Wage and Technology Dispersion

DARON ACEMOGLU
Massachusetts Institute of Technology

and

ROBERT SHIMER
Princeton University

First version received September 1997; final version accepted December 1999 (Eds.)

This paper explains why firms with identical opportunities may use different technologies and offer different wages. Our key assumption is that workers must engage in costly search in order to gather information about jobs (Stigler (1961)). In equilibrium, some firms offer high fixed cost, high productivity technologies, offer high wages, and fill job openings quickly. Other firms adopt less capital-intensive technologies and offer low wages, hiring mostly uninformed workers. In equilibrium, the amount of wage dispersion leaves workers indifferent about whether to gather information, and the fraction of informed workers leaves firms indifferent about their wage and technology choice. We show that worker search, which would appear to be a rent-seeking activity in partial equilibrium, may be efficiency-enhancing in general equilibrium.

1. INTRODUCTION

This paper embeds Stigler’s (1961) seminal search model in a general equilibrium framework. Firms adopt technologies and post wages to attract job applicants. Workers search in order to gather information about job openings, and then apply for the most attractive of these.

Our paper has three main contributions. First, we show that when search is costly, there will be wage dispersion among identical workers, even if all firms use the same technology, thus extending the results of Burdett and Judd (1983) to a general equilibrium labour market environment.

Second, we develop a framework for analysing firms’ technology choice and irreversible investments in a search environment. We show that the forces creating wage dispersion also lead to technology dispersion. Firms that offer higher wages fill job openings more rapidly, and so are willing to make larger irreversible investments in complementary inputs, such as capital.

Finally, we show that although in Stigler’s partial equilibrium model search is a costly rent-seeking behaviour, in general equilibrium it is likely to be efficiency enhancing. When a worker searches more intensively, she reduces firms’ monopsony power. This raises all wages towards the social cost of labour. Since other workers also benefit from the increase in wages and effectively free-ride on her search effort, there is too little search in equilibrium.

Because search is desirable, violations of the “law of one price” may improve the allocation of resources. If all firms offered the same wage, no worker would search across multiple firms, and so firms would offer the monopsony wage, as in the Diamond (1971) paradox. Cheap labour would make the labour market excessively tight and distort investments in other inputs. Search moderates these inefficiencies.
When search costs are small, wage and technology dispersion persists, and more productive firms continue to pay higher wages. But as search costs approach zero, the equilibrium converges to an efficient allocation. This contrasts with Grout's (1984) well-known result that when workers appropriate part of an investment through higher wages, firms underinvest. This argument does not generalize to our environment, because the upward sloping wage-investment schedule enables firms with high labour productivity to attract workers at a faster rate by offering higher wages. Without wage dispersion, workers would not search, and a high productivity firm would be unable to fill its vacancy more rapidly, discouraging investment.

We start in Sections 2 and 3 with a static model that illustrates the main contributions of our paper. In Section 4, we generalize our results to a dynamic environment, where search decisions of workers, and technology and wage decisions of firms, are forward looking. We discuss the related literature and draw empirical implications from our model and from existing related models in Section 5. Section 6 concludes.

2. THE STATIC ENVIRONMENT

This section describes a static economy that highlights the main elements of our model. We embed Stigler's (1961) partial equilibrium search model into a general equilibrium framework. Firms open jobs, make capital investments and post wages. Then workers search; they sample an optimal number of job openings, and observe the wage of those jobs. They then apply for a job at one of those firms.

There are decreasing returns to labour at the firm level. As a result, firms do not always hire all job applicants at a given posted wage, and some workers remain unemployed. At the same time, some firms will not manage to attract any applicants, and will be left with idle capital. The coexistence of unemployment and vacancies is due to this fundamental coordination problem in the absence of the Walrasian invisible hand (see also Peters (1991) and Montgomery (1991)).

2.1. Environment

There is a continuum 1 of risk-neutral workers and a much larger continuum $M \gg 1$ of risk-neutral firms. The sequence of events is as follows.

1. Each firm decides whether to create a job. If it creates a job, it chooses a level of capital $k \geq 0$ at cost $rk$ and posts a wage $w$. We also refer to $k$ as a firm's technology.
2. Each worker decides how many jobs to sample, her 'search intensity'. A worker who samples $n$ jobs learns the wage offered by $n$ randomly chosen active firms and applies to at most one of them. She pays a marginal search cost $c_n$ for sampling the $n$-th job.
3. Each firm that receives at least one application chooses randomly one of the applicants, pays her the posted wage, and produces. The remaining applicants are unemployed, while active firms that receive no applications are vacant. A worker's payoff is equal to her wage net of search costs if she is employed, and she loses her search cost otherwise. An active firm earns $f(k) - w - rk$ if it hires a worker, and loses its investment $rk$ if it is vacant. An inactive firm earns nothing.

In the dynamic model (Section 4), unemployed workers and vacancies may search again the following period. The static model abstracts from that continuation game.
Although workers rationally anticipate the equilibrium wage distribution, they do not know the wage offered by a particular firm. Because all firms appear to be identical, a worker randomly samples $n \equiv 1$ of them. We assume that all workers costlessly observe one wage, $c_1 = 0$. The marginal cost of sampling the $n$-th job, $c_n$, is positive and increasing in $n$.

Each firm $j$ makes a capital investment $k(j)$ and posts a wage rate $w(j)$. Let $x(j) = (k(j), w(j))$ denote its strategy. We adopt the convention that an inactive firm chooses $x(j) = (0, 0)$. The strategies of all firms are summarized by a mapping $x: [0, M] \to \mathbb{R}_+^2$.

Each worker $i$ chooses a sample size $n(i)$ and a preference function $p_n(i): \mathbb{R}^n \to \{0, 1, \ldots, n\}$ over arbitrary $n$-tuples of wages. Thus $p_n(w_1, \ldots, w_n) = s \in \{1, \ldots, n\}$ tells us that a worker who observes wages $(w_1, \ldots, w_n)$ will apply for the job offering $w_s$. $p_n(w_1, \ldots, w_n) = 0$ denotes that the worker chooses to remain unemployed. Workers' strategies are summarized by the mapping $y: [0, 1] \to \mathcal{S}$, where $\mathcal{S}$ represents the space of all preference functions with $n \equiv 1$. From this we can extract each worker's sample size and preference function.

### 2.2. Equilibrium

An equilibrium is a $(x^*, y^*)$ such that each firm $j$ chooses its capital–wage pair $x(j)$ to maximize its expected profit given the strategies of other players, $(x^*, y^*)$, and each worker $i$ chooses her sample size and preference function to maximize her expected wage net of search costs given $(x^*, y^*)$.

It is convenient in what follows to define an equilibrium in terms of aggregate variables rather than individual strategies. Let $\pi(k, w)$ be the expected profit of a firm choosing capital–wage pair $(k, w)$. Let $V$ be the measure of active firms, and $H$ denote the joint distribution of those firms' capital investments and wages, with support $X$. Also let $H^w$ and $H^k$ denote the marginal distribution of wages and capital, respectively, with supports $X^w$ and $X^k$. On the worker's side, let $R_n$ denote the expected return of a worker who chooses a sample size of $n$, and $z_n$ be the fraction of workers who sample $n \in \{1, 2, \ldots\}$ jobs, with $\sum_{n=1}^{\infty} z_n = 1$. Let $\rho(w)$ denote the expected return of a worker who applies for a job offering wage $w$.

**Definition 1.** An equilibrium consists of a measure of active firms $V$, a joint capital–wage distribution $H$ with support $X$, an expected profit function for firms $\pi$, expected return functions for workers $\rho$ and $R$, and sampling decisions $z_n$ and preference functions $\{p_n\}$ for workers such that:

1. (Profit Maximization) $\forall (k, w) \in X, \forall (k', w'), \pi(k', w') \equiv \pi(k, w) = 0$.
2. (Optimal Application) $\forall n \in \mathbb{N}, \forall (w_1, \ldots, w_n)$,

\[
p_n(w_1, \ldots, w_n)\neq s \text{ if } \rho(w_s) < \max_{s'} \rho(w_{s'}) \text{ or } \rho(w_s) < 0
\]

and

\[
p_n(w_1, \ldots, w_n)\neq 0 \text{ if } \max_{s'} \rho(w_{s'}) \equiv 0.
\]

3. (Optimal Sampling) $z_n = 0$ if $R_n < \max_{n'} R_{n'}$.

The first part of the definition ensures that all firms choose profit-maximizing strategies. Since firms always have the option of not entering the market, they must earn nonnegative
profits. We also assume that there are enough firms, \( M \gg 1 \), for the free-entry condition to be binding in equilibrium, so that active firms must earn exactly zero profits. The second and third parts of the definition ensure that workers use optimal strategies.

