$

Therefore, for every $j \in M(k)$ we must have $\delta \geq \max_{k' \in \mathcal{K}} \epsilon_j^{k'} - \epsilon_j^k$ as desired. □

Next, we focus on the stability conditions involving type k_*. For each $k \in \mathcal{K}$, the stability condition is:

$$\epsilon_{j_*}^{k_*} - \epsilon_{j_*}^{k} \geq \alpha_k - \alpha_{k_*} \geq \max_{j \in M(k)} \epsilon_j^{k_*} - \epsilon_j^k, \tag{36}$$

where j_* is the employer matched to i_*. For each $j \in \mathcal{E}$, let X_j be defined as $X_j = (\max_{k \in \mathcal{K}} \epsilon_j^k) - \epsilon_j^{k_*}$. The X_j are distributed i.i.d. with cumulative distribution $F(-1+\theta) = \theta^K/K$ for $\theta \in [0,1]$ (we shall not be concerned with the cumulative distribution for positive values). Let $\mathcal{B}$ be the event that exactly one of the X_j's is in $[-1, -1+1/\tilde{n}^{1/K}]$ (this will be X_{j_*}), and no X_j is in $[-1+1/\tilde{n}^{1/K}, -1+2/\tilde{n}^{1/K}]$. Under cumulative F, the probability of being in $[-1,-1+1/\tilde{n}^{1/K}]$ is $1/(K\tilde{n})$ and the probability of being in $[-1+1/\tilde{n}^{1/K}, -1+2/\tilde{n}^{1/K}]$ is $2^K/(K\tilde{n})$. It follows that

$$\Pr(\mathcal{B}) = \binom{\tilde{n}+1}{1, 0, \tilde{n}} \frac{1}{K\tilde{n}} \left(1 - 2^K/(K\tilde{n})\right)^{\tilde{n}} = \Omega(1)\,. \tag{37}$$

Clearly, under $\mathcal{B}$, we must have $j_* = \arg\min_{j \in \mathcal{E}} X_j$. Keeping this in mind, we state and prove our last claim.

Claim 5. *Suppose that $\mathcal{B}_3 \cap \mathcal{B}$ occurs. Take any core vector $(\alpha_{k_*}, (\alpha_k)_{k \in \mathcal{K}})$. Then*

$$\{\theta \in \mathbb{R} : (\alpha_{k_*}, (\alpha_k + \theta)_{k \in \mathcal{K}}) \text{ is in the core}\} \tag{38}$$

is an interval of length at least $1/n^{1/K} - 2\delta = \Omega(1/n^{1/K})$. In particular, $\mathcal{C} \geq \Omega(1/n^{1/K})$.

Proof. Define

$$\underline{\theta} = 1 - 2/\tilde{n}^{1/K} + \delta - \alpha_{\underline{k}} + \alpha_{k_*}$$
$$\overline{\theta} = 1 - 1/\tilde{n}^{1/K} - \alpha_{\bar{k}} + \alpha_{k_*}.$$

We claim that, under $\mathcal{B}_3 \cap \mathcal{B}$, we have that $\alpha(\theta) = (\alpha_{k_*}, (\alpha_k + \theta)_{k\in\mathcal{K}})$ is in the core for all $\theta \in [\underline{\theta}, \overline{\theta}]$. To establish this, we need to show that conditions (28) and (30) are satisfied. Since α belongs to the core, we immediately infer that (28) holds, and also (30) when $k_* \notin \{k, k'\}$ by the definition of $\alpha(\theta)$. That leaves us with (36). Now, for any $k \in \mathcal{K}$ and $\theta \in [\underline{\theta}, \overline{\theta}]$ we have

$$\alpha_k(\theta) = \alpha_k + \theta \leq \alpha_{\bar{k}} + \theta \leq \alpha_{\bar{k}} + \overline{\theta} = 1 - 1/\tilde{n}^{1/K} + \alpha_{k_*} \leq \epsilon_{j_*}^{k_*} - \epsilon_{j_*}^{k} + \alpha_{k_*} \,,$$

where we use the definitions of $\bar{k}$ and $\overline{\theta}$, and the fact that $\mathcal{B}$ occurs (hence $1-1/\tilde{n}^{1/K} \leq \epsilon_{j_*}^{k_*} - \epsilon_{j_*}^{k}$). This establishes the left inequality in (36). Similarly, for any $k \in \mathcal{K}$ we have

$$\begin{aligned}\alpha_k(\theta) &= \alpha_k + \theta \geq \alpha_{\underline{k}} + \theta \geq \alpha_{\underline{k}} + \underline{\theta} = 1 - 2/\tilde{n}^{1/K} + \delta + \alpha_{k_*} \\ &\geq \epsilon_j^{k_*} - \max_{k'\in\mathcal{K}} \epsilon_j^{k'} + \delta + \alpha_{k_*} \geq \epsilon_j^{k_*} - \epsilon_j^{k} + \alpha_{k_*} \qquad \forall j \in M(k)\,,\end{aligned}$$

where we use the definitions of $\underline{k}$ and $\underline{\theta}$ for the first two inqualities, and the fact that $\mathcal{B}$ occurs (so $1-2/\tilde{n}^{1/K} \geq \epsilon_j^{k_*} - \max_{k'\in\mathcal{K}} \epsilon_j^{k'}$, $\forall j \in M(k)$). Finally, the last inequality follows from $\mathcal{B}_3$ and Claim 4 (which implies that $-\max_{k'\in\mathcal{K}} \epsilon_j^{k'} + \delta \geq -\epsilon_j^{k}$ for $j \in M(k)$). This establishes the right inequality in (36). Thus, we have shown that $\alpha(\theta)$ is in the core for all $\theta \in [\underline{\theta}, \overline{\theta}]$. We have that the length of this interval is $1/\tilde{n}^{1/K} - (\alpha_{\bar{k}} - \alpha_{\underline{k}}) - \delta \geq 1/\tilde{n}^{1/K} - 2\delta = \Omega(1/\tilde{n}^{1/K})$, using (34). Therefore, $\mathrm{E}[\mathcal{C}] = \Omega(1/\tilde{n}^{1/K})$ under $\mathcal{B}_3 \cap \mathcal{B}$. □

Using Lemma 4 and Eq. (37) we have

$$\Pr(\mathcal{B}_3 \cap \mathcal{B}) = \Omega(1)\,.$$

Combining this with the claim above we obtain that $\mathrm{E}[\mathcal{C}] = \Omega(1/n^{1/K})$. □

E Proof of Theorem 2

We start by restating Theorem 2 and discussing the structure of the proof. Recall that we use the characterization of the size of the core from Lemma 1.

Theorem (Restatement of Theorem 2)**.** *Let idiosyncratic productivities be drawn i.i.d. from any fixed distribution F that is supported in an interval of the form $[0, C_u]$ or $[0, \infty)$ where $C_u < \infty$, and whose density f is strictly positive and continuous everywhere in the support.*[4] *In addition, suppose that $u(k, 1) \geq 0$ for all $k \in \mathcal{T}_\mathcal{L}$. Consider the setting in which $K \geq 2$, $Q = 1$, $n_\mathcal{E} > n_\mathcal{L}$ and let $m = n_\mathcal{E} - n_\mathcal{L}$. Under Assumption 2, we have, with high probability, that $\mathcal{C} \leq O^*\left(\frac{1}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}\right)$. Also, $E[\mathcal{C}] \leq O^*\left(\frac{1}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}\right)$.*

Note that Assumption 1 is automatically satisfied under the hypotheses of the theorem.

As in the proof of Theorem 1, we first prove the result for F being Uniform$(0, 1)$, and later extend our proof to more general F. The idea of the proof is as follows. First, we show a bound on the expectation of $\min_{k \in \mathcal{T}_\mathcal{L}} \{\alpha_k^{\max} - \alpha_k^{\min}\}$. In particular, we show that $\mathrm{E}\left[\min_{k \in \mathcal{T}_\mathcal{L}} \{\alpha_k^{\max} - \alpha_k^{\min}\}\right] = O^*\left(\frac{1}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}\right)$. To do so, we note that by condition (IM) in Proposition 2, we must have

$$\min_k \left(\alpha_k^{\max} - \alpha_k^{\min}\right) \leq \min_{k \in \mathcal{T}_\mathcal{L}} \left(\min_{j \in M(k)} \epsilon_j^k - \max_{j \in U} \epsilon_j^k\right).$$

Then, we consider two separate cases to prove the result, depending on the size of the imbalance. When $m \leq \log(n)$, the result is shown in Lemma 1, which we prove via an upper bound on $\min_{k \in \mathcal{T}_\mathcal{L}} \left(\min_{j \in M(k)} \epsilon_j^k\right)$. On the contrary, when $m \geq \log(n)$, the result is shown in Lemma 4. The proof of Lemma 4 relies mainly on the geometry of a core solution that (roughly) allows us to control the largest of the α's (all α's must be negative in the core since some employers are unmatched, and we control, roughly, the least negative α).

