Misallocation Inefficiency in Partially Directed Search*

Stanislav Rabinovich[†] Ronald Wolthoff[‡]

September 22, 2022

Abstract

We identify a misallocation inefficiency in search models, which is distinct from the aggregate entry distortion emphasized in the previous literature, and arises instead from partially directed search. We consider a framework in which workers differ in whether they can direct their search, and firms are heterogeneous in productivity. The main result is that too many workers apply to high-productivity firms, relative to the social optimum. This occurs because too many firms attract only random searchers, in order to extract more surplus from them. Because it is the low-productivity firms that do so, this induces all the directed searchers to concentrate at the high-productivity firms. A minimum wage can increase employment and welfare by reallocating workers across firms. With endogenous entry by either workers or firms, the misallocation inefficiency coexists with a standard entry externality; in this case, a proper combination of a tax or subsidy and a minimum wage can restore the efficient allocation.

Keywords: Directed search, random search, labor markets, minimum wage, misallocation, market power.

*A first draft of this paper was circulated under the title *Misallocation Effects of Labor Markets Frictions*. We thank the editor, associate editor and three anonymous referees for their valuable feedback. We are further grateful for an insightful discussion by Adrian Masters at the 2019 DC Search and Matching Workshop, and for helpful comments by Jim Albrecht, Manolis Galenianos, Cristina Lafuente-Martinez, Ricardo Lagos, Guido Menzio, Helene Turon, Susan Vroman, Xi Weng, Guillaume Wilemme, and seminar and conference participants at the National University of Singapore, the IZA Institute of Labor Economics, the Federal Reserve Bank of Cleveland, Drexel University, the Essex/RHUL SaM Webinar, the 2019 DC Search and Matching Workshop, the University at Albany, the 2019 European Search and Matching Workshop, Guanghua School of Management at Peking University, and the Fall 2019 Midwest Macroeconomics Meetings. Wolthoff gratefully acknowledges financial support from the Social Sciences and Humanities Research Council.

[†]University of North Carolina - Chapel Hill, 107 Gardner Hall, CB 3305, Chapel Hill, NC 27599, USA. Email: srabinov@email.unc.edu.

[‡]University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada. Email: ronald.p.wolthoff@gmail.com.

1 Introduction

Do markets with search frictions achieve efficient outcomes? This question is of obvious substantive importance, e.g. for policy interventions such as the minimum wage in the context of the labor market, which has been the canonical application. The theoretical literature on this question has focused overwhelmingly on a specific margin, namely the aggregate market tightness, which is in turn determined by agents' entry choices. In this paper, we depart from the focus on the *aggregate* entry efficiency by analyzing instead whether workers are allocated in a constrained-efficient manner across jobs. We identify a novel inefficiency, which arises from the combination of random and directed search and operates along the allocation margin.

We analyze the allocation of workers to firms in a specific model of the labor market that features partially directed search. The labor market is frictional: firms post wages, understanding that higher-wage job openings will be filled more quickly. The key features of the model are *partially directed search* and *firm heterogeneity*. The modeling of partially directed search follows Lester (2011): a fraction of the workers are directed searchers, who choose which job openings to apply for, understanding that higher posted wages attract more competing applicants. The remaining fraction are random searchers, who cannot control which firms they target and are assigned to vacancies randomly.[1] The environment differs from Lester (2011) by introducing firm heterogeneity in productivity. Our baseline model abstracts from entry decisions on either side of the market and thus abstracts from the key margin that has occupied much of the previous literature. The key question, instead, is the equilibrium allocation of existing workers across firms of differing productivities, and how it compares to the allocation chosen by the social planner.

The main result is that, in equilibrium, too many workers are employed at high-productivity firms, as compared to the social optimum. This contrasts with the intuitive prediction that frictions, such as mobility costs or lack of information, result in too many workers at low-productivity firms. Both the constrained-efficient allocation and the decentralized equilibrium are characterized by a threshold rule: firms above some productivity threshold attract both random and directed searchers, whereas firms with productivity below this threshold attract only random searchers. However, the equilibrium productivity threshold is always higher than the socially efficient one. The mechanism behind this inefficiency is akin to a monopsony distortion in the absence of wage-discrimination. In order to attract directed searchers, the firm must raise its wage, but then it must raise the wage to the random

[1] The framework thus nests the special cases of purely directed search, as surveyed in Wright et al. (2021), and purely random search.

searchers as well. This monopsony distortion makes firms inefficiently reluctant to attract directed searchers in equilibrium. Because this under-hiring incentive is stronger for low-productivity firms, equilibrium has too few firms attracting directed searchers; this means that directed searchers are concentrated at a smaller number of firms. As a result, there is a misallocation of labor toward the top of the productivity distribution.

The inefficiency we identify is distinct from the aggregate entry distortions emphasized in the previous literature e.g. (e.g. Hosios, 1990; Pissarides, 2000; Mangin and Julien, 2021): in our baseline framework, entry is fixed, and the inefficiency instead has to do with how existing workers are allocated across existing firms. It is also important to note that the inefficiency arises from the coexistence of random and directed search; indeed, the presence of random searchers leads to an inefficient allocation of directed searchers. This suggests more broadly that studying environments with partially directed search generates important insights that are absent from environments that consider either extreme separately.

We next show that a minimum wage can increase employment and welfare. It does so by inducing low-productivity firms to pay higher wages, thereby attracting workers away from the high-productivity firms. This reallocation of workers is efficiency-improving since it reduces the excess queue of workers at high-productivity firms; in fact, we show that an appropriately chosen minimum wage restores the constrained-efficient allocation. While the previous literature has focused overwhelmingly on the aggregate employment effects of the minimum wage, we emphasize its allocative effects. Notably, the minimum wage *does* affect employment in our framework, even in the absence of an extensive entry margin; in fact, the minimum wage just described raises employment. However, the reasons for this are distinct from the conventional narrative. According to the latter, mandating higher wages results in more workers willing to work because of an upward-sloping labor supply curve – an extensive-margin channel. In our model, mandating higher-wages results in a reallocation of workers from firms with many applicants to firms with few applicants; this raises employment in spite of total labor supply remaining constant.

Several extensions of our model both illustrate the robustness of our results and help relate them to the previous literature. First, we extend our baseline model to incorporate an extensive margin: either in the form of endogenous participation by workers, or in the form of endogenous entry by firms. In either case, we show that the misallocation inefficiency now coexists with a conventional entry inefficiency studied extensively in the literature. The entry inefficiency stems from the violation of the Hosios condition (Hosios, 1990; Pissarides, 2000; Mangin and Julien, 2021): aggregate entry by workers/firms is inefficient because the marginal entrant does not internalize the effect of their entry on the matching probability and average match surplus. The misallocation inefficiency operates in the same form as in our baseline model: a social planner who takes the aggregate numbers of workers and firms as given can still improve welfare by reallocating workers away from the high-productivity firms. Because such an environment is plagued by two distortions, a single minimum wage no longer suffices to restore constrained efficiency; instead, we show the constrained-efficient allocation can be achieved by a suitable combination of a tax/subsidy and a minimum wage.

Finally, we also extend our baseline model to allow for heterogeneous productivity on the worker side, as well as the firm side. We show that the direction of the misallocation is robust to this extension. The misallocation is still driven by a monopsony distortion, which causes too few firms to attract directed searchers. The same monopsony distortion leads to too few firms attracting high-productivity workers, conditional on attracting directed searchers at all. Both of these channels lead to too many workers concentrating at high-productivity firms.

Outline. This paper is organized as follows. In Section 2, we discuss the relationship between our paper and the existing literature. We describe the environment in Section 3 and characterize the planner's solution in Section 4. Section 5 presents the main results of the paper: we characterize the decentralized equilibrium (5.1), show that the equilibrium features misallocation (5.2), discuss limiting cases (5.3), and demonstrate that a minimum wage can be welfare-improving (5.4). In Section 6, we extend our baseline model to incorporate endogenous participation by workers (6.1) and endogenous entry by firms (6.2). We show that the misallocation inefficiency is robust to these extensions and coexists with a standard participation/entry distortion. In Section 7, we show that the misallocation inefficiency is robust to introducing worker heterogeneity. Finally, Section 8 concludes.

2 Relationship to literature

Partially directed search. Our paper builds most directly on the environment of Lester (2011) and falls more generally within the small but growing literature on partially directed search, such as Lentz and Moen (2017), Shi (2018), Cheremukhin et al. (2020), or Wu (2021). While the vast majority of previous work has focused on either purely random or purely directed search, the results from this recent literature – as well as our own – suggest that there are new insights to be gained from studying the interaction between the two, which would be missed when studying either extreme in isolation. The aforementioned papers on partially directed search differ widely in terms of how exactly they combine random and directed search and/or endogenize the degree to which search is targeted. Because of the variety of modeling strategies, the misallocation toward high-productivity firms that arises in our environment is absent from others. Below, we briefly elaborate on the feature of our environment driving this result.

As mentioned, our model environment follows Lester (2011), which we extend to allow for heterogeneity in firm productivity (as well as heterogeneity in worker productivity in Section 7).[2] The key feature of this environment is the co-existence of two distinct types of workers: random searchers who do not respond to posted wages and directed searchers who do. Its consequence is that firms face a labor supply curve that is kinked at the point where directed searchers are no longer willing to apply. As a result, firms may either attract only random searchers with a low wage or attract both types of searchers with a *discretely* higher wage. As we will demonstrate formally, this feature, akin to a monopsony distortion, makes firms inefficiently reluctant to attract directed searchers. The key contribution relative to Lester (2011) is to show that, with heterogeneous firms, firms at the bottom of the productivity distribution are the ones with the strongest incentive to attract random searchers only; that too many firms do so; and that this leads to inefficiently many workers trying to match at the most productive firms.

Search frictions and misallocation. As mentioned above, the majority of work on inefficiency driven by search frictions has focused on aggregate entry. Fewer papers have addressed how frictions can lead to an inefficient composition of jobs or inefficient allocation of workers among them; the ones that do focus on different mechanisms and typically reach the opposite result from ours. For example, in Bertola and Caballero (1994), Acemoglu (2001), and Davis (2001), too many workers are allocated to low-productivity firms: there, search is random, and the key friction leading to inefficiency is akin to an investment holdup problem in the labor market. In Acemoglu and Shimer (1999) and Golosov et al. (2013), search is directed, but workers are risk-averse; these papers show that there will likewise be misallocation towards low-productivity firms if workers cannot insure against the risk of not finding a job. In Galenianos et al. (2011), market power, driven by a finite number of firms, leads to misallocation of workers, namely too many workers being employed at low-productivity firms. Instead, in our paper, where market power is instead driven by imperfectly directed search, we show that there is misallocation toward *high-productivity* firms, contrary to the previous literature. This result stems from the co-existence of random and directed searchers: indeed, the presence of random searchers distorts firms' incentives to attract directed searchers, and hence how they are allocated across firms. We also share with Galenianos et al. (2011) the

[2] In other words, the model in Lester (2011) can be thought of as a special case of ours; we provide a formal discussion of this in Section 5.3. Note that Lester (2011) is set in the context of a product market rather than a labor market, and hence agents are buyers and sellers rather than workers and firms, but this distinction is inconsequential except for the interpretation.

prediction that a binding minimum wage reallocates workers towards low-productivity firms; however, in our setting this reallocation can be welfare-improving.

Relationship to the Hosios condition. Our baseline model abstracts from entry, and our main inefficiency result is distinct from the much-studied entry distortion. However, extensions of our baseline model in Section 6 introduce an extensive margin; in this case, we show that the conventional entry inefficiency emerges and coexists alongside the novel misallocation inefficiency. In these environments we show that, as usual, the entry inefficiency can be understood in terms of a violation of the Hosios condition (e.g. Hosios, 1990; Pissarides, 2000; Mangin and Julien, 2021). The combination of partially directed search and ex ante heterogeneity turns out to be of interest in this context. For example, when firms are heterogeneous, their entry decisions are endogenous, and search is not perfectly directed, an entering firm creates both a congestion externality (by affecting others' matching probability) and an output externality (by affecting the productivity composition of firms). The relevant optimality condition is therefore the generalized Hosios condition as defined by Mangin and Julien (2021). Examples of a similar two-fold externality in a random-search context are shown in previous work, e.g. Albrecht et al. (2010), Masters (2015), Julien and Mangin (2017), and Mukoyama (2019). We explicitly illustrate that the analogous generalized Hosios condition of Mangin and Julien (2021) is also applicable in our partially-directed search environment. The key innovation is that now the violation of this Hosios condition is not the only source of inefficiency: it exists alongside a misallocation of workers across existing jobs.

3 Environment

We consider a static model. There is a measure one of firms, indexed by $j \in [0, 1]$, each with one vacancy. Firms are heterogeneous in the output that they produce when matched with a worker. We denote the productivity of firm j by $y(j)$, where the function $y : [0, 1] \to \mathbb{R}_+$ is assumed to be continuous and strictly increasing.[3]

There is a measure one of workers, all initially unemployed. A fraction $\psi \in [0, 1]$ of workers are random searchers, who will be assigned randomly across all the vacancies. The remaining fraction $1 - \psi$ are directed searchers, who can choose which vacancy to target. Matching works as follows. If the searcher-vacancy ratio ('queue length') at a particular vacancy is λ, the vacancy gets filled with probability $m(\lambda)$, and each worker applying to

[3] The homogeneous-productivity economy can be dealt with by considering the limiting case as $y(0) \to y(1)$. We discuss the homogeneous-productivity case formally in Section 5.3.

that vacancy has a probability $m(\lambda)/\lambda$ of being matched. The matching function $m : \mathbb{R}_+ \to [0,1]$ satisfies the following standard assumptions (which are satisfied by common microfoundations of the meeting process, such as an urn-ball or geometric process): i) m is strictly increasing and concave, i.e. $m' > 0$ and $m'' < 0$; ii) $m(0) = 0$ and $m(\lambda) \leq 1$; and iii) the elasticity $\epsilon_m(\lambda) = \lambda m'(\lambda)/m(\lambda)$ is strictly positive and strictly decreasing in λ. Workers and firms who remain unmatched receive a payoff of zero.[4]

Note that if $\psi = 0$, the above is a standard competitive search environment. However, with $\psi > 0$, there will be random searchers at each vacancy. Hence, the queue length at each vacancy satisfies $\lambda \geq \psi$. At the other extreme, $\psi = 1$, we have a standard random search environment in which $\lambda = 1$ always.

There are multiple ways to interpret the partial randomness of search. One is in terms of differential information: random searchers could be viewed as workers uninformed about the posted wages of the various job openings. Alternatively, one could interpret some workers as more mobile across firms than others. The idea that frictions such as imperfect information or imperfect mobility lead to market power—and impede efficient allocation of labor—has a long tradition in economics; see e.g. Pigou (1932), Robinson (1933), Manning (2003). Crucially, information or mobility frictions that constrain individual workers will also constrain the social planner—an insight that will be important below.

4 Planner's problem

The planner chooses the distribution of workers across posted vacancies to maximize aggregate output. Thus, the planner's problem can be written as choosing $\lambda(j)$ for every $j \in [0,1]$ so as to maximize

$$\int_0^1 m(\lambda(j))\, y(j)\; dj. \tag{1}$$

The planner maximizes (1) subject to two constraints. First, there is a resource constraint, which says that the total measure of workers at all the vacancies must add up to one:

$$\int_0^1 \lambda(j)\, dj = 1. \tag{2}$$

Second, and crucially, the planner must respect the randomness of search for some workers. This means that the planner must assign the ψ random searchers randomly across all the

[4] It is straightforward to introduce a non-zero value of unemployment for workers, which would not change any of our results. An environment in which workers receive an unemployment benefit of b and firm j produces $Y(j)$ is equivalent to ours if one re-defines $y(j) = Y(j) - b$.

posted vacancies, thereby assigning a queue length of at least ψ to each vacancy. Hence, the random search constraint states

$$\lambda(j) \geq \psi \tag{3}$$

for all $j \in [0,1]$. Let η be the Lagrange multiplier on (2), and let $\mu(j)\,dj$ be the Lagrange multiplier on (3) for each j. The first-order condition for $\lambda(j)$ can be written as

$$\mu(j) = \eta - m'(\lambda(j))\,y(j)\,. \tag{4}$$

When constraint (3) does not bind, we have $\mu(j) = 0$ and thus $\lambda(j)$ solves $m'(\lambda(j))\,y(j) = \eta$. Whenever constraint (3) binds, we have $\lambda(j) = \psi$ and $\mu(j) = \eta - m'(\psi)\,y(j) > 0$. Moreover, $\eta - m'(\psi)\,y(j)$ is clearly decreasing in j. Therefore, the constraint binds for all j below some threshold j_p and does not bind for j above it. This threshold is given by

$$j_p = \inf\{j : m'(\psi)\,y(j) \geq \eta\}\,. \tag{5}$$

Given η, (5) uniquely determines j_p, which, together with η, determines the constrained-efficient queue length, denoted by $\lambda_p(j)$, for every j. The multiplier η is pinned down as the unique value for which the resource constraint holds with equality, i.e.

$$\int_0^1 \lambda_p(j)\; dj = 1. \tag{6}$$

This can be summarized as

Lemma 1. *The constrained-efficient allocation is characterized by a number η, a threshold $j_p \in [0,1]$ and a function $\lambda_p : [0,1] \to \mathbb{R}_+$ satisfying*

$$\begin{cases} \lambda_p(j) = \psi, & \text{if } j < j_p \\ m'(\lambda_p(j))\,y(j) = \eta, & \text{if } j \geq j_p \end{cases} \tag{7}$$

as well as (5) and (6). There exists a solution to this system, and it is unique. In particular, $\lambda_p(j)$ is strictly increasing in j for $j \geq j_p$.

Proof. See Appendix A.1. □

Note that the continuity of $y(j)$ implies the continuity of $\lambda_p(j)$, which will be important below in the comparison to the market equilibrium.

For our purposes, the interesting case is one where $j_p > 0$, i.e., the partial randomness of search is consequential. If this is the case, this interior j_p solves $m'(\psi)\,y(j_p) = \eta$. Next, we consider the conditions under which this is the case, i.e., the constraint (3) in fact binds for some j. The following result shows that this occurs when productivity dispersion and the fraction of random searchers are sufficiently large.

Lemma 2. *A necessary and sufficient condition for $j_p > 0$ is*

$$\int_0^1 \lambda(j)\, dj > 1, \tag{8}$$

where $\lambda : [0,1] \to \mathbb{R}_+$ solves $m'(\lambda(j))\, y(j) = m'(\psi)\, y(0)$ for each j.

