Ex-ante Selection and Ex-post Learning: Implications for labor market outcomes (Preliminary)

Gonul Sengul
European University Institute
February 15, 2010

Abstract

Firms want to hire the best worker for the job, and they choose strategies to select an employee accordingly. This paper analyzes how the speed of learning about the compatibility of an employee could affect these strategies. The longer it takes to learn, the harder it becomes to detect the mistake of hiring the "wrong" worker, thus, the more likely it is that the firm invests more in selection. The speed of learning would depend on the skill requirement of a job, so would the decision about hiring strategies. Hence, differences in speed of learning across skill groups would contribute towards inequalities in their labor market outcomes.

Keywords: Unemployment, Skill, Learning, Hiring Strategies, Job Search
JEL Code: E24, J64, J63.

The strategies that firms employ to decide whom to hire are far from being homogeneous across vacancies. Barron, Berger, and Black (1997) show that employers “search more” when they hire workers with more education and prior experience,
and for jobs with higher training requirements. As the education requirement of a vacancy increases, the number of interviews and the number of applicants per offer goes up as well as the number of hours interviewed per applicant and per offer. Bar- ron, Bishop, and Dunkelberg (1987) report that total time spent on hiring is longer for high-skill occupations than for low-skill occupations. What accounts for the differences in firms’ search strategies across vacancies with different skill requirements? This paper aims to give an answer to this question.

Firms want to hire the best worker as not all workers are equally productive at the same job. To find that best worker, firms choose a search strategy, i.e., how much resources to allocate to recruit and screen applicants. After hiring, firms learn about the quality of this employment relationship over time. The structure of this “ex-post learning” should affect the search strategies. The longer it takes for a firm to be sure about the compatibility of a worker, the harder it becomes to detect the mistake of hiring the “wrong” worker. Thus, the slower the “ex-post learning” the more likely it is that the firm will want to reduce the risk of mistake by investing more in search.

Could differences in the speed of ex-post learning across skill explain the differences in employers’ search strategies? We should expect that the speed of learning depends on the skill requirements of a job. Due to the complexity of tasks at higher-skilled jobs, it would be harder for firms to recognize whether the employee is a good-fit. As a result, we expect the speed of ex-post learning to be slower for high-skilled jobs. In return, one anticipates that firms would invest more in search of a worker for jobs with higher skill requirements.

There are other determinants of search strategies than the ex-post learning. One such determinant is proposed to be the length of employment contract. Pinoli (2008) uses a partial equilibrium model where firms can choose either ex-ante selection or ex-post learning to assess worker quality, and they can only offer either long term or short term contracts. Pinoli shows that if firms offer long term contracts, ex-ante selection is more likely to be the best strategy. She uses a UK employer-employee dataset to test the model and shows that “temporary contracts are associated with lower recruitment effort” for low-skill occupations. However, she finds that at higher skill level occupations the relationship between search and the length of contracts becomes insignificant.

While the main goal of this paper is to explore the relationship between the ex-post learning and employer search, it also provides an explanation for the skill differences in labor market outcomes. A vast literature looks at differences in wage
dynamics across skill. It is also well-documented that high-skilled workers have lower unemployment rates, lower job separation probabilities, and longer tenure compared to their low-skilled counterparts. Nagypal (2007) uses disparities in ex-post learning about match quality across skill to explain differences in their job separation probabilities. She assumes search strategies to be the same across skill. Sengul (2009), on the other hand, assumes the ex-post learning not to vary across skill and shows that differences in search strategies explain the same fact as well. This paper can identify the contributions and the relative importance of both ex-post learning and search on differences in job separation probabilities, as well as other labor market outcomes.

In addition to further our understanding of variations in labor market outcomes by skill, this paper also makes another important contribution to employer search literature by studying how the speed of ex-post learning influences firms’ search strategies. Employer search is an important determinant of unemployment, among other labor market outcomes. Although the literature documents employer search (Van Ours and Ridder (1992), for example) and analyzes how it affects the labor market outcomes (Pries and Rogerson (2005)), what shapes these strategies has been less explored. This paper aims to contribute towards filling that gap.

I use a labor search-matching model, which builds on the model used by Pries and Rogerson (2005), to analyze effects of ex-post learning speed on firms’ search strategies. Suppose there are two types of workers: high-skilled and low-skilled workers. Assume that the skill type is exogenous and workers are homogenous within a skill group. High-skilled workers are more productive than low-skilled workers. Moreover, assume that there are high-skilled and low-skilled jobs which can employ only the workers of the same skill type. Firms are free to choose what type of jobs they want to create. There is no other interaction across skill.

For each skill type, an employment relationship between a firm and a worker can be of either a good or a bad quality. Good-quality matches are expected to produce higher output than bad-quality matches within a skill group. I assume that the bad-quality matches are not profitable, thus firms and workers do not want to be in a bad-quality match, regardless of the skill type.

When a firm with a vacancy meets an unemployed worker, parties receive information regarding the likelihood of this match being of a good quality. Based on this information (likelihood), they decide whether to form the employment relationship. The information the vacant firm and the unemployed worker receive is the outcome
of the following process. The good-quality match likelihood either comes from a "basic" distribution, Ψ or from a "better" distribution, Ω. Better distribution is more likely to deliver higher values for a good-quality match likelihood, and it is more costly to use. Although firms cannot choose a distribution per se, they can increase their chances of getting a draw from the "better" distribution by investing in ex-ante selection. The more the firm invests, the higher is the likelihood they receive a higher likelihood of a good-quality match.