Next, we then show that, for every pair of types $k, q \in \mathcal{T}_\mathcal{L}$, we must have

$$\mathrm{E}\left[\min_{j \in M(k)} (\epsilon_j^k - \epsilon_j^q) - \max_{j \in M(q)} (\epsilon_j^k - \epsilon_j^q)\right] = O^*\left(\frac{1}{n}\right).$$

[4] If the lower limit of the support of F is positive, this can be absorbed into the u's, and hence this case is covered.

By condition (ST) in Proposition 2, this implies that for fixed $k, q \in \mathcal{T}_{\mathcal{L}}$, the expected maximum variation in $\alpha_k - \alpha_q$ in the core is bounded by $O^*\left(\frac{1}{n}\right)$.

Finally, we use the bounds in the first two steps to argue that, for every type $k \in \mathcal{T}_{\mathcal{L}}$,

$$\mathrm{E}\left[\alpha_k^{\max} - \alpha_k^{\min}\}\right] = O^*\left(\frac{1}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}\right),$$

which implies $\mathrm{E}[\mathcal{C}] = O^*\left(\frac{1}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}\right)$. This is done in the proof of Theorem 2.

We now show our bound on $\mathrm{E}\left[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}\right]$. To that end, let $Z_k = \min_{j\in M(k)} \epsilon_j^k$ and $U_k = \max_{j\in U} \epsilon_j^k$. By condition (IM) in Proposition 2, $\mathrm{E}\left[\min_k |\alpha_k^{\max} - \alpha_k^{\min}|\right] \leq \mathrm{E}[\min_k\{Z_k - U_k\}]$, and therefore we shall focus on bounding $\mathrm{E}[\min_k\{Z_k - U_k\}]$. As a reminder, we have defined $m = n_{\mathcal{E}} - n_{\mathcal{L}}$ and $\delta_n = \frac{\log(n)}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}$. Also, in all lemmas we are working under the assumptions of the theorem, that is, $K \geq 2, Q = 1$, $n_{\mathcal{E}} > n_{\mathcal{L}}$ and Assumption 2.

Lemma 1. *Suppose that $m \leq 6K\log(n_{\mathcal{E}})$. Then, there exists a constant $C_3 = C_3(K) < \infty$ such that $E\left[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}\right] \leq 2C_3 \frac{\log(n)}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}$.*

Proof. Let $Z_k = \min_{j\in M(k)} \epsilon_j^k$, $U_k = \max_{j\in U} \epsilon_j^k$, and $\delta_n = \frac{\log(n)}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}$. By condition (IM) in Proposition 2, $\mathrm{E}[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}] \leq \mathrm{E}[\min_k\{Z_k - U_k\}]$. As U_k is a nonnegative random variable, we have $\mathrm{E}[\min_k\{Z_k - U_k\}] \leq \mathrm{E}[\min_k\{Z_k\}]$. Therefore, we have

$$\mathrm{E}[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}] \leq \mathrm{E}\left[\min_k\{Z_k - U_k\}\right] \leq \mathrm{E}\left[\min_k\{Z_k\}\right] \leq C_3\delta_n + \Pr\left(\min_k Z_k \geq C_3\delta_n\right),$$

using $Z_k \leq 1$.

To finish the proof, it suffices to show that $\Pr(\min_k Z_k \geq C_3\delta_n) \leq C_3\delta_n$. Hence, our next step is to bound $\Pr(\min_k Z_k \geq C_3\delta_n)$. Now $\min_k Z_k \geq C_3\delta_n$ implies that all j such that $\epsilon_j \in [0, C_3\delta_n]^K$ are unmatched. But there are only m unmatched employers. It follows that

$$\Pr\left(\min_k Z_k \geq C_3\delta_n\right) \quad \leq \quad \Pr\left(\text{at most } m \text{ points in the hypercube } [0, C_3\delta_n]^K\right).$$

Let $X \sim \text{Bin}\left(n_{\mathcal{E}}, (C_3\delta_n)^K\right)$ be defined as the number of points, out of $n_{\mathcal{E}}$ in total, that fall in the hypercube $[0, C_3\delta_n]^K$. By assumption, $m \leq 6K\log(n)$ and thus $(C_3\delta_n)^K \geq (C_3 \log n/m)^K/n \geq 2^K \log n^K/n \geq 4(\log n)^2/n$ by defining $C_3 \geq 12K$ and using $K \geq 2$. Furthermore, using $n \leq 2n_{\mathcal{E}}$ we obtain $\text{E}[X] = n_{\mathcal{E}}\,(C_3\delta_n)^K \geq (n/2)4(\log n)^2/n = 2(\log n)^2$. It follows that

$$\Pr\left(\min_k Z_k \geq C_3\delta_n\right) \leq \Pr\left(X \leq 6K\log(n)\right) \leq \exp(-\Omega((\log n)^2)) \leq \frac{1}{n} \leq C_3\delta_n.$$

where the second inequality is obtained by applying the Chernoff bound. Hence, we have shown that

$$\text{E}[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}] \leq \text{E}\left[\min_k\{Z_k\}\right] \leq C_3\delta_n + \Pr\left(\min_k Z_k \geq C_3\delta_n\right) \leq 2C_3\delta_n,$$

which completes the proof. □

We now establish an upper bound for the case in which $m \geq 6K\log(n_{\mathcal{E}})$. For the following results up to Lemma 4 we shall assume that $m \geq 6K\log(n_{\mathcal{E}})$.

Before we move on, we briefly give some geometric intuition regarding the problem. For each agent $j \in \mathcal{E}$, let $\epsilon_j = (\epsilon_j^1, \ldots, \epsilon_j^K)$ denote the profile of values assigned by the K types of agents in $\mathcal{L}$ to agent j. Given our stochastic assumptions, all points ϵ_j will be distributed in the $[0,1]^K$ hypercube. Using Proposition 2, we can partition the $[0,1]^K$-hypercube into $K+1$ disjoint regions: K of them containing the n_k points corresponding to agents matched to type k ($1 \leq k \leq K$), and one region containing all unmatched agents. Furthermore, the region containing the unmatched agents is an orthotope[5] that has the origin as a vertex. This follows for the (IM) constraints in Proposition 2.

To establish an upper bound for the case in which m is greater than or equal to $6K\log(n_{\mathcal{E}})$, let $\mathcal{O}$ be the set of K-orthotopes contained in $[0,1]^K$ that have the origin as a vertex. Suppose that R is expanded by the same amount θ in each coordinate direction. Define $D(R)$ as the smallest value of θ such that an additional point ϵ_j is contained in the expanded orthotope. (If one of the side lengths becomes 1 before an additional point is reached, then define $D(R) = 0$. This will

[5] An orthotope (also called a hyperrectangle or a box) is the generalization of a rectangle for higher dimensions

never occur for R that contains only the unmatched agents.) As usual, let $Z_k = \min_{j \in M(k)} \epsilon_j^k$ and $U_k = \max_{j \in U} \epsilon_j^k$. We want to show that $\mathrm{E}\left[\min_k\{Z_k - U_k\}\right] \leq C_5 \delta_n$, for some constant $C_5 = C_5(K) < \infty$. To that end, note that $\min_k\{Z_k - U_k\}$ is equal to $D(R)$ for some orthotope $R \in \mathcal{O}$. In particular, $\min_k\{Z_k - U_k\}$ is equal to $D(R)$ when R is the orthotope that "tightly" contains all the m points in U.

For $R \in \mathcal{O}$, let $V(R)$ be defined as the volume of R. In addition, we define $|R|$ to be the number of points contained in R. We start by showing that, given that $m \geq 6K \log(n)$, an orthotope in $\mathcal{O}$ of volume less than $\frac{m}{4n_{\mathcal{E}}}$ in extremely unlikely to contain m points.