Proof. See Appendix A.2. □

Intuitively, the partial randomness of search captured by (3) is more likely to severely constrain the social planner when productivity is very dispersed. In fact, consider the limiting case with homogeneous productivity. The social planner would never want to assign different queue lengths to firms with the same productivity, due to the concavity of the matching function; with constant productivity, the social planner would therefore assign the same $\lambda = 1$ to all firms, and (3) clearly would not bind. Further, even when productivity is dispersed, the random search constraint (3) would not bind for a low enough ψ.

5 Equilibrium

We now analyze the decentralized equilibrium and show how and in what respects it differs from the planner's allocation. Each firm decides what wage to post. The $1 - \psi$ directed searchers observe all the posted wages and decide to which firm to apply. As is standard in competitive search theory, we restrict attention to symmetric applications strategies. The ψ random searchers are assigned to vacancies randomly. The combination of these choices determines the queue length at each firm, and hence its profits, as a function of the wage it posted.

The definition of equilibrium requires us to specify the queue length $\lambda^*(w)$ attracted by a firm as a function of the wage w it posts, even for wages that are not posted in equilibrium. This is specified as follows. If a firm posts a wage w and attracts a queue length λ, the utility of a directed searcher applying to that firm is $\frac{m(\lambda)}{\lambda} w$. Define the *market utility* $\mathcal{U}$ to be the maximum utility across all submarkets that a directed searcher can obtain:

$$\mathcal{U} \equiv \max_{w,\lambda} \frac{m(\lambda)}{\lambda} w, \tag{9}$$

where the maximization is performed over all the submarkets w, λ active in equilibrium. While $\mathcal{U}$ is an equilibrium object, each firm takes it as given when deciding what wage to post. In particular, each firm understands that, if the wage-queue combination it offers provides utility of less than $\mathcal{U}$, then it will not attract any directed searchers, and its queue length must therefore equal ψ. As for random searchers, their only decision is whether to accept or reject the posted wage of the firm they meet; this decision simply constrains wages to be non-negative due to individual rationality. Each firm j maximizes its profits taking this worker behavior into account.

Definition 1. *An equilibrium consists of a market utility $\mathcal{U}$, a function $\lambda^* : \mathbb{R}_+ \to [\psi, \infty)$ and a function $w : [0,1] \to \mathbb{R}_+$ that satisfy:*

1. *Worker optimization:*
$$\frac{m\left(\lambda^*\left(w\right)\right)}{\lambda^*\left(w\right)} w \leq \mathcal{U} \tag{10}$$
for all w, and
$$\frac{m\left(\lambda^*\left(w\right)\right)}{\lambda^*\left(w\right)} w < \mathcal{U} \implies \lambda^*\left(w\right) = \psi. \tag{11}$$

2. *Firm optimization: for each $j \in [0,1]$,*
$$w\left(j\right) \in \arg\max_{w' \geq 0} m\left(\lambda^*\left(w'\right)\right)\left(y\left(j\right) - w'\right). \tag{12}$$

3. *Market clearing:*
$$\int_0^1 \lambda^*\left(w\left(j\right)\right) dj = 1. \tag{13}$$

The first two items formalize the optimizing behavior of workers and firms described above, together with the restriction on the environment that ψ of the workers apply randomly. The market-clearing condition (13) is the analogue of (2), stating that the total measure of workers adds up to one. The resulting equilibrium *allocation* consists of an assignment of queue length to each firm, $\lambda_e\left(j\right)$, satisfying $\lambda_e\left(j\right) = \lambda^*\left(w\left(j\right)\right)$.

5.1 Characterization, existence and uniqueness

We now characterize the equilibrium allocation $\lambda_e\left(j\right)$. Consider a firm's choice of what wage to post. If firm j offers less than the market utility $\mathcal{U}$, it will attract random searchers only, therefore receiving profits $m\left(\psi\right)\left(y\left(j\right) - w\right)$. From this it is easy to conclude that a firm that chooses not to attract directed searchers will offer a wage of zero, and its profits are therefore

$$\pi^R(j) = m(\psi)\, y(j)\,. \tag{14}$$

On the other hand, if firm j would like to attract some directed searchers, it solves the problem

$$\pi^D(j) = \max_{w,\lambda} m(\lambda)\,(y(j) - w) \tag{15}$$

subject to

$$\frac{m(\lambda)}{\lambda} w \geq \mathcal{U}. \tag{16}$$

It is easy to see that (16) will bind. Solving for w using the binding constraint (16) and substituting into (15), we get the maximization problem

$$\pi^D(j) = \max_{\lambda} m(\lambda)\, y(j) - \lambda\mathcal{U}. \tag{17}$$

The solution, denoted $\lambda_e^D(j)$, satisfies the first-order condition

$$m'\left(\lambda_e^D(j)\right) y(j) = \mathcal{U}, \tag{18}$$

and is therefore continuous and strictly increasing in j. In other words, $\lambda_e^D(j)$ denotes the queue length that a type-j firm *would* optimally attract *if it were* to attract directed searchers, taking as given market utility $\mathcal{U}$. From substituting (18) into (17), we see that maximized profit equals

$$\pi^D(j) = \left(1 - \epsilon_m\left(\lambda_e^D(j)\right)\right) m\left(\lambda_e^D(j)\right) y(j)\,. \tag{19}$$

Comparing (19) to (14), we conclude that firm j weakly prefers to attract directed searchers if and only if

$$\left(1 - \epsilon_m\left(\lambda_e^D(j)\right)\right) m\left(\lambda_e^D(j)\right) \geq m(\psi)\,. \tag{20}$$

By the concavity of m, the left-hand side of (20) is strictly increasing in $\lambda_e^D(j)$ and thus in j. This means that there is a unique threshold

$$j_e = \inf\left\{j : \left(1 - \epsilon_m\left(\lambda_e^D(j)\right)\right) m\left(\lambda_e^D(j)\right) \geq m(\psi)\right\}, \tag{21}$$

which, at an interior ($j_e > 0$) equilibrium, solves (20) with equality. The equilibrium queue length of firm j, denoted by $\lambda_e(j)$, then satisfies $\lambda_e(j) = \psi$ for $j < j_e$ and $\lambda_e(j) = \lambda_e^D(j)$ for $j \geq j_e$. Finally, the market utility $\mathcal{U}$ of directed searchers is pinned down by the market clearing condition

$$\int_0^1 \lambda_e(j)\, dj = 1. \tag{22}$$

To summarize, we have

Lemma 3. *The decentralized equilibrium allocation is characterized by a market utility $\mathcal{U}$, a threshold j_e, and a function $\lambda_e : [0,1] \to \mathbb{R}_+$ satisfying*

$$\begin{cases} \lambda_e(j) = \psi, & \text{if } j < j_e, \\ m'(\lambda_e(j))\, y(j) = \mathcal{U}, & \text{if } j \geq j_e, \end{cases} \tag{23}$$

as well as (21) and (22). There exists a decentralized equilibrium, and it is unique. In particular, $\lambda_e(j)$ is strictly increasing in j for $j \geq j_e$.

Proof. See Appendix A.3. □

An immediate corollary, which will be important below, is that the equilibrium queue length $\lambda_e(j)$ is necessarily discontinuous at j_e, as long as $j_e > 0$. This follows since the queue length at the threshold, $\Lambda_e \equiv \lambda_e(j_e)$, is the solution to the indifference condition $(1 - \epsilon_m(\Lambda_e))m(\Lambda_e) = m(\psi)$. Since $\epsilon_m(\lambda) > 0$ for any λ, this indifference condition requires $\Lambda_e > \psi$.

Corollary 1. *If $0 < j_e < 1$, the equilibrium queue length $\lambda_e(j)$ is discontinuous at $j = j_e$.*

This result has a very clear and interesting economic interpretation. Intuitively, suppose that firm j_e is indifferent between posting a wage of zero (hence attracting a queue length of ψ) and posting a strictly positive wage high enough to attract directed searchers. Since the latter entails a discrete increase in the wage, indifference requires a discrete increase in the queue length. We come back to this point below in the discussion of inefficiency.

5.2 Constrained inefficiency

We can now state and prove our main result, namely that the equilibrium allocation is not constrained efficient in general. Moreover, the inefficiency involves a misallocation toward high-productivity firms.

Proposition 1. *Assume $\psi \in (0,1)$. Let $\eta, \lambda_p(\cdot)$ be the constrained-efficient allocation, with the corresponding threshold j_p. Let $\mathcal{U}, \lambda_e(\cdot)$ be the decentralized equilibrium allocation, with the corresponding threshold j_e. Then the following hold: (i) $\mathcal{U} \leq \eta$, (ii) $j_e \geq j_p$, and (iii) $\lambda_e(j) \geq \lambda_p(j)$ for all $j \geq j_e$. Moreover, if either (8) holds or $(1 - \epsilon_m(1))\, m(1) < m(\psi)$, then the allocation is constrained inefficient and (i)–(iii) hold strictly.*

Proof. See Appendix A.4. □

The result states that, relative to the social optimum, too many firms choose to attract random searchers only. This result is best interpreted as an example of a monopsony distortion. To build some intuition, consider the tradeoff faced by a social planner. Specifically, suppose that a firm of type j switches from attracting only random searchers to attracting a small measure, Δ, of directed searchers as well. The social marginal benefit of this perturbation, net of the shadow value of the workers, is

$$m(\psi+\Delta)\,y(j)-\Delta\eta-m(\psi)\,y(j)\approx\Delta m'(\psi)\,y(j)-\Delta\eta \tag{24}$$

Now, consider the same tradeoff for a firm in the decentralized equilibrium. As pointed out in Corollary 1, in order to raise its queue length from ψ to $\psi+\Delta$, the firm needs to raise its wage from zero to w, where w satisfies $\frac{m(\psi+\Delta)}{\psi+\Delta}w=\mathcal{U}$. The private marginal benefit of doing so is then

$$m(\psi+\Delta)\,y(j)-(\psi+\Delta)\,\mathcal{U}-m(\psi)\,y(j)\approx\Delta m'(\psi)\,y(j)-(\psi+\Delta)\,\mathcal{U}. \tag{25}$$

The firm treats the marginal cost of attracting Δ more workers to be $(\psi+\Delta)\mathcal{U}$ rather than $\Delta\mathcal{U}$, because it must raise the wage to the random searchers as well in order to attract the additional directed searchers. The social planner, on the other hand, treats the random searchers as "free" to attract. This mechanism, clearly similar to a monopsony distortion when firms cannot wage-discriminate, means that firms have an incentive to under-hire. It also explains the discontinuity alluded to earlier in Corollary 1. Equation (25) shows that the firm's profits are discontinuous in Δ at $\Delta=0$, because the wage bill to random searchers jumps up discretely the moment the firm attracts any directed searchers; this means that a firm cannot be indifferent between attracting a queue of exactly ψ and a queue of $\psi+\Delta$ unlesss Δ is discretely larger than zero. In other words, the firm effectively behaves like a monopsonist facing a kinked labor supply curve.

This monopsony distortion, combined with market clearing, implies that firms who do attract directed searchers attract too many of them relative to the social optimum, and that there are too few such firms relative to the social optimum. Firms' reluctance to attract directed searchers drives down the shadow value of a directed searcher: firms who do attract directed searchers are willing to do so because these workers are "too cheap" ($\mathcal{U}<\eta$) in equilibrium.

These results are illustrated in Figure 1, which plots the queue length as a function of j both in equilibrium and in the constrained-efficient allocation. The queue length solving the planner's problem, $\lambda_p(j)$, equals ψ for $j < j_p$ and is strictly increasing thereafter. The equilibrium queue length, $\lambda_e(j)$, equals ψ for $j < j_e$, where the threshold j_e is strictly higher than j_p: relative to the social optimum, too many firms attract random searchers only. At j_e, the queue jumps to $\Lambda_e > \psi$, and is strictly increasing thereafter, with $\lambda_e(j) > \lambda_p(j)$.

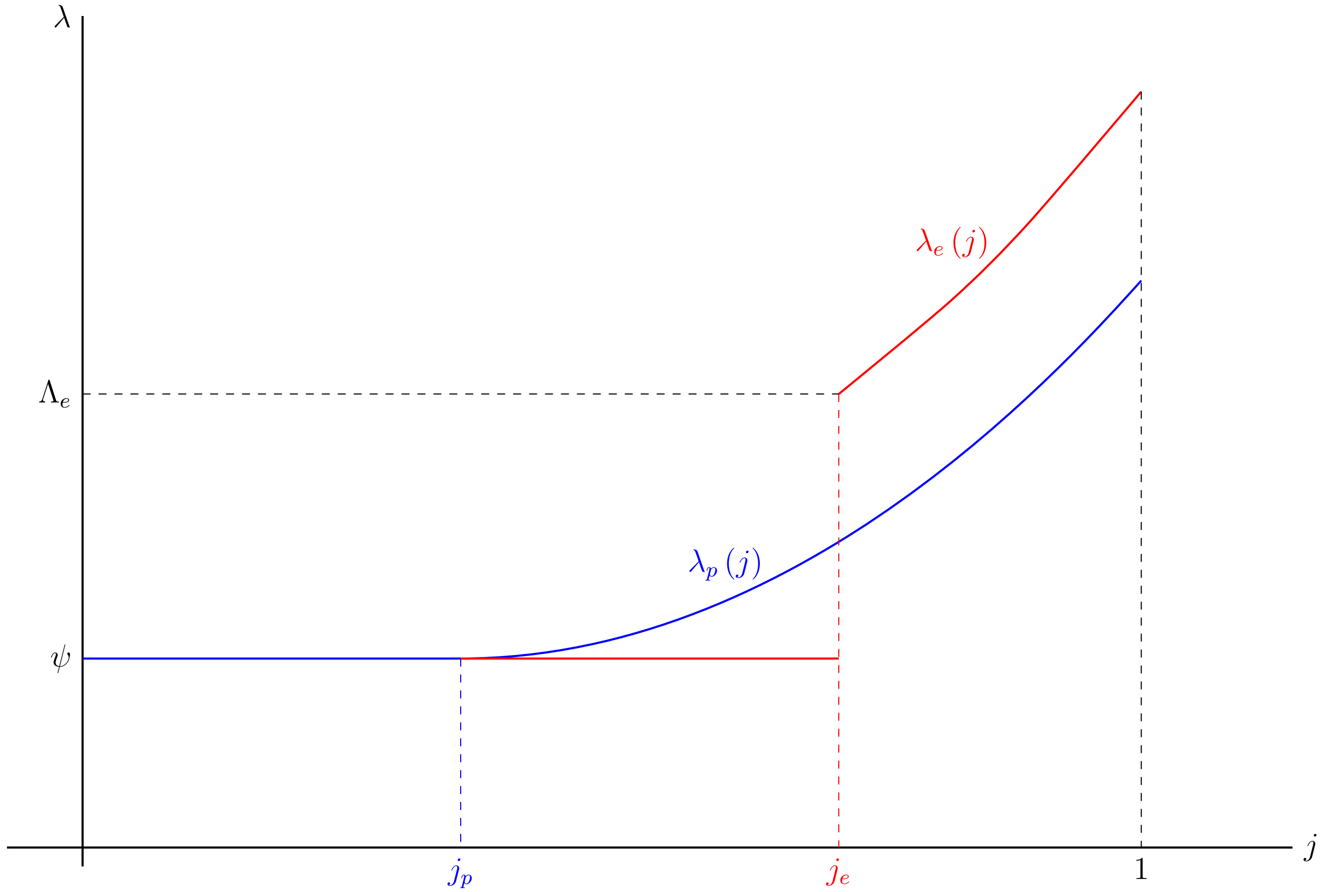

Figure 1: Equilibrium ($\lambda_e(j)$) vs. efficient allocation ($\lambda_p(j)$).

5.3 Special cases

The framework considered here has two key features: partially directed search and heterogeneous firms. It is instructive to show that the inefficiency emphasized here falls away in the extreme cases when either of these two features are absent.

Purely directed or purely random search. It is easy to see that partially directed search is essential for the misallocation inefficiency to occur: it disappears when search is either fully directed or fully random.

Corollary 2. *If either $\psi = 0$ or $\psi = 1$, the equilibrium allocation is constrained efficient.*

First, consider the special case when $\psi = 0$, i.e. search is purely directed. In this case, the equilibrium allocation is clearly constrained-efficient: we have $j_e = 0$, and $\mathcal{U}$ and $\lambda_e(j)$ are characterized by $m'(\lambda_e(j))\, y(j) = \mathcal{U}$ for each j, and $\int_0^1 \lambda_e(j)\, dj = 1$, which is exactly the solution to the social planner's problem.

More interestingly, the misallocation inefficiency also falls away when $\psi = 1$, i.e. search is purely random. In this case, we simply have $\lambda_e(j) = 1$ for all j. Workers are allocated randomly across firms of different productivities, but there is no inefficiency relative to the social planner's problem, because the inability to direct search *also constrains the social planner*. Partially directed search generates the misallocation inefficiency, because the presence of random searchers distorts the allocation of directed searchers relative to how the social planner would allocate them.

Homogeneous productivity. Next, we compare our results to an environment with homogeneous productivity, which is the special case handled in the previous literature, such as Lester (2011) and Bethune et al. (2020). Suppose that $y(j) = \overline{y}$ for all j. This is equivalent to a productivity distribution putting a probability of one on $\overline{y}$, and can be thought of as the limiting case of our model as $y(0) \to y(1)$.[5] In this case, the constrained-efficient allocation has $j_p = 0$ and $\lambda_p(j) = 1$ for all j (in particular, (8) trivially does not hold). This is because, by the concavity of the matching technology, the social planner would have all the workers searching in one submarket, and that submarket has queue length strictly larger than ψ. In other words, if all the firms are identical, the partial randomness of search is immaterial for the social planner.