To use this definition, we express the three payoff functions \( \pi, \rho \), and \( R \) in terms of aggregate variables. Let \( G_n(w) \) denote the probability that a worker who samples \( n \) random wages, including one offering \( w \), applies for the one offering \( w \). This function is defined by integrating workers' preference function over the wage distribution \( H^w \)

\[
G_n(w) = \int_{x^w} \ldots \int_{x^w} I(p_n(w, w_2, \ldots, w_n))dH^w(w_2) \ldots dH^w(w_n),
\]

where \( I \) is an indicator function, with \( I(1) = 1 \) and \( I(p) = 0 \) for all other \( p \); so it takes value 1 if a worker would choose to apply for the first wage. Let \( Q = 1/V \) denote the ratio of workers to vacancies. Each firm therefore expects to be sampled by \( Qnz_n \) workers who sample \( n \) firms. Then the expected number of applications received by a firm offering a wage \( w \) is

\[
q(w) = Q \sum_{n=1}^\infty nz_n G_n(w).
\]

Since each application decision is independent, the firm receives no applications with probability \( \exp(-q(w)) \).

The expected profit of a firm using strategy \( x = (k, w) \) is the probability that it receives an application, times its gross profit if it receives an application, \( f(k) - w \), minus the sunk cost of its capital investment, \( rk \)

\[
\pi(k, w) = (1 - \exp(-q(w)))(f(k) - w) - rk.
\]

Similarly, the probability that a worker applying for wage \( w \) is hired is equal to the probability that a firm offering wage \( w \) hires a worker, divided by the expected number of workers applying for that wage, \( q(w) \). Using this, the expected return to the application is

\[
\rho(w) = \frac{1 - \exp(-q(w))}{q(w)} w.
\]

Finally, the expected utility of a worker sampling \( n \geq 1 \) wages randomly drawn from \( H \) is

\[
R_n = n \int_{x^w} \rho(w)G_n(w)dH^w(w) - \sum_{i=1}^n c_i.
\]

For each of \( n \) wage draws, her expected payoff is equal to the probability that she applies for the wage, \( G_n(w) \), multiplied by her expected return, \( \rho(w) \), integrated over the density \( H^w \). Summing over the \( n \) independent wage draws and subtracting search costs, we obtain the worker's expected utility.

3. STATIC ANALYSIS

To highlight the main economic forces leading to wage dispersion, in Section 3.1 we specialize our model to abstract from firms' technology choices. Then in Section 3.2, we

\[1\] Suppose that there were \( m \) workers. A particular worker applies to a firm offering wage \( w \) with probability \( q(w)/m \), so the probability that the firm receives no applications is \((1 - q(w)/m)^m = \exp(-q(w))\). We have a continuum of firms here, so the last expression is exact.
turn to the main focus of this paper, a model with capital choices and technology dispersion. Wage and technology dispersion have the same source, although the interpretation of the model is quite different when firms choose to use different technologies. Section 3.3 considers the robustness of our conclusions, while Section 3.4 analyses the social efficiency of the decentralized allocations.

3.1. Wage dispersion without capital choice

Assume that \( f(k) = 1 \) for \( k \geq 1 \) and output is zero otherwise, so all active firms choose \( k = 1 \). Also assume that fixed costs are less than gross output, \( r < 1 \).

Workers prefer higher wages. The first step in our analysis is to show that in equilibrium, workers strictly prefer higher wages. Intuitively, workers must weakly prefer higher wages.\(^2\) Ruling out indifference is more difficult. It seems plausible that workers would be indifferent between two wages, and apply in sufficient numbers to the higher wage, so that the difficulty of obtaining it offsets the wage difference. To rule out this possibility, we use the fact that firms must also be indifferent between any wages offered in equilibrium:

Lemma 1. In equilibrium, \( \rho \) is strictly increasing on \( X^w \).

Proof. Any wage \( w \) must yield nonpositive profits, so \( (1 - \exp(-q(w)))(1 - w) \leq r \) for all \( w \), with equality if \( w \) is offered in equilibrium, \( w \in X^w \). Inverting this yields a lower bound on workers' expected return \( \rho(w) \)

\[
\rho(w) = \frac{1 - \exp(-q(w))}{q(w)} w \geq \frac{rw}{(1 - w)(\log(1 - w) - \log(1 - w - r))} \equiv \Phi(w),
\]

again with equality if \( w \in X^w \). Crucial is that \( \Phi \) is strictly quasiconcave with unique maximum \( w^* \in (0, 1 - r) \) solving

\[
(1 - w^* - r)(\log(1 - w^*) - \log(1 - w^* - r)) = rw^*.
\]

Next take any wage \( w \in X^w, w \neq w^* \), offered in equilibrium. Then \( \rho(w^*) \geq \Phi(w^*) > \Phi(w) = \rho(w) \). The first inequality follows from the construction of \( \Phi \); the second inequality follows because \( w^* \) maximizes \( \Phi(w^*) \); the equality uses the assumption that \( w \in X^w \). Then as \( \rho(w^*) > \rho(w) \), the “Optimal Application” part of the definition of equilibrium, combined with equations (1) and (2), ensures that workers are more likely to apply for \( w^* \): \( q(w^*) > q(w) \). And if \( w^* < w \), firms would make strictly more profits by offering \( w^* \), contradicting the assumption that \( w \) is offered in equilibrium. This proves that equilibrium wages are bounded above by \( w^* \).

Finally, since \( \Phi \) is quasiconcave and maximized at \( w^* \), it is strictly increasing on \( X^w \subseteq (-\infty, w^*]. \) As \( \rho(w) = \Phi(w) \) on \( X^w \), \( \rho \) is also strictly increasing on \( X^w \).

Characterizing the wage distribution. The knowledge that workers strictly prefer higher wages puts considerable structure on the wage distribution. We show that if all workers sample one job, firms have absolute monopsony power and always offer workers’

\(^2\) To see this, observe that if all workers strictly preferred one wage \( w_1 \) to another \( w_2 \), they would be more likely to apply for the preferred wage, so \( q(w_1) > q(w_2) \). But this could not be the case if \( w_2 > w_1 \): if a higher wage were easier to get, then workers would strictly prefer the higher wage.
reservation wage. Whereas, if some workers sample multiple jobs, the wage distribution is disperse.

**Lemma 2.** If \( z_1 = 1 \), all firms offer a wage of zero, i.e. \( X^w \equiv \{0\} \). If \( z_n > 0 \) for some \( n \geq 2 \), the support of the wage distribution consists of a convex, nonempty interval \([0, \bar{w}]\) and possibly the point \( w^* \equiv \bar{w} \) defined in (6). The wage distribution is atomless on \([0, \bar{w}]\), but may have an atom at \( w^* \).

**Proof.** First assume \( z_1 = 1 \). Workers who sample one job reject negative wages and accept positive ones, so (1) and (2) imply that \( q(w) = Q \) for all \( w \geq 0 \). Hence, profit-maximizing firms set \( w = 0 \).

Alternatively, assume \( z_1 < 1 \). Workers continue to reject negative wage offers, so we focus on \( w \in [0, w^*] \). First, a positive measure of firms cannot offer some wage \( w < w^* \). If they did, \( \rho(w') \equiv \Phi(w') > \Phi(w) = \rho(w) \) for all \( w' \in (w, w^*) \). This implies that a firm offering any higher wage would attract discretely more applicants. Therefore, there exists a sufficiently small increase in the wage which would be profitable.

Next, the wage distribution must not be degenerate at \( w^* \). For if it were, (5) implies \( R_n = \rho(w^*) - \sum_{i=1}^{n} c_i \), strictly decreasing in \( n \). Then all workers’ would optimally sample a single vacancy, contradicting \( z_1 < 1 \).

Finally, the proof that the support of the wage distribution is \([0, w] \cup w^* \) is standard. If there were a gap in the atomless part of the wage distribution, a firm offering a wage at the top of the gap could cut its wage without losing applicants, and so raise profits. Similarly, if no firm offered a wage below some threshold, the lowest-wage firm could costlessly reduce its wage to zero. \( \| \)

Lemma 2 captures Rothschild’s (1973) criticism of search models, but turns it on its head: if workers have a common reservation wage, all firms will offer this wage, and when all firms are offering the same wage, why should anyone sample multiple jobs? Lemma 2 shows that if some workers sample multiple jobs, the equilibrium wage distribution must not be degenerate. The result is also related to the information externality identified by Grossman and Stiglitz (1980). They argued that if asset prices transmit all available information, no trader can profit by learning additional information. Traders will invest in information only if there is sufficient noise in the system. Here workers will search only if there is sufficient wage dispersion.

