Learning Through Referrals

Manolis Galenianos*
Pennsylvania State University

February 2012

Abstract

This paper theoretically examines the firm’s choice to use different search channels in order to hire new workers. An equilibrium model is developed where the quality of a match is uncertain and firms search for workers through the market and through referrals. The intensity of use of each search channel is endogenized through the choice of channel-specific search effort. When referrals generate more accurate signals regarding match quality, the model predicts that referred workers have higher starting wages, higher productivity and lower separation rates than non-referred candidates and that these differentials decrease over time due to selection, which is consistent with the data. The model is extended by introducing productivity heterogeneity in firms and allowing the endogenous determination of signal quality. It is shown that high productivity firms choose greater accuracy of signals which diminishes the referral-market differential and leads to lower referral intensity, consistent with the data. This type of selection on the firm side explains why regressions that do not include firm fixed effects find a negative effect of referrals on wages in contrast to firm-level and other evidence.

*I would like to thank the National Science Foundation for financial support (grant SES-0922215).
1 Introduction

It is well-known that social networks are an important feature of labor markets (Granovetter, 1995). Most work that addresses the interaction of social networks with labor markets focuses on the effect of referrals on workers’ labor market outcomes. However, it seems desirable to also study the firms’ decision to use referrals as a method of finding workers, especially since there are strong regularities in the intensity of referral usage by different types of firm. Understanding what type of frictions are avoided by using referrals is a first step towards understanding the differentials in referral use across industries and, ultimately, what are the determinants of matching frictions across industries.\(^1\)

A number of papers have documented selection in the intensity of referral use by firm characteristics. The most complete empirical analysis is in Dustmann, Glitz and Schoenberg (2011, henceforth DGS) who, among other findings, show that high productivity firms use referrals to a lesser extent that low productivity firms. Additionally, and relatedly, Pellizzari (2010) and Holzer (1987) find that larger firms use referrals to a lower extent.\(^2\)

This paper presents a model that qualitatively accounts for the following facts about the use of referrals in the labor market. First, referred workers have lower separation rates than non-referred workers (Loury 1983; Simon and Warner 1992; Brown, Selten and Topa 2011; DGS); second, conditional on the firm, referred workers receive higher wages and have higher productivity (for productivity see Castilla 2005 and Pinkston 2011; for wages see Bayer, Ross and Topa 2008, Brown, Selten and Topa 2011 and DGS); third, an applicant is more likely to be hired conditional on his observable characteristics if he applies through a referral (Fernandez and Weinberg 1997; Castilla 2005; Brown, Setren and Topa 2011); fourth, the differentials in separation rates, wages and productivity decline with the workers’ tenure as documented in Castilla (2005), Brown, Setren and Topa (2011) and DGS.

\(^1\)See Galenianos (2011) for evidence on the differential in referral use across industries. See Davis, Faberman and Haltiwanger (2011) for evidence that matching frictions differ across industries.

\(^2\)DGS also develop a theoretical model which is closely related to the present model and is discussed below.
Fifth, referrals are used disproportionately by the less productive firms. Fact 5 is also related with the observation that the effect of finding a job through a referral on one’s wage is zero or negative when firm fixed effects are not controlled for (Pistaferri 1999; Pellizzari 2010; Bentolila, Michellaci and Suarez 2010) which directly contradicts the second fact and seems inconsistent with the first and third facts.

An equilibrium model with multiple hiring channels is developed to interpret these facts. The labor market is characterized by search frictions and by uncertain match quality. A firm and a worker meet through the market or through a referral. At the time of the meeting, the quality of the match is uncertain and the firm and worker observe a signal before deciding whether to consummate the match. The accuracy of the signal depends on the channel through which the firm and worker met. Match quality is revealed over the course of the employment relationship. The firm exerts effort in searching through each of the two channels. In the baseline model, firms are homogeneous.

Assuming that the signal accuracy is higher when a firm and a worker meet through a referral yields predictions that are consistent with facts (1)-(4). In equilibrium, a referred worker is more likely to be hired and, conditional on being hired, the posterior probability that the match is good is greater than that of a non-referred employee of equal tenure. This leads to higher wages for the referred worker and also greater probability of survival after match quality is revealed. Over time, as learning about match quality occurs, the posteriors of surviving workers converge and the differential in wages and separation rates declines.

The model is then extended in two ways: firms are heterogeneous in productivity and choose the accuracy of their signals subject to a cost. The equilibrium of the extended model is characterized and it is shown that high productivity firms choose higher levels of signal accuracy which diminishes the advantage of referrals. As a result, high productivity firms exert relatively more effort in hiring through the market than the low productivity firms and use referral at a lower rate. This prediction is consistent with fact (5) from the data.

---

3See Galenianos (2011) for a model where heterogeneity is introduced on the worker side.
Furthermore, if firm size is positively correlated with productivity, then it is also consistent with the finding that larger firms use referrals to a lesser extent than smaller firms.

DGS develop a theoretical model that has many similarities with the current paper: match quality is uncertain and referrals lead to more accurate signals regarding match quality. The present paper extends the analysis of DGS in three ways: first the equilibrium is analytically characterized while DGS resort to simulations; second, it allows for firm heterogeneity; third, it introduces the additional choices of search effort by channel of search and endogenous choice of signal accuracy. It is shown that the additional choices interact in interesting and non-trivial ways with the underlying heterogeneity. In particular they help interpret why high productivity firms choose lower levels of referral intensity despite the fact that referred workers appear preferable than non-referred ones.