To characterize equilibrium behavior, suppose first that the equilibrium has all the firms efficiently choosing identical queue lengths, $\lambda_e(j) = 1$. Each firm then attracts directed searchers, getting profits of $(1 - \epsilon_m(1))\, m(1)\, \overline{y}$. In order for this to be optimal, we must have $(1 - \epsilon_m(1))\, m(1) \geq m(\psi)$. If $(1 - \epsilon_m(1))\, m(1) < m(\psi)$, however, this cannot be an equilibrium, and the unique equilibrium must have firms randomizing between attracting directed searchers and not doing so. Because of homogeneous productivity, all the firms attracting directed searchers will have the same queue length, denoted Λ_e. Since the equilibrium has identical-productivity firms randomizing, we must have the indifference condition $(1 - \epsilon_m(\Lambda_e))\, m(\Lambda_e) = m(\psi)$. Without loss of generality, assume that firms with $j < j_e$ have $\lambda_e(j) = \psi$, and firms with $j \geq j_e$ have $\lambda_e(j) = \Lambda_e$; the threshold j_e must then satisfy the market-clearing condition $j_e \psi + (1 - j_e)\, \Lambda_e = 1$. This yields

[5] The analysis below closely follows Lester (2011); we include it for completeness and make no claims of originality here.

Corollary 3. *In the homogeneous-productivity case, the equilibrium takes one of two forms:*

1. *If $m(\psi) \leq (1 - \epsilon_m(1))\, m(1)$, the equilibrium allocation is constrained efficient: $j_e = 0$ and $\lambda_e(j) = 1$ for all j.*
2. *If $m(\psi) > (1 - \epsilon_m(1))\, m(1)$, the equilibrium allocation is constrained-inefficient; $\lambda_e(j) = \psi$ for $j < j_e$ and $\lambda_e(j) = \Lambda_e$ for $j \geq j_e$, where*

$$j_e = \frac{\Lambda_e - \psi}{\Lambda_e - 1} \tag{26}$$

and $(1 - \epsilon_m(\Lambda_e))\, m(\Lambda_e) = m(\psi)$.

With homogeneous firms, $(1 - \epsilon_m(1))\, m(1) \geq m(\psi)$ is both necessary and sufficient for constrained efficiency. Intuitively, when ψ is high, firms are indifferent between attracting directed searchers and exploiting random searchers; this leads to identical firms choosing different queue lengths, which is socially suboptimal. This can be thought of as a limiting case of our misallocation result.

5.4 Policy interventions: the minimum wage

The inefficiency identified in our analysis naturally raises the question of whether simple policy interventions can improve worker allocation and welfare. Our focus on the minimum wage is motivated by its prominence in policy debates surrounding growing employer market power. In particular, it is well known that employer market power can easily reverse theoretical predictions regarding the effects of the minimum wage on *total* employment (see e.g. Stigler, 1946; Bhaskar et al., 2002; Manning, 2011). Here, we identify a complementary mechanism through which the minimum wage may mitigate the *misallocation* of workers across firms.

Consider the effect of introducing a minimum wage $w_{\min}$. We will focus on parameters such that (8) holds. The definition of equilibrium is largely unchanged from Definition 1, except that firms' profit maximization in equation (12) is now performed subject to the constraint $w' \geq w_{\min}$. There are two cases to consider. A minimum wage below a threshold will bind for firms attracting only random searchers (who would otherwise pay a wage of zero), but will not bind for firms attracting directed searchers. In this case, the equilibrium queue length $\lambda^*(w)$ and market utility $\mathcal{U}$ are such that $\lambda^*(w_{\min}) = \psi$ and $\frac{m(\lambda^*(w_{\min}))}{\lambda^*(w_{\min})} w_{\min} < \mathcal{U}$. On the other hand, if the minimum wage is above the threshold, it will also bind for at least some firms attracting directed searchers. In this case, we have $\lambda^*(w_{\min}) \geq \psi$ (strict if the minimum wage is strictly above the threshold) and $\frac{m(\lambda^*(w_{\min}))}{\lambda^*(w_{\min})} w_{\min} = \mathcal{U}$. This is formalized in the following intuitive result:

Lemma 4. *A minimum wage $w_{\min}$ is binding for some firms attracting directed searchers if and only if $w_{\min} \geq \epsilon_m(\psi)\, y(j_p)$.*

Proof. See Appendix A.5. □

The threshold $\epsilon_m(\psi)\, y(j_p)$ is the "shadow" wage for the marginal firm attracting directed searchers in the constrained-efficient allocation. Intuitively, suppose that some of the firms attracting directed searchers pay the minimum wage; such firms would have a queue length strictly higher than ψ. Since a firm attracting only random searchers would still have to pay the minimum wage, there are no firms that attract only random searchers. In other words, the minimum wage is so high that the presence of ψ random searchers is immaterial for the equilibrium allocation.[6] In particular, this means that the minimum wage "overshoots" the planner's shadow wage.

For this reason, we focus on the case when $w_{\min} < \epsilon_m(\psi)\, y(j_p)$ and characterize the effects of a raise in the minimum wage in this range. We first observe that the equilibrium still has the cutoff property, whereby firms below some j_e attract random searchers only, obtaining a profit

$$\pi^R(j) = m(\psi)\left(y(j) - w_{\min}\right). \tag{27}$$

Since the minimum wage does not bind for any firm attracting directed searchers, the profit of such a firm continues to be $\pi^D(j)$ as specified in (19), where $\lambda_e^D(j)$ is determined by (18). The only difference compared to before is that the market utility $\mathcal{U}$ will have a different value.

Applying the envelope theorem to $\pi^D(j)$, we obtain

$$\frac{d}{dj}\left[\pi^D(j) - \pi^R(j)\right] = \left[m(\lambda_e(j)) - m(\psi)\right] y'(j) > 0. \tag{28}$$

This implies that $\pi^D(j) \geq \pi^R(j)$ if and only if $j \geq j_e$, where the unique cutoff j_e satisfies

$$m\left(\lambda_e^D(j_e)\right)\left(1 - \epsilon_m\left(\lambda_e^D(j_e)\right)\right) y(j_e) = m(\psi)\left(y(j_e) - w_{\min}\right). \tag{29}$$

The equilibrium is therefore pinned down by (18), (29), and the market-clearing condition.

We now consider the effect of raising the minimum wage. Taking as given the market utility, this has no effect on the expression for $\pi^D(j)$ in (19), but lowers the profits of a firm attracting only random searchers, given by (27). In equilibrium, this will induce more firms to attract directed searchers, lowering j_d and raising $\mathcal{U}$. This gives the following result:

[6] In fact, the equilibrium allocation under such a minimum wage coincides with the equilibrium allocation of an economy with $\psi = 0$.

Proposition 2. *Suppose there is a minimum wage $w_{\min} < \epsilon_m(\psi)\, y(j_p)$. The unique equilibrium is characterized by (18), (29), and the market-clearing condition (22). As long as it does not surpass $\epsilon_m(\psi)\, y(j_p)$, an increase in $w_{\min}$ (i) lowers j_d, (ii) raises $\mathcal{U}$, (iii) raises employment, and (iv) raises total welfare.*

Proof. See Appendix A.6. □

Intuitively, a small enough minimum wage is non-binding for firms already attracting directed searchers; however, it forces firms attracting solely random searchers to pay a higher wage. This lowers the opportunity cost of attracting directed searchers, inducing more firms to do so. The minimum wage thus reallocates some workers from firms with very high productivity to firms with medium productivity. Because there was too much congestion at high-productivity firms, this reallocation increases employment and welfare. It is worth noting that a minimum wage that is set too high will be efficiency-reducing. In particular, a minimum wage above $\epsilon_m(\psi)\, y(j_p)$ will bind for directed searchers, and therefore will result in inefficiently high queue lengths for low-productivity firms. In fact, the above analysis directly implies that the optimal minimum wage is precisely $\epsilon_m(\psi)\, y(j_p)$, which achieves the constrained-efficient outcome:

Corollary 4. *The constrained-efficient outcome is achieved by setting $w_{\min} = \epsilon_m(\psi)\, y(j_p)$. This minimum wage is increasing in ψ.*

Proof. See Appendix A.7. □

The last part of the claim, that the efficiency-restoring minimum wage is increasing in ψ, follows directly from the fact (shown in Appendix A.7) that $\epsilon_m(\psi)\, y(j_p)$ is increasing in ψ. Intuitively, a higher ψ raises the "shadow price" of an additional directed searcher, implying that the equilibrium wage paid to the marginal directed searcher needs to rise as well.

6 Extensive margin behavior

In this section, we extend our baseline environment by incorporating extensive-margin decisions: either through a participation decision by workers (Section 6.1) or an entry decision by firms (Section 6.2).[7] In doing so, we seek to connect our paper to the existing literature on Hosios-type distortions, where entry and/or participation *is* the key margin (Hosios, 1990;

[7]The key difference between the two is that, in the first case, the extensive-margin decision is made by the side of the market that is homogeneous, whereas in the second case it is made by the side that is heterogeneous. As we hope to show, the main message of both extensions is quite similar, and we include both for completeness.

Mangin and Julien, 2021). More specifically, we have a two-fold purpose in examining these extensions. First, we seek to show that the misallocation inefficiency we identify coexists alongside, and is distinct from, the conventional Hosios entry distortion. Second, we seek to understand how the Hosios entry distortion *itself* is manifested when search is partially directed.

6.1 Worker participation

We consider an environment in which worker participation in the market is endogenous. Suppose that workers have a cost of participating, z. We assume that each individual worker learns his/her status as random or directed searcher *after* making the participation decision.[8] Throughout, we assume that parameters are such that the solution is interior, i.e. there is a positive measure of workers participating, and a positive measure of firms attract only random searchers.[9]

Planner's problem. Letting u be the measure of participating workers, the planner's problem for this economy is now to choose both u and a function $\lambda : [0,1] \to \mathbb{R}_+$ to maximize

$$-zu + \int_0^1 m(\lambda(j))\, y(j)\, dj, \tag{30}$$

subject to the resource constraint

$$\int_0^1 \lambda(j)\, dj \leq u \tag{31}$$

and

$$\lambda(j) \geq \psi u \tag{32}$$

for all $j \in [0,1]$. The first constraint, (31), states that the measures of workers in all the submarkets cannot exceed the total measure of workers, u. The second constraint states that a fraction ψ of all the workers must be assigned to submarkets randomly, implying that the expected number of workers at each vacancy is at least ψu. The following lemma characterizes the solution to the planner's problem.

Lemma 5. *The constrained-efficient allocation is characterized by numbers u_p, η, and j_p*

[8] The opposite assumption would make the problem uninteresting, since random searchers would not participate, as they have zero bargaining power. At any rate, the assumption made presently is appropriate for illustrating the role of endogenous participation.

[9] This amounts to an assumption similar to condition (8) that we assumed in the baseline model.

and a function $\lambda_p : [0,1] \to \mathbb{R}_+$ *satisfying*

$$z = (1 - \psi j_p)\, \eta + \psi j_p m' (\psi u_p)\, \mathbb{E}\left[y(j) \,|j \leq j_p \right], \tag{33}$$

$$\int_0^1 \lambda_p(j)\; dj = u_p, \tag{34}$$

$$m' (\psi u_p)\, y(j_p) = \eta, \tag{35}$$

and

$$\begin{cases} \lambda_p(j) = \psi u_p, & \text{if } j < j_p, \\ m'(\lambda_p(j))\, y(j) = \eta, & \text{if } j \geq j_p. \end{cases} \tag{36}$$

Proof. See Appendix B.1. □

The social planner optimally chooses the measure of participating workers, u_p, and the threshold above which firms attract directed searchers, j_p. Note that conditions (34)–(36) are similar to our baseline model, as they characterize the constrained-efficient allocation of workers conditional on participating. Firms with $j < j_p$ attract random searchers only, and hence attract a queue of ψu_p. Firms with $j \geq j_p$ attract directed searchers in such a way that the marginal value of an extra worker, η, is equalized across $j \geq j_p$ submarkets.

The novel condition is (33), which determines the optimal level of participation. It equates the cost of participation, z, to the marginal benefit. With probability ψj_p, the new entrant is a random searcher *and* is allocated to a submarket in which firms attract random searchers only. In this case, the expected match surplus per matched worker in such a submarket is $\mathbb{E}\left[y(j)\,|j \leq j_p\right]$, and the entrant causes a congestion effect on the matching probability captured by $m'(\psi u_p)$. On the other hand, with probability $1 - \psi j_p$, the new entrant is allocated to a submarket with $j \geq j_p$, in which directed searchers are present. The marginal social value of adding such an entrant is η, which, as pointed out above, is endogenously equalized across directed-search submarkets.

Additional intuition, and connection to the Hosios condition. To better understand the conditions for constrained efficiency, and connect them to existing results in the literature, it is helpful to consider the two extreme cases. First, consider the case when search is purely random. In this case, $\psi = 1$ and therefore $j_p = 1$. All firms then attract the same queue length, $\lambda_p(j) = \psi u_p$, and the condition (33) pinning down u_p becomes $z = m'(u_p)\, \mathbb{E}[y(j)]$. The cost of participation equals the expected surplus per matched worker, multiplied by the effect of entry on the matching probability. As in Mangin and Julien (2021), it is instructive to rewrite this optimality condition in terms of the elasticity of the matching function:

$$\frac{z u_p}{m\left(u_p\right) \mathbb{E}\left[y\left(j\right)\right]} = \epsilon_m\left(u_p\right). \tag{37}$$

As expected, we obtain the standard Hosios condition. At the optimum, the worker's share in expected match surplus $\frac{m(u_p)}{u_p}\mathbb{E}\left[y\left(j\right)\right]$ must equal the elasticity of expected match surplus with respect to the number of entrants.

Now, consider the opposite extreme, in which search is purely directed. In this case, $\psi = 0$ and therefore $j_p = 0$. The shadow value of a worker must then equal the entry cost: $\eta = z$; for each j, the optimal queue is then given by $m'\left(\lambda\left(j\right)\right) y\left(j\right) = z$. We can again rewrite this in elasticity form as

$$\frac{z \lambda_p\left(j\right)}{m\left(\lambda_p\left(j\right)\right) y\left(j\right)} = \epsilon_m\left(\lambda_p\left(j\right)\right). \tag{38}$$

Thus, as emphasized by Mangin and Julien (2021), the constrained-efficient participation decision again satisfies the Hosios rule, which equates the elasticity of the matching technology to the cost of entry adjusted by expected match surplus (i.e. the worker's share in the expected match surplus). The difference between (37) and (38) is, of course, that in the random search case this optimality condition holds in expectation, whereas in the directed search case it holds separately for every j. In other words, under directed search with ex ante heterogeneity, the economy is endogenously divided into submarkets, with the appropriate optimal participation condition holding for each submarket. This result is well known from the directed search literature, such as Moen (1997); Shi (2009); Menzio and Shi (2010a,b, 2011). It is instructive to note here an obvious but important implication: the difference between (37) and (38) indicates that the inability to direct search constrains the social planner. The same intuition applies to the general formula in (33), which nests both (37) and (38) as special cases. Indeed, using the fact that $\eta = m'\left(\psi u_p\right) y\left(j_p\right)$, we can rewrite (33) in a similar form as

$$\frac{z \psi u_p}{m\left(\psi u_p\right)\left(\left(1 - \psi j_p\right) y\left(j_p\right) + \psi j_p \mathbb{E}\left[y\left(j\right) | j \leq j_p\right]\right)} = \epsilon_m\left(\psi u_p\right). \tag{39}$$

The social planner perceives the marginal benefit of an additional entrant to be expected match surplus for submarkets below j_p, but understands that it can allocate workers optimally, hence equalizing their marginal value, above j_p. Hence, it is still the case that the optimal participation decision satisfies the Hosios condition, whereby the expected surplus share of the marginal entrant equals the elasticity of the matching technology. What changes is the expected surplus itself, which depends on whether search is random, directed, or partially directed.

Equilibrium. We now turn to characterizing the decentralized equilibrium. For any given measure u_e of participating workers, the characterization of equilibrium is the same as in the baseline. Specifically, firms above some productivity cutoff j_e choose to attract directed searchers, whereas firms below this cutoff attract random searchers only. In such an equilibrium, a worker gets an endogenously determined level of market utility $\mathcal{U}$ if assigned to a submarket with $j \geq j_e$ and zero otherwise. Therefore, the ex-ante payoff of a participating worker (before their status as a random or directed searcher is revealed) is $(1-\psi j_e)\mathcal{U}$. In equilibrium, the measure of participating workers u_e is then such that this payoff equals the participation cost z, making the marginal worker indifferent between participating and staying out of the market. We therefore obtain the following result.

Lemma 6. *The decentralized equilibrium allocation is characterized by numbers u_e, $\mathcal{U}$, and j_e and a function $\lambda_e : [0,1] \to \mathbb{R}_+$ satisfying*

$$z = (1-\psi j_e)\mathcal{U}; \tag{40}$$

$$\int_0^1 \lambda_e\left(j\right)\, dj = u_e; \tag{41}$$

$$m\left(\lambda_e^D\left(j_e\right)\right)\left(1-\epsilon_m\left(\lambda_e^D\left(j_e\right)\right)\right) = m\left(\psi u_e\right), \tag{42}$$

where $\lambda_e^D\left(j\right)$ is the solution to $m'\left(\lambda_e^D\left(j\right)\right)y\left(j\right) = \mathcal{U}$; and

$$\begin{cases} \lambda_e\left(j\right) = \psi u_e, & \text{if } j < j_e, \\ m'\left(\lambda_e\left(j\right)\right)y\left(j\right) = \mathcal{U}, & \text{if } j \geq j_e. \end{cases} \tag{43}$$

Proof. See Appendix B.2. □

Conditions (41)–(43), which characterize the allocation of workers conditional on participating, are similar to the baseline model. The novel condition is (40), which characterizes the optimal participation decision.

Constrained inefficiency. Comparison of (40)–(43) to (33)–(36) shows that the equilibrium allocation now differs from the constrained-efficient one in two respects. First, comparison of (40) to (33) reveals that the equilibrium participation does not coincide with the social planner's optimal participation condition. Second, conditional on the number of participating workers, comparison of the remaining conditions demonstrates exactly the same misallocation inefficiency as uncovered in our baseline model. This is formalized in the following result.

Proposition 3. *Assume $\psi > 0$. Then equilibrium participation is too low: $u_e < u_p$. Furthermore, define $\{\lambda_o(j)\}$ as the solution to*

$$\max_{\{\lambda(j)\}} \int_0^1 m(\lambda(j))\, y(j)\, dj \quad s.t. \quad \int_0^1 \lambda(j)\, dj = u_e \quad and \quad \lambda(j) \geq \psi u_e. \tag{44}$$

There exists a threshold j_o such that

$$\begin{cases} \lambda_o(j) = \psi u_e, & if\ j < j_o, \\ m'(\lambda_o(j))\, y(j) = m'(\psi u_e)\, y(j_o), & if\ j \geq j_o. \end{cases} \tag{45}$$

If $\psi \in (0,1)$, we have $j_e \geq j_o$ and $\lambda_e(j) \geq \lambda_o(j)$ for $j > j_e$, with both inequalities strict if $j_o > 0$.

