Directed Search with Multiple Vacancies

Benjamin Lester∗†
University of Western Ontario

July 9, 2008

Abstract

Preliminary and incomplete: please do not circulate.

∗Contact: blester@uwo.ca.
†I would like to thank Audra Bowls, Braz Camargo, Philipp Kircher, Guido Menzio, and participants at the Penn Search and Matching workshop for helpful discussions. All errors are my own.
1 Introduction

A primary objective of labor economists is to understand the relationship between the economic environment, wages, and the market’s ability to allocate workers efficiently across firms. The literature on directed search offers an appealing framework for addressing many of these important issues, because it models explicitly the micro-foundations of both wage formation and matching. In this sense, the model can provide unique insights into the interaction between firm behavior and market efficiency: firms’ decisions affect the level of frictions in the labor market, and conversely the level of frictions in the labor market affects firms’ decisions.

The benchmark model of directed search, however, makes a variety of severe assumptions on the characteristics and behavior of workers and firms, thus limiting its scope as an analytical tool. One particularly restrictive assumption is that a firm can open at most a single position, so that there is no distinction between a vacancy and a firm, or between job creation along the intensive and extensive margins. In this paper, we relax this assumption and study the importance of firm size, firm growth, and market structure on the determination of wages, matching, and output. In particular, we solve a general equilibrium model in which profit-maximizing firms choose the number of vacancies to open and the wage associated with those vacancies, and workers choose the optimal application strategy to maximize their ex-ante expected payoff.

We fully characterize the (unique) equilibrium, and illustrate how the fundamentals of the economy affect the distribution of vacancies and wages across firms, and thus the strength of frictions in the matching process. At the individual level, these frictions will have important implications for the job-finding rate of workers and the job-filling rates of (potentially heterogeneous) establishments. At the aggregate level, of course, these frictions will determine the equilibrium level of employment and output. We will also illustrate how allowing firms to post multiple vacancies changes the implications for a variety of policy experiments that were previously analyzed under the assumption that the matching technology was independent of the economic environment. Finally, allowing for ex-ante heterogeneity across firms, we explore the extent to which such a model can generate the relationships between firms size, wages, prof-
itability, and hiring outcomes observed in the data.

1.1 Related Literature

As noted above, a distinguishing feature of the literature on directed search is that the processes of both wage formation and matching are endogenous. This stands in contrast to most alternative models of labor markets with search frictions, in which either wages or the matching technology is specified exogenously. For example, the large literature typified by Pissarides [24] and Mortensen and Pissarides [21] has been used extensively to study the behavior of wages, employment, and output. However, these models typically introduce frictions by employing an exogenously specified matching function that posits a relationship between the number of workers searching for a job, the number of vacant positions, and the number of resulting matches. Importantly, this matching function is generally assumed to be invariant to changes in the economic environment.

In response to this limitation, several authors have sought to establish micro-foundations for the matching function, such as Coles and Smith [8], Lagos [16], and Stevens [32]. However, the models used to generate endogenous matching functions have typically assumed a very limited role for wages in the matching process; wages play no allocative role in any of these models, and are often either fixed exogenously or determined via some ad hoc rule for the division of surplus between workers and firms.

Despite the potential advantages of the directed search paradigm, the “benchmark” model is limited in scope by a variety of severe assumptions on workers and firms.¹ Many of these assumptions have been relaxed in recent research. Albrecht, Gautier, and Vroman [3] and Galenianos and Kircher [10] allow workers to apply to multiple firms simultaneously. Julien, Kennes, and King [15], Coles and Eeckhout [7], and Shi [27] allow firms to post alternative pricing mechanisms. Acemoglu and Shimer [2] allow for risk-aversion and ex-ante capital investment, while Shi [25] allows for heterogeneous skills among workers. Shimer and Wright [31], Guerrieri [12], Menzio [18], and Lester [17] introduce informational frictions. However, little attention has

¹The literature on directed search can be traced back to Peters ([22], [23]) and Montgomery [20], and was further developed by Shimer [28], Moen [19], and Burdett, Shi, and Wright [6]. The “benchmark” model that is referred to here is closest in formulation to that in Burdett, Shi, and Wright [6].
been paid to the assumption that each firm can post only a single vacancy, and thus the model is silent on the importance of firm size, firm growth, or the distinction between job creation along the intensive and extensive margins.

Burdett, Shi, and Wright [6] first make the point that the distribution of vacancies across firms is a key determinant of wages and matching, but they do so by allowing for exogenous heterogeneity in the number of vacancies posted by each firm. Therefore, though they identify this type of heterogeneity as being potentially important in determining equilibrium outcomes, they are silent on the sources of firm heterogeneity and thus on the causes of changes in equilibrium wages, matching, and output. Shi [26] allows firms to grow by posting a single vacancy in consecutive periods, but since the distribution of vacancies in each period remains degenerate this has no consequences for the efficiency of the matching process. In contrast to these earlier works, the current paper characterizes the endogenously determined distribution of vacancies across firms and associated wages, allowing for firms to post multiple vacancies simultaneously.\footnote{Also see Hawkins [13], who considers multi-worker firms in a directed search environment in order to analyze the efficiency properties of the model.}

\section{The Model}

There is a fixed measure of both unemployed workers and firms. We denote by \( r \) the ratio of unemployed workers to firms, and assume that \( 0 < r < \infty \). The game proceeds in two stages. In stage one, firms face two decisions. The first decision is how many vacancies to post. For simplicity, we assume that they can either post one vacancy at cost \( C_1 \), or two vacancies at cost \( C_1 + C_2 \). A firm that is matched with one worker produces output \( y_1 \), and a firm that is matched with two workers produces output \( y_1 + y_2 \). If the cost function is concave (\( 0 \leq C_2 \leq C_1 \)) and the production function is convex (\( 0 \leq y_1 \leq y_2 \)), then in equilibrium all firms will either post two vacancies or remain inactive (i.e. post no vacancies). The more interesting case, in which some firms post a single vacancy and others post two vacancies, thus requires \textit{either} convexity in the costs of posting vacancies or concavity in the production function. As it is analytically more convenient, we will choose the former option, and assume that \( y_1 = y_2 \equiv y > 0, C_1 = 0, \)
and $0 < C_2 < y$. The second decision that firms face is the wage at which they are committed to paying their workers. We restrict our analysis to the case in which a firm with two vacancies sets the same wage for both positions.

In stage two, each worker observes the wage and the number of vacancies (hereafter referred to as the capacity) of all firms and applies to the firm that offers the highest expected return. We assume that workers can only apply to a single firm. If the number of workers that arrive at a particular firm exceeds capacity, the firm allocates the position(s) at random, with each worker receiving a job with equal probability. Therefore, the expected payoff from applying to each firm depends on both the posted wage and the probability of receiving the job.

2.1 Stage Two: Optimal Job Search

In stage two, we take the distribution of wages and capacities as given, and search for a symmetric strategy Nash equilibrium. As a first step, we must derive the expected value of a worker visiting a firm that has chosen capacity $k \in \{1, 2\}$ and has posted the wage $w$. In order to do so, it is convenient to consider the case of a finite number of workers and firms in fixed proportion, and then allow the number of workers (or firms) to tend to infinity. To that end, let $u$ denote the number of unemployed workers and $f$ denote the number of firms, with $r = u/f$.

Consider a firm $j$ who has one vacancy posted at wage $w$, and suppose that each worker applies to this firm with probability $\theta$. Let $\Omega_1$ denote the probability that a given worker $i$ is matched with this firm, conditional on visiting. Then the unconditional probability that worker $i$ gets matched with firm $j$ is equal to $\Omega_1 \theta$. This, of course, must equal the probability that

---

3 All of the results presented below remain true under the alternative assumptions that there are decreasing returns to scale in the production technology.

4 In reality, of course, workers can apply to many firms simultaneously. Allowing for multiple applications, even when firms’ capacity is specified exogenously, introduces considerable complexity. See Albrecht, Gautier, and Vroman [3] and Galenianos and Kircher [10]. Admittedly, it would be interesting to examine the case where firms can choose to post multiple vacancies (as in the current paper) and workers can apply to multiple positions (as in the aforementioned papers). However, to establish a benchmark, we maintain the assumption here that workers are constrained to one application.