Lemma 2. *Suppose that $m \geq 6K\log(n)$. For $R \in \mathcal{O}$ such that $V(R) < \frac{m}{4n_{\mathcal{E}}}$, we have $\Pr\left(|R| = m\right) \leq \frac{1}{n^{K+1}}$, where $V(R)$ denotes the volume and $|R|$ denotes the number of points in R.*

Proof. Let X denote the number of points in an orthotope in $\mathcal{O}$ of volume $\frac{m}{4n_{\mathcal{E}}}$. Then, $X \sim \mathrm{Bin}\left(n_{\mathcal{E}}, \frac{m}{4n_{\mathcal{E}}}\right)$. We have $\mu = \mathrm{E}[X] = m/4$. Using a Chernoff bound, we have,

$$\Pr(X \geq m) = \Pr(X \geq 4\mu) \leq (e^3/4^4)^{m/4} \leq \exp(-m/4).$$

Now $m/4 \geq 6K \log n/4 \geq (K+1)\log n$, using $K \geq 2$. Substituting back we obtain $\Pr(X \geq m) \leq \exp(-(K+1)\log n) = 1/n^{K+1}$. But $|X|$ stochastically dominates $|R|$ since $V(R) < \frac{m}{4n_{\mathcal{E}}}$. The result follows. □

Our next step will be to bound $\Pr\left(D(R) > C_4\delta \,\middle|\, E\right)$, for $R \in \mathcal{O}$ and some constant $C_4 = C_4(K) < \infty$, where E is the event defined as $E = \{|R| = m,\ V(R) \geq \frac{m}{4n_{\mathcal{E}}}\}$.

Lemma 3. *There exists some constant $C_4 = C_4(K) < \infty$ such that, for all $R \in \mathcal{O}$ with $V(R) \geq \frac{m}{4n_{\mathcal{E}}}$, we have that $P\left(D(R) > C_4\delta_n \,\middle|\, |R| = m\right) \leq \frac{1}{n^{K+1}}$, where $\delta_n = \frac{\log(n)}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}$.*

Proof. Conditioned on $|R| = m$, the remaining $n_{\mathcal{L}} = n_{\mathcal{E}} - m$ points are distributed uniformly i.i.d. in the complementary region of volume $(1 - V(R))$.

Let $F_{C_4\delta_n}$ denote the region swept when R is expanded by $C_4\delta_n$ along each coordinate axis. Clearly, $D(R) > C_4\delta_n$ if and only if region $F_{C_4\delta_n}$ contains no points.

Let X denote the number of points in $F_{C_4\delta_n}$, and let p denote the volume of $F_{C_4\delta_n}$. Then, $X \sim \text{Bin}(n_{\mathcal{L}}, p/(1-V(R)))$ and hence stochastically dominates $\text{Bin}(n_{\mathcal{L}}, p))$. Note that such a volume p is at least the volume obtained when expanding the hypercube of side $\ell = \sqrt[K]{\frac{m}{4n\varepsilon}}$ by $C_4\delta_n$ along each direction and, therefore, $p \geq K\ell^{(K-1)}C_4\delta_n$. Hence,

$$
\begin{aligned}
P(D(R) > C_4\delta_n) &= \Pr(X = 0) \leq (1-p)^{\mathcal{L}} \leq \exp\left\{-\Omega(np)\right\} \\
\leq \exp\{-\Omega(n(\frac{m}{n})^{(K-1)/K}C_4\delta_n)\} &= \exp\{-\Omega(C_4 \log n)\} \leq \frac{1}{n^{K+1}} \,,
\end{aligned}
$$

for appropriate C_4, where we have used Assumption 2.

□

Lemma 4. *Suppose that $m \geq 6K\log(n)$. Then, there exists a constant $C_5 = C_5(K) < \infty$, such that $E\left[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}\right] \leq C_5 \frac{\log(n)}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}$.*

Proof. Let $Z_k = \min_{j \in M(k)} \epsilon_j^k$, $U_k = \max_{j \in U} \epsilon_j^k$, and $\delta_n = \frac{\log(n)}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}$. By condition (IM) in Proposition 2, we know that $\alpha_k^{\max} - \alpha_k^{\min} \leq Z_k - U_k$. Then,

$$
\mathrm{E}\left[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}\right] \leq \mathrm{E}\left[\min_k\{Z_k - U_k\}\right].
$$

In addition, $\min_k\{Z_k - U_k\}$ is equal to $D(R)$ for some orthotope $R \in \mathcal{O}$. In particular, $\min_k\{Z_k - U_k\}$ is equal to $D(R)$ when R is the orthotope that "tightly" contains all the m points in U. Define $\mathcal{R} = \{R \in \mathcal{O} : \; |R| = m\}$. Then,

$$
\mathrm{E}\left[\min_k\{Z_k - U_k\}\right] \;\leq\; \mathrm{E}\left[\max_{R \in \mathcal{R}}\left\{\mathrm{D(R)}\right\}\right].
$$

To bound $\mathrm{E}\left[\max_{R\in\mathcal{R}}\left\{\mathrm{D(R)}\right\}\right]$, consider the grid that results from dividing each of the K coordinate axes in the hypercube into intervals of length $1/n$. Let Δ denote that grid. Suppose that we just consider orthotopes in the grid, that is, the orthotopes whose sides are multiples of

$\frac{1}{n}$. Let $\mathcal{R}_\Delta = \{R \in \mathcal{R} :\ R \in \Delta\}$. Then,

$$\max_{R\in\mathcal{R}} \left\{\mathrm{D(R)}\right\} \leq \max_{R\in\mathcal{R}_\Delta} \left\{\mathrm{D(R)}\right\} + \frac{1}{n},$$

and,

$$\mathrm{E}\left[\max_{R\in\mathcal{R}} \left\{\mathrm{D(R)}\right\}\right] \leq \mathrm{E}\left[\max_{R\in\mathcal{R}_\Delta} \left\{\mathrm{D(R)}\right\}\right] + \frac{1}{n}.$$

Hence, we just need a bound for $\mathrm{E}\left[\max_{R\in\mathcal{R}_\Delta} \left\{\mathrm{D(R)}\right\}\right]$. Let $V_* = \frac{m}{4n}$. Note that $\mathrm{D(R)} \leq 1$ for all $R \in \mathcal{O}$ and, therefore,

$$\mathrm{E}\left[\max_{R\in\mathcal{R}_\Delta} \left\{\mathrm{D(R)}\right\}\right] \quad \leq \quad \mathrm{E}\left[\max_{R\in\mathcal{R}'_\Delta} \left\{\mathrm{D(R)}\right\}\right] + \Pr\left(\min_{R\in\mathcal{R}_\Delta} V(R) < V_*\right)$$

where $\mathcal{R}'_\Delta = \{R \in \mathcal{R}_\Delta : V(R) \geq V_*\}$. Now, by union bound

$$\Pr\left(\min_{R\in\mathcal{R}_\Delta} V(R) < V_*\right) \leq \sum_{R\in\Delta:V(R)<V_*} \Pr(|R| = m) \leq n^K \cdot 1/n^{K+1} = 1/n\,,$$

using $|\{R \in \Delta : V(R) < V_*\}| \leq |\{R \in \Delta\}| = n^K$ and Lemma 2.

Furthermore,

$$\mathrm{E}\left[\max_{R\in\mathcal{R}'_\Delta} \left\{\mathrm{D(R)}\right\}\right] \leq \mathrm{E}\left[\max_{R\in\Delta:V(R)\geq V_*} \left\{\mathrm{D(R)}\mathbb{I}(|R| = m)\right\}\right]$$

Now we have

$$\begin{aligned}
&\Pr\left[\max_{R\in\Delta:V(R)\geq V_*} \left\{\mathrm{D(R)}\mathbb{I}(|R| = m)\right\} > C_4\delta_n\right] \\
\leq &\sum_{R\in\Delta:V(R)\geq V_*} \Pr(|R| = m) \Pr[\mathrm{D(R)} > C_4\delta_n \,|\, |R| = m] \\
\leq &\sum_{R\in\Delta:V(R)\geq V_*} 1 \cdot 1/n^{K+1} \leq n^K/n^{K+1} = 1/n,
\end{aligned}$$

using a union bound and Lemma 3 to bound the probability of $D(R) \geq C_4\delta_n$. It follows that

$$\mathrm{E}\left[\max_{R\in\mathcal{R}'_\Delta} \left\{\mathrm{D(R)}\right\}\right] \leq \ 1 \cdot \Pr\left[\max_{R\in\Delta:V(R)\geq V_*} \left\{\mathrm{D(R)}\mathbb{I}(|R| = m)\right\} > C_4\delta_n\right] + C_4\delta_n = 1/n + C_4\delta_n.$$

Substituting the individual bounds back, we obtain

$$\mathrm{E}\left[\max_{R\in\mathcal{R}}\{\mathrm{D(R)}\}\right] = C_4\delta_n + 2/n \le C_5\delta_n\,,$$

by defining $C_5 = C_4 + 2$ and using $1/n \le \delta_n$.