The output that a match can produce is the sum of a deterministic value, which depends on the quality of the match, and a noise term. The random component of the output is such that, if parties are not sure about the quality of the match, they either ascertain the true quality of the match, or continue with the same beliefs (i.e., all-or-nothing learning). How soon they learn depends on how large is the variation of the noise term. I assume that the noise is a uniformly distributed random variable and its boundaries are larger for high-skilled jobs. That implies that the speed of learning, which is a function of that noise term, is also exogenous and it is smaller for high-skilled.

If the worker-firm pair learns the match to be of a good quality, they stay attached until hit by an exogenous separation shock. If the match is revealed to be of a bad quality, parties unanimously decide to separate as bad-quality matches are not profitable. The separation decision is unanimous as parties share the net surplus from the match.

When deciding how much to invest in selection, firms compare the expected value of returns to using the "better" distribution to expected returns from using the "basic" distribution. Firms choose the investment amount such that the returns to a higher likelihood of a good-quality match is equal to the marginal cost of investment.

The rest of the paper is organized as follows. The following section of the paper lays out the model. The equilibrium of the model is defined and analyzed in section 2. Section 3 presents the quantitative results of the model and is followed by concluding remarks.

1 Model

To analyze the effects of ex-post learning speed on firms’ search strategies, I use a labor search-matching model that builds on the model used by Pries and Rogerson (2005). I assume that there are two types of workers: high-skilled and low-skilled.
The skill type is exogenous and each type has a total mass equal to one. Moreover, there are high-skilled and low-skilled jobs which can employ only the workers of the same skill type. Firms are free to choose the type of jobs they want to create. I assume that there is no other interaction between skill types. Thus, the rest of the model applies to workers and firms of both skill types.

Within each skill group, firms and workers are ex-ante identical. Workers can be unemployed or employed and there are filled and vacant jobs. A production unit in the model is a firm-worker pair (match hereafter). The output a match can produce depends on the skill type of the firm-worker pair and the fit between the employee and the job, i.e., the quality of the match. A match between a firm and a worker can be of either a good or a bad quality. Good-quality employment relationships are expected to produce higher output than bad-quality matches within a skill group. Let $y_j = y^g_j + \epsilon_j$ be the output produced by an $i$ quality match of $j$ skill type, where $i$ can be good ($g$) or bad ($b$) and $j$ can be high-skilled ($hs$) or low-skilled ($ls$). $y^g_j$ is the deterministic part of the output whereas $\epsilon_j$ is the stochastic part. Note that the assumption that good-quality matches produce higher output than bad-quality matches within a skill group implies $y^g_j > y^b_j$ for all $j$. Moreover, I assume that the productivity gap between good and bad quality matches is larger for high-skilled workers than it is for low-skilled workers, i.e., $y^g_{hs} - y^b_{hs} > y^g_{ls} - y^b_{ls}$. I assume that the bad-quality matches are not profitable, thus firms and workers do not want to be in a bad-quality match, regardless of the skill type.

The stochastic term $\epsilon_j$ is uniformly distributed over the interval $[-\bar{\epsilon}_j, \bar{\epsilon}_j]$. Moreover, assume that $\bar{\epsilon}_j > \frac{y^g_j - y^b_j}{2\bar{\epsilon}_j}$. Under this assumption, if parties are not sure about the quality of the match, they either ascertain the true quality of the match with probability $\pi_j = \frac{y^g_j - y^b_j}{2\bar{\epsilon}_j}$, or continue with the same beliefs (i.e., all-or-nothing learning). I assume that $\bar{\epsilon}_j$, and thus $\pi_j$, are exogenous. Observe that $\pi_j$ determines the speed of ex-post learning. If, for instance, $\pi_j = 1$, then parties learn the match quality after observing the first period of output as there would not be an output

---

1 Since I assume that there is no interaction across skill groups, the size of these groups does not matter per se.

2 Note that if $y_j < y^g_j - \bar{\epsilon}_j$, parties know that the match must be of bad quality while $y_j > y^b_j + \bar{\epsilon}_j$ indicates a good quality match. Hence, the probability that the match quality reveals, $\pi_j$, is

$$\pi_j = 1 - P(y^g_j - \bar{\epsilon}_j < y^g_j + \epsilon_j < y^b_j + \bar{\epsilon}_j) = 1 - \frac{y^b_j + \bar{\epsilon}_j - y^g_j + \epsilon_j}{2\bar{\epsilon}_j} = \frac{y^g_j - y^b_j}{2\bar{\epsilon}_j}.$$
value that both a good and a bad quality match could have produced. Note that the noise term is skill specific. I assume that 
\[ \frac{\epsilon_{hs}}{\epsilon_{ls}} > \frac{y^g_{hs} - y^b_{hs}}{y^g_{ls} - y^b_{ls}}, \]
which implies that \( \pi_{hs} < \pi_{ls} \).

In the model, unemployed workers and vacant jobs find each other via a function that endogenously determines workers’ and firms’ meeting probabilities. The \( M(u, v) \) function, generically named as matching function, takes the numbers of unemployed workers \( u \) and vacant jobs \( v \) and determines the number of meeting between these parties \( M \). The number of meetings determine the probability that an unemployed worker meets with a vacancy \( h_w = M(u, v)/u \) and the probability that a vacant job meets with a worker \( h_f = M(u, v)/v \). The matching function is such that \( h_w \) and \( h_f \) fall into the unit interval.

When a firm with a vacancy meets an unemployed worker, parties learn the likelihood of this match being of a good quality, \( \gamma \), and decide whether to form an employment relationship. The good-quality match likelihood, \( \gamma \), is drawn from either a “basic” distribution \( \Gamma \) or from a “better” distribution \( \Omega \). These distributions are defined over the unit interval that have standard properties of a cdf. Moreover, \( \Omega \) first order stochastically dominates \( \Gamma \), i.e., in expectation, the better distribution is more likely to yield a good-quality match than the basic distribution. Let \( \lambda \) be the probability that the good quality match likelihood is drawn from the better distribution, \( \Omega \). Firms choose \( \lambda \), with a cost of \( C(\lambda) \).