5 Restricting attention to symmetric strategies for workers is standard in this literature, and crucial for generating a coordination friction. This assumption is generally justified by assuming, as we do, that the labor market is large and workers are anonymous, thus making it difficult to coordinate on asymmetric strategies. See both Burdett, Shi, and Wright [6] and Shimer [29] for a more detailed discussion.

6 This follows the analysis of Burdett, Shi, and Wright [6].
firm $j$ is matched with worker $i$. The probability that the firm is matched with any worker (i.e. the probability that at least one worker applies) is given by $1 - (1 - \theta)^u$, and the probability of being matched with each particular worker is equal across $u$ workers. Therefore, we have that $\Omega_1 \theta = \frac{1 - (1 - \theta)^u}{u}$. As a result, the probability of each worker being matched with a firm that has one vacancy when all workers apply with probability $\theta$ is

$$\Omega_1(\theta) = \frac{[1 - (1 - \theta)^u]}{u \theta}. \quad (1)$$

Moreover, the expected payoff to each worker from applying to a firm with one vacancy at a posted wage $w$ when all workers are applying with probability $\theta$ is equal to $\Omega_1(\theta)w$.

Now consider the analogous problem of worker $i$ who is considering applying to a firm with two vacancies and wage $w$. If $u - 1$ other workers apply to this firm with probability $\theta$, then the probability that worker $i$ is hired is

$$\Omega_2(\theta) = (1 - \theta)^{u-1} + (u - 1)\theta(1 - \theta)^{u-2} + \sum_{i' = 2}^{u-1} \left\{ \frac{(u - 1)!}{i'!(u - 1 - i')!} \theta(1 - \theta)^{u-1-i'} \frac{2}{i' + 1} \right\} \quad (2)$$

$$= \frac{2}{u \theta} [1 - (1 - \theta)^u] - (1 - \theta)^{u-1}. \quad (3)$$

Naturally, the expected payoff from a worker applying to a firm with two vacancies at a posted wage $w$ when other workers are applying with probability $\theta$ is equal to $\Omega_2(\theta)w$.

Therefore, a symmetric strategy Nash equilibrium of the stage two game with a finite number of players is a strategy $\theta^* \equiv (\theta^*_1, ..., \theta^*_f)$ with the properties that

$$\theta^*_j > 0 \iff \Omega_{k_j}(\theta^*_j)w_j \geq \Omega_{k_{j'}}(\theta^*_j)w_{j'} \; \forall j' \neq j \quad (4)$$

for all firms $j$ with capacity $k_j$ and wage $w_j$, and

$$\sum_{j=1}^{f} \theta^*_j = 1. \quad (5)$$

A straight-forward extension of Peters [22] can be employed to establish that such an equilibrium exists and is unique.

We now consider an economy with $r = u/f$ fixed as $u \to \infty$. Our analysis is considerably simplified by the following conjecture about the equilibrium behavior of firms, which we confirm
to be true in the next section: all firms with the same capacity $k \in \{1, 2\}$ post the same wage $w_k$ in equilibrium. This implies that a worker is essentially choosing between two types of firms. Suppose $f_1 \leq f$ firms have a single vacancy and post wage $w_1$ (a type 1 firm), while $f_2 = f - f_1$ firms have two vacancies and post wage $w_2$ (a type 2 firm). Then a worker’s strategy can be characterized by the probability that he applies to some type 1 firm, which we denote $\sigma$, and the probability that he applies to some type 2 firm, $1 - \sigma$. Note that he applies to each type $k$ firm with equal probability, for $k \in \{1, 2\}$.

Let $\phi = f_1/f$ denote the fraction of type 1 firms. Moreover, if workers apply to each type 1 firm with equal probability $\theta_1$, then $\theta_1 = \sigma/f_1$. Using equation (1), we see that the expected payoff to a worker from applying to a type 1 firm as the number of agents tends to infinity is given by

$$
\lim_{u \to \infty} \left\{ \left[ 1 - \left( 1 - \left( r\sigma / \phi \right) \right)^u \right] \left( u\sigma / f\phi \right) \right\} w_1 = \left\{ \left[ 1 - \exp \left( -r\sigma / \phi \right) \right] \left( r\sigma / \phi \right) \right\} w_1.
$$

Similarly, if workers apply to all type 2 firms with equal probability $\theta_2$, then $\theta_2 = (1 - \sigma)/f_2$. Using equation (3), we find that the expected payoff from applying to a type 2 firm converges

$$
\lim_{u \to \infty} \left\{ \frac{2(1 - \phi)}{[(r(1 - \sigma))]} \left[ 1 - \left( 1 - \left( r(1 - \sigma) / u(1 - \phi) \right) \right)^u \right] - \left[ 1 - \left( r(1 - \sigma) / u(1 - \phi) \right)^{u-1} \right] \right\} w_2
$$

$$
= \left\{ \frac{2(1 - \phi)}{[(r(1 - \sigma))]} \left[ 1 - \exp \left( -r(1 - \sigma) / (1 - \phi) \right) \right] - \exp \left( -r(1 - \sigma) / (1 - \phi) \right) \right\} w_2.
$$

It will be convenient to define two variables

$$
q_1(\sigma, \phi) = \left( r\sigma / \phi \right) \quad (6)
$$

$$
q_2(\sigma, \phi) = \frac{r(1 - \sigma)}{(1 - \phi)}, \quad (7)
$$

which represent the expected number of workers (or expected queue length) at a type 1 and type 2 firm, respectively. In what follows, we will suppress the arguments of $q_1$ and $q_2$ for convenience, though the relationship between these values and the strategies $\sigma$ and $\phi$ should be understood.

In any equilibrium with strictly positive values of $q_1$ and $q_2$, it must be that the expected payoff to a worker is the same at type 1 and type 2 firms, so that

$$
\left\{ \left[ 1 - \exp(-q_1) \right] / q_1 \right\} w_1 = \left\{ \frac{2}{q_2} \left[ 1 - \exp(-q_2) \right] - \exp(-q_2) \right\} w_2 \equiv U. \quad (8)
$$
The variable $U$ has been introduced to denote the maximal value of workers’ expected payoffs, which is often referred to as the market utility in the existing literature.

### 2.2 Stage One: Profit Maximization

Now consider the problem facing firm $j \in [0,1]$. In stage one, firms take as given the optimal search behavior of unemployed workers in stage two. In particular, a firm takes the market utility as exogenous, and understands that, for any capacity choice $k \in \{1,2\}$ and wage $w_k$, the expected number of applications will be dictated by the relationship described in equation (8).

Therefore, one can show that a type 1 firm solves

$$
\pi_1(U) = \max_{q_1} \left[ \frac{1 - \exp(-q_1)(y - w_1)}{1 - \exp(-q_1)} \right] w_1 = U.
$$

One can substitute the constraint into the objective function, so that the problem of a type 1 firm can be re-written as

$$
\pi_1(U) = \max_{q_1} y[1 - \exp(-q_1)] - q_1 U.
$$

Note that the firm’s problem is strictly concave in $q_1$, and thus in $w$; this confirms the earlier conjecture that all firms with one vacancy will optimally post the same wage. More specifically, the optimal $q_1$ will satisfy $U = y[\exp(-q_1)]$, and the equilibrium wage must satisfy

$$
w_1 = \frac{y[\exp(1 + q_2)]}{1 - \exp(-q_1)}.
$$

Similarly, one can establish that a type 2 firm solves

$$
\Pi_2(U) = \max_{q_2} \left\{ 2[1 - \exp(-q_2)] - q_2 \exp(-q_2) \right\} (y - w_2) - C_2
$$

s.t. \quad \left\{ \frac{2[1 - \exp(-q_2)] - q_2 \exp(-q_2)}{q_2} \right\} w_2 = U.