Proof. See Appendix B.3. □

The first part of Proposition 3 establishes that worker participation is too low relative to the constrained-efficient level, as long as search is not purely directed. The second part states that, for a *given* level of worker participation, it is possible to reallocate directed searchers across submarkets so as to increase total output. The first inefficiency corresponds to the standard Hosios inefficiency considered in the previous literature. When making the participation decision, a worker does not internalize its effects on a firm's matching probability. The second inefficiency is the misallocation inefficiency identical to the baseline model. Given the number of participating workers, too many firms choose to attract only random searchers, as they exercise their monopsony power by doing so.

Policy interventions. The presence of a two-fold inefficiency in the endogenous participation case implies that a single policy instrument, such as a minimum wage, no longer suffices to restore efficiency. We confirm this in the result below, which shows that efficiency can instead be achieved through an appropriate combination of a minimum wage and a participation tax.

Corollary 5. *The constrained-efficient outcome is achieved by setting a minimum wage $w_{\min} = \epsilon_m(\psi u_p)\, y(j_p)$ and a tax on participation equal to*

$$\tau = \psi j_p m'(\psi u_p)\left(y(j_p) - \mathbb{E}\left[y(j)\,|j \leq j_p\right]\right). \tag{46}$$

Proof. We will show that, under the proposed minimum wage and tax, $u_e = u_p$ and $j_e = j_p$ indeed satisfy the equilibrium conditions. First, assuming that $u_e = u_p$, the minimum wage results in the constrained-efficient allocation of workers, by the same logic as in Corollary 4. A firm of type j_p is now indifferent between attracting random searchers only and attracting directed searchers, since $m(\psi u_p)(y(j_p) - w_{\min}) = m(\psi u_p)(1 - \epsilon_m(\psi u_p))\,y(j_p)$. Second, this means that the expected utility of both a directed searchers and a random searcher is $m'(\psi u_p)y(j_p)$, and so the optimal participation condition is now

$$z = m'(\psi u_p)\,y(j_p) - \tau. \tag{47}$$

Substituting in for τ from (46) gives us (33). $\square$

The above also makes clear why simply setting a minimum wage would not suffice. Setting the minumum wage to eliminate misallocation of workers, taking as given the efficient participation level, requires $w_{\min} = \epsilon_m(\psi u_p)\,y(j_p)$. In contrast, by (39), achieving efficient participation for given j_p would require a minimum wage of $\epsilon_m(\psi u_p)\left[(1 - \psi j_p)\,y(j_p) + \psi j_p \mathbb{E}\left[y(j)\,|j \le j_p\right]\right]$. Clearly, the same minimum wage cannot satisfy both conditions at the same time; introducing a second policy instrument, such as a tax on participation, fixes this discrepancy.[10]

6.2 Firm entry

Next, we extend our baseline environment with endogenous entry by firms, showing that analogous results apply. To do so, we continue to assume that there is a measure one of firms indexed by $j \in [0,1]$, with productivity $y(j)$, which is continuous and strictly increasing in j. After learning its productivity, each firm decides whether or not to operate, at cost $\kappa > 0$. This endogenously determines the measure of vacancies, v. The rest of the environment is unchanged relative to the baseline model. In particular, since a fraction ψ of the workers are random searchers, the queue length at any firm cannot be less than ψ/v. Throughout, we again focus on parameters such that the solution is interior; i.e., a positive measure of firms enter, and a positive measure of the entrants attract only random searchers.

Planner's problem. The planner chooses the set of firms that operate and the allocation of workers among those firms. It is simple to show that, with regard to entry, the planner will follow a threshold rule, where firm j enters if and only if $j \ge j_p^*$; the total measure of vacancies is therefore $v = 1 - j_p^*$ and the queue of random searchers is $\psi/\left(1 - j_p^*\right)$. Thus,

[10] The constrained-efficient allocation can also be implemented in other ways, e.g. using an ex-post tax on hiring rather than an ex-ante tax on participation; the proof is straightforward and is available on request.

the planner's problem is to choose j_p^* and $\lambda(j)$ for each $j \in \left[j_p^*, 1\right]$ to maximize

$$\int_{j_p^*}^{1} \left[m\left(\lambda\left(j\right)\right) y\left(j\right) - \kappa\right] dj \tag{48}$$

subject to the constraints

$$\int_{j_p^*}^{1} \lambda\left(j\right) dj = 1 \tag{49}$$

and

$$\lambda\left(j\right)\left(1 - j_p^*\right) \geq \psi \tag{50}$$

for all $j \geq j_p^*$. As before, the constraint (50) will bind for j below some threshold j_p and will not bind for j above it. We then obtain the following characterization of the constrained-efficient allocation:

Lemma 7. *The constrained-efficient allocation is characterized by numbers η, j_p^*, $j_p > j_p^*$, and a function $\lambda_p : \left[j_p^*, 1\right] \to \mathbb{R}_+$ satisfying:*

$$m'\left(\psi / \left(1 - j_p^*\right)\right) y\left(j_p\right) = \eta, \tag{51}$$

$$\begin{cases} \lambda_p\left(j\right) = \psi / \left(1 - j_p^*\right), & \text{if } j \in \left[j_p^*, j_p\right), \\ m'\left(\lambda_p\left(j\right)\right) y\left(j\right) = \eta, & \text{if } j \in \left[j_p, 1\right], \end{cases} \tag{52}$$

$$\int_{j_p^*}^{1} \lambda_p\left(j\right) dj = 1, \tag{53}$$

and

$$\begin{aligned} \kappa = \; & m\left(\lambda_p\left(j_p^*\right)\right)\left(1 - \epsilon_m\left(\lambda_p\left(j_p^*\right)\right)\right) y\left(j_p^*\right) \\ & - \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(\mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\} | j \geq j_p^*\right] - y\left(j_p^*\right)\right). \end{aligned} \tag{54}$$

Proof. See Appendix B.4. □

Given j_p^*, conditions (51)–(53) characterize the socially optimal $\lambda_p(j)$, j_p, and η, similarly to the exogenous entry case described in Lemma 1. Condition (51) pins down the threshold j_p for attracting directed searchers. Condition (52) characterizes the queue length for every firm that operates, and Condition (53) is the resource constraint.

The new equation is Condition (54), which states that the marginal benefit of adding an additional firm equals the marginal cost. The marginal cost of an additional entrant is κ. The marginal benefit is the expected match surplus of that extra firm, adjusted for the crowding-out of hiring by other firms due to congestion in the matching function. Notably, this marginal benefit consists of two terms. The first term, $m\left(\lambda_p\left(j_p^*\right)\right)\left(1 - \epsilon_m\left(\lambda_p\left(j_p^*\right)\right)\right) y\left(j_p^*\right)$, captures the fact that the marginal entrant mechanically adds to total output, but also crowds out hiring for other firms. The second term, $-\lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(\mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\}|j \geq j_p^*\right] - y\right.$ captures the fact that – as long as search is not perfectly directed – the marginal entrant crowds out firms that are *more productive*. This second term falls away if all the firms are homogeneous; in this case, only the standard congestion effect remains. This second term also falls away if search is purely directed, because then each productivity level is searching in a separate submarket, and hence only crowds out firms of the same productivity.

Additional intuition, and connection to the Hosios condition. As before, it is instructive to consider the two extreme cases of purely random and purely directed search, and to compare our results to what is known in the literature on the Hosios condition, e.g. Mangin and Julien (2021). First, consider the case when $\psi = 1$, so that search is purely random. In this case, the constrained-efficient allocation has $j_p = 1$, $\lambda_p(j) = 1/\left(1 - j_p^*\right) \equiv \lambda_p$ for every $j \geq j_p^*$, and the socially optimal entry condition characterizing j_p^* becomes

$$\kappa = m\left(\lambda_p\right)\left(1 - \epsilon_m\left(\lambda_p\right)\right) y\left(j_p^*\right) - \lambda_p m'\left(\lambda_p\right)\left(\mathbb{E}\left[y\left(j\right)|j \geq j_p^*\right] - y\left(j_p^*\right)\right). \tag{55}$$

Rearranging, we can rewrite this expression as

$$\frac{\kappa}{m(\lambda_p)\mathbb{E}\left[y\left(j\right)|j \geq j_p^*\right]} = \left(1 - \epsilon_m\left(\lambda_p\right)\right) - \left(1 - \frac{y\left(j_p^*\right)}{\mathbb{E}\left[y\left(j\right)|j \geq j_p^*\right]}\right). \tag{56}$$

This expression is an example of the generalized Hosios condition of Mangin and Julien (2021). The left-hand side is the entry cost adjusted by the expected match surplus. The right-hand side is the sum two elasticities. The first term is the elasticity of the matching probability with respect to the number of entering firms. The second term is the elasticity of expected match surplus with respect to the number of entering firms. The latter elasticity is nonzero because increasing the number of entrants amounts to lowering j_p^*, which lowers the average productivity in the market; i.e., by entering, a firm generates an output externality.[11]

On the other extreme, consider the case of purely directed search, i.e. $\psi = 0$. In this case, the constrained-efficient allocation has $j_p = j_p^*$, the queue length solves $m'\left(\lambda_p\left(j\right)\right) y\left(j\right) = \eta$ for every $j \geq j_p^*$, and the entry threshold solves

$$\kappa = m\left(\lambda_p\left(j_p^*\right)\right)\left(1 - \epsilon_m\left(\lambda_p\left(j_p^*\right)\right)\right) y\left(j_p^*\right). \tag{57}$$

[11] An output externality is present in the model with firm entry (i.e. (56)) but not with worker participation (i.e., (37)) because in the former, there is heterogeneity on the side of the market making the entry decision; i.e., when a firm decides to enter, it changes the productivity composition of firms in the market.

Each productivity level is now assigned a distinct submarket, and workers are allocated optimally among these submarkets, as is standard in the directed search literature (Moen, 1997; Shi, 2009; Menzio and Shi, 2010a,b, 2011). As a result, the social planner only needs to take into account that the marginal entrant crowds out peers of *its own* productivity. Rearranging equation (57), we can rewrite it in elasticity form as

$$\frac{\kappa}{m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right)} = 1 - \epsilon_m\left(\lambda_p\left(j_p^*\right)\right). \tag{58}$$

The worker's share in expected surplus equals the elasticity of the matching probability with respect to the number of entrants. Because search is directed, there is no output externality in this case, hence no second term to correct for it.

The same reasoning applies to the general case where search is partially directed. Condition (54) can be rewritten in the form of a generalized Hosios condition as

$$\frac{\kappa}{m\left(\lambda_p\left(j_p^*\right)\right) \mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\} | j \geq j_p^*\right]} = 1 - \epsilon_m\left(\lambda_p\left(j_p^*\right)\right) - \left[1 - \frac{y\left(j_p^*\right)}{\mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\} | j \geq j_p^*\right]}\right]. \tag{59}$$

There is an output externality, but only in the productivity range where no directed searchers are attracted, as captured by the term in brackets.

Equilibrium. We now consider the decentralized equilibrium. Each firm j decides whether or not to enter and, conditional on entering, whether to attract directed searchers or not. Since the profits of a firm are strictly increasing in j, firms will follow a threshold rule for entering: firms will enter if and only if $j \geq j_e^*$. Conditional on entering, the problem of a firm is the same as in the exogenous-entry case, except that the queue of random searchers is $\equiv \psi / \left(1 - j_e^*\right)$ instead of ψ. We can establish the following equilibrium characterization:

Lemma 8. *The equilibrium allocation is characterized by numbers $\mathcal{U}$, j_e^*, $j_e > j_e^*$, and a function $\lambda_e : [j_e^*, 1] \to \mathbb{R}_+$ satisfying:*

$$m\left(\lambda_e\left(j_e\right)\right)\left(1 - \epsilon_m\left(\lambda_e\left(j_e\right)\right)\right) = m\left(\psi / \left(1 - j_e^*\right)\right), \tag{60}$$

$$\begin{cases} \lambda_e\left(j\right) = \psi / \left(1 - j_e^*\right), & \text{if } j \in [j_e^*, j_e), \\ m'\left(\lambda_e\left(j\right)\right) y\left(j\right) = \mathcal{U}, & \text{if } j \in [j_e, 1], \end{cases} \tag{61}$$

$$\int_{j_e^*}^{1} \lambda_e\left(j\right)\, dj = 1, \tag{62}$$

and

$$\kappa = m\left(\lambda_e\left(j_e^*\right)\right) y\left(j_e^*\right). \tag{63}$$

Proof. See Appendix B.5. $\square$

Conditions (60)–(62) are analogous to the equilibrium characterization for the exogenous-entry case. Condition (63) is a free-entry condition, which says that the profit of the marginal firm equals the entry cost. One can solve these conditions in two steps. First, the free-entry condition determines the equilibrium threshold for entry j_e^*, and therefore $\psi/(1-j_e^*)$, in isolation from other equilibrium outcomes. Second, given this level of entry, the solution for $\mathcal{U}$, j_e and $\lambda_e(j)$ can be obtained in the same way as in Lemma 3.

Constrained inefficiency. Analogous to Section 6.1, comparison of the equilibrium conditions (60)–(63) to the efficiency conditions (51)–(54) reveals that the decentralized equilibrium is inefficient on two margins. First, there is a Hosios-like entry inefficiency: as long as search is note purely directed, entry is excessive relative to the social optimum, because an individual firm does not internalize either the crowding-out effect on other firms or the fact that those firms are more productive. Second, conditional on a level of entry, there is misallocation of labor in the same way as in our baseline model: because of the monopsony distortion, too few firms target random searchers, resulting in too many applicants at the higher-productivity firms. These results are formalized in the following proposition:

Proposition 4. *Assume $\psi > 0$. Then equilibrium entry is too high: $j_e^* < j_p^*$. Furthermore, define $\{\lambda_o(j)\}$ as the solution to*

$$\max_{\lambda(j)} \int_{j_e^*}^{1} m(\lambda(j))\, y(j)\, dj \quad s.t. \quad \int_{j_e^*}^{1} \lambda(j)\, dj = 1 \quad and \quad \lambda(j) \geq \psi/(1-j_e^*). \tag{64}$$

There exists a threshold j_o such that

$$\begin{cases} \lambda_o(j) = \psi/(1-j_e^*), & \text{if } j < j_o, \\ m'(\lambda_o(j))\, y(j) = m'(\psi/(1-j_e^*))\, y(j_o), & \text{if } j \geq j_o. \end{cases} \tag{65}$$

If $\psi \in (0,1)$, we have $j_e \geq j_o$ and $\lambda_e(j) \geq \lambda_o(j)$ for $j > j_e$, with both inequalities strict if $j_o > 0$.

Proof. See Appendix B.6. $\square$

Policy interventions. As in Section 6.1, the presence of a two-fold inefficiency in the environment with endogenous entry implies that a minimum wage alone is not sufficient to restore efficiency. Instead, the constrained-efficient allocation is implemented through a minimum wage combined with an appropriate entry subsidy. We formalize this in the result below.

Corollary 6. *The constrained-efficient outcome is achieved by setting* $w_{\min} = \epsilon_m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p\right)$ *and an entry subsidy*

$$\mathcal{X} = \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(y\left(j_p\right) - \mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\} | j \geq j_p^*\right]\right). \tag{66}$$

Proof. We will show that, under the proposed minimum wage and tax, $j_e^* = j_p^*$ and $j_e = j_p$ indeed satisfy the equilibrium conditions. First, assuming that $j_e^* = j_p^*$, the minimum wage results in the constrained-efficient allocation of workers, by the same logic as in Corollary 4. A firm of type j_p is now indifferent between attracting random searchers only and attracting directed searchers, since $m\left(\lambda_p\left(j_p^*\right)\right)\left(y\left(j_p\right) - w_{\min}\right) = m\left(\lambda_p\left(j_p^*\right)\right)\left(1 - \epsilon_m\left(\lambda_p\left(j_p^*\right)\right)\right) y\left(j_p\right)$. Second, the optimal entry condition is now

$$\begin{aligned}
\kappa = \; & m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right) - \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p\right) + \chi \\
= \; & m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right) - \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right) \\
& + \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(y\left(j_p^*\right) - \mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\} | j \geq j_p^*\right]\right),
\end{aligned} \tag{67}$$

which is equivalent to (54). □

The minimum wage aligns private and social incentives for a firm of type j_p, undoing the misallocation distortion. The entry subsidy undoes the Hosios inefficiency by aligning private and social incentives for a firm of type j_p^*. Setting only the former would have led to insufficient entry (intuitively, this would be forcing the marginal entrant, whose productivity is j_p^*, to pay the wage appropriate for a firm of productivity j_p). This illustrates once again that there are two distinct distortions, which require two policy instruments to correct for them.

7 Worker heterogeneity

The preceding analysis has assumed heterogeneous productivity on only one side of the market. In the Online Appendix, we show that our misallocation result carries over to an environment in which both firms and workers are heterogeneous in productivity. Analysis of such an environment is complicated, because it needs to address the question of sorting, i.e., which worker type matches with each firm type. As the literature has already shown in the context of purely directed search, the answer to this question generally depends on the details of the environment, including the specification of the matching function and the complementarities in production between workers and firms (see e.g. Eeckhout and Kircher, 2010). We introduce partially directed search into such an environment with two-sided heterogeneity, and examine its implications for the misallocation inefficiency. We establish the robustness of the two main inefficiencies of the baseline model: (1) too many firms target random searchers, and (2) too many workers apply to the firms at the top of the productivity distribution. Moreover, we argue that the intuition for these results, based on the monopsony distortion, carries over naturally to the model with two-sided heterogeneity. Below, we provide a brief summary of this analysis, with all the details relegated to the Online Appendix.

We assume that the surplus of a match is the product xy, where x is the productivity of the worker, and y is the productivity of the firm. Directed searcher workers can be either high or low productivity: $x \in \{x_1, x_2\}$, $x_1 < x_2$.[12] The mean productivity among random searcher workers is $\bar{x}$. Supposing that a firm of type y attracts directed searchers of type x and attracts, in total, a queue of size λ, its expected output is

$$m(\lambda)\left[\frac{\lambda-\psi}{\lambda}x + \frac{\psi}{\lambda}\bar{x}\right]y \tag{68}$$

The expression in brackets illustrates that λ now affects expected output conditional on a match. Attracting a higher queue length also results in attracting more workers of type x rather than $\bar{x}$. Because of two-sided heterogeneity, we need to characterize not only which firms attract directed searchers, but also – conditional on attracting directed searchers – which worker type each firm y attracts. In other words, we need to characterize the sorting of workers to firms conditional on attracting directed searchers. For the case when search is purely directed, existing literature (Eeckhout and Kircher (2010)) provides sufficient conditions on the production function and matching function such that sorting is positive (i.e. higher-y firms attract higher-x workers). We make analogous assumptions on the expression in (68) such that positive sorting also obtains in our, partially directed search, environment.