Overall,

$$\mathrm{E}\left[\min_k\{\alpha_k^{\max} - \alpha_k^{\min}\}\right] \;\le\; \mathrm{E}\left[\min_k\{Z_k - U_k\}\right] \;\le\; \mathrm{E}\left[\max_{R\in\mathcal{R}}\{\mathrm{D(R)}\}\right] \le C_5\delta_n\,,$$

as claimed. □

We now proceed to show that, for every pair of types $k, q \in \mathcal{T}_{\mathcal{L}}$, we have

$$\mathrm{E}\left[\min_{j\in M(k)}(\epsilon_j^k - \epsilon_j^q) - \max_{j\in M(q)}(\epsilon_j^k - \epsilon_j^q)\right] \le C_2\frac{\log(n_{\mathcal{E}})}{n_{\mathcal{E}}},$$

for appropriate $C_2 = C_2(K) < \infty$. This result is shown in Lemma 7. Along the way, we establish a couple of intermediate results.

Let $Z_k = \min_{j\in M(k)}\epsilon_j^k$ and $U_k = \max_{j\in U}\epsilon_j^k$. Note that Z_k is an upper bound for $-\alpha_k$. By the definition of Z_k, all the points corresponding to agents in $M(k)$ must be contained in the orthotope $[1 - Z_k, 1] \times [0,1]^{K-1}$. The following proposition establishes that Z_k cannot be arbitrarily close to 1.

Lemma 5. *Given a constant $c \in \mathbb{R}$, let the event E_c be defined as $E_c = \{\max_k \min_{j\in M(k)}\epsilon_j^k \le 1 - c\}$. Then, there exist constants $\theta = \theta(K) > 0$ and $C_6 = C_6(K) > 0$ such that, for large enough n, E_θ occurs with probability at least $1 - \exp(-C_6 n)$.*

Proof. Let $Z_k = \min_{j\in M(k)}\epsilon_j^k$. The proof follows from the previous observation that all the points corresponding to agents in $M(k)$ must be contained in the orthotope of volume $(1 - Z_k)$. Let $C < \infty$ be such that $\dfrac{n_{\mathcal{E}}}{n_{\mathcal{L}}} \le C$. By Assumption 2, such a C must exist. Furthermore, by

Assumption 2, there must exist $C_K \in \mathbb{R}$ such that $n_k \geq C_K n$ for all $k \in \mathcal{T}_{\mathcal{L}}$. Let $n_{\mathcal{E}}$ be the total number of points in the cube $[0,1]^K$. Let X denote the number of points out of the $n_{\mathcal{E}}$ ones that fall in the rectangle defined by $[1-\theta,1] \times [0,1]^{K-1}$. Then, $X \sim \text{Bin}(n_{\mathcal{E}}, \theta)$. Suppose that we set $\theta < \frac{C_K}{2C}$. Then, for large enough n and appropriate $C_6 > 0$, we have

$$\Pr(Z_k > 1-\theta) \leq \Pr(X \geq C_K n_{\mathcal{L}}) \leq \Pr\left(X \geq \frac{C_K n_{\mathcal{E}}}{C}\right) \leq \exp\left(-2C_6 n\right) \leq (1/K)\exp\left(-C_6 n\right),$$

where we have used a Chernoff bound, $2n_{\mathcal{E}} \geq n$, and $\exp(-C_6 n) \leq (1/K)$ for large enough n. The result follows from a union bound over possible k. □

Remark 1. *Let θ, E_θ, and C_6 be as defined in the statement of Lemma 5. Define $G_{k,q}$ as*

$$G_{k,q} = \left\{x \in [0,1]^K : \left(x_k \geq 1 - \frac{\theta}{2} \text{ or } x_q \geq 1 - \frac{\theta}{2}\right) \text{ and } x_r < \frac{\theta}{2} \text{ for all } 1 \leq r \leq K,\ r \neq k,q\right\}.$$

Under event E_θ, we must have $G_{k,q} \subseteq M(k) \cup M(q)$.

The above remark follows from Lemma 5 and the definition of $G_{k,q}$. If $j : \epsilon_j \in G_{k,q}$ were matched to a type $k' \notin \{k,q\}$, that would contradict the maximality of the matching as, by swapping the matches of $j' : j' \in M(k), \epsilon^k_{j'} = Z_k$ and j, the overall weight of the matching strictly increases. A similar argument rules out j being unmatched.

Lemma 6. *Let $G_{k,q}$ be as in Remark 1, and let θ be as defined in Lemma 5. Define $G'_{k,q}$ as follows:*

$$G'_{k,q} = G_{k,q} \cap \{x \in [0,1]^K, |x_k - x_q| \leq 1-\theta\}$$

Let $\mathcal{V}^{kq} = \{x : x = \epsilon^k_j - \epsilon^q_j,\ \epsilon_j \in G'_{k,q}\}$, and let

$$V^{kq} = \max(\textit{Difference between consecutive values in } \mathcal{V}^{kq} \cup \{-1+\theta, 1-\theta\}).$$

Then, there exists a function $f(n) = O^*(1/n)$ such that $\Pr\left(\overline{\mathcal{B}^{kq}}\right) \leq 1/n$ where $\mathcal{B}^{kq}$ is the event *that* $V^{kq} \leq f(n)$.

The proof of Lemma 6 is omitted as the required analysis is similar to (and much simpler than) that leading to Lemma 1. Essentially, V^{kq} consists of values taken by $\Theta(n)$ points distributed uniformly and independently in $[-1+\theta, 1-\theta]$, and so, with high probability, no two consecutive values are separated by more than $f(n) = O(\log n/n)$.

In the next lemma we bound the difference between every pair of α's.

Lemma 7. *Consider types* $k, q \in \mathcal{T}_{\mathcal{L}}$ *and let* f *be as defined in the statement of Lemma 6. Under event* $E_\theta \cap \mathcal{B}^{kq}$*, in every stable solution we must have that* $(\alpha_q^{\max} - \alpha_q^{\min}) \leq 2f(n) + (\alpha_k^{\max} - \alpha_k^{\min})$.

Proof. We claim that under E_θ, we must have that $\alpha_q - \alpha_k$ varies within a range of no more than V^{kq} within the core, where V^{kq} is as defined in the statement of Lemma 6. By Remark 1, under event E_θ we must have $G'_{kq} \subset M(k) \cup M(q)$, where G'_{kq} is as defined in the statement of Lemma 6. Suppose that G'_{kq} contains at least one vertex matched to type k and one to type q. Then, by condition (ST) in Proposition 2 we must have:

$$\begin{aligned}
(\alpha_q - \alpha_k)^{\max} - (\alpha_q - \alpha_k)^{\min} &\leq \min_{j \in M(k)} \{\epsilon_j^k - \epsilon_j^q\} - \max_{j \in M(q)} \{\epsilon_j^k - \epsilon_j^q\} \\
&\leq \min_{j \in M(k) \cap G'_{k,q}} \{\epsilon_j^k - \epsilon_j^q\} - \max_{j \in M(q) \cap G'_{k,q}} \{\epsilon_j^k - \epsilon_j^q\} \\
&\leq V^{kq}.
\end{aligned}$$

Next, consider the case in which all vertices in G'_{kq} are matched to type k (the analogous argument follows if they are all matched to type q). Under event E_θ, by Condition (IM) in Proposition 2 we must have $0 \leq -\alpha_k \leq 1-\theta$ and $0 \leq -\alpha_q \leq 1-\theta$. Therefore, $\alpha_q - \alpha_k \in [-1+\theta, 1-\theta]$. In addition, by Condition (ST) in Proposition 2 we must have $\alpha_q - \alpha_k \leq$

$\min_{j\in M(k)}\{\epsilon_j^k - \epsilon_j^q\}$. However,

$$\begin{aligned}(\alpha_q-\alpha_k)^{\max}-(\alpha_q-\alpha_k)^{\min} &\le \min_{j\in M(k)}\{\epsilon_j^k-\epsilon_j^q\}-(-1+\theta)\\ &\le \min_{j\in M(k)\cap G'_{k,q}}\{\epsilon_j^k-\epsilon_j^q\}-(-1+\theta)\\ &= \min_{j\in G'_{k,q}}\{\epsilon_j^k-\epsilon_j^q\}-(-1+\theta)\\ &\le V^{kq}.\end{aligned}$$

It follows that $(\alpha_q^{\max}-\alpha_q^{\min}) \le 2V^{kq} + (\alpha_k^{\max}-\alpha_k^{\min})$. By definition, under $\mathcal{B}^{kq}$ we have $V^{kq} \le f(n)$, which completes the proof. □

Finally, we complete the last step of the proof by showing the main theorem.