Repeating the same steps as above, it is clear that there exists a unique profit-maximizing wage for all firms with capacity equal to two. Specifically, $q_2$ will satisfy the first order condition $U = y[\exp(-q_2)(1 + q_2)]$ and the profit-maximizing wage is thus given by

$$
w_2 = \frac{y[\exp(1 + q_2)]}{2[1 - \exp(-q_2)] - q_2 \exp(-q_2)}.
$$
Of course, in any equilibrium in which both types of firms are active, it must be the case that profits are equal across types:

\[ [1 - \exp(-q_1)](y - w_1) = \{2[1 - \exp(-q_2)] - q_2 \exp(-q_2)\} (y - w_2) - C_2. \quad (11) \]

### 2.3 Equilibrium

We now define an equilibrium of the directed search model with endogenous capacity described above. We begin by characterizing an interior equilibrium, in which some firms post one vacancy and others post two, and then consider the conditions under which all firms post the same number of vacancies.

**Definition 1.** An equilibrium with heterogeneous firms is a pair of strategies \((\sigma^*, \phi^*) \in (0, 1)^2\), along with expected queue lengths \((q_1^*, q_2^*)\) and wages \((w_1^*, w_2^*)\) such that (i) firms earn equal profits, as in (11); (ii) the worker’s indifference condition (8) is satisfied; (iii) expected queue lengths are consistent with equilibrium strategies, as in (6) and (7); and (iv) wages are consistent with profit maximization, as in (9) and (10).

Substituting (9) and (10) into (8) and (11), the equilibrium conditions simplify to:

\[
\exp(-q_1) - \exp(-q_2)(1 + q_2) = 0 \quad (12)
\]

\[
\exp(-q_1)(1 + q_1) - \exp(-q_2)[2(1 + q_2) + q_2^2] + 1 - \frac{C_2}{y} = 0. \quad (13)
\]

Since \(q_1\) and \(q_2\) are simple functions of \(\sigma\) and \(\phi\), equilibrium is characterized by two equations in two unknowns. Let \(\Sigma_W(\phi) \leftarrow \sigma\) denote the relationship implied by the worker’s indifference condition (12), and \(\Sigma_F(\phi) \leftarrow \sigma\) denote the relationship implied by the firm’s equal profit condition (13).

We now present a series of results characterizing the existence, uniqueness, and regularity of interior equilibria. All of the proofs can be found in the appendix.

**Lemma 1.** The implicit functions \(\Sigma_W\) and \(\Sigma_F\) have the following properties:

1. \(\lim_{\phi \to 0} \Sigma_W(\phi) = 0\), \(\lim_{\phi \to 1} \Sigma_W(\phi) = 1\), and \(\lim_{\phi \to 1} \Sigma_F(\phi) = 1\).
2. $\Sigma_W$ and $\Sigma_F$ are increasing functions of $\phi$.

3. If there exists a $\phi^* \in (0, 1)$ such that $\Sigma_W(\phi^*) = \Sigma_F(\phi^*)$, then $\Sigma'_F(\phi^*) > \Sigma'_W(\phi^*)$. Therefore, there is at most one $\phi^* \in (0, 1)$ such that $\Sigma_W(\phi^*) = \Sigma_F(\phi^*)$.

Lemma 1 establishes that if an interior equilibrium exists, it is unique. To ensure existence, we derive conditions on the exogenous parameters $\{r, C_2, y\}$ that ensure $\Sigma_W(\phi) > \Sigma_F(\phi)$ for values of $\phi$ arbitrarily close to zero and $\Sigma_F(\phi) > \Sigma_W(\phi)$ for values of $\phi$ arbitrarily close to one. For convenience, we denote the cost of posting a vacancy relative to output by $c_2 \equiv C_2/y$, where $c_2 < 1$ by construction.

**Lemma 2.** Let $\bar{r}(c_2)$ denote the value of $r$ that satisfies

$$1 - c_2 = \exp(-r)[1 + (1 + r)\ln(1 + r)]. \quad (14)$$

Also let $\tilde{q}_2(r)$ denote the value of $q_2$ that satisfies

$$\exp(-r) = [1 + q_2] \exp[-q_2] \quad (15)$$

and $\underline{r}(c_2)$ denote the value of $r$ that satisfies

$$1 - c_2 = \exp[-\tilde{q}_2(r)]\{[1 + \tilde{q}_2(r)](1 - r) + \tilde{q}_2(r)^2\}. \quad (16)$$

Then:

1. there exists $\eta_1 \in (0, 1)$ such that $\Sigma_F(\phi) > \Sigma_W(\phi) \forall \phi \in (1 - \eta_1, 1) \iff r > \bar{r}(c_2)$.

2. there exists $\eta_2 \in (0, 1)$ such that $\Sigma_W(\phi) > \Sigma_F(\phi) \forall \phi \in (0, \eta_2) \iff r < \underline{r}(c_2)$.

From Lemmas 1 and 2, we can establish that there are three possible cases, corresponding to three possible types of equilibria. First, consider the second stage game when $r \geq \bar{r}(c_2)$ and some fraction $\phi' \in (0, 1)$ of firms have posted a single vacancy. Equilibrium in the second stage game requires that workers are indifferent between applying to type 1 and type 2 firms, so that the equilibrium strategy of the representative worker will be $\Sigma_W(\phi')$. Note that (i) $r \geq \bar{r}(c_2) \Rightarrow \Sigma_W(\phi') < \Sigma_F(\phi')$ for any $\phi' \in (0, 1)$, and (ii) that a firm will strictly prefer to post
two vacancies if \( \Sigma_W(\phi') < \Sigma_F(\phi') \). Therefore, in the parameter region in which \( r \geq \bar{r}(c_2) \), a firm will always find it preferable to post two vacancies, and thus the unique equilibrium is for all firms to post two vacancies. This case is illustrated in Figure 1. Similar reasoning reveals that when \( r \leq \underline{r}(c_2) \), \( \Sigma_W(\phi) > \Sigma_F(\phi) \) for all \( \phi \in (0, 1) \) and the unique equilibrium is for all firms to post a single vacancy. This case is illustrated in Figure 2. Only when \( \underline{r}(c_2) < r < \bar{r}(c_2) \) does there exist an equilibrium in which both types of firms are active, and this equilibrium is unique. This case is illustrated in Figure 3. Proposition 1 below formalizes this reasoning, and completely characterizes equilibrium in the model of directed search with endogenous capacity choice. The proof can be found in the appendix.

**Proposition 1.** For any \( 0 < r < \infty \) and \( C_2/y \equiv c_2 < 1 \), there exists a unique equilibrium. If \( r \geq \bar{r}(c_2) \), then \( \sigma^*(r, c_2) = \phi^*(r, c_2) = 0 \) and \( q_2^* = r \). If \( r \leq \underline{r}(c_2) \), then \( \sigma^*(r, c_2) = \phi^*(r, c_2) = 1 \) and \( q_1^* = r \). If \( \underline{r}(c_2) < r < \bar{r}(c_2) \), then

\[
\sigma^*(r, c_2) = \frac{r(c_2) [\bar{r}(c_2) - r]}{r [\bar{r}(c_2) - \underline{r}(c_2)]}, \quad (17) \\
\phi^*(r, c_2) = \frac{\bar{r}(c_2) - r}{\bar{r}(c_2) - \underline{r}(c_2)} \quad (18)
\]

with \( q_1^* = \underline{r}(c_2) \) and \( q_2^* = \bar{r}(c_2) \). In this equilibrium, \( \partial \phi^*/\partial c_2 > 0 \), \( \partial \sigma^*/\partial c_2 > 0 \), \( \partial \phi^*/\partial r < 0 \), and \( \partial \sigma^*/\partial r < 0 \).