We show that the equilibrium allocation, as well as the constrained-efficient allocation, is characterized by two cutoffs for firm productivity. We denote the constrained-efficient cutoffs by $j_{p,1}$, $j_{p,2}$ and the equilibrium cutoffs by $j_{e,1}$, $j_{e,2}$. Firms below the first cutoff attract only random searchers. Firms above the first cutoff but below the second attract directed searchers of type x_1. Firms above the second cutoff attract directed searchers of type x_2. In other words, there continues to be a cutoff rule for attracting directed searchers, and there is positive assortative matching conditional on doing so.

[12] The model can be extended easily to more than two worker types: the results do not depend on the assumption of two types, which we make for simplicity.

We next show that in equilibrium, both cutoffs are too high relative to the constrained-efficient allocation: $j_{e,1} > j_{p,1}$ and $j_{e,2} > j_{p,2}$. This has two implications. First, in equilibrium too few firms attract directed searchers. Second, too few firms attract directed searchers of type x_2. The latter implies that the firms who do attract workers of type x_2 attract too many of them, and hence we again have too many applicants at the top of the productivity distribution.

It is useful to elaborate on the mechanism driving the results. The baseline model above has made clear that the misallocation inefficiency is related to a monopsony distortion. The logic of the baseline model extends to the environment with worker heterogeneity as follows. The same monopsony distortion – conditional on attracting directed searchers at all – makes firms inefficiently reluctant to attract high-productivity workers rather than low-productivity workers. Consider a firm deciding whether to attract type-x_1 or type-x_2 workers. In order to attract the higher-productivity workers, the firm must raise the offered wage. But this requires it to also raise the wage for the random searchers, whose productivity is unaffected. In other words, the possibility of attracting a random searcher dampens the firm's expected benefit (in terms of increased productivity) from raising the wage. This makes firms inefficiently reluctant to attract type x_2 rather than type x_1. In turn, all the high-productivity workers now concentrate at the very high-productivity firms, leading to even more crowding at the top of the firm productivity distribution.

8 Conclusion

This paper identifies a novel inefficiency in search models, which is distinct from from the much-studied entry distortion and operates instead on the allocation margin. This novel inefficiency arises when random and directed search coexist, and goes in an arguably unexpected direction: relative to the constrained-efficient allocation, too many workers apply to high-productivity firms. This result serves as a cautionary note against treating misallocation of resources as synonymous with low productivity. The misallocation highlighted here manifests itself instead as suboptimally low employment. A minimum wage can increase employment and welfare, but in our framework it does so by reallocating workers across firms rather than by drawing more workers into the labor market.

There are a number of directions for future research. First, we have assumed that the fraction of directed searchers is exogenously given. If workers can decide, at a cost, whether or not to direct their search, potential strategic complementarities can arise between workers' investments in information and firms' wage posting decisions, possibly leading to multiple equilibria. Moreover labor market policies can now affect employment, output and welfare not only directly, but also by changing the incentives to acquire information. Second, our analysis has been theoretical: to this end, the model is deliberately parsimonious and stylized. Quantifying the model's implications for employment, welfare, measured matching efficiency, and the effects of policies is a promising but challenging research agenda, which would surely require a dynamic model, most likely extended to allow on-the-job search. A key challenge is identifying the degree to which search is directed; Lentz and Moen (2017) represent progress in this dimension.

References

Acemoglu, D. (2001). Good Jobs versus Bad Jobs. *Journal of Labor Economics*, 19(1):1–21.

Acemoglu, D. and Shimer, R. (1999). Efficient unemployment insurance. *Journal of Political Economy*, 107(5):893–928.

Albrecht, J., Navarro, L., and Vroman, S. (2010). Efficiency in a search and matching model with endogenous participation. *Economics Letters*, 106(1):48–50.

Bertola, G. and Caballero, R. J. (1994). Cross-sectional efficiency and labour hoarding in a matching model of unemployment. *The Review of Economic Studies*, 61(3):435–456.

Bethune, Z., Choi, M., and Wright, R. (2020). Frictional goods markets: Theory and applications. *The Review of Economic Studies*, 87(2):691–720.

Bhaskar, V., Manning, A., and To, T. (2002). Oligopsony and monopsonistic competition in labor markets. *Journal of Economic Perspectives*, 16(2):155–174.

Cheremukhin, A., Restrepo-Echavarria, P., and Tutino, A. (2020). Targeted search in matching markets. *Journal of Economic Theory*, 185:1–43.

Davis, S. J. (2001). The quality distribution of jobs and the structure of wages in search equilibrium. Technical report, National Bureau of Economic Research.

Eeckhout, J. and Kircher, P. (2010). Sorting and decentralized price competition. *Econometrica*, 78:539–574.

Galenianos, M., Kircher, P., and Virag, G. (2011). Market power and efficiency in a search model. *International Economic Review*, 52:85–104.

Golosov, M., Maziero, P., and Menzio, G. (2013). Taxation and redistribution of residual income inequality. *Journal of Political Economy*, 121(6):1160–1204.

Hosios, A. K. (1990). On the efficiency of matching and related models of search and unemployment. *Review of Economic Studies*, 57:279–298.

Julien, B. and Mangin, S. (2017). Efficiency of job creation in a search and matching model with labor force participation. *Economics Letters*, 150:149–151.

Lentz, R. and Moen, E. R. (2017). Competitive or random search?

Lester, B. (2011). Information and prices with capacity constraints. *American Economic Review*, 101(4):1591–1600.

Mangin, S. and Julien, B. (2021). Efficiency in search and matching models: A generalized Hosios condition. *Journal of Economic Theory*, 193(C).

Manning, A. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Manning, A. (2011). Imperfect competition in the labor market. *Handbook of Labor Economics*, 4:973–1041.

Masters, A. (2015). Efficiency in a search and matching model with participation policy. *Economics Letters*, 134:111–113.

Menzio, G. and Shi, S. (2010a). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, 145(4):1453–1494.

Menzio, G. and Shi, S. (2010b). Directed search on the job, heterogeneity and aggregate fluctuations. *American Economic Review*, 100(2):327–332.

Menzio, G. and Shi, S. (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, 119:468–510.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of Political Economy*, 105:385–411.

Mukoyama, T. (2019). Heterogeneous jobs and the aggregate labour market. *The Japanese Economic Review*, 70(1):30–50.

Pigou, A. (1932). *The Economics of Welfare*. Palgrave Classics in Economics. Palgrave Macmillan.

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory*. Oxford University Press.

Robinson, J. (1933). *The Economics of Imperfect Competition*. London, Macmillan and Co., Ltd.

Shi, S. (2009). Directed search for equilibrium wage-tenure contracts. *Econometrica*, 77:561–584.

Shi, S. (2018). Sequentially mixed search and equilibrium price dispersion.

Stigler, G. (1946). The economics of minimum wage legislation. *American Economic Review*, 36:358–365.

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. (2021). Directed search: A guided tour. *Journal of Economic Literature*, 59(1):90–148.

Wu, L. (2021). Partially Directed Search in the Labor Market.

A Proofs

A.1 Proof of Lemma 1

The fact that the solution satisfies (5), (6), and (7) is proved in the text. It remains to prove existence and uniqueness. Because (5) and (7) define j_p and $\lambda_p(j)$ as functions of η, the left-hand side of (6) is a function of a single variable, η. Existence follows because the left-hand side of (6) is continuous in η, approaches ψ as $\eta \to \infty$, and approaches infinity as $\eta \to 0$. To prove uniqueness, we will show that the left-hand side of (6) is strictly decreasing in η. Suppose $\widehat{\eta} > \eta$; the corresponding $\widehat{j}_p$ and j_p then satisfy $\widehat{j}_p \geq j_p$, while the corresponding $\widehat{\lambda}_p(j)$ and $\lambda_p(j)$ satisfy $\widehat{\lambda}_p(j) \leq \lambda_p(j)$, with stricty inequality when $j > j_p$. Hence,

$$\begin{aligned}
\widehat{j}_p\psi + \int_{\widehat{j}_p}^1 \widehat{\lambda}_p(j)\, dj &< \widehat{j}_p\psi + \int_{\widehat{j}_p}^1 \lambda_p(j)\, dj \\
&= j_p\psi + \int_{j_p}^1 \lambda_p(j)\, dj + \int_{j_p}^{\widehat{j}_p} (\psi - \lambda_p(j))\, dj \\
&\leq j_p\psi + \int_{j_p}^1 \lambda_p(j)\, dj,
\end{aligned} \tag{A.1}$$

where the last inequality used the fact that $\psi < \lambda_p(j)$ for $j > j_p$.

A.2 Proof of Lemma 2

Recall that $\lambda(j)$ solves $m'(\lambda(j))\, y(j) = m'(\psi)\, y(0)$ for each j while $\lambda_p(j)$ satisfies (7). Since m is strictly concave, $\lambda(j)$ is strictly increasing. Therefore, $\lambda(j) \geq \psi$ for any j. Since the left-hand side of the resource constraint (6) is strictly decreasing in η, this implies that

$$\begin{aligned}
1 &= \int_0^1 \max\{\psi, \lambda_p(j)\}\, dj \\
&< \int_0^1 \max\{\psi, \lambda(j)\}\, dj \\
&= \int_0^1 \lambda(j)\, dj,
\end{aligned} \tag{A.2}$$

if and only if $\eta > m'(\psi)\, y(0)$, which in turn is equivalent to $j_p > 0$.

A.3 Proof of Lemma 3

The proof is similar to the proof of Lemma 1. Because $\lambda_e(j)$ satisfies (23) and j_e satisfies (21), the left-hand side of the market-clearing condition (22) is a function of a single variable,

$\mathcal{U}$. Existence follows because the left-hand side is continuous in $\mathcal{U}$, approaches ψ as $\mathcal{U} \to \infty$, and approaches infinity as $\mathcal{U} \to 0$. To prove uniqueness, we will show that the left-hand side of (22) is strictly decreasing in $\mathcal{U}$. Note that, from (23), we can write (22) as

$$j_e \psi + \int_{j_e}^{1} \lambda_e(j)\, dj = 1 \tag{A.3}$$

For $j \geq j_e$, $\lambda_e(j)$ is a decreasing function of $\mathcal{U}$ by (18), while j_e is increasing in $\mathcal{U}$ by (21). Suppose $\widehat{\mathcal{U}} > \mathcal{U}$; the corresponding $\widehat{j}_e$ and j_e satisfy $\widehat{j}_e \geq j_e$, while the corresponding $\widehat{\lambda}_e(j)$ and $\lambda_e(j)$ satisfy $\widehat{\lambda}_e(j) \leq \lambda_e(j)$, with strict inequality when $j > j_e$. Hence,

$$\begin{aligned}
\widehat{j}_e \psi + \int_{\widehat{j}_e}^{1} \widehat{\lambda}_e(j)\, dj &< \widehat{j}_e \psi + \int_{j_e'}^{1} \lambda_e(j)\, dj \\
&= j_e \psi + \int_{j_e}^{1} \lambda_e(j)\, dj + \int_{j_e}^{\widehat{j}_e} (\psi - \lambda_e(j))\, dj \\
&\leq j_e \psi + \int_{j_e}^{1} \lambda_e(j)\, dj.
\end{aligned} \tag{A.4}$$

The last inequality has used the fact that, for $j \geq j_e$, we have $(1 - \epsilon_m(\lambda_e(j)))\, m(\lambda_e(j)) > m(\psi)$; since $\epsilon_m(\lambda) > 0$ for all λ, this implies $m(\lambda_e(j)) > m(\psi)$ and therefore $\lambda_e(j) > \psi$.

A.4 Proof of Proposition 1

We first consider the case of an interior solution. Suppose that (8) holds, so that $j_p > 0$. We will show that $j_e > j_p$. The proof is by contradiction, in three steps. First, we show that $j_e \leq j_p$ implies $\mathcal{U} < \eta$. Suppose that $j_e \leq j_p$. Defining $\Lambda_e = \lambda_e(j_e)$ as in the text, it follows that $(1 - \epsilon_m(\Lambda_e))m(\Lambda_e) \geq m(\psi)$ and therefore $\Lambda_e > \psi$ strictly (this is just a re-statement of Corollary 1). But then, we have

$$\begin{aligned}
\mathcal{U} &= m'(\Lambda_e)\, y(j_e) \\
&< m'(\psi)\, y(j_e) \\
&\leq m'(\psi)\, y(j_p) \\
&= \eta.
\end{aligned} \tag{A.5}$$

Second, we show that $\mathcal{U} < \eta$ implies $\lambda_e(j) > \lambda_p(j)$ for all $j \geq j_e$. For $j \in [j_e, j_p]$, $\lambda_e(j) > \psi = \lambda_p(j)$. Furthermore, for all $j \geq j_p$, $\lambda_e(j)$ is given by the solution to $m'(\lambda_e(j))y(j) = \mathcal{U}$, and $\lambda_p(j)$ is given by the solution to $m'(\lambda_e(j))y(j) = \eta$, so that $\mathcal{U} < \eta$ implies $\lambda_e(j) > \lambda_p(j)$ by the concavity of m. Third, we show that $\lambda_e(j) > \lambda_p(j)$ for $j \geq j_e$ together with $j_e \leq j_p$

implies a violation of the resource constraint. We have

$$
\begin{aligned}
1 &= j_p\psi + \int_{j_p}^{1} \lambda_p\left(j\right)\, dj \\
&= j_e\psi + (j_p - j_e)\psi + \int_{j_p}^{1} \lambda_p\left(j\right)\, dj \\
&< j_e\psi + \int_{j_e}^{1} \lambda_e\left(j\right)\, dj.
\end{aligned}
\tag{A.6}
$$

This is a contradiction, so we must have $j_e > j_p$. Next, we show that $\mathcal{U} < \eta$. If $\mathcal{U} \geq \eta$, we have $\lambda_e\left(j\right) \leq \lambda_p\left(j\right)$ for all $j \geq j_e$, and so

$$
\begin{aligned}
1 &= j_p\psi + \int_{j_p}^{1} \lambda_p\left(j\right) dj \\
&= j_p\psi + \int_{j_p}^{j_e} \lambda_p\left(j\right) dj + \int_{j_e}^{1} \lambda_p\left(j\right) dj \\
&> j_p\psi + \int_{j_e}^{1} \lambda_e\left(j\right) dj,
\end{aligned}
\tag{A.7}
$$

contradicting market clearing once again. This implies that $\mathcal{U} < \eta$ and therefore $\lambda_e\left(j\right) > \lambda_p\left(j\right)$ for $j \geq j_e$, as had to be shown.

Finally, consider the case when (8) does not hold, and so the constrained-efficient allocation has $j_p = 0$. In this case, a necessary and sufficient condition for the equilibrium to be constrained efficient is $\left(1 - \epsilon_m\left(\lambda_p\left(0\right)\right)\right) m\left(\lambda_p\left(0\right)\right) \geq m\left(\psi\right)$. If this is violated, then in equilibrium we must have $\lambda_e\left(0\right) = \psi$ and therefore $j_e > 0$. This implies $\mathcal{U} < \eta$, since, as already argued above, $\mathcal{U} \geq \eta$ together with $j_e > j_p$ would violate the resource constraint. Since $\mathcal{U} < \eta$, the definitions of $\lambda_p\left(j\right)$ and $\lambda_e\left(j\right)$ imply $\lambda_e\left(j\right) > \lambda_p\left(j\right)$ for all $j \geq j_e$. Finally, note that the resource constraint (6) requires $\lambda_p\left(0\right) < 1$. Since $\left(1 - \epsilon_m\left(\lambda\right)\right) m\left(\lambda\right)$ is increasing in λ, $\left(1 - \epsilon_m\left(1\right)\right) m\left(1\right) < m\left(\psi\right)$ implies $\left(1 - \epsilon_m\left(\lambda_p\left(0\right)\right)\right) m\left(\lambda_p\left(0\right)\right) < m\left(\psi\right)$ and hence constrained inefficiency.

A.5 Proof of Lemma 4

Step 1: If $w_{\min}$ does not bind for firms attracting directed searchers, then $w_{\min} < \epsilon_m\left(\psi\right) y\left(j_p\right)$. Suppose that $w_{\min}$ does not bind for firms attracting directed searchers. This implies that the inefficiency result of Proposition 1 carries over, so $j_e > j_p$ and $\mathcal{U} < \eta$. This means that firm j_p strictly prefers not to attract directed searchers in equilibrium. Denoting again by

$\lambda_e^D(j)$ the solution to the first-order condition (18), we then have

$$\begin{aligned} m(\psi)(y(j_p)-w_{\min}) &> m\left(\lambda_e^D(j_p)\right)\left(1-\epsilon_m\left(\lambda_e^D(j_p)\right)\right)y(j_p) \\ &> m(\psi)(1-\epsilon_m(\psi))\,y(j_p)\,, \end{aligned} \tag{A.8}$$

where the second line follows because $m(\lambda)(1-\epsilon_m(\lambda))$ is increasing in λ. The above implies $w_{\min} < \epsilon_m(\psi)\,y(j_p)$.