Since workers strictly prefer higher wages (Lemma 1), equation (1) implies \( G_n(w) = (H^w(w))^n - 1 \), the probability that the other \( n - 1 \) wage offers are less than \( w \), assuming the wage distribution \( H^w \) is atomless at \( w \). This enables us to rewrite the expected utility of a worker sampling \( n \) jobs, \( R_n \) in equation (5), directly in terms of the equilibrium wage distribution. Her utility is equal to the payoff \( \Phi(w) = \rho(w) \) if the highest wage \( w \in X^w \) is drawn from the atomless part of the wage distribution, integrated over the wage density, plus the payoff if her sample includes \( w^* \), times the probability of this event, minus search costs:

\[
R_n = n \int_0^w \Phi(w)(H^w(w))^{n-1} dH^w(w) + \Phi(w^*)(1 - (H^w(\bar{w}))^n) - \sum_{i=1}^{n} c_i.
\]

Note that \( \int_{\bar{w}}^{w^*} (1 - (H^w(w))^n)\Phi'(w)dw = (1 - (H^w(\bar{w}))^n)(\Phi(w^*) - \Phi(\bar{w})) \), since \( H^w(w) = H^w(\bar{w}) \) for \( w \in (\bar{w}, w^*) \). Then using integration by parts on the previous expression, we
obtain a useful alternative representation of $R_n$

$$R_n = \int_0^w (1 - (H^n(w))^n) \Phi'(w) dw - \sum_{i=1}^n c_i.$$  \hspace{1cm} (7)

Equation (7) states that the expected gross value of sampling $n$ jobs is the probability that at least one sampled wage is greater than $w$, $1 - (H^n(w))^n$, integrated over the marginal return to a higher wage $\Phi'(w)$.

**Workers’ sample size.** We now show that in equilibrium, no worker ever samples more than two jobs. Sampling is costly. For search to be worthwhile, there needs to be a sufficiently disperse distribution of wages. However, firms will only offer a disperse wage distribution when some workers take the first job that comes along. Therefore, irrespective of the costs of search, a number of workers must sample only one job. These uninformed workers facilitate a distribution of wages, making it worthwhile for others to search. In other words, in equilibrium, some workers *free-ride* on the search of others.

**Lemma 3.** In any equilibrium, $z_1 + z_2 = 1$.

**Proof.** Equation (7) implies there are decreasing returns to fixed sample size search: $R_n - R_{n-1} > R_{n+1} - R_n$ for all $n \geq 2$. We omit this result, as it is well-known from the partial equilibrium search literature, *e.g.* Footnote 4 of Stigler (1961).

Next, we show that $z_1 > 0$. We know from Lemma 2 that if $z_1 < 1$, there is an atomless distribution of firms offering wages in the nonempty interval $[0, \bar{w}]$. If $z_1 = 0$, a firm offering a zero wage never hires any worker, since any worker who contacts it will have also contacted at least one other firm, offering a positive wage with probability 1. That is, $q(0) = 0$. This implies $\pi(1, 0) = -r < 0$, so it would not be profitable to offer a zero wage in equilibrium, a contradiction that proves $z_1 > 0$.

Finally, as $z_1 > 0$, the third part of Definition 1, Optimal Sampling, implies $R_1 \equiv R_n$ for all $n$, and in particular $R_1 \equiv R_2$. Then, decreasing returns to search implies $R_1 \equiv R_2 > R_3 > R_4 \ldots$. If $R_1 > R_2$, $z_1 = 1$. Otherwise, $R_1 = R_2$ and $z_1 + z_2 = 1$.  \hspace{1cm} \|}

At first glance, it may seem non-generic for workers to be exactly indifferent between sampling one job and sampling two. However, the wage distribution is endogenously determined by profit-maximizing firms, and so we will show that there will be an equilibrium in which $R_1 = R_2$. Moreover, we argue in Section 3.3 that there are natural forces likely to lead an economy towards the point where workers are indifferent between different search strategies.

**Characterizing the equilibrium.** We can now give a complete characterization of search equilibria. Since the wage distribution is degenerate at zero when $z = 1$ (Lemma 2), the characterization of an equilibrium without search is simple. We therefore focus on $z_1 = 1 - z_2 < 1$.

Substitute for $q(w) = Q((1 - z_2) + 2z_2 G_2(w))$ in (3) to give

$$\pi(1, w) = (1 - \exp (-Q((1 - z_2) + 2z_2 G_2(w))))(1 - w) - r = 0 \text{ for } w \in X^w.$$  

Assuming that $z_2 > 0$, we can invert this to get the probability that a worker who samples $w$ and one other job chooses to apply for $w$:

$$G_2(w) = \frac{\log (1 - w) - \log (1 - w - r) - Q(1 - z_2)}{2Q z_2}.$$  \hspace{1cm} (8)
As noted above, $G_2(w) = H^w(w)$ for all $w \in [0, \bar{w}]$, so equation (8) gives the wage distribution except at the point of the atom, $w^*$. 

Now by Lemma 2, $0 \in X^w$, so $\pi(1, 0) = 0$. Since we also know that $G_2(0) = 0$, equation (8) pins down $Q$ in terms of $z_2$:

$$Q = \frac{-\log(1-r)}{1-z_2}.$$  

(9)

From (9), $Q$ is positive and increasing in $z_2$. When more workers search, the labour market is less tight. Firms are induced to pay higher wages to attract workers with more options, shifting rents to workers, and therefore discouraging entry. Expressing this differently, there are rents in the economy that must be dissipated. This either happens by firms entering in larger numbers until the fixed cost of entry $r$ exhausts the rents, or it takes the form of workers searching for high wage jobs, inducing firms to offer higher wages. Next, substitute (9) into (8) to obtain:

$$G_2(w) = \frac{1-z_2}{2z_2} \left( \log \left( \frac{1-w-r}{1-w} \right) - 1 \right) = H^w(w),$$  

(10)

where the second equality holds if we are on the atomless part of the wage distribution, $w \leq \bar{w}$.

To find $\bar{w}$, suppose a fraction $\mu \in [0, 1]$ of firms offer $w^*$. Then the top of the atomless part of the wage distribution satisfies $G_2(\bar{w}) = H^w(\bar{w}) = 1 - \mu$, or

$$\bar{w} = 1 - \frac{r}{1 - (1-r)^{(1-z_2(2\mu-1))(1-z_2)}},$$  

(11)

Finally, to calculate $\mu$, use the fact that a firm offering $w^*$ will fail to receive an application from a worker who has sampled this job, only when the worker has also sampled another firm with $w^*$. In this case, each firm will get an application with probability $1/2$. That is, $G_2(w^*) = 1 - \mu/2$. Combining this with (10) pins down $\mu$ in terms of $z_2$

$$\mu = \max \left( \frac{1+z_2}{z_2}, \frac{1-z_2}{z_2} \left( \log \left( \frac{1-w^*-r}{1-w^*} \right) - 1 \right) \right).$$  

(12)

The “max” operator takes care of the case where the wage distribution is atomless.

This gives a complete description of the wage distribution in terms of the endogenous search intensity $z_2$. An increase in $z_2$ raises the wage distribution in the sense of first-order stochastic dominance—i.e. $H^w(w)$ decreases for all $w$. This captures the free-rider problem in our model. When a worker samples more jobs, she improves the distribution of wages for all other workers, but only captures a small share of the wage gains herself.

If search intensity is sufficiently high, all firms offer a wage of $w^*$, so $\mu = 1$. In fact, this happens before the point where all workers sample two jobs, at

$$z = 1 - \frac{\log(1-r)}{\log(1-w^*-r) - \log(1-w^*)}.$$  

There can never be an equilibrium with $z_2 \geq z$, because if the wage distribution would degenerate at $w^*$, no worker would have an incentive to sample a second job. To pin down $z_2 \in (0, z)$, we use the restriction that workers are indifferent about sampling one or two
jobs when \( z_2 \in (0, \bar{z}) \), \( R_1 = R_2 \). Using (7), we obtain

\[
\int_0^{w^*} H''(w)(1-H''(w))\Phi'(w)dw = c_2. \tag{13}
\]

**Proposition 1.** Assume \( f(k) = 1 \) for \( k \geq 1 \) and 0 otherwise. Then there exists a \( \bar{c} \) such that iff \( \bar{c} > c_2 \), there are at least two equilibria in which the wage distribution is \( H''(w) \) on \( X'' = [0, \bar{w}] \cup w^* \), defined by (6), (10), (11), and (12); and \( z_2 \) satisfies (13). Also, there always exists an equilibrium in which \( z_2 = 0 \) and \( H'' \) is degenerate at a zero wage. In any equilibrium, \( Q = 1/V \) satisfies (9). There are no other equilibria.