We now discuss the conditions on \( r \), \( C_2 \), and \( y \) that determine which of the three potential equilibria are attained.\(^7\) If the ratio of unemployed workers to firms is sufficiently large, the cost of posting a second vacancy is sufficiently small, or the output from a match is sufficiently high, so that \( r > \bar{r}(c_2) \), the unique equilibrium is \( (\sigma^*, \phi^*) = (0, 0) \); all firms post two vacancies and workers apply to these types of firms with probability one. The wage is given by (10), with \( q_2 = r \). In this region of the parameter space, the probability of receiving at least two workers is sufficiently high to justify the relatively low costs of posting a vacancy, independent of the number of other type two firms. On the other hand, if the ratio of unemployed workers to firms is sufficiently small, the cost of posting a vacancy is sufficiently high, or the output from a match

\(^7\)For the following reasoning, it is useful to note that both \( \underline{r}(c_2) \) and \( \bar{r}(c_2) \) are increasing functions of \( c_2 \).
is sufficiently low, so that \( r < r(c_2) \), the unique equilibrium is \((\sigma^*, \phi^*) = (1, 1)\); all firms post a single vacancy and workers apply to these types of firms with probability one. Here the wage is given by (9), with \( q_1 = r \). In this region of the parameter space, the probability of finding a second worker is not high enough to justify the relative cost of the vacancy, independent of the strategies of other firms. When \( r(c_2) < r < \bar{r}(c_2) \), the unique equilibrium is a mixed strategy equilibrium; some firms post one vacancy, some firms post two vacancies, and workers apply to each type of firm with strictly positive probability. Wages at a type 1 firm are given by (9), with \( q_1 = r(c_2) \), and wages at a type 2 firm are given by (10), with \( q_2 = \bar{r}(c_2) \). Figure 4 illustrates the disjoint regions of the parameter space in which each of these type of equilibria arise.
There are several interesting features of the interior equilibria characterized above. To start, these equilibria exhibit wage dispersion despite our assumptions that workers and firms are homogenous and workers have access to perfect information about firms’ wages and capacity.\footnote{An analogous result is derived in Shi \cite{Shi2012}.} Indeed, one can show that in equilibrium $w_1$ is greater than $w_2$; type 2 firms offer workers a greater chance of being hired, and hence extract a larger portion of the surplus in each match. Of course, this prediction is counter-factual: it is well-documented that, ceteris paribus, larger firms pay higher wages (see Brown and Medoff \cite{Brown1983}). We attempt to reconcile the model with this fact in section 3.

Also note that a change in the ratio of workers to firms only has effects along the extensive margin in an interior equilibrium. More specifically, the expected queue lengths, wages, and profits at type 1 and type 2 firms, as well as the market utility of workers, are independent of $r$ when $r \in [\underline{\tau}(c_2), \bar{\tau}(c_2)]$; an increase in $r$ causes a proportional increase in the fraction of type 2 firms, leaving all agents equally well off. Outside of this region, queue lengths and profits are increasing in $r$ while wages and market utility are decreasing.

### 2.4 Matching

We will now characterize the aggregate number of matches formed in equilibrium. Unlike standard exogenous matching functions, or even the endogenous matching function generated by the benchmark model of directed search, the number of unemployed workers and vacancies will \textit{not} be sufficient statistics to determine the equilibrium number of matches in the current model. Instead, in order to characterize the number of matches that arise in equilibrium, one would also need to know the distribution of vacancies across firms, and thus the state of the economic environment.

Let $m_k(q_k)$ denote the expected number of matches at a firm with capacity $k$ when the expected queue length is $q_k$. We established earlier that

\begin{align*}
    m_1(q_1) &= [1 - \exp(-q_1)] \\
    m_2(q_2) &= \{2[1 - \exp(-q_2)] - q_2 \exp(-q_2)\}. \tag{20}
\end{align*}
Then the matching technology can be summarized by

\[
M(u, f, \sigma, \phi) = \begin{cases} 
    f \left\{ \phi m_1 \left[ \frac{u}{f \sigma} \right] \right\} + (1 - \phi) m_2 \left[ \frac{u(1 - \sigma)}{f(1 - \phi)} \right] : \phi \in (0, 1) \\
    f \left\{ \phi m_2 \left[ \frac{u}{f} \right] \right\} : \phi = 0 
\end{cases}
\]

where \( M \) denotes the aggregate number of matches formed when a measure \( f \) of firms post a single vacancy with probability \( \phi \) (and post two vacancies with probability \( 1 - \phi \)), and a measure \( u \) of unemployed workers apply to a single-vacancy firm with probability \( \sigma \) (and apply to a two-vacancy firm with probability \( 1 - \sigma \)).

Naturally, we have that \( \frac{\partial M}{\partial u} > 0 \) and \( \frac{\partial M}{\partial f} > 0 \), which are standard results for any matching technology. However, also note that the number of total matches is decreasing in the fraction of type 1 firms, given equilibrium behavior of workers at the second stage; that is, \( \frac{\partial M}{\partial \sigma} \bigg|_{\sigma = \Sigma_w(\phi)} < 0 \).

The reasoning behind this result is that, for any values \( u > 0 \) and \( f > 0 \), \( m_1[u/(2f)] < m_2(u/f) \).

In words, given the same number of unemployed workers and vacancies, more matches will be formed if \( f \) firms are posting two vacancies than if \( 2f \) firms are posting a single vacancy: the higher the concentration of vacancies, the more efficient the matching function. This concept is crucial to understanding the dependence of the matching function on state variables. A higher concentration of vacancies reduces coordination frictions. This is best understood by considering the limiting case: if a single firm posts all of the vacancies, matching is most efficient as the short side of the market will be perfectly matched.

Therefore, those parameter values that induce a high concentration of vacancies will be associated with more efficient matching. Let us define the number of matches formed in equilibrium by

\[
M^* = M(u, f, \sigma^*, \phi^*).
\]

Since \( \bar{r}, \tilde{r}, \sigma^* \), and \( \phi^* \) are increasing functions of \( c_2 \), the aggregate matching function is more efficient when the ratio of workers to firms \( r \) is high, the costs of posting a vacancy \( C_2 \) are low, or the level of aggregate productivity \( y \) is high.

Let us also define the function

\[
V(f, \phi) = f [\phi + 2 (1 - \phi)], \quad (21)
\]
so that $V$ represents the total number of vacancies posted when a fraction $\phi$ of a measure $f$ of firms post a single vacancy, and the remainder post two vacancies. Again, we let

$$V^* = V(f, \phi^*),$$

where $\phi^*$ represents the equilibrium strategy of a representative firm when the ratio of workers to firms is $r = u/f$ and the cost of posting a second vacancy relative to the productivity of a match is $c_2$.

The job-finding rate is the probability that a randomly selected worker is matched with a firm. Since workers are homogeneous and strategies are symmetric, the equilibrium job-finding rate is simply $M^*/u$, the aggregate number of matches divided by the number of workers searching for a job. Similarly, the job-filling rate is the probability that a randomly selected vacancy is filled, which is simply $M^*/V^*$.

2.4.1 Implications of an Endogenous Matching Function

To contrast the implications of the endogenously generated matching function derived here with a typical, exogenously specified matching function, consider the effects of increasing the number of firms in an industry, holding constant the number of unemployed workers, the cost of posting a vacancy, and the productivity of a match. Figure 5 illustrates the job-finding rate generated from the current model, as well as the number of vacancies that are being posted in equilibrium.\(^9\)

Had we assumed that a firm is a fixed number of vacancies, the entry of an additional firm should unambiguously increase the job-finding rate. However, when firms choose the number of vacancies to post, the entry of an additional firm implies more competition for workers and thus a decrease in the incentive of other firms to post a second vacancy. Therefore, there are two effects on the job-finding rate from the entry of an additional firm. The first, positive effect will be an increase in the aggregate number of vacancies.\(^10\) The second, negative effect is that the distribution will skew towards single-vacancy firms, thus decreasing the efficiency of

---

\(^9\)In this example, we set $u = 1$, $y = 1$, and $C_2 = .25$, though there is nothing special about this parameterization; the phenomena discussed above are robust features of the model.

\(^10\)It is not obvious that the number of vacancies is increasing in $r$, but it can be shown analytically.
the matching technology. As the figure clearly indicates, this second effect may dominate, thus causing a non-monotonicity in the job-finding rate as a function of the number of workers or, more generally, as a function of the ratio of workers to firms. This finding could have important policy implications. Since many policies affect the number of workers in the labor market (such as unemployment insurance) and the number of active firms (such as anti-trust laws), a model’s failure to account for changing frictions in the labor market could lead to misguided policy recommendations.