Step 2: If $w_{\min} < \epsilon_m(\psi)\,y(j_p)$, it does not bind for firms attracting directed searchers.
Suppose that the minimum wage binds for at least some firms who attract directed searchers. This means that the smallest queue length obtained in equilibrium, which we denote by λ_0, must satisfy

$$\frac{m(\lambda_0)}{\lambda_0}w_{\min} = \mathcal{U} \tag{A.9}$$

and $\lambda_0 \geq \psi$. It then follows that $\lambda_e(j) = \max\left\{\lambda_0, \lambda_e^D(j)\right\}$ must satisfy

$$\lambda_e(j) = \begin{cases} \lambda_0, & j \leq j_0 \\ \lambda_e^D(j)\,, & j > j_0 \end{cases} \tag{A.10}$$

where the threshold j_0 satisfies $m'(\lambda_0)\,y(j_0) = \mathcal{U}$. Finally, $\mathcal{U}$ must satisfy the modified market-clearing condition

$$\int_0^1 \max\left\{\lambda_0, \lambda_e^D(j)\right\}dj = 1. \tag{A.11}$$

Note that (A.9) defines $\mathcal{U}$ as a decreasing function of λ_0, and (A.11) defines $\mathcal{U}$ as an increasing function of λ_0, so that the equilibrium $\mathcal{U}$ and λ_0, and hence j_0, are uniquely determined for any $w_{\min}$. Furthermore, since $\lambda_0 \geq \psi$, comparing (A.11) to (6) and (7) implies that $\mathcal{U} \geq \eta$; but then the corresponding minimum wage must satisfy

$$\frac{\lambda_0}{m(\lambda_0)}\mathcal{U} \geq \frac{\psi}{m(\psi)}\eta = \epsilon_m(\psi)\,y(j_p) \tag{A.12}$$

where the inequality transpires because $\lambda/m(\lambda)$ is increasing in λ. This proves that a minimum wage strictly less than $\epsilon_m(\psi)\,y(j_p)$ cannot be binding for firms attracting directed searchers.

A.6 Proof of Proposition 2

Consider a minimum wage $w_{\min} < \epsilon_m(\psi)\,y(j_p)$, which binds for firms attracting random searchers only, and does not bind for firms attracting any directed searchers. The equilibrium is fully characterized by a market utility $\mathcal{U}$ for directed searchers, and a threshold j_e above which firms attract directed searchers, satisfying the indifference condition (29) and the market clearing condition (22).

Proof of (i) and (ii). We first analyze the effect of $w_{\min}$ on j_e and $\mathcal{U}$. Totally differentiating (29) with respect to $w_{\min}$, we obtain

$$-m(\psi) = (m(\Lambda_e) - m(\psi))\, y'(j_e)\, \frac{dj_e}{dw_{\min}} - \Lambda_e \frac{d\mathcal{U}}{dw_{\min}}, \tag{A.13}$$

where $\Lambda_e \equiv \lambda_e(j_e)$ is the solution to (29). Totally differentiating the market clearing condition (22) with respect to $w_{\min}$, we obtain

$$\frac{dj_e}{dw_{\min}} = \frac{d\mathcal{U}}{dw_{\min}} \times \frac{1}{\Lambda_e - \psi} \int_{j_e}^{1} \frac{1}{y(j)\, m''(y(j))}\, dj. \tag{A.14}$$

Combining (A.13) with (A.14) gives

$$\frac{d\mathcal{U}}{dw_{\min}} = m(\psi) \left[\Lambda_e - \left(\frac{m(\Lambda_e) - m(\psi)}{\Lambda_e - \psi}\right) y'(j_e) \int_{j_e}^{1} \frac{1}{y(j)\, m''(y(j))}\, dj\right]^{-1} > 0, \tag{A.15}$$

from which (A.14) immediately implies $\frac{dj_e}{dw_{\min}} < 0$.

Proof of (iii). We now turn to characterizing the effect of the minimum wage on aggregate employment, which is given by

$$\mathcal{E} = \int_0^1 m(\lambda_e(j))\, dj = j_e m(\psi) + \int_{j_e}^{1} m(\lambda_e(j))\, dj. \tag{A.16}$$

Differentiating (A.16) with respect to $w_{\min}$ and using (A.14), we get

$$\begin{aligned} \frac{d\mathcal{E}}{dw_{\min}} &= \frac{dj_e}{dw_{\min}} \times (m(\psi) - m(\Lambda_e)) + \frac{d\mathcal{U}}{dw_{\min}} \times \int_{j_e}^{1} \frac{m'(\lambda_e(j))}{y(j)\, m''(\lambda_e(j))}\, dj \\ &= \frac{d\mathcal{U}}{dw_{\min}} \times \int_{j_e}^{1} \frac{1}{y(j)\, m''(\lambda_e(j))} \left[m'(\lambda_e(j)) - \frac{m(\Lambda_e) - m(\psi)}{\Lambda_e - \psi}\right] dj \\ &> 0. \end{aligned} \tag{A.17}$$

The last line follows from the concavity of m, since $m'' < 0$ and, for all $j \geq j_e$,

$$m'(\lambda_e(j)) \leq m'(\Lambda_e) < \frac{m(\Lambda_e) - m(\psi)}{\Lambda_e - \psi}. \tag{A.18}$$

Proof of (iv). A similar argument applies to welfare, which is given by

$$\mathcal{W} = \int_0^1 m\left(\lambda_e\left(j\right)\right) y\left(j\right)\, dj = m\left(\psi\right) \int_0^{j_d} y\left(j\right)\, dj + \int_{j_e}^1 m\left(\lambda_e\left(j\right)\right) y\left(j\right)\, dj. \tag{A.19}$$

Differentiating (A.19) with respect to $w_{\min}$, we get

$$\begin{aligned} \frac{d\mathcal{W}}{dw_{\min}} &= \frac{dj_e}{dw_{\min}} \times \left(m\left(\psi\right) - m\left(\Lambda_e\right)\right) y\left(j_e\right) + \frac{d\mathcal{U}}{dw_{\min}} \times \int_{j_e}^1 \frac{m'\left(\lambda_e\left(j\right)\right)}{m''\left(\lambda_e\left(j\right)\right)}\, dj \\ &= \frac{d\mathcal{U}}{dw_{\min}} \times \int_{j_e}^1 \frac{1}{y\left(j\right) m''\left(\lambda_e\left(j\right)\right)} \left[m'\left(\lambda_e\left(j\right)\right) y\left(j\right) - \frac{m\left(\Lambda_e\right) - m\left(\psi\right)}{\Lambda_e - \psi} y\left(j_e\right)\right] dj \\ &> 0. \end{aligned} \tag{A.20}$$

The last inequality follows since $m'' < 0$ and

$$m'\left(\lambda_e\left(j\right)\right) y\left(j\right) = \mathcal{U} = m'\left(\Lambda_e\right) y\left(j_e\right) < \frac{m\left(\Lambda_e\right) - m\left(\psi\right)}{\Lambda_e - \psi} y\left(j_e\right) \tag{A.21}$$

by the concavity of m and the definition of $\lambda_e\left(j\right)$.

A.7 Proof of Corollary 4

The result that the minimum wage $w_{\min} = \epsilon_m\left(\psi\right) y\left(j_p\right)$ implements the constrained-efficient allocation follows directly from the proof of Lemma 4. In particular, from equation (A.11), it is immediate that constrained efficiency requires $\lambda_0 = \psi$, which, by (A.12), transpires when $w_{\min} = \epsilon_m\left(\psi\right) y\left(j_p\right)$. It remains to show that $\epsilon_m\left(\psi\right) y\left(j_p\right)$ is increasing in ψ. Note that $\epsilon_m\left(\psi\right) y\left(j_p\right) = \frac{\psi}{m(\psi)}\eta$, where η is the solution to the resource constraint It then follows that η is increasing in ψ by the resource constraint, and $\frac{\psi}{m(\psi)}$ is increasing in ψ by the assumptions on m. This completes the proof.

B Proofs for Section 6: Extensive margin behavior

B.1 Proof of Lemma 5

Let η be the Lagrange multiplier on (31), and let $\mu\left(j\right)\, dj$ be the Lagrange multiplier on (32) for each j. The first-order conditions for $\lambda\left(j\right)$ and u, respectively, are

$$\mu\left(j\right) = \eta - m'\left(\lambda\left(j\right)\right) y\left(j\right) \tag{B.1}$$

and

$$z = \eta - \psi \int_0^1 \mu(j)\ dj. \tag{B.2}$$

From (32) and (B.1), we obtain the characterization of the optimal $\lambda(j)$. If (32) does not bind, we have $\mu(j) = 0$ and $\lambda(j)$ is then given by $m'(\lambda(j))\, y(j) = \eta$. If the constraint (32) binds, then we have $\lambda(j) = \psi u_p$, and $\mu(j) = \eta - m'(\psi u_p)\, y(j)$. As before, this implies a threshold rule where the constraint binds for $j < j_p$ and does not bind for $j \geq j_p$ for some j_p. Assuming an interior j_p, it must satisfy (35). Next, we can write $\mu(j) = \max\{0, \eta - m'(\psi u_p)\, y(j)\}$. Substituting this into (B.2) and rearranging gives (33).

B.2 Proof of Lemma 6

The derivation of (41)–(43) is identical to the baseline model. Condition (40) follows directly from optimal participation and the fact that the utility of a random searcher is zero.

B.3 Proof of Proposition 3

The second part is identical to the baseline model. We show here that $u_e < u_p$. Suppose, for a contradiction, that $u_e \geq u_p$. Letting Λ_e be the solution to $m'(\Lambda_e)\, y(j_e) = \mathcal{U}$, we have

$$\begin{aligned}
& m'(\psi u_p)\left[(1-\psi j_p)\, y(j_p) + \psi \int_0^{j_p} y(j)\ dj\right] \\
= \;& z \\
= \;& (1-\psi j_e)\mathcal{U} \\
= \;& (1-\psi j_e)\, m'(\Lambda_e)\, y(j_e) \\
< \;& (1-\psi j_e)\, m'(\psi u_e)\, y(j_e) \\
\leq \;& (1-\psi j_e)\, m'(\psi u_p)\, y(j_e) \\
< \;& m'(\psi u_p)\left[(1-\psi j_e)\, y(j_e) + \psi \int_0^{j_e} y(j)\ dj\right].
\end{aligned} \tag{B.3}$$

The expression in brackets is easily verified to be increasing in j_e; therefore, $j_p < j_e$, and so

$$\begin{aligned}
(1-\psi j_e)\mathcal{U} &= z \\
&= (1-\psi j_p)\,\eta + \psi m'(\psi u_p) \int_0^{j_p} y(j)\ dj \\
&> (1-\psi j_p)\,\eta \\
&> (1-\psi j_e)\,\eta,
\end{aligned} \tag{B.4}$$

so that $\mathcal{U} > \eta$. In turn, this means that $\lambda_e(j) < \lambda_p(j)$ for all $j > j_e$. Finally, we argue that $j_e > j_p$ and $\lambda_e(j) < \lambda_p(j)$ together imply that $u_p > u_e$. We have

$$\begin{aligned} u_p &= j_p \psi u_p + \int_{j_p}^{1} \lambda_p(j)\, dj \\ &= j_p \psi u_p + \int_{j_p}^{j_e} \lambda_p(j)\, dj + \int_{j_e}^{1} \lambda_p(j)\, dj \\ &> j_p \psi u_p + (j_e - j_p)\, \psi u_p + \int_{j_e}^{1} \lambda_e(j)\, dj \\ &= j_e \psi u_p + \int_{j_e}^{1} \lambda_e(j)\, dj. \end{aligned} \tag{B.5}$$

By rearranging, we obtain

$$(1 - \psi j_e)\, u_p > \int_{j_e}^{1} \lambda_e(j)\, dj. \tag{B.6}$$

We must also have

$$(1 - \psi j_e)\, u_e = \int_{j_e}^{1} \lambda_e(j)\, dj. \tag{B.7}$$

Together, these imply $u_p > u_e$, which is the desired contradiction.

B.4 Proof of Lemma 7

The social planner is maximizing (48) subject to the constraints (49) and (50). Let η be the Lagrange multiplier on (49), and let $\mu(j)\, dj$ be the Lagrange multiplier on (50) for each j. The first-order condition for $\lambda(j)$ is then $\mu(j)(1 - j_p^*) = \eta - m'(\lambda(j))\, y(j)$, which leads, as in the exogenous entry case, to the optimal queue length, as characterized in equation (52). Define $j_p = \inf\left\{j : m'\left(\psi/\left(1 - j_p^*\right)\right) y(j) \geq \eta\right\}$. Since $j_p > j_p^*$, we immediately obtain (51)–(53); in particular, we must have $\lambda_p\left(j_p^*\right) = \psi/\left(1 - j_p^*\right)$ and $\eta = m'\left(\lambda_p\left(j_p^*\right)\right) y(j_p)$. Moreover, this gives

$$\mu(j) = \frac{1}{1 - j_p^*} m'\left(\lambda_p\left(j_p^*\right)\right) \max\left\{0, y(j_p) - y(j)\right\}. \tag{B.8}$$

Next, the first-order condition for j_p^* reads

$$\kappa = m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right) - \lambda_p\left(j_p^*\right) \eta + \int_{j_p^*}^{1} \lambda_p(j)\, \mu(j)\, dj. \tag{B.9}$$

Substituting (51) and (B.8) into (B.9), we get

$$\begin{aligned}
\kappa = \; & m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right) - \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p\right) \\
& + \frac{1}{1-j_p^*} \int_{j_p^*}^{j_p} \lambda_p\left(j\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(y\left(j_p\right) - y\left(j\right)\right) dj \\
= \; & (m\left(\lambda_p\left(j_p^*\right)\right) - \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)) y\left(j_p^*\right) \\
& + \frac{1}{1-j_p^*} \int_{j_p^*}^{1} \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(y\left(j_p^*\right) - y(j_p)\right) dj \\
& + \frac{1}{1-j_p^*} \int_{j_p^*}^{j_p} \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(y\left(j_p\right) - y\left(j\right)\right) dj \\
= \; & (m\left(\lambda_p\left(j_p^*\right)\right) - \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)) y\left(j_p^*\right) \\
& - \frac{1}{1-j_p^*} \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right) \int_{j_p^*}^{1} [\min\{y\left(j\right), y\left(j_p\right)\} - y\left(j_p^*\right)] \, dj,
\end{aligned} \tag{B.10}$$

which is equivalent to (54).

B.5 Proof of Lemma 8

For a given j_e^*, by previous results, the equilibrium is characterized by a market utility $\mathcal{U}$ and a threshold rule such that only firms with $j \geq j_e$ attract directed searchers. In an interior solution, $j_e > j_e^*$, so that the marginal entrant attracts only random searchers. This immediately gives conditions (60)–(62) similar to Lemma 3 for the exogenous entry case. The zero profit condition for the marginal entrant then yields (63).

B.6 Proof of Proposition 4

First, we show that entry is inefficiently high in equilibrium. From (54) and (63), we have

$$\begin{aligned}
m\left(\lambda_p\left(j_p^*\right)\right) y\left(j_p^*\right) - \kappa & > m\left(\lambda_p\left(j_p^*\right)\right)\left(1 - \epsilon_m\left(\lambda\left(j_p^*\right)\right)\right) y\left(j_p^*\right) - \kappa \\
& = \lambda_p\left(j_p^*\right) m'\left(\lambda_p\left(j_p^*\right)\right)\left(\mathbb{E}\left[\min\left\{y\left(j\right), y\left(j_p\right)\right\} | j \geq j_p^*\right] - y\left(j_p^*\right)\right) \\
& > 0 \\
& = m\left(\lambda_e\left(j_e^*\right)\right) y\left(j_e^*\right) - \kappa.
\end{aligned} \tag{B.11}$$

Moreover, $\lambda_p\left(j_p^*\right) = \psi/(1-j_p^*)$ and $\lambda_e\left(j_e^*\right) = \psi/(1-j_e^*)$. Since $m\left(\psi/\left(1-j\right)\right) y\left(j\right)$ is increasing in j, this establishes that $j_e^* < j_p^*$. Next, consider (64). The solution is analogous to Lemma 1 and give rise to a threshold j_o. When $\psi \in (0,1)$, the proof that $j_e > j_o$ and $\lambda_e\left(j\right) > \lambda_o\left(j\right)$ for $j > j_e$ is then identical to the proof of Proposition 1 in Appendix A.4.

Online Appendix

C Worker Heterogeneity

C.1 Environment

The environment is identical to the baseline model except for worker heterogeneity. There is a unit measure of firms and a unit measure of workers.[13] We assume that a fraction ϕ_i of directed searchers has productivity x_i, where $i \in \{1, 2\}$ and $0 < x_1 < x_2$.[14] For random searchers, we simply specify their expected productivity, which we denote by $\overline{x}$; the exact distribution of their productivity is irrelevant.

Worker and firm productivity jointly determine the surplus from a match, which we assume to be equal to xy. We want to emphasize that this assumption serves to ensure a particular pattern of sorting (namely that higher-x workers match to higher-y firms) among directed searchers. While providing analytical convenience, it is not crucial for the actual inefficiency result.

We make no assumptions on the correlation between workers' productivity and their status as a directed or random searcher (in particular, we do not impose that $\overline{x} = \phi_1 x_1 + \phi_2 x_2$). We only bound the expected productivity of random searchers by assuming that i) $\overline{x} < x_2$ and ii) $x_1 > m(\psi)\,\overline{x}$. The first assumption states that a random searcher is less productive in expectation than the high-productivity directed searcher. The second states that the expected output from hiring random searchers only is less than the output from hiring a low-productivity directed searcher with certainty. The latter assumption is largely without loss of generality: if it is violated, low-type directed searchers will never be hired, so that the equilibrium is the same as in a world in which all directed searchers have productivity x_2.

The presence of random searchers implies that firms cannot perfectly control the productivity of the worker that they meet. Consider a firm of type y attracting a total queue length $\lambda \geq \psi$ and attracting directed searchers of type i. The expected output of such a firm is

$$N(x_i, \lambda)\, y \equiv m(\lambda) \left(\frac{\lambda - \psi}{\lambda} x_i + \frac{\psi}{\lambda} \overline{x} \right) y. \tag{C.1}$$

To understand (C.1), note that a fraction ψ/λ of workers in the firm's queue are random searchers, whose expected productivity is $\overline{x}$. The remaining fraction $(\lambda - \psi)/\lambda$ are directed

[13] In particular, we abstract from the entry margins considered in Section 6.

[14] Thus, we assume that there are only two productivity types. This is done for simplicity of exposition; all the results here carry over to the case with an arbitrary number of worker types.

searchers, whose productivity is x_i.

In case of either purely random or purely directed search, (C.1) assumes the familiar form of either $m(\lambda)\, x_i y$ (for purely directed search) or $m(\lambda)\, \overline{x} y$ (for purely random search). For the intermediate case, (C.1) illustrates the key technical complication that arises in this environment: expected output is non-separable in queue length and worker productivity, or, to put it differently, the expected output conditional on hiring now depends on the queue length. This complicates the subsequent analysis of sorting relative to e.g. Eeckhout and Kircher (2010). To make progress, we make the following additional assumption on the function N defined by (C.1):

Assumption 1. *The function $N(x,\lambda)$ defined by (C.1) is strictly increasing and concave in λ and satisfies*

$$N_x - \frac{N_\lambda}{N_{\lambda\lambda}} \left(N_{x\lambda} - \frac{N_x}{\lambda - \psi} \right) > 0 \tag{C.2}$$

for $\lambda > \psi$ and $x \in [x_1, x_2]$.