Once the match is formed, the firm and the worker observe the level of output and learn the match quality with probability \( \pi \). If the worker-firm pair learns the match to be of a good quality, they stay attached until hit by an exogenous separation shock. If the match is revealed to be of a bad quality, parties unanimously decide to separate as bad-quality matches are not profitable. I assume that \( y^b \leq b \) and \( y^g > b \). Under this assumption, bad quality matches are undesirable in equilibrium.

### 1.1 Firms’ Bellman Equations

I start with formalizing firms’ decision problem. A firm needs to form beliefs about the search strategies of other firms in its optimization problem as others’ actions will affect wages. Let \( \bar{X} \) be the belief about the other firms’ search effort. Let me start with the value of a firm with a posted vacancy, \( V(\lambda, \bar{X}) \). The value of a vacancy is
the discounted value of expected profits, net of cost of the vacancy.

\[
V(\lambda, \lambda) = -c + \max_{\lambda \in [0,1]} \left\{ -C(\lambda) + \beta(1 - h_f)V(\lambda, \lambda) + \beta h_f \left[ \lambda \int J(\gamma, \lambda, \lambda) d\Omega + (1 - \lambda) \int J(\gamma, \lambda, \lambda) d\Gamma \right] \right\},
\]

where \( h_f \) is the likelihood that the firm meets with a worker, which is a function of the number of vacancies and unemployment, and \( J(\gamma, \lambda, \lambda) \) is the value of the firm when in a match which is of good quality with \( \gamma \) probability. Note that the firm pays a vacancy cost regardless of its search strategy. The firm's search choice depends on the cost of the search and how this choice affects \( J(\gamma, \lambda, \lambda) \). Equation (2) formalizes the problem of a firm that is in a match with a worker.

\[
J(\gamma, \lambda, \lambda) = \max \left\{ V(\lambda, \lambda), \ E(y|\gamma) - w(\gamma, \lambda, \lambda) + \beta \delta V(\lambda, \lambda) \right. \\
\left. + \beta (1 - \delta) \left[ \pi(\gamma J(1, \lambda, \lambda) + (1 - \gamma) J(0, \lambda, \lambda)) + (1 - \pi)J(\gamma, \lambda, \lambda) \right] \right\}
\]

where \( \delta \) is the exogenous probability of job destruction, \( w(\gamma, \lambda, \lambda) \) is the wage the firm pays, and \( E(y|\gamma) \) is the expected value of output, which is defined as \( E(y|\gamma) = y^h \gamma + y^l (1 - \gamma) \).

The firm compares the expected discounted value of profits from producing output with the current worker to the discounted present value of separating from the worker (being vacant). The value firm gets from producing is the sum of the current period profit, which is the expected value of output produced net of the wage paid to the worker, and the discounted value of being in a match with the same worker in the subsequent period, if the match survives.

### 1.2 Workers' Bellman Equations

Workers face a choice problem that is similar to firms, except for the choice of search strategies. Let \( U(\lambda) \) be the value of unemployment to a worker, and \( W(\gamma, \lambda, \lambda) \) be the value of being in a match with a firm to a worker where \( \gamma \) is the probability that the match is good quality and \( \lambda \) is the workers' belief about the firms' choice. If a worker is unemployed, she gets the unemployment income, \( b \), at the current period. With probability \( 1 - h_w \) the worker does not meet with any firms, thus continues to be unemployed in the subsequent period. The worker meets with a firm with
probability $h_w$ and gets an expected value from being in a match with a firm. The equation below states these sequence of events:

$$U(\lambda) = b + \beta(1 - h_w)U(\lambda) + \beta h_w \left[ \lambda \int W(\gamma, \lambda, \lambda)d\Omega \right] + (1 - \lambda) \int W(\gamma, \lambda, \lambda)d\Gamma$$

(3)

Note that the expected value the worker gets from being in a match with the firm depends on the selection strategy the firm has chosen. The formal statement of a worker’s decision problem when she is in a match is as follows:

$$W(\gamma, \lambda, \lambda) = \max \left\{ U(\lambda), w(\gamma, \lambda, \lambda) + \beta \delta U(\lambda) + \beta(1 - \delta) \left[ \pi (\gamma W(1, \lambda, \lambda)) + (1 - \gamma)W(0, \lambda, \lambda) \right] + (1 - \pi) W(\gamma, \lambda, \lambda) \right\}.$$  

(4)

The worker needs to decide between being unemployed and being in a match which is of a good quality with $\gamma$ probability. If the workers chooses to be in the match, she receives a wage and continues to get a value which depends on the quality of the match in the subsequent period and whether the match survives the risk of destruction.

### 1.3 Wage Determination and Flows Across Employment States

Wages are determined according to the Nash bargaining rule, where workers’ bargaining power is $\mu$. The wage rate that solves the bargaining problem is such that a worker gets a constant ($\mu$) fraction of the net value generated by the worker-firm union. The wage is a weighted average of the expected output the match can produce and the worker’s outside option, which is the value of unemployment.\[3\]

After some 

\[\tilde{W}(\gamma) = w(\gamma, \lambda, \lambda) + \beta \delta U(\lambda) + \beta(1 - \delta) \left[ \pi (\gamma W(1, \lambda, \lambda) + (1 - \gamma)W(0, \lambda, \lambda)) + (1 - \pi) W(\gamma, \lambda, \lambda) \right] \]

\[\tilde{J}(\gamma) = E(y|\gamma) - w(\gamma, \lambda, \lambda) + \beta \delta V(\lambda, \lambda) + \beta(1 - \delta) \left[ \pi (\gamma J(1, \lambda, \lambda) + (1 - \gamma)J(0, \lambda, \lambda)) + (1 - \pi) J(\gamma, \lambda, \lambda) \right] \]