Now consider the sensitivity of equilibrium matching to perturbations in $y$. The solid line in Figure 6 illustrates the equilibrium job-finding rate over different values of productivity $y$, given a fixed number of firms, unemployed workers, and the posting cost $C_2$.\(^{11}\) Naturally, the job-finding rate is increasing in $y$: as the productivity of a match increases, firms post additional vacancies and there are more vacancies per worker. Indeed, any sensible matching function

\(^{11}\)For this figures, we have chosen the values $f = u = 1$ and $C_2 = .25$. Again, there is nothing special about these values: the general result holds for all values of $f > 0$, $u > 0$, and $C_2 < y$. 

\[\text{Figure 5: Job-Finding (r)}\]
should deliver a job-finding rate that is increasing in the number of vacancies. The distinction here is that the degree to which the job-finding rate is increasing in $y$ is affected not only by the number of additional vacancies created but also by the change in the distribution of vacancies across firms. In particular, a greater fraction of firms are posting two vacancies, and as a result there is a decrease in the coordination frictions and an increase in the efficiency of the matching process.

We plot two additional lines in attempt to illustrate the effect of the changing distribution of vacancies across firms, controlling for the number of vacancies. The dotted line is a plot of $m_1[u/V(u, f, c_2)]/u$: this is the job-finding rate if we assume that the number of vacancies posted for each value of $y$ is equal to the number of vacancies posted by firms in the equilibrium of our model, but that each vacancy is posted by a single firm. That is, the dotted line represents the job-finding rate if all vacancies are created by the entry of single-vacancy firms. The dashed line is a plot of $m_2\{u/[V(u, f, c_2)/2]\}/u$: this is the job-finding rate if we assume that the number of vacancies posted for each value of $y$ is equal to the number of vacancies posted by firms in the equilibrium of our model, but that each firm is posting two vacancies. That is, the dashed line represents the job-finding rate if all vacancies are created by the entry of two-vacancy firms.

Figure 6: Job-Finding ($y$)

Figure 7: Job-Filling ($y$)

It is apparent from this figure that accounting for the change in the distribution of vacancies across firms amplifies the response of the job-finding rate to changes in the level of aggregate productivity. An economist who exogenously assumed that each vacancy was created by a sin-
gle firm would underestimate the job-finding rate for large values of $y$, while an economist who exogenously assumed that every two vacancies were created by a single firm would overestimate the job-finding rate for small values of $y$. Indeed, any assumption that shuts down the distributional effects discussed above will underestimate the volatility in the job-finding rate in response to aggregate shocks; exogenously assuming that some fraction $\chi \in (0, 1)$ of vacancies are created by one-vacancy firms and the remaining fraction $1 - \chi$ of vacancies are created by two-vacancy firms will both underestimate the job-finding rate for large values of $y$ and overestimate the job-finding rate for small values of $y$.

Similar analysis can be applied to the job-filling rate, which is represented by the solid line in Figure 7. Again, we plot two additional lines: the dotted line is $m_1[u/V(u, f, c_2)]/V(u, f, c_2)$ and the dashed line is $m_2[u/[V(u, f, c_2)/2]]/V(u, f, c_2)$. These represent the job-filling rate assuming that all firms post one vacancy (the dotted line) or two vacancies (the dashed line), when the number of vacancies for each value of $y$ coincides with the number of vacancies created by firms in the equilibrium of our model. Notice that an economist who exogenously assumed that each vacancy was created by a single firm would underestimate the job-filling rate for large values of $y$. Similarly, an economist who exogenously assumed that every two vacancies were created by a single firm would overestimate the job-filling rate for small values of $y$. Again, it follows that any assumption that shuts down the distributional affects of the matching process will underestimate the volatility of the job-filling rate in response to aggregate shocks.

Though the sensitivity analysis above is generated from a static model, it has potentially interesting implications for a dynamic setting. In particular, if standard models are underestimating (over-estimating) the job-finding and the job-filling rates in good (bad) states of the world, then they are also potentially under-estimating (over-estimating) the incentives of workers to search for jobs and firms to post vacancies in these states, respectively. Therefore, the model developed here suggests that the assumption of a state-independent matching function may contribute to the well-known inability of standard models to capture the volatility of vacancy creation and unemployment over the business cycle.\(^{12}\) However, such suggestions at this stage

\(^{12}\)See Andalfatto [4] and Shimer [30].
are purely heuristic; a more convincing argument would require both a dynamic model and a proper calibration.\textsuperscript{13}

3 Heterogeneous Firms

An impetus for early models of directed search, in particular that of Montgomery [20], was to explain wage differentials amongst homogeneous workers. The mechanism in such models was straightforward: by posting a higher wage, a firm could increase the expected number of applications it would receive, and thus increase the probability of filling its vacancy. Therefore, if firms were heterogeneous with respect to the productivity of a match, those firms with high productivity (i.e. a high cost of not filling a vacancy) would optimally choose to post high wages. With this mechanism, Montgomery [20] was able to generate wage differentials amongst homogeneous workers that were positively correlated with firms’ profitability and job-filling (or vacancy yield) rates, as in the data.\textsuperscript{14} In the spirit of Montgomery’s initial experiment, we now incorporate firm heterogeneity into the model with endogenous capacity, and study the relationship between wage differentials, profitability, vacancy yields, and firm size.

3.1 Extending the Model

Suppose now that we allow firms to be heterogeneous with respect to their level of productivity. We normalize the measure of firms to 1, and assume that each firm \(i \in [0, 1]\) produces output \(y^i\) for each employed worker. This idiosyncratic productivity is distributed according to the cumulative density function \(F(y^i)\), which is assumed to be continuously differentiable with full support over the interval \([0, \bar{y}] \subset \mathbb{R}_+\). We maintain the assumption that there exists a measure of homogeneous workers, and denote this measure by \(r\) to preserve the interpretation of \(r\) as the ratio of workers to firms. A firm can post a single vacancy at cost \(C_1 = 0\), or two vacancies at cost \(C_2 > 0.\textsuperscript{15}

\textsuperscript{13}Such an exercise is in progress.
\textsuperscript{14}For evidence that high wage firms are more profitable, see Abowd, Kramarz, and Margolis [1], among others. For evidence that wages are positively correlated with the queue lengths of applicants, and thus positively correlated with vacancy yield rates, see Holzer, Katz, and Krueger [14].
\textsuperscript{15}The assumption that \(C_1 = 0\) is not entirely innocuous here. We point out below the implications of this assumption.
3.1.1 Profit Maximization

Consider the problem of the firm, taking as given the market utility $U$ that workers can attain by optimally applying to other firms. As we showed in the previous section, should firm $i$ choose to open a single vacancy, it will post a wage $w_i^1$ that solves the profit maximization problem

$$\begin{align*}
\max & \quad [1 - \exp(-q_i^1)](y^i - w_i^1) \\
\text{s.t.} & \quad \left\{ [1 - \exp(-q_i^1)]/q_i^1 \right\} w_i^1 = U.
\end{align*}$$

Again, this problem has a unique solution in which the queue length at firm $i$ under the optimal wage-posting strategy is given by

$$U = y^i \exp(-q_i^1). \quad (22)$$

Let us denote by $\hat{q}_i(y^i; U)$ the value of $q_i^1$ that satisfies equation (22), and let $q_i(y^i; U) = \max\{0, \hat{q}_i(y^i; U)\}$. An equivalent way to write this is that $q_i(y^i; U) = \hat{q}_i(y^i; U)$ if $y^i \geq U$ and zero otherwise. For values of $y_i < U$, the firm cannot offer a worker expected payoff $U$ without offering either $w_i > y_i$ (which is not profitable) or $q_i^1 < 0$ (which is not feasible). For this reason, we say that those firms with $y^i < U$ are inactive.\footnote{We assume, without loss of generality, that these firms do not post a wage. Note, however, that this assumption is not without loss of generality in a game with a finite number of agents. In this case, the wages posted by inactive firms will affect the profitability of deviations by individual firms; see Galenianos and Kircher [11] for an example of this.}