Proof of Theorem 2. By definition, $\mathcal{C} = \sum_{k=1}^{K} \frac{N(k)|\alpha_k^{\max}-\alpha_k^{\min}|}{n_{\mathcal{L}}}$, where $N(k)$ is defined to be the number of agents of type k that are matched. For a given instance, let $k^* = \mathrm{argmin}_k\{\alpha_k^{\max}-\alpha_k^{\min}\}$. Let $\mathcal{B} = E_\theta \cap (\cap_{k,q}\mathcal{B}^{k,q})$. Note that using Lemmas 1 and 6 and a union bound, we obtain that

$$\Pr(\overline{\mathcal{B}}) \le \Pr(\overline{E_\theta}) + \sum_{k,q\in\mathcal{K}:k\neq q} \Pr(\overline{\mathcal{B}^{kq}}) = O(1/n)\,.$$

By Lemma 7, under $\mathcal{B}$, for every $k \in \mathcal{T}_{\mathcal{L}}$ we have

$$\alpha_k^{\max}-\alpha_k^{\min} \le 2f(n) + \alpha_{k^*}^{\max}-\alpha_{k^*}^{\min}.$$

Therefore,

$$\begin{aligned}\mathrm{E}[\mathcal{C}] &\le \mathrm{E}\left[\alpha_{k^*}^{\max}-\alpha_{k^*}^{\min}\right] + 2f(n) + \Pr(\overline{\mathcal{B}})\cdot O(1)\\ &\le O\left(\frac{\log(n)}{n^{\frac{1}{K}}m^{\frac{K-1}{K}}}\right) + O^*(1/n) + O(1/n)\\ &= O^*\left(\frac{1}{n^{\frac{1}{K}}m^{\frac{K-1}{K}}}\right)\end{aligned}$$

where the first inequality follows from the above together with using the upperbound of $O(1)$

for the core size; the second inequality is obtained by using the bound on $\mathrm{E}\left[\alpha_{k^*}^{\max} - \alpha_{k^*}^{\min}\right]$ from Lemma 1 for $m \le 6K\log(n)$ and Lemma 4 for $m \ge 6K\log(n)$, as well as the definition of $f(n)$ and $\Pr(\overline{\mathcal{B}}) = O(1/n)$ shown above. Finally, we note that each of the steps yields a bound that holds with high probability (instead of a bound on expected value) if we multiply by an additional $\log n$ factor. This yields that, with high probability,

$$\mathcal{C} \le O^*\left(\frac{1}{n^{\frac{1}{K}} m^{\frac{K-1}{K}}}\right).$$

This completes the proof for F being Uniform$(0,1)$.

Now consider general F satisfying the conditions stated in the theorem. We extend our proof just as we extended the proof of Theorem 1 to general F. Lemma 2 gives us a lower bound of $-U \le \alpha_{kq}$ for any core α, whereas we already know that $-C_u \ge \alpha_{kq} \le 0$ since some employers are unmatched and others are matched. Overall, we know $\alpha_{kq} \in [-\min(U, C_u), 0]$. We scale utilities down by a factor $\min(U, C_u)$. Now the density of firm productivities is uniformly lower-bounded in $[0,1]$, and firm productivity values in $[0,1]$ are the only relevant ones. Moreover, $\Omega(n)$ firms have all productivities lying in $[0,1]$. Hence, our bound on the gap between consecutive order statistics and their linear combinations in the unit hypercube implies the same bound (up to constant factors) for general F, and hence we get the same bound on core size for general F. □

F Proof of Theorem 3

In this appendix, we state and prove a more detailed version of Theorem 3.

Theorem 2. *Fix K and Q and a distribution F. Also fix the fraction ρ_k of each agent type k, where k can be a type of worker or a type of employer. We draw a market with n agents by independently drawing the type of each agent, and then independently drawing the idiosyncratic productivities from the distribution F. We obtain the following bounds as a function of n.*

- **Limit characterization of α.** *There exists $\alpha^* = \alpha^*(K, Q, \rho, F)$ such that as $n \to \infty$, we have that both $\alpha^{\max}$ and $\alpha^{\min}$ converge (in probability and almost surely) to α^*. In fact,* *for any $h_n = \omega(1)$ we have that with high probability, for every core outcome (M, α),*

$$\|\alpha - \alpha^*\| \leq h_n/\sqrt{n}\,. \tag{39}$$

- **Size of the core.** *There exists $C = C(K, Q, \rho, F) < \infty$ such that with high probability, the size of the core is bounded as*

$$\mathcal{C} \leq C \log n/n\,. \tag{40}$$

We formally establish the claims stated as part of the proof at the end of this appendix.

Proof of Theorem 2. We define a discrete tatonnement operator $\mathsf{T} : \mathbb{R}^{K+Q} \to \mathbb{R}^{K+Q}$ that takes a price vector α to an updated price vector $\mathsf{T}\alpha$ in a limit market (we do not formally define the limit market, but it guides our intuition). The operator T will be monotone, as a result of a gross substitutes property on both sides of the market. Its fixed point α^* will be turn out to be the limiting core solution. We also define a corresponding operator T_n for the n-th market to relate core outcomes α in that market to α^*. Our approach draws on the order-theoretic/lattice-theoretic approach that has yielded rich dividends in matching theory (Gale and Shapley 1962, Kelso and Crawford 1982, Hatfield and Milgrom 2005) and in the study of equilibrium more broadly.

We start by defining the "limit" operator T. The idea is that when T acts on a price vector α to produce an updated price vector $\mathsf{T}\alpha$, it computes each price $(\mathsf{T}\alpha)_{kq}$ as follows: treating all other prices as given, it computes the limiting "demand curve" $D_{kq}(p)$ of the number of type k workers who want to match to type q employers as a function of the price between type-pair (k, q) (here p is the dummy variable for this price), and the limiting "supply curve" $S_{kq}(p)$ of the number of type q employers who want to match to type k workers. Then $(\mathsf{T}\alpha)_{kq}$ is set to the value of p at which the demand and supply curves intersect. (There will be a unique such point, since $D_{kq}(p)$ will be strictly decreasing in p and $S_{kq}(p)$ will be strictly increasing in p.)

We define

$$D_{kq}(p;\alpha) = \rho_k \int_0^\infty f(z + p - u(k,q)) \prod_{q' \neq q} F(z + \alpha_{kq'} - u(k,q'))\ \mathrm{d}z \tag{41}$$

(where z serves as a dummy variable for the utility earned by an agent of type k, who finds type q at price p more attractive to match to than any other type q' at price $\alpha_{kq'}$) and

$$S_{kq}(p;\alpha) = \rho_q \int_0^\infty f(z - p) \prod_{k' \neq k} F(z - \alpha_{k'q})\ \mathrm{d}z\,. \tag{42}$$

We now suppress the α dependence, simply writing $D_{kq}(p)$ and $S_{kq}(p)$. It is easy to verify that $D_{kq}(\cdot)$ is strictly monotone decreasing and that $S_{kq}(\cdot)$ is strictly monotone increasing, given that $f(\cdot)$ is positive everywhere. Furthermore, both $D_{kq}(\cdot)$ and $S_{kq}(\cdot)$ are positive everywhere, and $\lim_{p\to\infty} D_{kq}(\cdot) = 0$ and $\lim_{p\to-\infty} S_{kq}(\cdot) = 0$. Combining, we deduce that there is a unique price where the demand curve $D_{kq}(\cdot)$ and the supply curve $S_{kq}(\cdot)$ intersect, and we define $(\mathsf{T}\alpha)_{kq}$ to be equal to this price.