The first part of Assumption 1 states that N inherits the monotonicity and concavity properties of m; the last part, (C.2), is the equivalent of the assumption that the matching function has a decreasing elasticity. It can be shown that, given the assumptions already made on m, Assumption 1 amounts to assuming that $\overline{x}$ is not too large.[15] Again, we want to emphasize that this technical assumption gets used mainly in the proof of positive assortative matching (which generalizes the technique of Eeckhout and Kircher (2010) to the partially directed search environment), but not in the proof of the actual inefficiency result. In addition, we again focus on parameters such that the solution is interior; i.e., a positive measure of firms attract only random searchers.

C.2 Planner's problem

The social planner chooses how to allocate directed-searcher workers of each productivity across firms. The social planner is constrained by the total measure of workers of each type, as well as the constraint that the random searchers must be allocated randomly, so that the expected number of workers at each firm is at least ψ. The same constraint, due to heterogeneous productivity, will also imply that the expected output at a firm is affected by the allocation of workers as well, via (C.1). We index submarkets by the type j of the firm advertising in that submarket and the type i of directed searcher (if any) that it attracts. For each submarket (i,j), the social planner chooses the probability $h_i(j)$ that firm j enters

[15] In fact, for a specific functional form of m, e.g. the commonly used matching function $m(\lambda) = \lambda/(1+\lambda)$, Assumption 1 is implied by the earlier assumptions that $\overline{x} < x_2$ and $x_1 > m(\psi)\, \overline{x}$.

that submarket[16] and the queue length $\lambda_i(j)$ in that submarket. The planner's problem can then be written as maximizing

$$\int_0^1 \sum_i N(x_i, \lambda_i(j))\, y(j)\, h_i(j)\; dj, \tag{C.3}$$

subject to a resource constraint on the measure of firms of each type,

$$\sum_i h_i(j) \leq 1 \quad \text{for all } j \in [0,1]; \tag{C.4}$$

a resource constraint on the measure of directed searchers of each type,

$$\int (\lambda_i(j) - \psi)\, h_i(j)\; dj \leq (1-\psi)\,\phi_i \quad \text{for } i = 1,2; \tag{C.5}$$

a non-negativity constraint on the measure of firms in each submarket,

$$h_i(j) \geq 0, \quad \text{for all } j \in [0,1]; \tag{C.6}$$

and a non-negativity constraint on the measure of directed-searcher workers in each submarket,

$$\lambda_i(j) \geq \psi. \tag{C.7}$$

In words, a firm of type j attracts some directed searchers of type i if both $h_i(j) > 0$ and $\lambda_i(j) > \psi$ are true.

Lemma 9. *The constrained-efficient allocation is characterized by numbers $j_{p,1}$, $j_{p,2}$, η_1, η_2, and a function $\lambda_p : [0,1] \to \mathbb{R}_+$, such that*

1. *Firms with $j < j_{p,1}$ attract random searchers only. Firms with $j \in [j_{p,1}, j_{p,2})$ attract directed searchers of type 1 but not of type 2. Firms with $j \in [j_{p,2}, 1]$ attract directed searchers of type 2 but not of type 1.*

2. *The queue length $\lambda_p(j)$ satisfies*

$$\lambda_p(j) = \begin{cases} \psi, & \text{if } j < j_{p,1} \\ \lambda^D_{p,1}(j), & \text{if } j \in [j_{p,1}, j_{p,2}) \\ \lambda^D_{p,2}(j), & \text{if } j \in [j_{p,2}, 1] \end{cases} \tag{C.8}$$

[16] Writing the planner's problem this way gives an additional degree of freedom, as it allows the firms to randomize between submarkets. We will see that this option is never used at the optimum.

where, for $i = 1, 2$, $\lambda^D_{p,i}(j)$ is defined by

$$N_\lambda\left(x_i, \lambda^D_{p,i}(j)\right) y(j) = \eta_i. \tag{C.9}$$

3. *The thresholds $j_{p,1}$ and $j_{p,2}$ satisfy*

$$N_\lambda(x_1, \psi)\, y(j_{p,1}) = \eta_1. \tag{C.10}$$

and

$$N\left(x_1, \lambda^D_{p,1}(j_{p,2})\right) - \left(\lambda^D_{p,1}(j_{p,2}) - \psi\right)\eta_1 = N\left(x_2, \lambda^D_{p,2}(j_{p,2})\right) - \left(\lambda^D_{p,2}(j_{p,2}) - \psi\right)\eta_2. \tag{C.11}$$

4. *Resource constraints on each type of worker are satisfied:*

$$\begin{aligned} \int_{j_{p,1}}^{j_{p,2}} \left(\lambda_p(j) - \psi\right) dj &= (1-\psi)\,\phi_1, \\ \int_{j_{p,2}}^{1} \left(\lambda_p(j) - \psi\right) dj &= (1-\psi)\,\phi_2. \end{aligned} \tag{C.12}$$

Proof. To characterize the solution, let $\pi_p(j)$, η_i, $\xi_i(j)$ and $\mu_i(j)$ be the Lagrange multipliers on (C.4), (C.5), (C.6), and (C.7). Note that the multiplier $\pi_p(j)$ is interpreted as the shadow value for a type-j firm; similarly, η_i is the shadow value of a type-i worker. The necessary first-order conditions with respect to $h_i(j)$ and $\lambda_i(j)$, respectively, can then be written as

$$N(x_i, \lambda_i(j))\, y(j) - (\lambda_i(j) - \psi)\,\eta_i - \pi_p(j) + \xi_i(j) = 0 \tag{C.13}$$

and

$$\left[N_\lambda(x_i, \lambda_i(j))\, y(j) - \eta_i\right] h_i(j) + \mu_i(j) = 0. \tag{C.14}$$

Since $\xi_i(j) \geq 0$, equation (C.13) implies that $\pi_p(j) = \max\{\pi_{p,1}(j), \pi_{p,2}(j)\}$, where

$$\pi_{p,i}(j) = \max_{\lambda \geq \psi} N(x_i, \lambda)\, y(j) - (\lambda - \psi)\,\eta_i. \tag{C.15}$$

Denote by $\lambda^D_{p,i}(j)$ the solution to (C.15) for each i. We first show a preliminary result:

Claim 2. *If $h_1(j) > 0$, $h_2(j) > 0$ and $\lambda^D_{p,2}(j) = \psi$, then $\lambda^D_{p,1}(j) = \psi$.*

Proof. To establish this result, notice that $h_1(j) > 0$ and $h_2(j) > 0$ imply that $\pi_{p,1}(j) =$

$\pi_{p,2}(j)$. Further, $\lambda^D_{p,2}(j) = \psi$ implies that

$$\pi_{p,2}(j) = N(x_2, \psi)\, y(j) = m(\psi)\, \overline{x} y(j)\,. \tag{C.16}$$

Now, suppose that $\lambda_{p,1}(j) = \lambda > \psi$. By (C.14) and (C.15), this means that

$$\pi_{p,1}(j) = \left[N(x_1, \lambda) - (\lambda - \psi)\, N_\lambda(x_1, \lambda)\right] y(j)\,. \tag{C.17}$$

The right-hand side of (C.17) reduces to $m(\psi)\, \overline{x} y(j)$ when $\lambda = \psi$. A contradiction then follows, because the derivative of the right-hand side of (C.17) with respect to λ equals

$$-(\lambda - \psi)\, N_{\lambda\lambda}(x_i, \lambda)\, y(j) > 0, \tag{C.18}$$

which implies that $\pi_{p,1}(j) > \pi_{p,2}(j)$. □

This shows that a firm that attracts any directed searchers does not randomize between submarkets.

Next, we show that the decision to attract directed searchers follows a threshold rule, similarly to the homogeneous-worker case.

Claim 3. *If $\pi_p(j) > m(\psi)\, \overline{x} y(j)$, then $\pi_p(j') > m(\psi)\, \overline{x} y(j')$ for all $j' > j$.*

Proof. To see this, note that $\pi_p(j) > m(\psi)\, \overline{x} y(j)$ means that $\pi_{p,i}(j) > m(\psi)\, \overline{x} y(j)$ for either $i = 1$ or $i = 2$ or both. By the envelope theorem,

$$\frac{d}{dj}\pi_{p,i}(j) = N\left(x_i, \lambda^D_{p,i}(j)\right) y'(j) \geq N(x_i, \psi) = \frac{d}{dj} m(\psi)\, \overline{x} y(j)\,. \tag{C.19}$$

□

This result implies that we can define the cutoff

$$j_{p,1} = \inf\{j : \pi_p(j) \geq m(\psi)\, \overline{x} y\}, \tag{C.20}$$

such that the planner only instructs firms with $j < j_{p,1}$ to target random searchers. In an interior solution (i.e., $j_{p,1} > 0$), the first-order conditions (C.13) and (C.14) immediately imply (C.10), similarly to the homogeneous-worker model.

Finally, we show that there is positive assortative matching above $j_{p,1}$.

Claim 4. *Let $j_{p,2}$ satisfy $\pi_{p,2}(j_{p,2}) = \pi_{p,1}(j_{p,2})$. Then $\pi_{p,2}(j_{p,2}) > \pi_{p,1}(j_{p,2})$ for all $j > j_{p,2}$.*

Proof. It is sufficient to prove that that $\frac{d}{dj}\pi_{p,2}(j_{p,2}) > \frac{d}{dj}\pi_{p,1}(j_{p,2})$; by the envelope theorem, this holds if and only if

$$N\left(x_2, \lambda_{p,2}^D(j_{p,2})\right) > N\left(x_1, \lambda_{p,1}^D(j_{p,2})\right). \tag{C.21}$$

To prove this, note that the indifference condition $\pi_{p,2}(j_{p,2}) = \pi_{p,1}(j_{p,2})$ is equivalent to

$$\begin{aligned} & N\left(x_2, \lambda_{p,2}^D(j_{p,2})\right) - (\lambda_{p,2}^D(j_{p,2}) - \psi) N_\lambda\left(x_2, \lambda_{p,2}^D(j_{p,2})\right) \\ = \; & N\left(x_1, \lambda_{p,1}^D(j_{p,2})\right) - (\lambda_{p,1}^D(j_{p,2}) - \psi) N_\lambda\left(x_1, \lambda_{p,1}^D(j_{p,2})\right). \end{aligned} \tag{C.22}$$

We can therefore define a function $\ell_p(x) : [x_1, x_2] \to \mathbb{R}_+$ such that $\ell_p(x_i) = \lambda_{p,i}^D(j_{p,2})$ and $N(x, \ell_p(x)) - (\ell_p(x) - \psi) N_\lambda(x, \ell_p(x))$ is constant for $x \in [x_1, x_2]$. Total differentiation then implies that

$$\ell_p'(x) = \frac{N_x(x, \ell_p(x)) - (\ell_p(x) - \psi) N_{x\lambda}(x, \ell_p(x))}{(\ell_p(x) - \psi) N_{\lambda\lambda}(x, \ell_p(x))} \tag{C.23}$$

for any $x \in (x_1, x_2)$. Differentiating $N(x, \ell_p(x))$ with respect to x and using (C.23) then yields

$$\frac{d}{dx} N(x, \ell_p(x)) = \frac{(\ell_p(x) - \psi) N_{\lambda\lambda} N_x + N_\lambda N_x - (\ell_p(x) - \psi) N_\lambda N_{x\lambda}}{(\ell_p(x) - \psi) N_{\lambda\lambda}}. \tag{C.24}$$

The right-hand side is positive by condition (C.2). Hence, $N\left(x_2, \lambda_{p,2}^D(j_{p,2})\right) > N\left(x_1, \lambda_{p,1}^D(j_{p,2})\right)$, establishing the desired single-crossing condition. □

The above shows that sorting is positive: if, at some $j_{p,2}$, the social planner is indifferent between having the firm attract type-1 or type-2 workers, then type-2 workers are strictly preferred at any higher j. The same result immediately implies that the $j_{p,2}$ satisfying such indifference is unique and is given by (C.11). It also implies that, if $j_{p,1} > 0$, then it satisfies (C.10); in other words, indifference between random searchers and directed searchers must occur for directed searchers of type 1. This completes the construction of the cutoffs $j_{p,1}$ and $j_{p,2}$. The characterization of the queue length $\lambda_p(j)$ for any j then follows immediately from the first-order conditions. □

Lemma 9 makes two key statements, both captured in part 1. First, as in the homogeneous-worker case, the decision to attract directed searchers follows a threshold rule. Firms with j below the cutoff $j_{p,1}$ attract random searchers only, whereas firms with j above the cutoff attract directed searchers. Second, there is positive assortative matching among the firms who do attract directed searchers, as evidenced by the second cutoff, $j_{p,2}$. Next, in part

2, we proceed to characterize the optimal queue in each region. The numbers η_1 and η_2 are the shadow values of directed searchers of type 1 and type 2, respectively. It is useful to define the function $\lambda_{p,i}^{D}(j)$, which, by (C.9), denotes the optimal queue length for firm j *conditional* on attracting directed searchers of type i. Part 3 characterizes the optimal cutoffs. Equation (C.10) is an indifference condition between attracting random searchers and attracting directed searchers of type 1, while (C.11) is an indifference condition between attracting directed searchers of type 1 and attracting directed searchers of type 2. Finally, part 4 states that the resource constraint for each type of worker holds.

C.3 Equilibrium

In a decentralized equilibrium, each firm j chooses, in addition to a wage, whether to attract directed searchers and, if so, which type of directed searcher to attract. In doing so, it takes as given the market utilities $\mathcal{U}_1$ and $\mathcal{U}_2$ of each type of directed searcher. The following lemma characterizes the equilibrium.

Lemma 10. *The equilibrium allocation is characterized by numbers $j_{e,1}$, $j_{e,2}$, $\mathcal{U}_1$, $\mathcal{U}_2$, and a function $\lambda_e : [0,1] \to \mathbb{R}_+$, such that*

1. *Firms with $j < j_{e,1}$ attract random searchers only. Firms with $j \in [j_{e,1}, j_{e,2})$ attract directed searchers of type 1 but not of type 2. Firms with $j \in [j_{e,2}, 1]$ attract directed searchers of type 2 but not of type 1.*

2. *The queue length $\lambda_e(j)$ satisfies*

$$\lambda_e(j) = \begin{cases} \psi, & j < j_{e,1} \\ \lambda_{e,1}^{D}(j), & j \in [j_{e,1}, j_{e,2}) \\ \lambda_{e,2}^{D}(j), & j \in [j_{e,2}, 1] \end{cases} \tag{C.25}$$

   where, for $i = 1, 2$, $\lambda_{e,i}^{D}(j)$ is defined by

$$N_\lambda\left(x_i, \lambda_{e,i}^{D}(j)\right) y(j) = \mathcal{U}_i. \tag{C.26}$$

3. *The thresholds $j_{e,1}$ and $j_{e,2}$ satisfy*

$$m(\psi)\,\overline{x} = N\left(x_1, \lambda_{e,1}^{D}(j_{e,1})\right) - \lambda_{e,1}^{D}(j_{e,1})\,\mathcal{U}_1 \tag{C.27}$$

and

$$N\left(x_1, \lambda_{e,1}^D\left(j_{e,2}\right)\right) - \lambda_{e,1}^D\left(j_{e,2}\right)\mathcal{U}_1 = N\left(x_2, \lambda_{e,2}^D\left(j_{e,2}\right)\right) - \lambda_{e,2}^D\left(j_{e,2}\right)\mathcal{U}_2. \tag{C.28}$$

4. *Resource constraints on each type of worker are satisfied:*

$$\begin{aligned} \int_{j_{e,1}}^{j_{e,2}} \left(\lambda_e\left(j\right) - \psi\right) dj &= \left(1-\psi\right)\phi_1, \\ \int_{j_{e,2}}^{1} \left(\lambda_e\left(j\right) - \psi\right) dj &= \left(1-\psi\right)\phi_2. \end{aligned} \tag{C.29}$$

Proof. In equilibrium, firm j can either target random searchers, low-type directed searchers, or high-type directed searchers. Similar to the baseline model, a firm targeting random searchers will offer a wage of 0 and therefore receive an expected payoff equal to

$$\pi_e^R\left(j\right) = m\left(\psi\right)\overline{x}y\left(j\right). \tag{C.30}$$

In contrast, a firm targetting directed searchers with productivity x_i must provide them with their market utility $\mathcal{U}_i$. Such a firm therefore solves

$$\pi_{e,i}^D\left(j\right) = \max_{\lambda,w} N\left(x_i, \lambda\right)y\left(j\right) - m\left(\lambda\right)w, \tag{C.31}$$

subject to

$$\frac{m\left(\lambda\right)}{\lambda}w \geq \mathcal{U}_i. \tag{C.32}$$

As in the baseline model, the constraint will bind and can be substituted into the objective to obtain

$$\pi_{e,i}^D\left(j\right) = \max_{\lambda} N\left(x_i, \lambda\right)y\left(j\right) - \lambda\mathcal{U}_i. \tag{C.33}$$

The solution $\lambda_{e,i}^D\left(j\right)$ solves the first-order condition

$$N_\lambda\left(x_i, \lambda\right)y\left(j\right) = \mathcal{U}_i. \tag{C.34}$$

The maximized profit from targeting directed searchers of type x_i therefore equals

$$\pi_{e,i}^D\left(j\right) = \left[N\left(x_i, \lambda_{e,i}^D\left(j\right)\right) - \lambda_{e,i}^D\left(j\right)N_\lambda\left(x_i, \lambda_{e,i}^D\left(j\right)\right)\right]y\left(j\right). \tag{C.35}$$

Clearly, firms will choose who to target by comparing the associated payoffs; their equilibrium payoff therefore equals $\pi_e\left(j\right) = \max\left\{\pi_e^R\left(j\right), \pi_{e,1}^D\left(j\right), \pi_{e,2}^D\left(j\right)\right\}$.

The characterization of the equilibrium allocation is then straightforward and follows closely the process in the proof of Lemma 9. We first establish the following result.