It is easy to show that $q_i(y^i; U)$ is a well defined function for any value of $U \in (0, \bar{y})$. From (9), the firm’s optimal wage conditional on posting a single vacancy is given by

$$w_1(y^i; U) = y^i \frac{q_i(y^i; U) \exp[-q_i(y^i; U)]}{1 - \exp[-q_i(y^i; U)]} \quad (23)$$

and profits are thus

$$\pi_1(y^i; U) = y^i \left\{ 1 - [1 + q_1(y^i; U)] \exp[-q_1(y^i; U)] \right\}. \quad (24)$$

Now consider the problem of firm $i$ should it choose to open up two vacancies. Again, as we showed in the previous section, the firm will post a wage $w_2^i$ that solves the profit maximization
This problem also has a unique solution in which the queue length at firm $i$ under the optimal wage-posting strategy is given by

$$U = y^i [(1 + q_2^i) \exp (-q_2^i)].$$

Let us denote by $\hat{q}_2(y^i; U)$ the value of $q_2^i$ that satisfies equation (25), and let $q_2(y^i; U) = \max\{0, \hat{q}_2(y^i; U)\}$. This, too, is a well defined function that could equivalently be written as $q_1(y^i; U) = \hat{q}_1(y^i; U)$ if $y^i \geq U$ and zero otherwise. From (10), the optimal wage and resulting profit for firm $i$ are given by

$$w_2(y^i; U) = \frac{y^i q_2(y^i; U) [1 + q_2(y^i; U)] \exp [-q_2(y^i; U)]}{2[1 - \exp [-q_2(y^i; U)]] - q_2(y^i; U) \exp [-q_2(y^i; U)]},$$

$$\pi_2(y^i; U) = y^i [2 - \exp [-q_2(y^i; U)] \{2[1 + q_2(y^i; U)] + q_2(y^i; U)^2\} - C_2].$$

### 3.2 Equilibrium Characterization

In order to characterize the equilibrium, we first establish that firms’ optimal behavior can be summarized by a cut-off rule: for any level of market utility $U$, there will exist at most one level of productivity, which we denote $\tilde{y}$, such that those firms with $y^i \in [\tilde{y}, \tilde{y}]$ post two vacancies, those with $y^i \in [U, \tilde{y})$ post a single vacancy, and those with $y^i \in (0, U)$ are inactive. The proof is relegated to the appendix.

**Lemma 3.** For any $U \in (0, \tilde{y})$, the gains from posting a second vacancy, $\pi_2(y^i; U) - \pi_1(y^i; U)$, are weakly increasing in $y^i$, and strictly increasing for any $y^i > U$. Moreover, there exists a $\hat{y} \in (0, \tilde{y}]$ such that $\pi_2(y^i; U) - \pi_1(y^i; U) < 0$ for all $y < \hat{y}$.

Lemma 3 establishes that there exists at most one value $\tilde{y} \in (0, \tilde{y}]$ such that $\pi_2(\tilde{y}; U) = \pi_1(\tilde{y}; U)$, and that if such a $\tilde{y}$ does not exist then $\pi_2(y^i; U) < \pi_1(y^i; U)$ for all $y^i$. Thus, there are two potential types of equilibrium: either all active firms post a single vacancy, or there is a
cutoff \( \tilde{y} \in (0, \bar{y}) \) such that firms post two vacancies if and only if \( y \geq \tilde{y} \). In the lemma below, we show that the gains from posting a second vacancy are decreasing in the level of market utility. This result will later help us to partition the set of equilibria according to the values of the underlying parameters.

**Lemma 4.** For any \( y^i \in [0, \bar{y}] \) and \( U', U'' \in (0, \bar{y}) \), \( U'' > U' \) \( \Rightarrow \) \( \pi_2(\bar{y}; U') - \pi_1(\bar{y}; U') \geq \pi_2(\bar{y}; U'') - \pi_1(\bar{y}; U'') \), with strict inequality for \( y^i > U' \).

Let us denote by \( \bar{U} \) the value of \( U \) such that \( \pi_1(\bar{y}; U) = \pi_2(\bar{y}; U) \). In words, \( \bar{U} \) is the largest value of \( U \) such that some firms would at least weakly prefer to post two vacancies; we know from Lemma 4 that for any \( U < \bar{U} \) there exists a strictly positive measure of firms that prefer to post two vacancies, and for any \( U > \bar{U} \) all firms strictly prefer to post only a single vacancy. In addition, let \( U_1 \) denote the value of \( U \) that satisfies

\[
\int_0^{\bar{y}} q_1(y^i; U_1) dF(y^i) = r.
\]

In words, \( U_1 \) is the market utility of workers when all firms choose to post a single vacancy. Note that the critical values \( \bar{U} \) and \( U_1 \) depend only on exogenous parameters such as \( r, C_2 \) and the distribution \( F(y^i) \). We now define the equilibrium concept, and characterize the regions of the parameter space in which each of these two types of equilibria exist.

**Definition 2.** An equilibrium is a cut-off \( \tilde{y}^* \in (0, \bar{y}) \) and a market utility \( U^* \) such that

\[
\pi_1(\tilde{y}^*; U^*) \geq \pi_2(\tilde{y}^*; U^*), \quad \text{with equality if } \tilde{y}^* < \bar{y}, \tag{28}
\]

and

\[
\int_0^{\tilde{y}^*} q_1(y^i; U^*) dF(y^i) + \int_{\tilde{y}^*}^{\bar{y}} q_2(y^i; U^*) dF(y^i) = r. \tag{29}
\]

The first condition in the definition above simply requires that firms behave optimally. Note that if \( \tilde{y}^* = \bar{y} \), all firms post a single vacancy. The second condition requires aggregate consistency: the market utility \( U^* \) must be such that the induced distribution of queue lengths across

\[\footnote{Had we allowed \( 0 < C_1 < C_2 \), there could exist a third type of equilibrium, in which all active firms post two vacancies. However, doing so adds considerable complexity to the equilibrium characterization without providing considerable insight.} \]
firms is consistent with the number of workers in the economy. We now characterize equilibrium in this environment. Again, the proof can be found in the appendix.

**Proposition 2.** If \( U_1 \geq \bar{U} \) then there exists a unique equilibrium in which all firms post a single vacancy, so that \( \bar{y}^* = \bar{y} \) and \( U^* = U_1 \). If \( U_1 < \bar{U} \) then there exists a unique equilibrium in which some firms post a single vacancy and others post two vacancies, so that \( \bar{y}^* < \bar{y} \) and \( U^* < \bar{U} \).

To provide some intuition for Proposition 2, note that \( \bar{U} \) is decreasing in \( C_2 \) and increasing in \( \bar{y} \), while \( U_1 \) is decreasing in \( r \). Therefore, an equilibrium with only single-vacancy firms is likely when \( C_2 \) is large, \( \bar{y} \) is small, and/or \( r \) is small. This is natural: if it costly to post a second vacancy, if there are few very productive firms, or if there are few workers, then the benefits of posting a second position will not justify the costs. Alternatively, if \( C_2 \) is small, \( \bar{y} \) is large, and/or \( r \) is large, we have an equilibrium in which some firms post a single vacancy and others post two vacancies.

### 3.3 Productivity, Wages, Vacancy Yield, and Firm Size

We now explore the model’s implications for the relationships between productivity, wages, vacancy yield rates, and firm size. The key findings can be easily illustrated via a simple numerical example. Suppose that \( y^i \) is distributed uniformly across the interval \([0,1]\), so that \( F(y^i) = y^i \) and \( \bar{y} = 1 \), and also that \( r = 1 \) and \( C_2 = .1 \). Figures 8 and 9 below depict the equilibrium values of wages and the vacancy yield (or the job-filling rate, as defined earlier). In this equilibrium, \( \bar{y}^* = .5 \) and \( U^* = .316 \).