Strict gross substitutes property on both sides of the limit market. Write $\alpha \preceq \alpha'$ if $\alpha_{kq} \leq \alpha'_{kq}$ for all $k \in \mathcal{T}_\mathcal{L}, q \in \mathcal{T}_\mathcal{E}$, and $\alpha \prec \alpha'$ if at least one of these inequalities is strict. Now, viewing $D_{kq}(p;\cdot)$ as a function of α, it is easy to verify that $D_{kq}(p;\cdot)$ is monotone increasing with respect to the defined ordering on α's (the demand for type q increases if the price of other employer types increases); i.e., if $\alpha \preceq \alpha'$, then $D_{kq}(p;\alpha) \leq D_{kq}(p;\alpha')$ for all p. In other words, the demand from each type of worker satisfies strict gross substitutes. Similarly, $S_{kq}(p;\cdot)$ is monotone decreasing by the strict gross substitutes property for each type of employer.

Thus, increasing α causes the demand curve to move up and the supply curve to move down, and hence it causes the price at which they intersect to move up. It follows that T is monotone increasing in α.

Fact 8. *If $\alpha \preceq \alpha'$ then $\mathsf{T}\alpha \preceq \mathsf{T}\alpha'$.*

We now show existence of a fixed point of T, thereby overcoming the challenge that the set of possible α's is not compact. To do this, we show that there is a (large) price P such that with $\alpha_P := (P, P, \ldots, P)$ we have $\mathsf{T}\alpha_P \preceq \alpha_P$, and another (large negative) price P' such that with

$\alpha_{P'} := (P', P', \ldots, P')$ we have $\mathsf{T}\alpha_{P'} \succeq \alpha_{P'}$. Then, starting with α_P and iteratively applying T, we get a fixed point α^*, using monotonicity.

Set all prices to be equal to P. Now P is the only variable. We have

$$D_{kq}(P; \alpha_P) = \rho_k \int_P^\infty f(w - u(k,q)) \prod_{q' \neq q} F(w - u(k,q'))\, \mathrm{d}w\,. \tag{43}$$

Note that $D_{kq}(P; \alpha_P)$ is positive everywhere, decreasing in P, and $\lim_{P\to\infty} D_{kq}(P; \alpha_P) = 0$. We also have that $\min_{k,q} D_{kq}(P; \alpha_P)$ is positive everywhere, decreasing in P, and

$$\lim_{P\to\infty} \min_{k,q} D_{kq}(P; \alpha_P) = 0\,.$$

Similarly,

$$S_{kq}(P; \alpha_P) = \rho_q \int_{-P}^\infty f(w) \prod_{k' \neq k} F(w)\, \mathrm{d}w\,. \tag{44}$$

Note that $S_{kq}(P; \alpha_P)$ is positive everywhere, increasing in P, and $\lim_{P\to-\infty} S_{kq}(P; \alpha_P) = 0$. We also have $\min_{k,q} S_{kq}(P; \alpha_P)$ is positive everywhere, increasing in P, and

$$\lim_{P\to-\infty} \min_{k,q} S_{kq}(P; \alpha_P) = 0\,.$$

It follows that for large enough P we have $\mathsf{T}\alpha_P \preceq \alpha_P$ and for small enough P' we have $\mathsf{T}\alpha_{P'} \succeq \alpha_{P'}$. We deduce that T is a monotone self-mapping of $[P', P]^{K+Q}$, and it follows that it has a fixed point, and that the set of fixed points forms a complete lattice. (A fixed point can be computed by starting with α^P and iteratively applying T until convergence.) Combining this with the fact that any fixed point must balance supply and demand, we obtain uniqueness (proofs of all claims are at the end of this appendix).

Claim 1. *The operator T has a unique fixed point α^*.*

Consider the fixed point α^* of T. Let $g(n) = h_n/\sqrt{n}$ for any $h_n = \omega(1)$. In the argument below we use $h_n = \sqrt{\log n} \Rightarrow g(n) = \sqrt{\log n/n}$, but this is only an example of a possible h_n that can be used. Define α^{lb} by $\alpha^{\mathrm{lb}}_{kq} = \alpha^*_{kq} - g(n)$ for all k, q and α^{ub} by $\alpha^{\mathrm{ub}}_{kq} = \alpha^*_{kq} + g(n)$.

We shall show that, for appropriate C, there is a core outcome in the n-th market satisfying $\alpha^{\text{lb}} \preceq \alpha \preceq \alpha^{\text{ub}}$, and that all core outcomes are within $C \log n/n$ of α. Define $\nu^*_{kq} = D_{kq}(\alpha^*_{kq}; \alpha^*)$, and show that $N(k,q)/n$ approaches ν^*_{kq}.

We now define an operator T_n such that its fixed point will capture a core solution in the n-th market. Analogous to Eq. (43) and Eq. (44), we define demand and supply for the n-the market to be[6]

$$\hat{D}_{kq}(p;\alpha) = \big|\big\{i : \tau(i) = k, u(k,q) - p + \eta_i^q \geq u(k,q') - \alpha_{kq'} + \eta_i^{q'} \text{ for all } q' \neq q\big\}\big| \tag{45}$$

and

$$\hat{S}_{kq}(p;\alpha) = \big|\big\{j : \tau(j) = q, p + \epsilon_j^k \geq \alpha_{k'q} + \epsilon_j^{k'} \text{ for all } k' \neq k\big\}\big| \tag{46}$$

for the n-th market. We define $(\mathsf{T}_n\alpha)_{kq} = \inf\{p : \hat{S}_{kq}(p;\alpha) > \hat{D}_{kq}(p;\alpha)\}$. As before, D_{kq} is (weakly) monotone decreasing in p and (weakly) monotone increasing in α, whereas S_{kq} is (weakly) monotone increasing in p and (weakly) monotone decreasing in α. (The monotonicities in α constitute a weak gross substitutes property for each type of agent on both sides of the market.) We deduce monotonicity of T_n.

Fact 9. *If $\alpha \preceq \alpha'$ then $\mathsf{T}_n\alpha \preceq \mathsf{T}_n\alpha'$.*

Our next claim implies that T_n has a fixed point that lies between α^{lb} and α^{ub}.

Claim 2. *With high probability, we have $\mathsf{T}_n\alpha^{\text{ub}} \preceq \alpha^{\text{ub}}$ and $\mathsf{T}_n\alpha^{\text{lb}} \succeq \alpha^{\text{lb}}$. Hence, T_n has a fixed point α satisfying $\alpha^{\text{lb}} \preceq \alpha \preceq \alpha^{\text{ub}}$.*

The proof of this claim uses the convergence of the empirical distribution of (type, productivity vector) for agents to the limiting distribution. The main part of the proof is to show that w.h.p., at prices α^{ub}, the demand is less than the supply for each type pair (whereas the opposite is true at α^{lb}). This is accomplished by showing that the realized demand (supply) at

[6] To be perfectly precise, we should ensure that each worker counts toward exactly one unit of total demand by using a mix of strict and weak inequalities in the definition of D_{kq}, based on the direction in which ties between q and q' are broken in each case. To make all the details work out, we require agents to break ties in favor of the type they are matched to under the maximum-weight matching M.

α^{ub} is less (more) than the limiting demand (supply) at α^*. In turn, the definition of α^{ub} helps us show this, because the relative prices of different types are the same as under α^*, but each of them is now more expensive relative to remaining unmatched.

It is not hard to show that the fixed points of T correspond to core price vectors.

Claim 3. *Let α be a fixed point of T_n. Then (M, α) is a core outcome, where M is the maximum-weight matching.*

Though we don't need the converse, we remark that it holds: for any core outcome (M, α), the prices α are a fixed point of T_n.

Combining Claims 2 and 3 immediately gives a core solution α close to α^*, i.e., the bound (1) for *some* core solution. We are also in a position to prove (2), leveraging the following claim that controls demand and supply in the n-th market for all price vectors between α^{lb} and α^{ub}. In particular, the bounds in the claim apply to the core solution α, where $\hat{D}_{kq}(\alpha_{kq}; \alpha) = \hat{S}_{kq}(\alpha_{kq}; \alpha) = N(k, q)$, thus yielding (2).

Claim 4. *Let $\nu^*_{kq} = D_{kq}(\alpha^*_{kq}; \alpha^*)$. Then, there exists $f(n) = O^*(\sqrt{n})$ such that, with high probability, for all $\alpha^{\mathrm{lb}} \preceq \alpha \preceq \alpha^{\mathrm{ub}}$ we have, for all $(k, q) \in \mathcal{T}$, that*

$$\left|\hat{D}_{kq}(\alpha_{kq}; \alpha) - n\nu^*_{kq}\right| \le f(n)\,,$$
$$\left|\hat{S}_{kq}(\alpha_{kq}; \alpha) - n\nu^*_{kq}\right| \le f(n)\,.$$

For the second part of the theorem (bound (3)), we prove the following claim, and use it together with the fact that weight$(M) = \Omega(n)$ with high probability, which follows from $\rho_k > 0$ for all k, and $f(x) > 0$ for all x. The claim also completes the proof of (1) for *all* core solutions.