The panels in Figure 1 plot the gross return to searching for a second job as a function of the endogenous variable \( z_2 \leq \bar{z} \), for four different values of \( r \). Search intensity \( z_2 \) is determined in equilibrium so that this gross return is equal to the marginal cost of sampling a second job. The top row shows the limit as \( r \) converges to zero and \( r = 0.1 \). The bottom row shows \( r = 0.5 \) and \( r = 0.9 \). The shape of this figure is essentially independent of \( r \).

**Proof.** The text before the proposition justifies the characterization of equilibrium. The only remaining issue is existence. Observe that the left-hand side of equation (13) is implicitly a continuous function of the endogenous variable \( z_2 \). Moreover, the left-hand side of equation (13) is nonnegative, and evaluates to 0 at \( z_2 = 0 \) and \( z_2 = \bar{z} \) (see Figure 1). Therefore, this expression must have an interior maximum, whose value we denote by \( \bar{c} \). If \( c_2 > \bar{c} \), there is no solution to (13), and so no way that workers can be indifferent between sampling one and two jobs. For smaller values of \( c_2 \), continuity ensures that there are at least two solutions to (13), and so at least two equilibria with search. \( \blacksquare \)

An equilibrium without search always exists in this model. This reflects the Diamond (1971) paradox: if workers sample one job, firms will all offer the monopsony wage, and
it does not pay workers to search harder. If $c_1$ were positive, no worker would choose to participate and the market would shut down.

Our innovation is to show that for any moderate cost of sampling a second job, there is a wage distribution such that workers are indifferent between sampling one and two jobs. In equilibrium, some workers choose to gather more information than others. This endogenous informational heterogeneity induces firms to offer a distribution of wages. Because some workers will apply for any positive wage, some firms choose to reduce their hiring probability in return for a lower wage bill.

3.2. Technology dispersion

It is often claimed that wage dispersion exists because firms have different productivities. We would like to understand whether and why firms choose to use different technologies, and why there would be a link between firms’ wage and technology choices in such an environment. To address these questions, we reintroduce firms’ capital choice by assuming that the production function $f$ is increasing, strictly concave, and continuously differentiable. We also assume that there is a minimum efficient scale of production, $k_0 > 0$ with $f(k_0) = k_0 f'(k_0)$. This requires that $f(k)$ is negative for small values of $k$, which can be interpreted as a fixed cost of job creation or wage posting. Also assume there is a $k_1 > k_0$ with $f'(k_1) = r$. This gives an upper bound on firms’ capital investment.

Profit maximizing firms choose a capital–wage pair ($k, w$) only if $k$ maximizes $\pi$ for a given value of $w$. Equation (3) then yields a first order condition satisfied by any $(k, w) \in X$, that is by any capital–wage pair offered in equilibrium:

$$\frac{\partial \pi(k, w)}{\partial k} = (1 - \exp(-q(w))) f'(k) - r = 0.$$  

(14)

Since $f$ is concave, this defines a unique capital associated with any wage. We also know that if $(k, w) \in X$, firms offering this combination make zero profits, $\pi(k, w) = 0$. We combine this with (14) and eliminate $q(w)$ to get a simple monotonic relationship between any $(k, w) \in X$:

$$w = f(k) - kf'(k) \equiv W(k).$$  

(15)

Strict concavity of $f(k)$ implies that $W(\cdot)$ is strictly increasing, with $W(k_0) = 0$. Imposing the wage-investment schedule (15) allows us to reduce the dimension of the firms’ problem by focusing on their capital investments only. We indirectly define workers’ return over capital rather than wages, and firms’ profit over capital alone. Once we determine the marginal distribution of capital $H^k$, it is straightforward to back out the joint distribution of wages and capital using (15).

Workers prefer more capital. In Section 3.1, we showed that workers strictly prefer to apply to jobs offering higher wages. We now show that workers prefer to apply to jobs using more capital. This requires an additional assumption on the production function. Let $\phi(k) \equiv \log f'(k) - \log (f'(k) - r)$, defined on $[k_0, k_1)$, where $f'(k_1) = r$. Then define $\Phi: [k_0, k_1) \to \mathbb{R}$ by

$$\Phi(k) = \frac{r W(k)}{f'(k) \phi(k)},$$  

(16)
where \( W(k) \) is defined by (15). \( \Phi \) is strictly positive, and limits to zero at \( k_0 \) and \( k_1 \). In order to ensure that workers strictly prefer firms using more capital, we impose

Assumption 1. \( \Phi \) is strictly quasiconcave on its support and is maximized at \( k^* \in (k_0, k_1) \).

This is a restriction on the production function \( f \) alone, and can be easily verified for particular production functions.3

Assumption 1 enables us to prove an analogue of Lemma 1: in equilibrium, workers prefer more capital intensive jobs. The reasoning is unchanged. \( \Phi \) defines the return to a worker applying for a job with capital \( k \) if the firm offering that job makes zero profits. No firm will use more than \( k^* \) units of capital, so \( \Phi(k) = \rho(W(k)) \) is strictly increasing on the support of the capital distribution \( X^k \).

Characterizing the equilibrium. Once we know that workers strictly prefer jobs using more capital, the analogue of Lemma 2 obtains immediately. If \( z_1 = 1 \), all firms offer a zero wage, and so must invest in \( k_0 = W^{-1}(0) \) units of capital. If \( z_1 < 1 \), the support of the capital distribution \( X^k \) is a convex, nonempty interval \([k_0, \bar{k}]\) and possibly the point \( k^* \), which is the maximizer of \( \Phi \). It is atomless on \([k_0, \bar{k}]\), but may have an atom at \( k^* \). Lemma 3 also carries over to this environment. In particular, we always have \( z_1 + z_2 = 1 \). We omit the proofs of these results, as they are identical to those in the previous subsection.

Building on these lemmas, we once again obtain \( q(w) = Q((1 - z_2) + 2z_2G_2(w)) \). Since firms’ gross profit for hiring a worker is \( f(k) - W(k) = kf'(k) \), the zero profit condition is

\[
\pi(k, W(k)) = (1 - \exp(-q(w)))kf'(k) - rk = 0.
\]

Invert this:

\[
G_2(W(k)) = \frac{\phi(k) - Q(1 - z_2)}{2z_2}, \tag{17}
\]

where \( \phi(k) = \log(f'(k)) - \log(f'(k) - r) \). Again, \( G_2(W(k)) = H^k(k) \) for \( k \in [k_0, \bar{k}] \), the atomless part of the wage distribution. We can pin down the worker–firm ratio \( Q \) from the fact that a firm using capital \( k_0 \) will offer a zero wage, and so will not hire a worker who samples another firm, that is \( G_2(k_0) = 0 \):

\[
Q = \frac{\phi(k_0)}{1 - z_2}. \tag{18}
\]

Substituting (18) into (17):

\[
G_2(W(k)) = \frac{1 - z_2}{2z_2} \left( \frac{\phi(k)}{\phi(k_0)} - 1 \right) = H^k(k), \tag{19}
\]

3. One can prove that \( \Phi \) is strictly quasiconcave if \( kf''(k) \) is increasing and \( \phi(k) \) is convex. Also, numerical simulations show that for generalized Cobb–Douglas production functions of the form \( f(k) = Ak^a - k \) or \( f(k) = A(k - k)^a \), \( \Phi \) is always strictly concave, and hence strictly quasiconcave. If \( \Phi \) were not quasiconcave, there could be multiple atoms in the capital distribution, each corresponding to a local maximum of \( \Phi \). This introduces the possibility, for example, of a two-point capital distribution, which would prevent us from extending Lemma 3 to this environment. Even if \( z_1 = 0 \), firms at the bottom of the capital distribution could earn a profit by hiring workers who only sample other firms at the bottom of the capital distribution. While this possibility is interesting, we have been unable to find a concave production function that violates assumption (16). For this reason, we impose Assumption 1 in the remainder of the paper.
where the second equality holds for $k \in [k_0, \tilde{k}]$. The top of the atomless parts of the capital distribution, $\tilde{k}$, satisfies $H^k(\tilde{k}) = 1 - \mu$, where $\mu$ is the fraction of firms offering $k^*$, determined as before

$$\mu = \max \left\{ \frac{1 + z_2}{z_2} \frac{1 - z_2}{z_2} \frac{\phi(k^*)}{\phi(k_0)} ; 0 \right\}. \quad (20)$$

There is again a cap on the fraction of searchers $z_2$, since we know that $\mu < 1$ in a search equilibrium: $z_2 < z \equiv 1 - \phi(k_0)/\phi(k^*)$.