First, note that the wage function is non-monotonic; more productive firms do not necessarily offer higher wages when they can also offer a greater hiring probability through posting multiple vacancies. However, consistent with the data, the average wage offered by firms with two vacancies (the dashed line) is greater than the average wage offered by firms with one vacancy (the dotted line). Second, notice that the vacancy yield jumps when firms switch from posting a single vacancy to two vacancies. In comparison to a model that assumed an exogenous number of vacancies per firm, the model here predicts a greater disparity between the vacancy yield for
firms with low levels of productivity and firms with high levels of productivity.\footnote{This observation could prove helpful in understanding the highly non-linear relationship between the vacancy yield and employer growth documented in Davis, Faberman, and Haltiwanger \cite{9}.}

4 Conclusion

To be completed.
Appendix

Proof of Lemma 1

Proof. (i) Suppose, towards a contradiction, that \( \lim_{\phi \to 0} \Sigma_W(\phi) = \delta \) for some \( \delta > 0 \). Then \( \lim_{\phi \to 0} q_1 = \lim_{\phi \to 0} r \Sigma_W(\phi)/\phi = \infty \) while \( \lim_{\phi \to 0} q_2 = \lim_{\phi \to 0} r [1 - \Sigma_W(\phi)]/(1 - \phi) = r(1 - \delta) < \infty \), which clearly contradicts the equality in (12). Similarly, suppose \( \lim_{\phi \to 1} \Sigma_h(\phi) = 1 - \delta \) for \( h \in \{W,F\} \) and \( \delta > 0 \). Then \( \lim_{\phi \to 1} q_1 = \lim_{\phi \to 1} r \Sigma_h(\phi)/\phi = [r(1 - \delta)] < \infty \) while \( \lim_{\phi \to 1} q_2 = \lim_{\phi \to 1} r [1 - \Sigma_h(\phi)]/(1 - \phi) = \infty \). Since \( \lim_{q_2 \to \infty} \exp(-q_2)(1 + q_2) = \lim_{q_2 \to \infty} \exp(-q_2)[2(1 + q_2) + q_2^2] = 0 \) and \( 1 - c_2 > 0 \), clearly the equalities of (12) and (13) cannot hold.

(ii) Implicit differentiation of \( \Sigma_W \) and \( \Sigma_F \) yields

\[
\frac{\partial \Sigma_W(\phi)}{\partial \phi} = \frac{q_2 \exp(-q_2) \left[ \frac{1 - \sigma}{(1 - \phi)^2} \right] + \exp(-q_1) \left( \frac{q}{\phi^2} \right)}{q_2 \exp(-q_2) \left( \frac{1}{1 - \phi} \right) + \exp(-q_1) \left( \frac{1}{\phi} \right)} \geq 0 \tag{30}
\]

\[
\frac{\partial \Sigma_F(\phi)}{\partial \phi} = \frac{q_2^2 \exp(-q_2) \left[ \frac{1 - \sigma}{(1 - \phi)^2} \right] + q_1 \exp(-q_1) \left( \frac{q}{\phi^2} \right)}{q_2^2 \exp(-q_2) \left( \frac{1}{1 - \phi} \right) + q_1 \exp(-q_1) \left( \frac{1}{\phi} \right)} \geq 0 \tag{31}
\]

(iii) To prove that \( \Sigma_W \) and \( \Sigma_F \) cross at most once on the domain \( \phi \in (0,1) \), I will show that at any intersection between the two curves, \( \frac{\partial \Sigma_F(\phi)}{\partial \phi} > \frac{\partial \Sigma_W(\phi)}{\partial \phi} \). In particular, consider an intersection at some point \((\phi', \sigma')\) with \( \sigma' = \Sigma_W(\phi') = \Sigma_F(\phi') \) and corresponding expected queue lengths \( q_1' \) and \( q_2' \). One can show that

\[
\frac{\partial \Sigma_F(\phi)}{\partial \phi} \geq \frac{\partial \Sigma_W(\phi)}{\partial \phi} \quad \text{if} \quad q_2'(\phi' - \sigma') > q_1'(\phi' - \sigma'). \tag{32}
\]

This condition is true by the definition of \( q_1' \) and \( q_2' \). ■

Proof of Lemma 2

Proof. We sketch the proof establishing that there exists \( \eta_1 \in (0,1) \) such that \( \Sigma_F(\phi) > \Sigma_W(\phi) \) \( \forall \phi \in (1 - \eta_1, 1) \) \( \iff r > \bar{r}(c_2) \). The symmetric argument can be used to show that there exists \( \eta_2 \in (0,1) \) such that \( \Sigma_W(\phi) > \Sigma_F(\phi) \) \( \forall \phi \in (0, \eta_2) \) \( \iff r < \bar{r}(c_2) \).
Consider a value of $\phi$ arbitrarily close to 1. We know that $\Sigma_W$ is continuous, and that $\lim_{\phi \to 1} \Sigma_W = 1$. Therefore, using (6) and (12), the values of $q_1$ and $q_2$ implied by $\Sigma_W$ become arbitrarily close to $r$ and $\tilde{q}_2(r)$, respectively, as $\phi \to 1$. We also know that $\Sigma_F$ is continuous, and that $\lim_{\phi \to 1} \Sigma_F = 1$. Therefore, using (6) and (13), the values of $q_1$ and $q_2$ implied by $\Sigma_W$ become arbitrarily close to $r$ and $\hat{q}_2(r)$ as $\phi \to 1$, where $\hat{q}_2(r)$ satisfies

$$2 - \exp(-\hat{q}_2(r))[2(1 + \hat{q}_2(r)) + \hat{q}_2(r)^2] - c_2 = 1 - \exp(-r)(1 + r). \quad (33)$$

The left (right) hand side of the equality in (33) is the expected profits of a type two (one) firm. The final step of the proof comes from the fact that, for this value of $\phi$ arbitrarily close to 1, $\Sigma_F(\phi) > \Sigma_W(\phi)$ if and only if $\tilde{q}_2(r) > \hat{q}_2(r)$ by (7), which is true if and only if

$$2 - \exp(-\tilde{q}_2(r))[2(1 + \tilde{q}_2(r)) + \tilde{q}_2(r)^2] - c_2 > 1 - \exp(-r)(1 + r)$$

$$\Leftrightarrow 2 - \exp(-\tilde{q}_2(r))[2(1 + \tilde{q}_2(r)) + \tilde{q}_2(r)^2] - c_2 > 1 - \exp(\tilde{q}_2(r))(1 + \tilde{q}_2(r))(1 + r)$$

$$\Leftrightarrow \exp(-\tilde{q}_2(r))[1 + \tilde{q}_2(r)(1 - r) + \tilde{q}_2(r)^2] < 1 - c_2,$$

which is equivalent to $r > r(c_2)$. \hfill \blacksquare

**Proof of Proposition 1**

*Proof.* As discussed in the text, the conditions on $\{r, C_2, y\}$ that determine the three types of equilibrium follow directly from Lemmas 1 and 2. In order to characterize equilibrium strategies in the interior case and perform comparative statics, we require the intermediate result established in Lemma 5 below.