Claim 5. *$\pi_e(j) > \pi_e^R(j)$ then $\pi_e(j') > \pi_e^R(j')$ for all $j' > j$.*

Proof. Note that $\pi_e(j) > \pi_e^R(j)$ means that $\pi_{e,i}^D(j) > \pi_e^R(j)$ for some $i \in \{1,2\}$. By the envelope theorem,

$$\frac{d}{dj}\pi_{e,i}^D(j) = N\left(x_i, \lambda_{e,i}^D(j)\right) y'(j) \geq N(x_i, \psi)\, y'(j) = \frac{d}{dj}\pi_e^R(j), \tag{C.36}$$

where the inequality follows from $N_\lambda(x_i, \lambda) \geq 0$. Hence, $\pi_e(j') > \pi_e^R(j')$ for all $j' > j$. $\square$

We can therefore define the cutoff

$$j_{e,1} = \inf\left\{j : \pi_e(j) > \pi_e^R(j)\right\}, \tag{C.37}$$

such that only firms with $j \leq j_{e,1}$ target random searchers.

Next, we show that there is positive sorting above $j_{e,1}$ by establishing a single-crossing condition: if a particular firm $j_{e,2} \in (j_{e,1}, 1)$ is indifferent between targeting low-type directed-searchers and targeting high-type directed searchers, then firms with higher productivity strictly prefer targeting high-type directed searchers.

Claim 6. *Let $j_{e,2}$ satisfy $\pi_{e,2}^D(j_{e,2}) = \pi_{e,1}^D(j_{e,2})$. Then $\pi_{e,2}^D(j_{e,2}) > \pi_{e,1}^D(j_{e,2})$ for all $j > j_{e,2}$.*

Proof. To prove single crossing, it suffices to show that the indifference condition $\pi_{e,2}^D(j_{e,2}) = \pi_{e,1}^D(j_{e,2})$ implies that $\frac{d}{dj}\pi_{e,2}^D(j_{e,2}) > \frac{d}{dj}\pi_{e,1}^D(j_{e,2})$, which, by the envelope theorem, is equivalent to

$$N\left(x_2, \lambda_{e,2}^D(j_{e,2})\right) > N\left(x_1, \lambda_{e,1}^D(j_{e,2})\right). \tag{C.38}$$

To prove this, note that $\pi_{e,2}^D(j_{e,2}) = \pi_{e,1}^D(j_{e,2})$ is equivalent to

$$\begin{aligned} &N(x_2, \lambda_{e,2}^D(j_{e,2})) - \lambda_{e,2}^D(j_{e,2}) N_\lambda(x_2, \lambda_{e,2}^D(j_{e,2})) \\ =&N(x_1, \lambda_{e,2}^D(j_{e,1})) - \lambda_{e,1}^D(j_{e,2}) N_\lambda(x_1, \lambda_{e,1}^D(j_{e,2})) \end{aligned} \tag{C.39}$$

We can therefore define a function $\ell_e(x) : [x_1, x_2] \to \mathbb{R}_+$ such that $\ell_e(x_i) = \lambda_{e,i}^D(j_{e,2})$ and $N(x, \ell_e(x)) - \ell_e(x) N_\lambda(x, \ell_e(x))$ is constant for $x \in [x_1, x_2]$. Total differentiation then implies that

$$\ell_e'(x) = \frac{N_x(x, \ell_e(x)) - \ell_e(x) N_{x\lambda}(x, \ell_e(x))}{\ell_e(x) N_{\lambda\lambda}(x, \ell_e(x))} \tag{C.40}$$

for any $x \in (x_1, x_2)$. Using this result, we can then consider how $N(x, \ell_e(x))$ varies with x, which yields

$$\frac{d}{dx} N(x, \ell_e(x)) = \frac{\ell_e(x) N_{\lambda\lambda} N_x + N_\lambda N_x - \ell_e(x) N_\lambda N_{x\lambda}}{\ell_e(x) N_{\lambda\lambda}}. \tag{C.41}$$

The right-hand side of (C.41) is positive by condition (C.2). Hence, $N\left(x_2, \lambda^D_{e,2}(j_{e,2})\right) > N\left(x_1, \lambda^D_{e,1}(j_{e,2})\right)$, which implies $\frac{d}{dj}\pi^D_{e,2}(j_{e,2}) > \frac{d}{dj}\pi^D_{e,1}(j_{e,2})$ and thus sorting is positive. □

Finally, the fact that $j_{e,1} < j_{e,2}$ implies that $j_{e,1}$ must satisfy the indifference condition (C.27). This completes the construction of the thresholds, and the characterization of $\lambda_e(j)$ for every j then follows from the firms' first-order conditions. □

The decentralized equilibrium allocation thus takes the same form as the constrained-efficient allocation: the choice to attract directed searchers follows a cutoff rule, and there is positive assortative matching above the cutoff. Of particular interest are conditions (C.27) and (C.28). Condition (C.27) states that the firm of type $j_{e,1}$ is indifferent between attracting random searchers only and attracting directed searchers of type 1. Condition (C.28) states that the firm of type $j_{e,2}$ is indifferent between attracting directed searchers of type 1 and attracting directed searchers of type 2.

C.4 Constrained inefficiency

We can now compare the equilibrium allocation to the planner's solution. In the next proposition, we establish that the equilibrium is inefficient and that the key inefficiencies are the same as in the baseline model: at the bottom of the productivity distribution, too many firms target random searchers, and too many workers apply to firms at the top of the productivity distribution.

Proposition 5. *Assume $\psi \in (0,1)$. Let $j_{p,1}$, $j_{p,2}$, η_1, η_2 and $\lambda_p(\cdot)$ be the constrained-efficient allocation, and let $j_{e,1}$, $j_{e,2}$, $\mathcal{U}_1$, $\mathcal{U}_2$ and $\lambda_e(\cdot)$ be the decentralized equilibrium allocation. Then the following hold: (i) $j_{e,1} > j_{p,1}$, (ii) $j_{e,2} > j_{p,2}$, and (iii) $\lambda_e(j) > \lambda_p(j)$ for all $j \geq j_{e,2}$.*

Proof. We establish that the equilibrium allocation deviates from the planner's solution. Specifically, (i) too few firms target directed searchers; (ii) too few firms target directed searchers of type 2; and, as a consequence (iii) too many workers apply to the firms with the highest productivity. As a result, the direction of the misallocation is the same as in the baseline model.

(i) Too few firms target directed searchers. We establish that $j_{e,1} > j_{p,1}$ by showing that $j_{e,1} \leq j_{p,1}$ leads to a contradiction. We do so in three steps. First, we show that $j_{e,1} \leq j_{p,1}$

implies $j_{e,2} < j_{p,2}$. Second, we show that $j_{e,2} < j_{p,2}$ implies $\lambda_{e,2}^{D}(j) > \lambda_{p,2}^{D}(j)$. Finally, we show that $j_{e,2} < j_{p,2}$ together with $\lambda_{e,2}^{D}(j) > \lambda_{p,2}^{D}(j)$ violates the resource constraint.

For the first step, note that the indifference condition defining $j_{e,1}$ implies

$$\begin{aligned} N(x_1, \psi) &= m(\psi)\,\overline{x} \\ &= N\left(x_1, \lambda_{e,1}^{D}(j_{e,1})\right) - \lambda_{e,1}^{D}(j_{e,1})\, N_\lambda\left(x_1, \lambda_{e,1}^{D}(j_{e,1})\right) \\ &< N\left(x_1, \lambda_{e,1}^{D}(j_{e,1})\right). \end{aligned} \tag{C.42}$$

This immediately implies that $\lambda_{e,1}^{D}(j_{e,1}) > \psi$. We also know that $N_\lambda\left(x_1, \lambda_{e,1}^{D}(j_{e,1})\right) y(j_{e,1}) = \mathcal{U}_1$. Moreover, given the assumption $j_{e,1} \leq j_{p,1}$, we have $N_\lambda(x_1, \psi)\, y(j_{e,1}) = \eta_1$. Together, these imply that $U_1 < \eta_1$, since $N_{\lambda\lambda} < 0$. For the same reason, it then follows that

$$\lambda_{e,1}^{D}(j) > \lambda_{p,1}^{D}(j). \tag{C.43}$$

The resource constraints then imply

$$\begin{aligned} \int_{j_{p,1}}^{j_{p,2}} \lambda_{p,1}^{D}(j)\; dj = \phi_1 &= \int_{j_{e,1}}^{j_{e,2}} \lambda_{e,1}^{D}(j)\; dj \\ &> \int_{j_{e,1}}^{j_{e,2}} \lambda_{e,1}^{D}(j)\; dj \\ &\geq \int_{j_{p,1}}^{j_{e,2}} \lambda_{e,1}^{D}(j)\; dj, \end{aligned} \tag{C.44}$$

where the first inequality follows from $\lambda_{e,1}^{D}(j) > \lambda_{p,1}^{D}(j)$ and the second from our assumption that $j_{e,1} \leq j_{p,1}$. This inequality can only hold if $j_{e,2} < j_{p,2}$.

In the second step, we show that $j_{e,2} < j_{p,2}$ implies that $\lambda_{e,2}^{D}(j) > \lambda_{p,2}^{D}(j)$. Note that $j_{e,2} < j_{p,2}$ means that

$$\begin{aligned} & N\left(x_1, \lambda_{p,1}^{D}(j_{e,2})\right) - \left(\lambda_{p,1}^{D}(j_{e,2}) - \psi\right) N_\lambda\left(x_1, \lambda_{p,1}^{D}(j_{e,2})\right) \\ > \;& N\left(x_2, \lambda_{p,2}^{D}(j_{e,2})\right) - \left(\lambda_{p,2}^{D}(j_{e,2}) - \psi\right) N_\lambda\left(x_2, \lambda_{p,2}^{D}(j_{e,2})\right). \end{aligned} \tag{C.45}$$

Subtracting $\psi N_\lambda\left(x_1,\lambda_{p,1}^D(j_{e,2})\right)$ from both sides then yields

$$\begin{aligned}
& N(x_1,\lambda_{p,1}^D(j_{e,2}))-\lambda_{p,1}^D(j_{e,2})\,N_\lambda\left(x_1,\lambda_{p,1}^D(j_{e,2})\right)\\
> & \; N\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)-\lambda_{p,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)\\
& +\psi\left(N_\lambda\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)-N_\lambda\left(x_1,\lambda_{p,1}^D(j_{e,2})\right)\right)\\
= & \; N\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)-\lambda_{p,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)+\psi\left(\frac{\eta_2}{y(j_{e,2})}-\frac{\eta_1}{y(j_{e,2})}\right)\\
> & \; N\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)-\lambda_{p,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{p,2}^D(j_{e,2})\right).
\end{aligned} \tag{C.46}$$

We already know from the first step that $\lambda_{e,1}^D(j)>\lambda_{p,1}^D(j)$ for any j. Together with

$$\frac{d}{d\lambda}\left(N(x_i,\lambda)-\lambda N_\lambda(x_i,\lambda)\right)=-\lambda N_{\lambda\lambda}(x_i,\lambda)>0, \tag{C.47}$$

this implies

$$\begin{aligned}
& N\left(x_1,\lambda_{e,1}^D(j_{e,2})\right)-\lambda_{e,1}^D(j_{e,2})\,N_\lambda\left(x_i,\lambda_{e,1}^D(j_{e,2})\right)\\
> & \; N\left(x_1,\lambda_{p,1}^D(j_{e,2})\right)-\lambda_{e,1}^D(j_{e,2})\,N_\lambda\left(x_i,\lambda_{p,1}^D(j_{e,2})\right)\\
> & \; N\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)-\lambda_{p,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{p,2}^D(j_{e,2})\right),
\end{aligned} \tag{C.48}$$

where the second inequality follows from (C.46). By the indifference condition defining $j_{e,2}$, we have

$$\begin{aligned}
& N\left(x_2,\lambda_{e,2}^D(j_{e,2})\right)-\lambda_{e,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{e,2}^D(j_{e,2})\right)\\
= & \; N\left(x_1,\lambda_{e,1}^D(j_{e,2})\right)-\lambda_{e,1}^D(j_{e,2})\,N_\lambda\left(x_i,\lambda_{e,1}^D(j_{e,2})\right).
\end{aligned} \tag{C.49}$$

It is then immediate that

$$\begin{aligned}
& N\left(x_2,\lambda_{e,2}^D(j_{e,2})\right)-\lambda_{e,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{e,2}^D(j_{e,2})\right)\\
> & \; N\left(x_2,\lambda_{p,2}^D(j_{e,2})\right)-\lambda_{p,2}^D(j_{e,2})\,N_\lambda\left(x_2,\lambda_{p,2}^D(j_{e,2})\right),
\end{aligned} \tag{C.50}$$

and therefore $\lambda_{e,2}^D(j_{e,2})>\lambda_{p,2}^D(j_{e,2})$. The latter requires $\mathcal{U}_2<\eta_2$, which in turn implies that $\lambda_{e,2}^D(j)>\lambda_{p,2}^D(j)$ for all j.

The final step is to show that the combination of $j_{e,2}<j_{p,2}$ and $\lambda_{e,2}^D(j)>\lambda_{p,2}^D(j)$ violates the resource constraint. We have

$$\begin{aligned}
\int_{j_{p,2}}^1 \lambda_{p,2}^D(j)\;dj=\phi_2 &= \int_{j_{e,2}}^1 \lambda_{e,2}^D(j)\;dj\\
&> \int_{j_{e,2}}^1 \lambda_{p,2}^D(j)\;dj\\
&> \int_{j_{p,2}}^1 \lambda_{p,2}^D(j)\;dj,
\end{aligned} \tag{C.51}$$

which is the desired contradiction. Hence, $j_{e,1} > j_{p,1}$.

(ii) Too few firms target directed searchers of type 2. Next, we establish that $j_{e,2} > j_{p,2}$. Towards a contradiction, suppose that $j_{e,2} \leq j_{p,2}$. The combination of $j_{e,1} > j_{p,1}$ and $j_{e,2} \leq j_{p,2}$ implies that $\lambda^D_{e,1}(j) > \lambda^D_{p,1}(j)$ and $\lambda^D_{e,2}(j) \leq \lambda^D_{p,2}(j)$ by the resource constraints. Since $N(x_i, \lambda) - \lambda N_\lambda(x_i, \lambda)$ is increasing in λ, we then have

$$\begin{aligned}
& N\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) - \lambda^D_{p,2}(j_{e,2}) N_\lambda\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) \\
\geq \;& N\left(x_2, \lambda^D_{e,2}(j_{e,2})\right) - \lambda^D_{e,2}(j_{e,2}) N_\lambda\left(x_2, \lambda^D_{e,2}(j_{e,2})\right) \\
= \;& N\left(x_1, \lambda^D_{e,1}(j_{e,2})\right) - \lambda^D_{e,1}(j_{e,2}) N_\lambda\left(x_1, \lambda^D_{e,1}(j_{e,2})\right) \\
> \;& N\left(x_1, \lambda^D_{p,1}(j_{e,2})\right) - \lambda^D_{p,1}(j_{e,2}) N_\lambda\left(x_1, \lambda^D_{p,1}(j_{e,2})\right),
\end{aligned} \tag{C.52}$$

where the second line follows from $\lambda^D_{p,2}(j) \geq \lambda^D_{e,2}(j)$, the third line follows from the indifference condition at $j_{e,2}$, and the fourth line follows from $\lambda^D_{e,1}(j) > \lambda^D_{p,1}(j)$. Adding $\psi N_\lambda\left(x_1, \lambda^D_{p,1}(j_{e,2})\right)$ to both sides and using (C.9), we then obtain

$$\begin{aligned}
& N(x_1, \lambda^D_{p,1}(j_{e,2})) - \left(\lambda^D_{p,1}(j_{e,2}) - \psi\right) N_\lambda\left(x_1, \lambda^D_{p,1}(j_{e,2})\right) \\
< \;& N\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) - \left(\lambda^D_{p,2}(j_{e,2}) - \psi\right) N_\lambda\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) \\
& - \psi\left(N_\lambda\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) - N_\lambda\left(x_1, \lambda^D_{p,1}(j_{e,2})\right)\right) \\
= \;& N\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) - \left(\lambda^D_{p,2}(j_{e,2}) - \psi\right) N_\lambda\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) - \psi\left(\frac{\eta_2}{y(j_{e,2})} - \frac{\eta_1}{y(j_{e,2})}\right) \\
< \;& N\left(x_2, \lambda^D_{p,2}(j_{e,2})\right) - \left(\lambda^D_{p,2}(j_{e,2}) - \psi\right) N_\lambda\left(x_2, \lambda^D_{p,2}(j_{e,2})\right).
\end{aligned} \tag{C.53}$$

In other words, the social planner prefers for a firm with $j = j_{e,2}$ to attract type-2 workers rather than type-1 workers. This implies $j_{e,2} > j_{p,2}$, which is the desired contradiction. Hence, $j_{e,2} > j_{p,2}$.

(iii) Too many applicants at high-productivity firms. Since $j_{e,2} > j_{p,2}$, the resource constraint for high-type workers,

$$\int_{j_{e,2}}^{1} \lambda^D_{e,2}(j)\, dj = \phi_2 = \int_{j_{p,2}}^{1} \lambda^D_{p,2}(j)\, dj, \tag{C.54}$$

requires $U_2 < \eta_2$. This in turn implies $\lambda^D_{e,2}(j) > \lambda^D_{p,2}(j)$ for all $j > j_{e,2}$, i.e., the firms at the top of the productivity distribution attract too many applicants. □

As before, too few firms attract directed searchers. This, in turn, implies that the directed searchers are concentrated among a smaller subset of firm types. Moreover, within these firms, part (ii) of Proposition 5 states that too few firms attract the type-2 workers, and, as a consequence, (iii) there are too many workers at the top of the firm productivity distribution. The economic mechanism for both results is the same and is driven by the monopsony distortion discussed earlier. Consider the choice of whether or not to attract directed searchers. In order to attract any directed searchers as opposed to none, a firm must raise its wage. However, because of the inability to wage-discriminate, it must also raise the wage for the random searchers. As a result, the firm perceives the cost of attracting directed searchers as larger than that for the social planner. This component is the same as for the homogeneous-worker case. However, and crucially, the same mechanism also distorts the firm's choice between type-1 and type-2 workers conditional on attracting directed searchers. In order to attract directed searchers of type 2 rather than type 1, the firm must raise the wage sufficiently. But, because a fraction of the workers are random searchers of lower productivity, the cost of doing so perceived by the firm is larger than that for the social planner. In other words, conditional on attracting directed searchers, firms are inefficiently reluctant to attract type 2.