An increase in search intensity $z_2$ now raises the capital distribution (and hence also the wage distribution) in the sense of first-order stochastic dominance. When more workers sample two jobs, firms are induced to offer higher wages. The number of active firms must fall in order to restore profitability to active firms. The labour market becomes less tight, so firms are willing to make larger capital investments.

Finally, we pin down the fraction of workers sampling two jobs using the condition that workers are indifferent between sampling one or two jobs, i.e. $R_1 = R_2$. Updating equation (13) using the notation of this subsection, we require

$$\int_{k_0}^{k^*} H^k(k)(1 - H^k(k))\Phi'(k)dk = c_2. \quad (21)$$

**Proposition 2.** Assume $f(k)$ is increasing, strictly concave, and continuously differentiable, with $f(k_0) = k_0 f''(k_0)$ and $f'(k_1) = r$ for some $0 < k_0 < k_1$, and $\Phi$ satisfies Assumption 1. Then there exists a $\tilde{\epsilon}$ such that if $\tilde{\epsilon} > c_2$, there are at least two equilibria in which the capital distribution is $H^k(k)$ on $X^k \equiv [k_0, \tilde{k}] \cup k^*$, defined by (19) and (20); $w$ satisfies (15); and $z_2$ satisfies (21). Also, there always exists an equilibrium in which $z_2 = 0$ and $H$ is degenerate at $(k, w) = (k_0, 0)$. In any equilibrium, $Q = 1/V$ satisfies (18). There are no other equilibria.

We omit the existence proof, which parallels the proof of Proposition 1.

### 3.3. Robustness

One might be concerned that search equilibria are quite fragile. We require firms to offer a particular distribution of wages and capital, and the correct fraction of workers to search. How could a decentralized economy arrive at such an allocation?

We imagine a (possibly fictitious) infinite repetition of the game. Although a complete theory of evolutionary stability in games with a continuum of agents goes beyond the scope of this paper, it appears that market forces should push agents towards one of the search equilibria described in Propositions 1 and 2.

To see why, consider first the behaviour of firms for a given value of $z_2$. Suppose for some reason that the capital–wage distribution is depressed, so too many firms choose $x < \hat{x}$, for some $\hat{x} \equiv (\hat{k}, W(\hat{k}))$. An individual firm offering a low wage will realize that if it raises its capital–wage pair to $\hat{x}$, it will not lose many applicants, and thus earn supernormal profits. This creates a natural equilibrating force. Which firm will do this? Here the literature makes two suggestions. First, in the spirit of our fictitious play with a large number of players, it may be that only a small fraction of firms can change their policy in a given period. In this case, it will be whichever firms are fortunate enough to change their policy. Second, as in Harsanyi’s (1973) purification argument, if there were a small
amount of heterogeneity across firms, then it would be the firms with the lowest cost of capital that would choose to offer the highest capital–wage pairs.

Next consider the behaviour of workers in the neighbourhood of a given equilibrium. Recall that there are generally two equilibria with search. In a neighbourhood of the search equilibrium with the higher value of \( z_2 \), workers' search decisions are strategic substitutes. If a few more workers sample two jobs, raising \( z_2 \), firms will offer a less disperse wage distribution, reducing the incentive to sample two jobs (see Figure 1). That is, because the relevant measure of wage dispersion, the left-hand side of equation (13) or (21), decreases in \( z_2 \) in the neighbourhood of an equilibrium, the equilibrium is stable in terms of fictitious play. A decentralized economy is likely to reach such an equilibrium.

On the other hand, in a neighbourhood of the search equilibrium with the lower value of \( z_2 \), workers' search decisions are strategic complements. If a few more workers sample two jobs, firms respond by making the wage distribution more disperse. This raises the incentive to sample a second job, moving the economy away from the equilibrium. This equilibrium is unstable; a decentralized economy is unlikely to happen upon it.

3.4. Social efficiency

One might expect that free-rider problems will make the equilibrium allocation highly inefficient. This section shows to the contrary that the stable search equilibrium comes surprisingly close to decentralizing an efficient allocation, particularly when search costs \( c_2 \) are low. We use the usual notion of efficiency in the search literature (e.g. Hosios (1990), Pissarides (1990)), looking at the choice of a “social planner” who maximizes total output by choosing the number of active firms \( V=1/Q \) and the capital intensity of each active firm \( k \),

4  but who is subject to the same coordination and search problems as the decentralized economy.

Since search is purely rent-seeking from the social planner's perspective, she would have each worker sample only one job. Total output is

\[
\max_{k,Q} \left( 1 - \exp(-Q) \right) f(k) - rk,
\]

The numerator is the fraction of firms that hire workers times the output per active firm, minus the fixed investment cost \( rk \). This is multiplied by the number of active firms, or equivalently divided by \( Q \). Note that the social planner does not care about wages, so only output appears in the objective function.

To characterize the efficient allocation \( \hat{k}, \hat{Q} \), observe that for fixed \( Q \), the objective is concave in \( k \). Thus the first-order condition with respect to \( k \) is necessary for a maximum, and yields \( (1 - \exp(-Q)) f'(k) = r \), or equivalently \( Q = \phi(k) \). Using this to eliminate \( Q \) from the objective, we find that the planner chooses \( k \) to maximize

\[
\max_k r \frac{f(k) - kf'(k)}{f'(k)\phi(k)} = \max_k \Phi(k).
\]

Then it follows from the definition of \( k^* \) that \( \hat{k} = k^* \) and \( \hat{Q} = \phi(k^*) \).

Although the decentralized economy never achieves this allocation when there are positive search costs, it approaches it in a stable search equilibrium as search costs \( c_2 \)

4. One could also look at the behaviour of a social planner in the environment of Section 3.1, with no capital choice. As our conclusions would be unchanged, we focus on the more general environment here.
become small. In that equilibrium, a reduction in search costs leads more workers to search. This increases $z_2$ towards $\bar{z} = 1 - \phi(k_0)/\phi(k^*)$, raising the equilibrium wage and capital distribution in the sense of first-order stochastic dominance. The mass of firms that choose $x^* = (k^*, W(k^*))$ increases, and wage dispersion declines.

In the limit when $c_2 = 0$, all firms choose $x^*$. Wage dispersion disappears, but workers are willing to sample two jobs, because it is costless. Using (18), the worker–vacancy ratio, $Q$, is driven to $\phi(k_0)/(1 - \bar{z}) \equiv \phi(k^*)$. The efficient allocation is therefore decentralized. We conclude that search and wage/technology dispersion are efficient ways of allocating workers to firms and of providing firms with correct investment incentives. Despite the free-rider problem and other externalities, if only a fraction $\bar{z}$ of agents sample two wages, the equilibrium is efficient.

When search costs are positive, workers spend resources on search, which from a partial equilibrium perspective is purely rent-seeking and socially wasteful. Nevertheless, in general equilibrium, search is useful. First, given the existence of wage and technology dispersion, search improves sorting; it helps allocate workers to more capital-intensive firms. Second, search limits firms’ monopsony power and drives up wages. This mitigates entry and forces firms to use more efficient technologies with higher labour productivity. Thus the existence of wage and technology dispersion, a necessary condition for search, enhances efficiency in a search economy.

This result contrasts with Acemoglu (1996), who finds that ex ante investments are always distorted if firms and workers bargain over wages after matching, because firms that make larger investments are forced to share part of their profits with workers through an upward-sloping wage-investment schedule (see also Grout (1984)). With wage commitments, however, firms are the residual claimants on any additional returns generated by a superior technology, after conditioning on workers’ application decision. The limiting equilibrium is efficient even though the upward sloping wage-investment schedule (15) does not disappear.

Our results also generalize Moen (1997) and Shimer (1996)’s finding that if workers could costlessly observe all posted wages in an economy without exogenous capital intensity, the equilibrium would be efficient. Acemoglu and Shimer (1999) extend that result to an economy with endogenous capital intensity, but maintain the assumption that workers costlessly observe all posted wages.