**Lemma 5.** In any interior equilibrium, $\phi^* > \sigma^*$.

*Proof.* We prove this by establishing that $\Sigma_W(\phi) < \phi \forall \phi \in (0,1)$. Suppose towards a contradiction that $\Sigma_W(\phi) \geq \phi$ for some $\phi \in (0,1)$. Then

$$q_1 = \frac{r \Sigma_W(\phi)}{\phi} \geq \frac{r[1 - \Sigma_W(\phi)]}{1 - \phi} > 0.$$ 

But then

$$\exp(-q_1) \leq \exp(-q_2) < (1 + q_2) \exp(-q_2),$$

26
a contradiction. ■

Now, to derive the equilibrium strategies \((\sigma^*, \phi^*)\), let us differentiate \(\Sigma_F\) and \(\Sigma_W\) and employ Cramer’s rule to find that

\[
\begin{align*}
\frac{\partial \phi^*}{\partial r} &= \frac{\phi^*(1 - \phi^*)}{r(\sigma^* - \phi^*)} < 0 \\
\frac{\partial \sigma^*}{\partial r} &= \frac{\sigma^*(1 - \sigma^*)}{r(\sigma^* - \phi^*)} < 0 \\
\frac{\partial \phi^*}{\partial c_2} &= \left\{\frac{[q_2^* \exp(-q_2^*)]}{[\exp(-q_1^*)/\phi^*]} + \frac{[\exp(-q_1^*)/\phi^*]}{[\phi^2(1 - \phi^*)^2]}\right\} > 0 \\
\frac{\partial \sigma^*}{\partial c_2} &= \left\{\frac{[q_2^* \exp(-q_2^*)]}{[\exp(-q_1^*)]} + \frac{[\exp(-q_1^*)/\phi^*]}{[\phi^2(1 - \phi^*)^2]}\right\} > 0.
\end{align*}
\]

This implies that

\[
\begin{align*}
\frac{\partial q_1^*}{\partial r} &= \frac{\sigma^* + \frac{r}{\phi^*} \left(\frac{\partial \sigma^*}{\partial r}\right) - \frac{r\sigma^*}{\phi^2} \left(\frac{\partial \phi^*}{\partial r}\right)}{1 - \sigma^* - \frac{r}{\phi^*} \left(\frac{\partial \sigma^*}{\partial r}\right) + \frac{r(1 - \sigma^*)}{(1 - \phi^*)^2} \left(\frac{\partial \phi^*}{\partial r}\right)} = 0 \\
\frac{\partial q_2^*}{\partial r} &= \frac{\sigma^* + \frac{r}{\phi^*} \left(\frac{\partial \sigma^*}{\partial r}\right) - \frac{r\sigma^*}{\phi^2} \left(\frac{\partial \phi^*}{\partial r}\right)}{1 - \sigma^* - \frac{r}{\phi^*} \left(\frac{\partial \sigma^*}{\partial r}\right) + \frac{r(1 - \sigma^*)}{(1 - \phi^*)^2} \left(\frac{\partial \phi^*}{\partial r}\right)} = 0.
\end{align*}
\]

Therefore, for any \(\underline{r}(c_2) < r < \bar{r}(c_2)\), we have that (i) \(\partial q_k^*/\partial r = 0\) for \(k \in \{1, 2\}\), (ii) \(\lim_{\phi \to 0} q_1^* = \underline{r}(c_2)\), and (iii) \(\lim_{\phi \to 1} q_2^* = \bar{r}(c_2)\). This implies that \(q_1^* = \underline{r}(c_2)\) and \(q_2^* = \bar{r}(c_2)\) for all \(\phi^* \in (0, 1)\). The expressions in (17) and (18) follow directly from the observation that in all interior equilibria

\[
\begin{align*}
\underline{r}(c_2) &= \frac{r\sigma^*}{\phi^*} \\
\bar{r}(c_2) &= \frac{r(1 - \sigma^*)}{1 - \phi^*}.
\end{align*}
\]

**Proof of Lemma 3.** To economize on notation, we let \(q_j^* \equiv q_j(y_i; U)\) for \(j \in \{1, 2\}\). First note that

\[
\begin{align*}
\frac{\partial}{\partial y_i} \left[\pi_2(y_i; U) - \pi_1(y_i; U)\right] &= \left[2 - \exp(-q_2^*)(2 + q_2^*)\right] - \left[1 - \exp(-q_1^*)\right] \\
&= 1 - \exp(-q_2^*) - \left[\exp(-q_2^*)(1 + q_2^*) + \exp(-q_1^*)\right] \\
&= 1 - \exp(q_2^*) \geq 0 \forall q_2^* \geq 0, \text{ with strict inequality if } q_2^* > 0(36)
\end{align*}
\]

where we’ve used the fact that \(\exp(-q_2^*)(1 + q_2^*) = \exp(-q_1^*) = U/y_i\). This establishes the first fact, and the second fact follows immediately since \(\pi_2(0; U) = -C_2 < 0 = \pi_1(0, U)\). ■
Proof of Lemma 4. This follows immediately from the fact that

\[
\frac{\partial}{\partial y^i} [\pi_2(y^i; U) - \pi_1(y^i; U)] = q_i^1 - q_i^2 \leq 0, \text{ with strict inequality if } q_i^j > 0.
\]  

(37)

Proof of Proposition 2. If \( U_1 \geq \bar{U} \), then \( \pi_2(\bar{y}; U) - \pi_1(\bar{y}; U) \leq 0 \) from Lemma 4. Therefore \( \pi_2(y^i; U) \leq \pi_1(y^i; U) \forall y^i \in [0, \bar{y}] \) by Lemma 3, so that \( \bar{y}^* = \bar{y} \) satisfies (28), and by construction \( U^* = U_1 \) satisfies (29). To see that this is the only equilibrium, suppose towards a contradiction that there exists another equilibrium with \( \bar{y}^* < \bar{y} \). We know that \( \bar{y}^* = \bar{y} \) yields market utility \( U^* = U_1 \). We can implicitly differentiate (29) using Leibniz’s rule to get that

\[
\frac{\partial U^*}{\partial \bar{y}^*} = \frac{-[q_2(\bar{y}^*; U^*) - q_1(\bar{y}^*; U^*)]}{\int_0^{\bar{y}} \frac{1}{y^i \exp[q_1(y^i; U^*)]} dF(y^i) + \int_{\bar{y}}^{\bar{y}} \frac{1}{y^i \exp[q_2(y^i; U^*)]q_2(y^i; U^*)} dF(y^i)} < 0.
\]  

(38)

Therefore, if \( \exists \bar{y}^* < \bar{y} \), the associated market utility \( U^* \geq U_1 \geq \bar{U} \). But \( U^* \geq \bar{U} \Rightarrow \pi_2(y^i; U^*) < \pi_1(y^i; U^*) \forall y^i \in (0, \bar{y}) \Rightarrow \# \bar{y}^* \in (0, \bar{y}) \) such that \( \pi_2(\bar{y}^*; U^*) = \pi_1(\bar{y}^*; U^*) \), a contradiction.

Now suppose that \( U_1 < \bar{U} \). We will show that there exists a pair \( (\bar{y}^*, U^*) \in (0, \bar{y}) \times (0, \bar{U}) \) that satisfies the equilibrium conditions (28) and (29), and that this pair is unique. To do so, it will be convenient to define \( \Upsilon_F : U \mapsto \bar{y} \) as the implicit function in equation (28), which represents the marginal firm’s indifference condition. Likewise, we define \( \Upsilon_A : U \mapsto 0 \) as the implicit function in equation (29), which represents the aggregate consistency requirement. We will show that \( U_1 < \bar{U} \Rightarrow \exists! U^* \text{ such that } \Upsilon_F(U^*) = \Upsilon_A(U^*) = \bar{y}^* < \bar{y} \).

Note that (i) \( \lim_{U \to 0} \Upsilon_F(U) = C_2 \), (ii) \( \lim_{U \to 0} \Upsilon_A(U) = \bar{y} \), and (iii) \( \Upsilon_F \) is increasing in \( U \):

\[
\frac{\partial \Upsilon_F(U)}{\partial U} = \frac{q_2(\bar{y}; U) - q_1(\bar{y}; U)}{1 + \exp[-q_1(\bar{y}; U)] - \exp[-q_2(\bar{y}; U)][2 + q_2(\bar{y}; U)]} > 0.
\]  

(39)

Also note that (i) \( \lim_{U \to U_1} \Upsilon_A(U) = \bar{y} \) by definition, and (ii) \( \Upsilon_A \) is decreasing in \( U \), as established in (38) above. Since \( F \) is assumed to be continuous, we know that both \( \Upsilon_F \) and \( \Upsilon_A \) are continuous as well. Therefore, since \( U_1 < \bar{U} \), it is clear that \( \Upsilon_F \) and \( \Upsilon_A \) intersect at a unique value \( U^* \in (U_1, \bar{U}) \) and that \( \Upsilon_F(U^*) = \Upsilon_A(U^*) < \bar{y} \).
References