Claim 5. *There exists $C = C(K, Q, \rho, f) < \infty$ such that for all k, q we have*

$$|\alpha^{\max}_{kq} - \alpha^{\min}_{kq}| \le C\frac{\log n}{n}\,. \tag{47}$$

The proof of this claim relies on the fact that all core α_{kq}'s must lie between some consecutive order statistics of ϵ_i^q's for workers i who are type k, and are either unmatched or matched to type q. Furthermore, these order statistics are close together in the vicinity of α^*, and we already know that it is sufficient to consider this vicinity. □

We now provide proofs of all the claims above that facilitated our proof of Theorem 2.

Proof of Claim 1. The set of $\alpha \in [P', P]^{K+Q}$ is a complete sublattice. Since T is a monotone self-mapping of this set, it has a nonempty set of fixed points that themselves form a complete lattice. Take the two extreme fixed points $\alpha^{\max}(T)$ and $\alpha^{\min}(T)$. Suppose that $\alpha^{\max}(T) \succ \alpha^{\min}(T)$. Fix a price vector α. A direct calculation yields the following intuitive expression for the total demand from worker type k:

$$\sum_q D_{kq}(\alpha_{kq}; \alpha) = \rho_k \Big(1 - \prod_q F(\alpha_{kq} - u(k,q))\Big) . \tag{48}$$

Similarly, the total supply from an employer of type q is

$$\sum_k S_{kq}(\alpha_{kq}; \alpha) = \rho_q \Big(1 - \prod_k F(-\alpha_{kq})\Big) . \tag{49}$$

For ease of notation, we write $D_{kq}^{\max} = \sum_q D_{kq}(\alpha_{kq}^{\max}(\mathsf{T}); \alpha^{\max}(\mathsf{T}))$, and so on. It follows from $\alpha^{\max}(T) \succ \alpha^{\min}(T)$ and Eqs. (48) and (49) that

$$\sum_k \sum_q D_{kq}^{\max} < \sum_k \sum_q D_{kq}^{\min}$$
$$\sum_k \sum_q S_{kq}^{\max} > \sum_k \sum_q S_{kq}^{\min}.$$

At any fixed point of T, we know that $D_{kq} = S_{kq} \Rightarrow \sum_k \sum_q D_{kq} = \sum_k \sum_q S_{kq}$. Hence,

$$\sum_k \sum_q D_{kq}^{\min} = \sum_k \sum_q S_{kq}^{\min}$$
$$\Rightarrow \sum_k \sum_q D_{kq}^{\max} < \sum_k \sum_q S_{kq}^{\max} ,$$

a contradiction. $\square$

Proof of Claim 2. We shall use the fact that $\mathrm{E}[\hat{D}_{kq}(p;\alpha)] = nD_{kq}(p;\alpha)$ for any p and α, along with concentration bounds, to obtain the claim. We shall show that w.h.p., we have $\hat{D}_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}) < \hat{S}_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}})$, which will immediately imply that $(\mathsf{T}_n\alpha^{\mathrm{ub}})_{kq} \le \alpha^{\mathrm{ub}}_{kq}$, i.e., $\mathsf{T}_n\alpha^{\mathrm{ub}} \preceq \alpha^{\mathrm{ub}}$. (The proof of $\mathsf{T}_n\alpha^{\mathrm{lb}} \succeq \alpha^{\mathrm{lb}}$ is analogous and we omit it.) We show existence of the fixed point by starting with α^{ub} and iteratively using T_n until (monotone) convergence to α is obtained.

Note that $\hat{D}_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}})$ is distributed as Binomial$(n, D_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}))$. It follows that

$$\Pr[\hat{D}_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}) < n(D_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}) + c_n)] \ge 1 - \exp(-2nc_n^2)\,, \tag{50}$$

for any c using a standard Chernoff bound. Furthermore, a direct calculation[7] gives

$$D_{kq}(\alpha^*_{kq};\alpha^*) - D_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}) = \rho_k \int_0^{g(n)} f(z + \alpha^*_{kq} - u(k,q)) \prod_{q'\neq q} F(z + \alpha^*_{kq'} - u(k,q'))\,\mathrm{d}z$$

$$\Rightarrow \lim_{n\to\infty} \frac{D_{kq}(\alpha^*_{kq};\alpha^*) - D_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}})}{g(n)} = \rho_k f(\alpha^*_{kq} - u(k,q)) \prod_{q'\neq q} F(\alpha^*_{kq'} - u(k,q')),$$

using the continuity of $f(\cdot)$ and that $\lim_{n\to\infty} g(n) = 0$. Define

$$c_n = \frac{g(n)}{2} \min_{k,q} \Big[\rho_k f(\alpha^*_{kq} - u(k,q)) \prod_{q'\neq q} F(\alpha^*_{kq'} - u(k,q'))\Big].$$

Note that $nc_n^2 = \Omega((\log n)^2)$. Then, for large enough n, we have

$$D_{kq}(\alpha^*_{kq};\alpha^*) \le D_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}) + c_n \qquad \text{for all } k,q\,.$$

Substituting in Eq. (50), we obtain

$$\Pr[\hat{D}_{kq}(\alpha^{\mathrm{ub}}_{kq};\alpha^{\mathrm{ub}}) < nD_{kq}(\alpha^*_{kq};\alpha^*)] \ge 1 - \exp\big(-\Omega(\log n)\big) = 1 - o(1) \qquad \text{for all } k,q\,; \tag{51}$$

[7] It is important here that α^{ub} and α^* differ by the same amount in each coordinate, meaning that the relative attractiveness of types on the other side is unchanged.

i.e., with high probability,

$$\hat{D}_{kq}(\alpha^{\rm ub}_{kq};\alpha^{\rm ub}) < nD_{kq}(\alpha^*_{kq};\alpha^*) \qquad \text{for all } k,q\,.$$

Using an analogous argument, we establish that with high probability,

$$\hat{S}_{kq}(\alpha^{\rm ub}_{kq};\alpha^{\rm ub}) > nS_{kq}(\alpha^*_{kq};\alpha^*) \qquad \text{for all } k,q\,.$$

Now, by definition of α^* we have $D_{kq}(\alpha^*_{kq};\alpha^*) = S_{kq}(\alpha^*_{kq};\alpha^*)$ for all k,q. We deduce that, w.h.p.,

$$\hat{S}_{kq}(\alpha^{\rm ub}_{kq};\alpha^{\rm ub}) > \hat{D}_{kq}(\alpha^{\rm ub}_{kq};\alpha^{\rm ub}) \qquad \text{for all } k,q\,,$$

as needed. □

Proof of Claim 3. Let α be a fixed point of T_n. Then α clears the market up to tie-breaking, and using the fact that this is a two-sided matching setting, the tie-breaking can be done so as to obtain a matching M' such that (M',α) is a core outcome. This also means that $M' = M$, the maximum-weight matching. □

Proof of Claim 4. Consider $\hat{S}_{kq}(\alpha_{kq};\alpha)$. Let $\hat{S}^{\max}_{kq}$ and $\hat{S}^{\min}_{kq}$ be the largest and smallest values, respectively, of $\hat{S}_{kq}$ for α in the specified range. Using the monotonicities of $\hat{S}_{kq}$ in its arguments, we have

$$\hat{S}^{\max}_{kq} = \hat{S}_{kq}(\alpha^{\rm ub}_{kq};\alpha^{\rm lb})\,,$$
$$\hat{S}^{\min}_{kq} = \hat{S}_{kq}(\alpha^{\rm lb}_{kq};\alpha^{\rm ub})\,.$$

Using the definition (46) of $\hat{S}_{kq}$, we have that

$$\hat{S}^{\max}_{kq} - \hat{S}^{\min}_{kq} \le \big|\{j : \tau(j) = q, |\alpha^*_{kq} + \epsilon^k_j| \le g(n)\}\big| + \sum_{k' \neq k} \big|\{j : \tau(j) = q, |\epsilon^k_j - \epsilon^{k'}_j + \alpha^*_{kq} - \alpha^*_{k'q}| \le g(n)\}\big| . \quad (52)$$