Then, the wage is such that

$$\tilde{W}(\gamma) - U(\lambda) = \mu \left( \tilde{W}(\gamma) - U(\lambda) + \tilde{J}(\gamma) - V \right)$$
algebra, one can show that

\[ w(\gamma, \lambda, \bar{\lambda}) = \mu \left[ y^h \gamma + y^l (1 - \gamma) \right] + (1 - \mu)(1 - \beta)U(\bar{\lambda}) \quad (5) \]

The Nash bargaining assumption guarantees the unanimity of the separation or match formation decision. That is because parties bargain over the net surplus of match surplus, and if the surplus is positive (negative) they decide to form the match (separate). Hence there is no inconsistency across parties in decision making.

Note that at any period there will be two types of employment relationships; those that are known to be good quality matches and those are with unknown quality. Let \( e^g_t \) and \( e^n_t \) denote the number of workers with good and unknown quality matches, respectively. Moreover, let

\[ E(\gamma | \gamma^*) = \frac{\int_{\gamma^*} \gamma d\Omega}{\int_{\gamma^*} d\Omega} + (1 - \lambda) \frac{\int_{\gamma^*} \gamma d\Gamma}{\int_{\gamma^*} d\Gamma} \quad (6) \]

be the conditional expected value of \( \gamma \).

The number of workers who have known (good) quality matches is the sum of the current period good quality matches that survive and the current period unknown quality matches that survive and reveal to be good quality.

\[ e^g_{t+1} = (1 - \delta)e^g_t + e^n_t (1 - \delta) \pi E(\gamma | \gamma^*) \quad (7) \]

The number of unknown quality matches of the subsequent period are the current period unknown quality matches that survive and stay unknown quality and the newly formed matches.

\[ e^n_{t+1} = (1 - \delta)e^n_t (1 - \pi) + f(\theta)u_t \quad (8) \]

The unemployment evolves according to the following equation:

\[ u_{t+1} = (1 - f(\theta))u_t + e^n_t (\delta + (1 - \delta) \pi (1 - E(\gamma | \gamma^*))) + e^g_t \delta \quad (9) \]

Substituting \( \tilde{W}(\gamma) \) and \( \tilde{J}(\gamma) \) into one of these last two equations and solving for \( w(\gamma, \lambda, \bar{\lambda}) \) gives the wage equation.
1.4 Equilibrium

The steady state equilibrium, for each sector, is a list \{v, u, \lambda^*, \lambda, w(\gamma), J(\gamma), V, W(\gamma), U, h_w, h_f\} such that

- \{J(\gamma), V, W(\gamma), U\} satisfy equations (3), (4), (1), and (2).

- \(V = 0\).

- \(w(\gamma, \lambda)\) is the solution to the Nash bargaining.

- \(\gamma^*\) solves
  \[
  U(\lambda) = w(\gamma^*, \lambda, \lambda) + \beta \delta U(\lambda) + \beta (1 - \delta) \left[ \pi \gamma W(1, \lambda, \lambda) + (1 - \pi) W(\gamma^*, \lambda, \lambda) \right]
  \]
  \[
  V(\lambda, \lambda) = E(y|\gamma^*) - w(\gamma^*, \lambda) + \beta \delta V(\lambda, \lambda) + \beta (1 - \delta) \left[ \pi \gamma J(1, \lambda, \lambda) + (1 - \pi) J(\gamma^*, \lambda, \lambda) \right]
  \]

- employment stocks are such that
  \[
  \delta e^g = e^n (1 - \delta) \pi E(\gamma|\gamma^*)
  \]
  \[
  e^n = (1 - \delta) e^n (1 - \pi) + f(\theta) u
  \]
  \[
  u = (1 - f(\theta)) u + e^n (\delta + (1 - \delta) \pi (1 - E(\gamma|\gamma^*)) + e^g \delta
  \]

- \(\lambda^*\) is selected optimally, and firms do not want to deviate from it (i.e. \(\lambda = \lambda^*\)).

1.4.1 Optimal Threshold Value (Hiring Standards)

It is more convenient to work with surplus equations, total net value of a match, rather than value of a match to the firm and the worker. Let \(S(\gamma, \lambda, \lambda)\) be the net value a match generates.

\[
S(\gamma, \lambda, \lambda) = W(\gamma, \lambda, \lambda) - U(\lambda) + J(\gamma, \lambda, \lambda) - V(\lambda, \lambda)
\]  
(10)

Substituting equations (4) and (2) into equation (10) for \(W(\gamma, \lambda, \lambda)\) and \(J(\gamma, \lambda, \lambda)\), respectively, yields

\[
S(\gamma, \lambda, \lambda) = \max \left\{ 0, \frac{E(y|\gamma) + \beta (1 - \delta) \pi \gamma S(1, \lambda, \lambda) - (1 - \beta) U(\lambda) - (1 - \beta) V(\lambda, \lambda)}{1 - \beta (1 - \delta) (1 - \pi)} \right\}
\]  
(11)
The threshold likelihood $\gamma^*$ for firms and workers form a match is such that $S(\gamma^*, \lambda, \overline{\lambda}) = 0$, which implies

$$
\gamma^*(\lambda, \overline{\lambda}) = \frac{(1 - \beta)U(\overline{\lambda}) + (1 - \beta)V(\lambda, \overline{\lambda}) - y^l}{y^h - y^l + \beta(1 - \delta)\pi S(1, \lambda, \overline{\lambda})}
$$

(12)

Assumptions $y^b = b$ and $y^g h > b$ suffice for the existence of a threshold value\(^4\). First, notice that, how a firm’s investment decision affects its threshold depends on whether that investment decision increases or decrease the value of posting a vacancy. If the investment increases the value, the threshold also goes up.