Finally, because of the free-rider problem with costly search, there is too little search in a decentralized equilibrium. A small search subsidy, financed by lump-sum taxation, will raise output. To see why, first observe that firms dissipate profits through entry, so net output is simply equal to the wage of a representative worker, net of search expenditures, $z_2 c_2$:

$$ (1 - z_2) \int_{k_0}^{k^*} (1 - H^k(k)) \Phi'(k)dk + z_2 \left( \int_{k_0}^{k^*} (1 - (H^k(k))^2) \Phi'(k)dk - c_2 \right). $$

In a stable search equilibrium, a small search subsidy, which reduces the cost of search $c_2$ by $\epsilon$, leads to a marginal increase in search intensity $z_2$. This has three effects: first, it raises the capital distribution $H^k$, and so raises the expected wage of all workers conditional on their search intensity; second, it discretely raises the expected wage of those workers who search more intensively by the amount on the left-hand side of equation (21); and third, it raises search expenditures of those workers who sample an additional job by $c_2$, which is socially wasteful. Equation (21) shows that the last two effects cancel out when starting from an economy without a search subsidy, i.e. with $\epsilon = 0$. Thus the marginal effect of a
small search subsidy is to raise social welfare by

\[-(1 - z_2) \int_{k_0}^{k^*} \frac{\partial H^k(k)}{\partial z_2} \Phi'(k) dk - 2z_2 \int_{k_0}^{k^*} \frac{\partial H^k(k)}{\partial z_2} H^k(k) \Phi'(k) dk,\]

which is strictly positive since \(H^k(k)\) is decreasing in \(z_2\). An immediate corollary of this is that the stable search equilibrium yields more output than any other equilibrium, including that with no search, because for a given search cost \(c_2\), welfare is higher when search intensity is higher.

4. THE DYNAMIC MODEL

Workers may prefer to engage in a \textit{sequential} search strategy, rather than the \textit{“fixed sample size”} search that our static model imposes. Sequential search allows a worker to condition future search decisions on current outcomes. Nevertheless, this comes at some cost, since impatient workers prefer to obtain a higher wage earlier rather than later. A larger sample size facilitates the timely acquisition of information (Morgan and Manning (1985)). To address these intertemporal optimization issues, this section extends our model to a dynamic environment.

We consider a discrete time, infinite horizon search economy. Workers and firms are risk-neutral, and maximize the present discounted value of their income net of search costs using a common discount factor \(\beta \in (0, 1]\). At the start of every period, workers are either unemployed or employed. Likewise, firms are either inactive, maintain an open vacancy, or have a filled job. Associated with every active firm is a capital intensity \(k\). During each period, the sequence of events is as follows:

1. Inactive firms decide whether to create a job vacancy by purchasing capital \(k > 0\) at cost \(vk\). Those that do not remain inactive throughout the period.
2. Firms with unfilled vacancies from the previous period and those creating vacancies this period post a wage.
3. Unemployed workers choose their search intensity \(n \geq 1\) at marginal cost \(c_n\), and observe the wage of \(n\) vacancies. They then apply to at most one of the sampled vacancies. For simplicity, employed workers may not search.
4. Each vacancy that receives an application hires one of the applicants. The firm has a filled job and the applicant is employed. The rest of the applicants remain unemployed, and the remaining vacancies stay vacant.
5. All firms with filled jobs produce and pay the promised wage.
6. Each active firm, vacant or filled, faces an independent and exogenous probability \(\delta > 0\) that its capital stock is destroyed. At the start of the following period, the firm is inactive. If the firm had a filled job, its employee becomes unemployed.

When \(\beta = 0, \delta = 1, \) and \(v = r\), this is equivalent to an infinite repetition of the static model in Section 2. To maintain comparability with the static model for more general parameter values, we define \(r = v(1 - \beta(1 - \delta))\), so \(r\) is the rental rate of capital, and the price of capital accounts both for discounting and depreciation.

Although we could once again define an equilibrium in terms of workers’ strategies, to save space we define it directly in terms of aggregate variables. Since we are only interested in steady state equilibria, we suppress time subscripts. Otherwise, our notation is unchanged. Let \(\pi(k, w)\) denote the present discounted value of a vacancy using \(k\) units of capital and offering a wage \(w\), net of its capital cost \(vk\). \(V\) denotes the measure of
vacancies, and $H$ is the joint distribution of capital and wages across firms with vacancies, with support $X$.

On the worker's side, let $R_n$ denote the expected (lifetime) utility of an unemployed worker who samples $n$ vacancies, and $\rho(w)$ denote the expected utility of an unemployed worker who applies for wage $w$. Also define $R^* = \max_n R_n$ as the expected utility of an unemployed worker. Let $z_n$ be the fraction of unemployed workers who sample $n$ vacancies, and $p_n(w_1, \ldots, w_n)$ be the unemployed workers' preference function over arbitrary $n$-tuples of wages.

Using these, we can define a steady state equilibrium of the dynamic model.

**Definition 2.** A steady state equilibrium consists of a measure of vacancies $V$, an unemployment rate $U$, a joint capital–wage distribution $H$ with support $X$, an expected profit function for firms $\pi$, expected return functions for workers $\rho$ and $R$, and sampling decisions $z_n$ and preference functions $\{p_n\}$ for workers such that:

1. (Profit Maximization) $\forall (k, w) \in X$, $\forall (k', w')$, $\pi(k', w') \leq \pi(k, w) = 0$.  
2. (Optimal Application) $\forall n \in \mathbb{N}$, $\forall (w_1, \ldots, w_n)$,  
   $$p_n(w_1, \ldots, w_n) \neq i \text{ if } \rho(w_i) < \max_j \rho(w_j) \text{ or } \rho(w_i) < \beta R^*$$

   and

   $$p_n(w_1, \ldots, w_n) \neq 0 \text{ if } \max_j \rho(w_j) \geq \beta R^*. \ldots$$

3. (Optimal Sampling) $z_n = 0$ if $R_n < R^*$.  
4. (Steady State) $z_n = 0$ if $R_n < R^*$.

There are two significant changes in the definition of equilibrium. First, we require that the measure of vacancies and unemployed workers are constant over time. Second, we recognize that workers' reservation utility is $\beta R^* \geq 0$, so workers may reject jobs offering a positive wage.

To characterize a steady state equilibrium, first observe that $q(w)$, still given by equations (1) and (2), defines the fraction of unemployed workers who apply to a firm offering a wage $w$, conditional on sampling such a firm. Next, consider a vacancy with $k$ units of capital offering a wage of $w$. If offered in equilibrium, i.e. $(k, w) \in X$, the vacancy must have value $vk$, according to the free entry condition. Moreover, it must satisfy the Bellman equations

$$vk = (1 - \exp(-q(w)))F(k, w) + \exp(-q(w))\beta(1 - \delta)vk,$$

$$F(k, w) = f(k) - w + \beta(1 - \delta)F(k, w).$$

With probability $1 - \exp(-q(w))$, the vacancy hires a worker, yielding a filled job with value $F(k, w)$. Otherwise, and unless the firm's capital is destroyed in the mean time, the vacancy remains the following period. Similarly, the value of the filled job $F(k, w)$ comes from net profits $f(k) - w$ until the match ends. Combining these two equations and using the fact that $v = r/(1 - \beta(1 - \delta))$

$$rk = (1 - \exp(-q(w)))\left(\frac{f(k) - w + \beta(1 - \delta)rk}{1 - \beta(1 - \delta)}\right).$$
Since a small change in $k$ holding $w$ fixed must not affect profits, we obtain the standard wage-investment schedule (15), $w = W(k) \equiv f(k) - kf'(k)$ if $(k, w) \in X$. Use this to eliminate the wage from the previous equation

$$r = (1 - \exp(-q(W(k)))) \left( \frac{f'(k) - \beta(1 - \delta)r}{1 - \beta(1 - \delta)} \right).$$

(22)

This gives a simple relationship between a vacant firm's capital intensity and the rate at which it hires workers, if the zero profit condition binds. Invert this:

$$q(W(k)) = \log \left( \frac{f'(k) - \beta(1 - \delta)r}{w} \right) - \log \left( \frac{f'(k) - r}{\phi(k)} \right),$$

which is a generalization of the definition of $\phi(\cdot)$ in the previous section.