Now we obtain

$$\begin{aligned}\big|\{j : \tau(j) = q, |\alpha^*_{kq} + \epsilon^k_j| \le g(n)\}\big| &= \text{Binomial}(n\rho_q, F(-\alpha^*_{kq} + g(n)) - F(-\alpha^*_{kq} - g(n))) \\ &\le O^*(\sqrt{n}) ,\end{aligned}$$

w.h.p., using that $g(n) = \sqrt{\log n/n}$. Let F_1 be the distribution of the difference between two i.i.d. draws from F. The F_1 is also a distribution with positive and continuous density everywhere in $(-\infty, \infty)$. We can bound each of the other terms in Eq. (52) as

$$\begin{aligned}&\big|\{j : \tau(j) = q, |\epsilon^k_j - \epsilon^{k'}_j + \alpha^*_{kq} - \alpha^*_{k'q}| \le g(n)\}\big| \\ &= \text{Binomial}(n\rho_q, F_1(-\alpha^*_{kq} + \alpha^*_{k'q} + g(n)) - F_1(-\alpha^*_{kq} + \alpha^*_{k'q} - g(n))) \\ &\le O^*(\sqrt{n}) .\end{aligned}$$

Combining, we obtain that, with high probability,

$$\begin{aligned}\hat{S}^{\max}_{kq} - \hat{S}^{\min}_{kq} &\le O^*(\sqrt{n}) \\ \Rightarrow |\hat{S}_{kq}(\alpha_{kq}; \alpha) - \hat{S}_{kq}(\alpha^*_{kq}; \alpha^*)| &\le O^*(\sqrt{n}) \quad (53)\end{aligned}$$

for all $\alpha^{\text{lb}} \preceq \alpha \preceq \alpha^{\text{ub}}$.

Recall that $\nu^*_{kq} = D_{kq}(\alpha^*_{kq}; \alpha^*) = S_{kq}(\alpha^*_{kq}; \alpha^*)$ and $\mathrm{E}[\hat{S}_{kq}(\alpha^*_{kq}; \alpha^*)] = nS_{kq}(\alpha^*_{kq}; \alpha^*)$. A Chernoff bound gives that, with high probability,

$$|\hat{S}_{kq}(\alpha^*_{kq}; \alpha^*) - n\nu^*_{kq}| \le O^*(\sqrt{n}) . \quad (54)$$

Combining Eqs. (53) and (54) yields

$$|\hat{S}_{kq}(\alpha_{kq};\alpha) - n\nu^*_{kq}| \le O^*(\sqrt{n}), \tag{55}$$

for all $\alpha^{\mathrm{lb}} \preceq \alpha \preceq \alpha^{\mathrm{ub}}$.

The proof of the corresponding bound on $\hat{D}_{kq}(\alpha_{kq};\alpha)$ is analogous.

□

Proof of Claim 5. Having found a core solution α (the fixed point of T_n) that moreover satisfies $\alpha^{\mathrm{lb}} \preceq \alpha \preceq \alpha^{\mathrm{ub}}$, we now use these bounds on α to control the size of the core.

Recall that there is (with probability 1) a unique[8] maximum-weight matching M. We focus on a subset of agents S that are of type k and either matched to type q or unmatched under M. We show that, w.h.p., the largest value of η_i^q among $i \in S \backslash M$ is within $C \log n/n$ of the smallest value of η_i^q among $i \in S \cap M$, yielding the desired bound on $|\alpha^{\max}_{kq} - \alpha^{\min}_{kq}|$.

Define

$$S = \{i : \tau(i) = k, \eta_i^{q'} < \alpha^*_{kq'} - g(n) - u(k,q') \ \forall\, q' \ne q\}. \tag{56}$$

If $\alpha \succeq \alpha^{\mathrm{lb}}$ (which holds w.h.p.), we know that no agent in S is matched to $q' \ne q$. Now, we get that the likelihood of an agent being in S is

$$\rho_k \prod_{q' \ne q} F(\alpha^*_{kq'} - g(n) - u(k,q')) \ge 0.9 \rho_k \prod_{q' \ne q} F(\alpha^*_{kq'} - u(k,q'))$$

for large enough n, using continuity of F and $\lim_{n\to\infty} g(n) = 0$. (Note that we do not need to reveal η_i^q to compute whether $i \in S$.) Now divide $[\alpha^*_{kq} - u(k,q) - 2g(n), \alpha^*_{kq} - u(k,q) + 2g(n)]$ into intervals of length $\Delta = (C/3)\log n/n$, where

$$C = 4 \Big/ \min_{k,q} \Big[\rho_k f(\alpha^*_{kq} - u(k,q)) \prod_{q' \ne q} F(\alpha^*_{kq'} - u(k,q'))\Big].$$

[8] We do not need uniqueness of M to bound the core size, but we use it to simplify the argument.

Using continuity of $f(\cdot)$, we have that

$$\inf_{x \in [\alpha^*_{kq} - u(k,q) - 2g(n), \alpha^*_{kq} - u(k,q) + 2g(n)]} f(x) \geq 0.9 f(\alpha^*_{kq} - u(k,q)) \,.$$

Using the previous three equations, and independence between membership in a subinterval and membership in S, it follows that the number of agents in S who belong to any individual subinterval is Binomial(n, p_1) for

$$p_1 \geq \Delta \cdot 0.9 f(\alpha^*_{kq} - u(k,q)) \cdot 0.9 \rho_k \prod_{q' \neq q} F(\alpha^*_{kq'} - u(k,q')) \geq \log n / n \,,$$

and hence each subinterval has at least one agent in S with probability at least $1 - 1/n$. We deduce using a union bound that w.h.p., all subintervals have at least one member of S. Now, in any core solution, each i such that $\eta_i < \alpha_{kq} - u(k,q)$ must remain unmatched, and $\eta_i > \alpha_{kq} - u(k,q)$ must be matched to type q. Furthermore, we have that w.h.p., $\alpha_{kq} \in (\alpha^*_{kq} - g(n), \alpha^*_{kq} + g(n))$. Considering the subintervals on each side of the one in which α_{kq} occurs, since each of them has at least one agent (who also belongs to S), and considering that M is unique, we deduce that

$$|\alpha^{\max}_{kq} - \alpha^{\min}_{kq}| \leq 3\Delta = C \frac{\log n}{n} \,.$$

Using a union bound, we deduce that w.h.p., this holds simultaneously for all k, q. □

G Proof of Lemma 2

Proof of Lemma 2. Let (M, α) be a core solution. Suppose that $\alpha_{k^*q^*} = \max_{k,q} \alpha_{kq}$. Consider worker type q^*. Since $\alpha_{k^*q^*} \geq \alpha_{kq^*}$ for all $k \neq k^*$, we obtain that each agent i of type q^* has likelihood at least $1/K$ of counting toward $\hat{S}_{k^*q^*} = \hat{S}_{k^*q^*}(\alpha_{k^*q^*}; \alpha)$; see Eq. (46). We have that w.h.p., $\hat{S}_{k^*q^*} \geq (\#\text{ agents of type } q^*)/K - O^*(\sqrt{n}) > Cn/2$, using Assumption 2. At prices α, an agent of type k^* prefers being unmatched to matching with type q^* with likelihood $F(\alpha_{k^*q^*} - u(k^*, q^*))$. Hence, w.h.p., we have $\hat{D}_{k^*q^*} = \hat{D}_{k^*q^*}(\alpha_{k^*q^*}; \alpha) < (\text{number of agents of type } k^*)(1 - F(\alpha_{k^*q^*} - u(k^*, q^*))) + O^*(\sqrt{n}) \leq n(1 - F(\alpha_{k^*q^*} - u(k^*, q^*))) + O^*(\sqrt{n}) < Cn/2$ if $\alpha_{k^*q^*} > U$

for some $U < \infty$. (Here we used that $\lim_{P \to \infty} 1 - F(P - u(k^*, q^*)) = 0$.) But $\hat{D}_{k^*q^*} = \hat{S}_{k^*q^*}$, a contradiction. It follows that $\alpha_{k^*q^*} \leq U$, and hence $\alpha_{kq} \leq U$ for all k, q. The lower bound of $-U \leq \alpha_{kq}$ can be similarly established. □

References

Durrett, Rick. 2010. *Probability: Theory and Examples*. Cambridge University Press.

Gale, David, Lloyd S. Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* **69**(1) 9–15.

Hatfield, John William, Paul R. Milgrom. 2005. Matching with contracts. *The American Economic Review* **95**(4) 913–935.

Kelso, Jr., Alexander S., Vincent P. Crawford. 1982. Job matching, coalition formation, and gross substitutes. *Econometrica* **50**(6) 1483–1504.