In equilibrium, given that all the firms choose $\overline{\lambda}$ investment level in selection, the threshold value of a good quality match likelihood goes up with investment. This is because when a firm with vacancy and an unemployed worker are in a meeting, what matters is the worker’s outside option (in addition the parameters of the model) since the firm’s outside option is zero. The worker’s outside option depends on other firms’ search behavior. As other firms increase their investment in selection, since now they face better prospects for matches, they choose to have higher hiring standards (threshold). Moreover, the higher the market tightness, the higher the cutoff. Also note that if the speed of learning about the match quality increases, the cutoff value decreases.

1.4.2 Optimal Distribution Choice, Selection

We can rewrite a firm’s investment choice in selection using surplus functions as follows:

$$
\max_{\lambda \in [0, 1]} \left\{ -C(\lambda) + \beta h f (1 - \mu) \left[ \lambda \int S(\gamma, \lambda, \overline{\lambda}) d\Omega + (1 - \lambda) \int S(\gamma, \lambda, \overline{\lambda}) d\Gamma \right] \right\}
$$

\(^4\)Note that

$$
S(1, \lambda, \overline{\lambda}) = \max \left\{ 0, \frac{y^g - (1 - \beta)U(\overline{\lambda}) - (1 - \beta)V(\lambda, \overline{\lambda})}{1 - \beta(1 - \delta)} \right\},
$$

$$
S(0, \lambda, \overline{\lambda}) = \max \left\{ 0, \frac{y^b - (1 - \beta)U(\overline{\lambda}) - (1 - \beta)V(\lambda, \overline{\lambda})}{1 - \beta(1 - \delta)} \right\}
$$

As long as $y^b < (1 - \beta)U(\overline{\lambda})$ and $y^g > (1 - \beta)U(\overline{\lambda})$, good quality matches are desirable while bad quality ones are undesirable (as $V(\lambda, \overline{\lambda}) = 0$ in equilibrium). Thus, there will be a positive value for $\gamma^*(\lambda, \overline{\lambda})$, such that firms and workers are indifferent between forming a match vs. not forming one. Moreover, since $(1 - \beta)U(\overline{\lambda}) = b + \frac{\theta_\mu (c + C(\overline{\lambda}))}{(1 - \mu)}$, assumptions $y^b = b$ and $y^g > b$ guarantee the existence (for sufficiently large $y^g - b$).
The first order condition for the selection of the technology is:

$$C'(\lambda) = \beta h_f(1 - \mu) \left[ \int_{\gamma(\lambda)} S(\gamma, \lambda, \lambda) d\Omega - \int_{\gamma(\lambda)} S(\gamma, \lambda, \lambda) d\Gamma \right]$$

(13)

Everything else being the same, firms would choose a higher investment in selection for a higher value of good quality output. Hence, ex-ante, high-skill firms are more likely to do more selection. However, notice that, the choice of $\lambda$ also depends on the threshold value and the market tightness, in addition to the speed of learning.

How does the hiring choice, $\lambda$, change with the speed of learning, $\pi$? When the speed of learning increases, the returns to a match at every level of $\gamma$ increases, but less so for higher values of gamma;

$$\frac{\partial S(\gamma, \lambda)}{\partial \pi} = \frac{\beta(1 - \delta)((1 - \beta)U(\lambda) - y')}{(1 - \beta(1 - \delta)(1 - \pi)^2)} (1 - \gamma).$$

(14)

Hence, faster learning is more valuable for matches with higher likelihood of being of a bad quality. This relationship implies that how the speed of learning affects the investment in hiring depends on the threshold likelihood value.$^5$ More specifically, firms compare the unconditional probability of getting a bad quality match across distributions. The do so since bad-quality match likelihood determines the gain from an increase in speed of learning. For low levels of threshold the basic distribution is more likely to deliver a bad quality match. Hence, an increase in learning speed makes returns from using the basic distribution go up more, compared to returns from the better distribution. This makes the difference between using the better distribution vs. the basic distribution relatively less important. Hence, firms reduce the amount they invest in the better distribution. On the other hand, when the threshold is sufficiently high, the unconditional probability of getting a match is low for any distribution, but even more so for the basic one. That happens since for high threshold values, the basic distribution is less likely to deliver an acceptable match. In this case an increase in the speed of learning increases the returns to using the basic distribution less than the returns for the better distribution. That makes firms

$^5$The derivative of the right hand side of equation (14) is:

$$\beta h_f(1 - \mu)\frac{\beta(1 - \delta)[ - y' + (1 - \beta)U(\lambda)]}{(1 - \beta(1 - \delta)(1 - \pi))^2} \left[ \int_{\gamma(\lambda)} (1 - \gamma)d\Omega - \int_{\gamma(\lambda)} (1 - \gamma)d\Gamma \right]$$

12
choose to invest more in $\Omega$. That analysis suggests that, high-skilled firms, who have slower learning speed and higher output, are more likely to choose higher investment. That does not guarantee an equilibrium in which they actually choose a higher $\lambda$ than the low-skilled firms.