Similarly, an unemployed worker who applies for a job offering a wage of $w$ gets

$$\rho(w) = \frac{1 - \exp(-q(w))}{q(w)} E(w) + \left( 1 - \frac{1 - \exp(-q(w))}{q(w)} \right) \beta R^*,$$

$$E(w) = w + \beta(1 - \delta)E(w) + \beta \delta R^*.$$  

The worker is hired with probability $(1 - \exp(-q(w)))/q(w)$ and moves to the employed state with value $E(w)$. Otherwise, she remains unemployed, obtaining the continuation value $\beta R^*$. In the employed state, she earns a wage of $w$, but she becomes unemployed with probability $\delta$ in any period, receiving continuation value $R^*$. Combining these equations, we obtain

$$\rho(w) = \frac{1 - \exp(-q(w))}{q(w)} \left( \frac{w - \beta(1 - \beta)(1 - \delta)R^*}{1 - \beta(1 - \delta)} \right) + \beta R^*.$$  

Now if $(k, w) \in X$, then we can use (22) to pin down $q(w)$:

$$\rho(W(k)) = r \frac{W(k) - \beta(1 - \beta)(1 - \delta)R^*}{(f'(k) - \beta(1 - \delta)r)\phi(k)} + \beta R^* = \Phi(k).$$

(23)

Again, this is a generalization of the definition of $\Phi$ in equation (16), and collapses to (16) when $\beta = 0$. We again impose:

**Assumption 2.** $\Phi$ is strictly quasiconcave for all values of $R^*$ and is maximized at $k^*$.

Note that $k^*$ now depends on $R^*$, the endogenous value of an unemployed worker.

Assumption 2 immediately leads to the analogues of Lemmas 1–3. All workers sample either one or two jobs in any period, i.e. $z_1 + z_2 = 1$, and apply to the highest posted wage/capital. The support of the capital distribution is a convex interval $[\tilde{k}_0, \tilde{k}]$ and possibly the point $k^*$.

We can no longer assert that the bottom of the capital distribution is $k_0$ with $W(k_0) = 0$, as workers will reject a small positive wage if the value of unemployment is positive. The bottom of the wage distribution is given by the condition that workers are indifferent about accepting a job: $\Phi(\tilde{k}_0) \equiv \beta R^*$, or using (23):

$$f(\tilde{k}_0) - \tilde{k}_0 f'(\tilde{k}_0) \equiv \beta(1 - \beta)(1 - \delta)R^*. $$

(24)
Otherwise, the derivation of the capital distribution replicates equations (18)–(20):

\[ Q = \frac{\phi(k_0)}{1 - z_2}, \]  

\[ H^k(k) = \frac{1 - z_2}{2z_2} \left( \frac{\phi(k)}{\phi(k_0)} - 1 \right), \quad k \in [k_0, k], \]  

\[ \mu = \max \left( \frac{1 + z_2}{z_2} \left( \frac{1 - z_2}{2z_2} \frac{\phi(k^*)}{\phi(k_0)} \right), 0 \right). \]

Similarly, workers are indifferent between sampling one and two jobs under a generalized version of (21):

\[ \int_{k_0}^{k^*} H^k(k)(1 - H^k(k))\Phi'(k)dk = c_2. \]

The dynamic model has one additional unknown variable, the value of an unemployed worker \( R^* \). We pin this down using the fact that unemployed workers are willing to sample only one job, \( R^* = R_1 \):

\[ R^* = \int_{k_0}^{k^*} \Phi(k)dH^k(k). \]

In summary:

**Proposition 3.** Assume \( f(k) \) is increasing, strictly concave, and continuously differentiable, with \( f(k_0) = k_0 f'(k_0) \) and \( f'(k_1) = r \) for some \( 0 < k_0 < k_1 \), and \( \Phi \) satisfies Assumption 2. Then there exists a \( \overline{c} \) such that if \( c > c_2 \), there are at least two equilibria in which the capital distribution is \( H^k(k) \) on \( X^k \equiv [k_0, k] \cup k^* \), defined by (24), (26), and (27); \( w \) satisfies (15); \( z_2 \) satisfies (28); and \( R^* > 0 \) satisfies (29). Also, there always exists an equilibrium in which \( z_2 = 0 \) and \( H \) is degenerate at \( (k, w) = (k_0, 0) \). In any equilibrium, \( Q \equiv U/V \) satisfies (25); and the steady state unemployment rate \( U = \delta Q/(\delta Q + (1 - \delta)\varepsilon) \), where \( \varepsilon \) is the average probability that a vacancy is filled in a given period,

\[ \varepsilon \equiv r \int_{X^k} \left( \frac{1 - \beta(1 - \delta)}{f'(k) - \beta(1 - \delta)r} \right) dH^k(k). \]

There are no other equilibria.

This is a straightforward generalization of Proposition 2, which justifies our detailed analysis of the static model.

The dynamic environment introduces two complexities, both addressed in the appendicized proof. First, the equilibrium capital–wage distribution depends on the value of an unemployed worker. This introduces another equation and creates a fixed point problem in solving for the equilibrium. Second, the steady state unemployment rate is endogenous. This issue is easily addressed, as it is possible to solve for \( U \) after calculating the distribution \( H \) and the unemployment vacancy ratio \( Q \).

5. The non-steady state behaviour of this economy can also be characterized. There exists an equilibrium path along which \( Q \) and \( H \) immediately jump to their steady state values, while the unemployment rate converges slowly. Since there are multiple steady state equilibria, we expect that there are infinitely many non-stationary Rational Expectations equilibria. Solving for them goes beyond the scope of this paper.
The comparative statics and welfare results from Section 3.3 generalize to the dynamic environment. The efficient allocation, which maximizes the present discounted value of output, is again the limit of the stable search equilibrium when $c_2$ converges to zero. When search costs are positive, some firms underinvest and too many firms are active given the unemployment rate. No firm ever overinvests compared to the efficient investment level $k^*$.

5. EMPIRICAL IMPLICATIONS AND RELATED LITERATURE

This paper has developed and explored a novel theory of wage and technology dispersion. A number of other explanations for these phenomena exist in the literature, however, and so this section confronts the primary implications of our theory and competing theories with existing empirical evidence. While we do not believe that our mechanism explains all wage and technology dispersion, the evidence suggests that it may represent an important part of the story.

A large empirical literature documents wage dispersion among observationally identical workers. Murphy and Topel (1986) argue that part of the dispersion is due to unobserved heterogeneity, but this appears not to be the whole story (e.g. Krueger and Summers (1988), Reynolds (1951), Abowd, Kramarz, and Margolis (1999), and Groshen (1991)). For example, Krueger and Summers (1988) and Gibbons and Katz (1992) find that workers who move from a high to a low wage industry lose approximately the wage differential between the two industries. Holzer, Katz and Krueger (1991) find that high wages attract more applicants.

Motivated by this evidence, a theoretical literature explains wage dispersion among ex ante identical workers. One class of models presumes that firms differ in the labour productivity, and introduce bargaining or efficiency wage mechanisms that ensure more productive firms pay higher wages. In Mortensen and Pissarides (1994), productivity differences are due to firm-, technology-, or match-specific shocks. Montgomery (1991), Acemoglu (2001), and Pissarides (1994) consider exogenous differences in technology across industries. Thus in these papers, wage dispersion is a consequence of technology dispersion. In contrast, our theory generates endogenous technology dispersion within or between industries as a consequence of wage dispersion.

We believe that theories linking wages and technology are promising because there is empirical evidence of a systematic relationship between wage and technology differentials both at the industry level and within industries. Katz and Summers (1989) find a correlation between industry wage levels and measures of technology like the capital–labour ratio and R&D intensity. A number of other papers find a similar correlation between a firm’s wage level and its technology choice relative to other firms in the industry. (Doms, Dunn and Troske (1997), Autor, Katz and Krueger (1998), Chennells and Van Rennen (1998) and Entorf and Kramarz (1998)). For example, Doms, Dunn and Troske (1997) find that production workers in plants with the fewest advanced technologies receive wages 20% lower than production workers in plants with the most advanced technologies. Similarly, Abowd, Kramarz and Margolis (1999) find that it is the high productivity and high capital firms that pay higher wages in French data.

This line of empirical research can be used to distinguish between the exogenous technology models discussed above and our model of the joint determination of technology and wages. Chennells and Van Rennen (1998) investigate the determination of wages

6. Many of these papers offer explanations of price dispersion for identical goods. We reexpress these models in terms of labour markets to ease comparability.
and technology using an instrumental variables approach. They instrument a plant's technology using lagged R&D intensity and number of patents; and they use the gender composition of the plant's employees as an instrument for its wages. They find no evidence that technology affects wages, but a strong effect of wages on technology. Doms, Dunn and Torske (1997) exploit a longitudinal data set to shed some additional light on this question. Plants that adopt more advanced technologies do not pay higher wages to their existing employees. However, it is precisely the plants that pay higher wages that adopt the new technologies. They conclude that "... compared with plants that adopt fewest technologies..., plants that adopt the most technologies... paid production workers 14.6\% higher wages in 1977..." (p. 279). This evidence is quite encouraging for us. It suggests at a minimum that wages and technologies are jointly determined, and that the relationship between them is more complex than previously assumed.