One point worth mentioning before closing this section is the effect of other firms’ selection decisions on a firm’s choice. For equilibrium, we need, $\lambda^* = \bar{\lambda}$, i.e., $V(\bar{\lambda}, \bar{\lambda}) \geq V(\tilde{\lambda}, \bar{\lambda})$, $\forall \tilde{\lambda} \in [0, 1]$, where

$$V(\bar{\lambda}, \bar{\lambda}) = -c - C(\bar{\lambda}) + \beta V(\bar{\lambda}, \bar{\lambda}) + \beta h_f(1 - \mu)[\int_{\gamma^*} S(\gamma, \bar{\lambda}, \bar{\lambda})d\Omega + (1 - \bar{\lambda}) \int_{\gamma^*} S(\gamma, \bar{\lambda}, \bar{\lambda})d\Gamma]$$

and

$$V(\tilde{\lambda}, \bar{\lambda}) = -c - C(\tilde{\lambda}) + \beta V(\bar{\lambda}, \bar{\lambda}) + \beta h_f(1 - \mu)[\int_{\gamma^*} S(\gamma, \bar{\lambda}, \bar{\lambda})d\Omega + (1 - \bar{\lambda}) \int_{\gamma^*} S(\gamma, \bar{\lambda}, \bar{\lambda})d\Gamma]$$

That implies, a firm will not deviate if

$$\beta h_f(1 - \mu)(\bar{\lambda} - \tilde{\lambda}) \left\{ \int_{\gamma^*} S(\gamma, \bar{\lambda}, \bar{\lambda})d\Omega - \int_{\gamma^*} S(\gamma, \bar{\lambda}, \bar{\lambda})d\Gamma \right\} \geq C(\bar{\lambda}) - C(\tilde{\lambda}) \quad (15)$$

2 Implications for Labor Market Outcomes

2.1 Wages

Workers will start to work for a wage that depends on the likelihood of the match being of good quality. As production takes place, if parties learn the true match quality the worker either looses her job or gets a wage raise. Note that workers will not get any further raise once they learn to be in a good quality match. Recall that wage equation (equation (5)):

$$w(\gamma, \bar{\lambda}) = \mu(\gamma y^h + (1 - \gamma)y^f) + (1 - \mu)(1 - \beta)U(\bar{\lambda})$$

Note that there is no direct effect of ex-post learning speed on a wage at a particular period. However, it will affect the distribution of wages. This is because as ex-post learning increases, so does the fraction of matches that are of a good quality. Selection, i.e., $\bar{\lambda}$, also affects the distribution of wages, as well as the levels.

An increase in $\bar{\lambda}$ will increase the average wage rate as both workers’ outside option and the conditional expected likelihood of good quality matches increase
with $\lambda$. The effect on variance of wages is ambiguous because both values of $\gamma$ and the conditional mean goes up. Whether the variation will be larger or not depends mainly on the shape of distributions.

### 2.2 Implications for Employment Dynamics

#### 2.2.1 Unemployment

Job finding probability and vacancy filling probability ($f$ and $q$) are

$$f = h_w(\theta) \left( \lambda \int_{\gamma^*(\lambda)} w(\gamma, \lambda) d\Omega + (1 - \lambda) \int_{\gamma^*(\lambda)} \gamma d\Psi \right)$$

$$q = h_f(\theta) \left( \lambda \int_{\gamma^*(\lambda)} w(\gamma, \lambda) d\Omega + (1 - \lambda) \int_{\gamma^*(\lambda)} \gamma d\Psi \right)$$

Note that $\left( \lambda \int_{\gamma^*(\lambda)} \gamma d\Omega + (1 - \lambda) \int_{\gamma^*(\lambda)} \gamma d\Psi \right)$ is decreasing in $\lambda$. Thus, we expect an increase in selection to reduce the probability of filling a vacancy (an increase in vacancy duration). However, since which direction the market tightness goes is ambiguous, the overall effect of ex-post learning on job finding and vacancy filling probabilities is ambiguous.

---

6Let $w_s$ denote the starting wage of a new hire. Then

$$E w_s = \mu (y^h - y^l) E(\gamma|\gamma^*) + \mu y^l + (1 - \mu)(1 - \beta) U(\lambda)$$

$$\sigma_s^2 = \frac{\lambda \int_{\gamma^*(\lambda)} (w(\gamma, \lambda) - Ew)^2 d\Omega}{\int_{\gamma^*(\lambda)} d\Gamma} + \frac{(1 - \lambda) \int_{\gamma^*(\lambda)} (w(\gamma, \lambda) - Ew)^2 d\Gamma}{\int_{\gamma^*(\lambda)} d\Gamma}$$

$$\sigma^2 = (\mu(y^h - y^l))^2 \left[ E(\gamma^2|\gamma^*) - E(\gamma|\gamma^*)^2 \right]$$

The average and the standard deviation of overall wages are:

$$E w = \frac{\delta \mu ((y^h - y^l) E(\gamma|\gamma^*) + y^l)}{\delta + (1 - \delta) \pi E(\gamma|\gamma^*)} + \frac{(1 - \delta) \pi E(\gamma|\gamma^*)\mu y^h}{\delta + (1 - \delta) \pi E(\gamma|\gamma^*)} + (1 - \beta) U(\lambda)$$

$$\sigma^2 = \mu^2 (y^h - y^l)^2 \left[ E(\gamma^2|\gamma^*) - 2 \frac{\delta + (1 - \delta) \pi E(\gamma|\gamma^*)^2}{\delta + (1 - \delta) \pi E(\gamma|\gamma^*)} + \frac{(\delta + (1 - \delta) \pi E(\gamma|\gamma^*))^2}{(\delta + (1 - \delta) \pi E(\gamma|\gamma^*))^2} \right]$$
Total separation rate $s$ is

\[
s = \delta \frac{\delta + (1 - \delta)\pi}{\delta + (1 - \delta)\pi E(\gamma | \gamma^*)}
\]

One can show that increase in speed of learning increases job separation probability. Moreover, increase in investment in search decreases this probability.

Unemployment rate $u$ will be

\[
u = \frac{\delta}{f + \delta} \frac{\delta + (1 - \delta)\pi}{\delta + (1 - \delta)\pi E(\gamma | \gamma^*)} = \frac{s}{f + s}.
\]

Since the change in $f$ is ambiguous, how unemployment responses to change in $\pi$ is also ambiguous.

### 2.2.2 Tenure

A higher ex-post learning will affect tenure directly as most of the bad-quality matches are sorted out at the earlier periods of employment. Thus one would expect the empirical hazard rate to be higher at the beginning, then decline faster. It would also affect tenure through changing the $\bar{X}$ and the threshold value. A lower threshold value means a higher number of bad quality matches in expectation, thus increases separations at all tenure levels for a given $\pi$. A higher investment in selection reduces the number of bad quality matches, resulting in a longer expected tenure.

An employed worker is identified by her tenure and the beliefs about the match quality. Let $n^u(\tau, \gamma)$ be the number of workers with $\tau$ periods of tenure and with $\gamma$ probability of good quality match in period $t$ where $\gamma < 1$. Let $n^k(\tau)$ be the number of known (good) quality matches with tenure $\tau$. Moreover, let $X(\gamma) \in \{0, 1\}$ reflect the decision whether to form the match.

\[
n^u_t(\tau, \gamma) = \left\{
\begin{array}{ll}
u_{t-1} h_u(\theta_t) \left( X X(\gamma) d\Omega + (1 - X) X(\gamma) d\Gamma \right) & \tau = 1 \\
(1 - \delta)(1 - \pi) n^u_{t-1}(\tau - 1, \gamma) & \tau > 1 
\end{array}
\right.
\]

\[
n^k_t(\tau, \gamma) = (1 - \delta) n^k_{t-1}(\tau - 1) + (1 - \delta) \int_\gamma \gamma n^u_{t-1}(\tau - 1, \gamma) d\gamma
\]

To see the effects on tenure, let us look at evolution of employment. Let $\Delta(t, \gamma)$ be the fraction of $\gamma$ quality matches among all the matches with $t$ periods of tenure.
After some algebra, one can show that, ∀τ ≥ 1
\[ n(τ + 1, γ) = (1 − δ)(1 − π)Δ(τ, γ)n(τ), \]
\[ n(τ + 1, 1) = (1 − δ)n(τ)\left(1 − \int_γ Δ(τ, γ)dγ + π \int_γ γΔ(τ, γ)dγ\right). \]
\[ n(τ + 1) = (1 − δ)n(τ)\left(1 − \int_γ Δ(τ, γ)dγ + π \int_γ γΔ(τ, γ)dγ + (1 − π) \int_γ Δ(τ, γ)dγ\right). \]
Thus, the expression for ∆(τ, γ) becomes
\[ ∆(τ + 1, γ) = (1 − π)Δ(τ, γ) \]
\[ \frac{1}{1 − π \int_γ Δ(τ, γ)dγ + π \int_γ γΔ(τ, γ)dγ}. \]

By iterations, one can show that
\[ \Delta(τ+1, γ) = \frac{(1 − π)\Delta(τ, γ)}{(1 − π)^{τ}( \bar{X}dΩ(γ) + (1 − \bar{X})dΓ(γ) + \pi \sum_{j=0}^{τ-1} (1 − π)^j (\bar{X}f_{γ^*} γdΩ + (1 − \bar{X})f_{γ^*} γdΓ)} \]

Observe that one can write the probability of separating from a job at period τ + 1, given that the employment has survived τ periods as follows\(^7\)
\[ n(τ + 1) = (1 − δ)n(τ)\left(1 − π \int_γ Δ(τ, γ)dγ + π \int_γ γΔ(τ, γ)dγ\right). \]
\[ h(τ + 1|τ) = 1 − \frac{n(τ + 1)}{n(τ)} = 1 − (1 − δ)\left(1 − π \int_γ Δ(τ, γ)dγ + π \int_γ γΔ(τ, γ)dγ\right), \]

### 3 A Quantitative Analysis

As the discussion above demonstrates, although we can explore the direct effects of change in speed of learning on labor market outcomes, the total effects are mostly
ambiguous. To explore the general equilibrium effects and equilibrium outcomes, I carry on the following numerical exercises.

Table 1: Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.996</td>
<td>4% discount factor</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.5</td>
<td>workers’ bargaining power</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1</td>
<td>Matching function parameter</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.006</td>
<td>Exogenous fraction of job destructions</td>
</tr>
<tr>
<td>$c$</td>
<td>1.8</td>
<td>Vacancy creation cost</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>2.15</td>
<td>Search cost parameter</td>
</tr>
<tr>
<td>$\beta_b$</td>
<td>3.62</td>
<td>Parameter value of basic distribution</td>
</tr>
<tr>
<td>$\beta_v$</td>
<td>1.62</td>
<td>Parameter value of better distribution</td>
</tr>
<tr>
<td>$b_{ls}$</td>
<td>1</td>
<td>unemployment income, low-skilled</td>
</tr>
<tr>
<td>$y_{ls}$</td>
<td>$b_{ls}$</td>
<td>bad-quality output, low-skilled</td>
</tr>
<tr>
<td>$y_{ls}$</td>
<td>2.5</td>
<td>good-quality output, low-skilled</td>
</tr>
<tr>
<td>$\tau_{ls}$</td>
<td>1/6</td>
<td>low-skilled learning speed</td>
</tr>
<tr>
<td>$\tau_{hs}$</td>
<td>1/18</td>
<td>high-skilled learning speed</td>
</tr>
<tr>
<td>$b_{hs}$</td>
<td>1</td>
<td>unemployment income, high-skilled</td>
</tr>
<tr>
<td>$y_{hs}$</td>
<td>$b_{hs}$</td>
<td>bad-quality output, high-skilled</td>
</tr>
<tr>
<td>$y_{hs}$</td>
<td>2.5</td>
<td>good-quality output, high-skilled</td>
</tr>
</tbody>
</table>