A second set of papers assumes that all jobs are equally productive, and generates wage dispersion because workers have different reservation wages (e.g. Salop and Stiglitz (1982), Albrecht and Axell (1984), Sattinger (1991)). Since the evidence discussed above suggests an important link between technology and wages, we believe that models that allow for productivity heterogeneity across firms are empirically more promising. Still, the mechanism used to generate wage and technology dispersion in our papers builds on the mechanisms that generate wage dispersion in those models. In particular, a number of these papers consider informational heterogeneity as a source of wage dispersion. Butters (1997), Lang (1991), Robert and Stahl (1993), Stahl (1994), and Stiglitz (1985) assume that the number of job openings that a worker learns about depends on a random advertising technology. Workers who are fortunate to receive a lot of advertisements have more opportunities, and hence a higher reservation wage on average. Salop and Stiglitz (1977) and Varian (1980) assume that a fraction of the population is uninformed and will accept a low wage job, while informed workers seek out and get high wage jobs.

Using a mechanism very similar to ours, Burdett and Judd (1985) endogenize this informational heterogeneity. They prove the existence of an equilibrium with wage dispersion, in which a fraction of the population chooses to learn about two wages, while the remainder stay uninformed. The amount of wage dispersion exactly offsets the cost of sampling two firms. The main difference between our work and theirs is that Burdett and Judd work in a model without unemployment. Workers are always hired when they apply for a job, because the marginal product of labour is constant. Although allowing for a decreasing marginal product of labour (i.e. firms can only hire one worker) complicates our analysis considerably, it also enables us to introduce a second factor of production, capital, and to close the model with a free-entry condition. We could not achieve these goals under Burdett and Judd's linear technology, and thus could not explore the complementarities between wage and technology choice, nor conduct our welfare analysis.

Another important paper by Burdett and Mortensen (1998) obtains wage dispersion in a model of on-the-job search. Employed workers have a higher reservation wage than unemployed workers, because they have the option of staying at their old job. Some firms adopt higher wages to help attract workers from low wage firms, while simultaneously reducing quits.

This type of model has a strong implication for the amount of wage dispersion across cohorts. As a cohort ages, most of its members will move up towards the top of the wage distribution. This implies that within-cohort wage dispersion should decline with age. Our theory, in contrast, predicts that wage dispersion should be approximately constant across cohorts. This first-order prediction of on-the-job search models has not been investigated directly. However, the data reported in Farber and Gibbons (1996) from the NLSY sheds
some light on this issue. In that data set, wage dispersion increases slightly as a cohort ages. For example, the standard deviation of wages normalized by the mean wage of those with no labour market experience is 0.41. The same number is 0.47 for those with eleven years of experience. Future empirical research on the behaviour of within- and between-cohort wage dispersion can further distinguish between Burdett and Mortensen's approach and ours.

Finally, recent attempts to structurally estimate equilibrium search models of wage dispersion suggest that our approach is important for capturing salient features of the labour market. For example, van den Berg and Ridder (1998) estimate Burdett and Mortensen's (1988) model using Dutch data. However, they are forced to introduce exogenous technology dispersion in order to match the observed wage distribution. In response, Robin and Roux (1998) and Postel-Vinay and Robin (1999) build on our paper to generate technology dispersion endogenously, and estimate a general equilibrium search model of this type with French data. We believe that the co-determination of wages and technology will continue to prove fruitful in this line of research and enhance our understanding of the labour market.

6. CONCLUSION

This paper proposed a model in which firms make technology (capital) choices and post wages. Workers engage in costly search to gather information, and make their application decisions using this information. The equilibrium of this model involves wage dispersion among identical workers, even when all firms have the same level of labour productivity. The most original contribution of our analysis, however, was to endogenize firms' technology choice, and hence the productivity of labour. The same forces leading to wage dispersion also create endogenous technology dispersion, and higher productivity firms pay higher wages to identical workers.

Perhaps our most surprising result is the efficiency of wage and technology dispersion. In competitive markets, the law of one price is necessary for efficiency. In our frictional environment, deviations from the law of one price are required for an efficient functioning of the market economy. Moreover, as the costs of sampling additional wages decline, the wage dispersion equilibrium of our economy approaches the constrained efficient allocation, despite the many informational and pecuniary externalities. The source of these results is the novel interaction identified in this paper. When there is no wage dispersion, workers do not search enough, and there is too little competition for labour. This distorts capital–labour ratios and the entry margin of firms. Wage dispersion encourages search and enables the economy to reach a more efficient allocation of resources. In fact, subsidizing search improves welfare in this economy.

The contribution of this paper is theoretical. Nevertheless, because it is motivated by the empirical observation that wage and technology dispersion are important, we provided a number of implications that distinguish our approach from others. Additional empirical work is necessary to investigate these implications, and establish which approach best approximates the causes of wage dispersion.

APPENDIX

Proof of Proposition 3. Existence. In the other existence proofs, we showed that the left-hand side of equations (13) and (21) is a continuous function of search intensity $z_2$, equal to zero when $z_2 = 0$ or $z_2 = z$, and positive otherwise. Now we show that the left-hand side of equation (28) is a continuous function of the welfare
of unemployed workers $R^*$, equal to zero when $R^* = 0$ or $R^* = \bar{R} > 0$, a value to be defined in the proof. To do this, we first show that equation (29) implicitly defines a continuous and strictly increasing relationship between $z_2$ and $R^*$, $z_2 = Z(R^*)$. Then we mimic the proof of Proposition 1.

Using (23) to eliminate $\Phi$, rewrite (29) as

$$
(1 - \beta)R^* = \int_{k_0}^{k^*} r \left[ W(k) - \beta(1 - \beta)(1 - \delta)R^* \right] dH^e(k).
$$

(31)

The only direct effect of an increase in search intensity $z_2 \in [0, 2]$ on this equation is to continuously raise the capital distribution in the sense of first order stochastic dominance (equations (26) and (27)). Using Assumption 2, the integrand of the right-hand side of (31), $\Phi - BR^*$, is increasing in $k$ on the support of the integral $[k_0, k^*]$, so an increase in $z_2$ continuously raises that integral. Equality in (31) is restored through an increase in $R^*$. This proves that the equation defines a strictly increasing and continuous relationship between $z_2 \in [0, 2]$ and $R^*$.

When $z_2 = 0$, the capital distribution degenerates at $k_0$ satisfying $W(k_0) = \beta(1 - \beta)(1 - \delta)R^*$. Thus the right-hand side of (31) evaluates to zero, so $R^* = 0$. Inverting this, $Z(0) = 0$. When $z_2 = \bar{z} \equiv 1 - \phi(\bar{R}_0)/\phi(R^*)$, the capital distribution degenerates at $k^*$. We can again use (31) to implicitly define $\bar{z} = Z(\bar{R})$ for some $\bar{R} > 0$. For $R^* \in (0, \bar{R})$, $Z(R^*) \in (0, 2)$, so the capital distribution is non-degenerate.

Finally, turn to the condition that workers are indifferent between sampling one and two jobs, (28). The left-hand side is directly and indirectly (through $Z$) a continuous function of $R^*$. It is strictly positive for $R^* \in (0, \bar{R})$, since the capital distribution is non-degenerate, but evaluates to zero for $R^* = 0$ or $R^* = \bar{R}$. Thus it obtains a maximal value $\bar{c} > 0$ at some $R^* \in (0, \bar{R})$. Using standard arguments, for all $c_2 < \bar{c}$, there are at least two values of $R$ satisfying the workers’ indifference condition, each of which corresponds to an equilibrium.

**Steady state unemployment.** The fraction of vacancies that hire a worker in any given period, $\bar{v}$, can be calculated by integrating the probability that a vacant firm using $k$ units of capital hires a worker in any period, $1 - \exp(-q(W(k)))$ over the density of vacant firms using capital $k$, $dH^e(k)$. Use equation (22) to eliminate $1 - \exp(-q(W(k)))$ from this expression and derive (30). Finally, the steady state unemployment rate satisfies a “job creation equals job destruction” equation, $(1 - \delta)U/\bar{Q} = \delta(1 - U)$, which is easily solved for $U$. ||

**Acknowledgements.** This is a revision of our 1997 paper “Efficient Wage Dispersion”. We are grateful to two anonymous referees, Patrick Bolton, Peter Diamond, Bengt Holmstrom, Hyun Shin, and seminar participants at Brown, Harvard–MIT, Northwestern, Northwestern Summer Institute, Nuffield College, Maryland, Pennsylvania, Princeton and University College London. Financial support from the National Science Foundation Grant SBR-9602116 for the first author and SBR-9709881 for the second author are gratefully acknowledged.

**REFERENCES**