Let me start with assigning values to parameters that are common across skill. Model period is a month. I choose exogenous destruction rate $\delta = 0.006$, discount factor $\beta = 0.9967$, and bargaining parameter $\mu = 0.5$. I use the following functional form for the matching function: $M = \frac{uv}{(u^{\alpha} + v^{\alpha})^{\frac{1}{\alpha}}}$ where $\alpha = 1$. I normalize $b_{ls}$ for low-skilled workers to 1. I assume the distributions from which the good-quality match likelihood comes are Beta distributions with the first parameter normalized to 1. I assume the second parameters are 1.6 and 3.6 for the better and the basic distributions, respectively. Moreover I assume $C(\lambda) = \kappa \lambda^2$. I assume that the speed of learning for low-skilled and high-skilled jobs are such that, on average, the match quality is revealed within 6 and 18 periods of employment for low-skilled and high-skilled, respectively. The values of all parameters are reported on Table 1.

When I run the model for high-skilled and low-skilled workers, I get the following results: High-skilled workers have lower unemployment rates, vacancy filling and job finding probabilities, as well as lower job separation probabilities. Table 2 shows
Table 2: Results

<table>
<thead>
<tr>
<th></th>
<th>Low-skilled</th>
<th>High-Skilled</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda )</td>
<td>0.284</td>
<td>0.663</td>
</tr>
<tr>
<td>unemployment rate*</td>
<td>0.0627</td>
<td>0.045</td>
</tr>
<tr>
<td>job finding probability</td>
<td>0.226</td>
<td>0.208</td>
</tr>
<tr>
<td>job separation probability</td>
<td>0.015</td>
<td>0.0098</td>
</tr>
<tr>
<td>vacancy filling probability</td>
<td>0.388</td>
<td>0.179</td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.583</td>
<td>1.16</td>
</tr>
<tr>
<td>( \gamma^* )</td>
<td>0.151</td>
<td>0.362</td>
</tr>
<tr>
<td>( \frac{1}{1-u} )</td>
<td>0.088</td>
<td>0.16</td>
</tr>
<tr>
<td>average output</td>
<td>2.365</td>
<td>4.407</td>
</tr>
<tr>
<td>average wage</td>
<td>2.284</td>
<td>4.45</td>
</tr>
<tr>
<td>std. dev. of wage</td>
<td>0.449</td>
<td>0.787</td>
</tr>
<tr>
<td>average starting wage</td>
<td>1.856</td>
<td>3.73</td>
</tr>
<tr>
<td>std. dev. of starting wage</td>
<td>0.136</td>
<td>0.31</td>
</tr>
</tbody>
</table>

these values for each skill type. Moreover, Figure 3 shows how the likelihood of losing one’s job changes over tenure for each skill group. These differences are in line with the data.

Note that there are both productivity differences and ex-post learning difference across skill. To see their relative contributions to outcomes, I first run the same model where there is no productivity differences, and then I look at what would have happened if the only difference across skill was in productivity. Results are displayed on Table 3.

First two columns of Table 3 are labor market outcomes for low- and high-skilled workers, respectively, when high-skilled workers are more productive and have slower ex-post learning. The third column is the labor market outcomes for high-skilled workers when the only thing that separates them from their low-skilled counterparts is the difference in learning speed. If we compare the first and the third columns, we see the effects of ex-post learning. Note that it makes firms invest more in search. It also increases threshold for hiring and reduces the relative number of vacancies. As a results, job finding and job separation probabilities are lower. But unemployment rate is higher, which contradicts the data.

If, on the other hand, there was not a difference in the speed of learning, but the only difference was the productivity gap, high-skilled firms would still invest in search
Job Hazard Rates

Job Hazard Rates, Counterfactual
more than the low skilled firms. They would also require higher minimum standards (threshold) to hire. Note that, unemployment rate and separation probabilities would have been lower. However, job finding probability would have been much higher, which contradicts the data.

As figure 3 shows, differences in productivity and speed of learning have different implications for tenure profiles. If the only difference across skill was the productivity gap, then their empirical likelihoods of losing a job, conditional on tenure, would be changing at a similar fashion over time. So, productivity just shifts that hazard down without affecting the slope much. On the other hand, if the only difference across skill were learning, the slope of the empirical hazard would be more flat. Hence one can conclude that, the ex-post learning mostly affects the slope of hazard function while the productivity affects the level.

### 4 Concluding Remarks

Could differences in the speed of ex-post learning across skill explain the differences in employers’ search strategies? We should expect that the speed of learning depends
on the skill requirements of a job. Due to the complexity of tasks at higher-skilled jobs, it would be harder for firms to recognize whether the employee is a good-fit. As a result, we expect the speed of ex-post learning to be slower for high-skilled jobs. In return, one anticipates that firms would invest more in search of a worker for jobs with higher skill requirements.

This paper shows that, slower ex-post learning affects firms’ ex-ante investment decision in employee selection. I use a labor search-matching model, which builds on the model used by Pries and Rogerson (2005), with two types of workers: high-skilled and low-skilled workers. For each skill type, an employment relationship between a firm and a worker can be of either a good or a bad quality. Good-quality matches are expected to produce higher output than bad-quality matches within a skill group. I assume that the bad-quality matches are not profitable, thus firms and workers do not want to be in a bad-quality match, regardless of the skill type.

In this environment, high-skilled firms, the ones that have higher productivity and slower speed of learning, are more likely to invest more in ex-ante hiring. The numerical exercises show that the differences in ex-post learning not only explains the differences in employers’ selection strategies across skill, but also inequalities in these skill groups’ labor market outcomes.

References