2 The Labor Market With Homogeneous Firms

This section describes and characterizes the baseline model and its main predictions.

2.1 The Model

Time runs continuously, the horizon is infinite and the labor market is in steady state. There is a measure $n$ of workers who are ex ante homogeneous, risk-neutral, maximize expected discounted utility and discount the future at rate $r$. A worker is either employed or unemployed. The flow utility of unemployment is $z$ and the flow utility of employment is equal to the wage.

There is an endogenous measure of firms which is determined through free entry. Each firm hires one worker, is risk-neutral, maximizes expected discounted profits and discounts the future at rate $r$. A firm is either filled and producing or vacant and searching. The flow profit when vacant is $-K$ and the flow profit when producing is equal to output minus the wage.
The output of a match is given by $y\bar{x}$ where $y$ is the firm’s productivity and $\bar{x}$ is the expected match quality. Match quality can take one of two values, “good” or “bad”: $x \in\{x_B, x_G\}$ with $x_G > x_B$. Therefore, when the probability that a match is good is $p$, the expected match quality is given by $\bar{x} = px_G + (1 - p)x_B$. For simplicity, it is assumed that $x_G = 1$ and $x_B = 0$. We will assume:

**Assumption 1**: $y > 2z$.

Match quality is determined when a worker and a firm first meet and it is good with probability $\gamma \in (0,1)$. At the time of the meeting, the worker and firm do not observe the realization of match quality but they observe an informative signal which is described in detail below. The worker and the firm are symmetrically informed about the probability that their match is good.

During employment, match quality is perfectly revealed at rate $\lambda$. It is clear that after learning match quality, the match persists if it is revealed to be good and it is terminated if it revealed to be bad. See Pries (2004) for a micro-foundation of this specification.

Matches are destroyed exogenously at rate $\delta$ and, possibly, endogenously after learning about match quality occurs. There is no on the job search. The surplus is split through Nash bargaining where the workers’ bargaining power is $\beta$.

The following assumption will be maintained:

**Assumption 2**: $\gamma \geq \max\left\{ \frac{\delta}{2\delta + \lambda}, \frac{r + \delta}{2(\delta + \lambda)} \right\}$.

A firm and a worker meet through two different channels: the market, $M$, or a referral, $R$. The two channels differ in the information that is transmitted about match quality. When a worker and a firm meet through channel $i \in\{M, R\}$, a binary signal $s_i \in\{g_i, b_i\}$ is generated regarding match quality. The signal $s_i$ is correct with probability $q_{Gi}$ if the match is good and $q_{Bi}$ if it is bad, which will be referred to as the signal’s accuracy. Without loss of generality, it is assumed that $q_{Gi} \geq 1/2$ and $q_{Bi} \geq 1/2$. 


When a firm and a worker meet they observe a signal and decide whether to match. Let $d_i(s_i) \in \{0, 1\}$ denote the decision rule of whether to match ($d_i(s_i) = 1$) or not ($d_i(s_i) = 0$). Attention will be restricted to equilibria where the match is never formed after a bad signal: $d_M(b_M) = d_R(b_R) = 0$.

Denote the unconditional probability that the worker generates a good signal by $\pi_i$ and the posterior probability that the match is good conditional on a good signal by $p_i$.\(^4\) Therefore,

\[
\begin{align*}
\pi_i &= \gamma q_{Gi} + (1 - \gamma)(1 - q_{Bi}) \\
p_i &= \frac{\gamma q_{Gi}}{\pi_i}
\end{align*}
\]

The aggregate flow of meetings between unemployed workers and vacancies through the two channels is given by a standard Cobb-Douglas function:

\[
m(v, u) = \mu v^n u^{1-\eta}
\]

where $u$ denotes the number of unemployed workers and $v$ denotes the number of vacancies.

This flow of meeting is distributed through the market or referral channels depending on the effort exerted by vacancies in searching through each channel. Let $E_M$ and $E_R$ denote the aggregate effort exerted in searching through the market and referrals, respectively. The flow of meetings through the two channels are given by:

\[
\begin{align*}
m_M(v, u, E_M, E_R) &= \frac{E_M}{E_M + E_R} m(v, u) \\
m_R(v, u, E_M, E_R) &= \frac{E_R}{E_M + E_R} m(v, u)
\end{align*}
\]

A vacancy chooses its level of search effort for each of the two channels $e_M$ and $e_R$. The cost of search effort is given by $C_M(e_M) + C_R(e_R)$ where $C_i(e) = c_i e^2/2$. In principle, the

\(^4\)Since we focus on the case where a hire only occurs after a good signal, there is no need to condition on the signal.
marginal cost of exerting additional effort need not be equal across the two channels.

The rate at which vacancy $j$ meets a worker through channel $i$ is:

$$\alpha^j_{Fi} = \frac{e_i}{E_i E_M + E_R} \frac{m(v,u)}{v}$$

Attention is restricted to symmetric equilibria where $e_i = E_i$ for all $j$ and $i$.

A worker meets a vacancy through channel $i$ at rate:

$$\alpha_{Wi} = \frac{E_i}{E_M + E_R} \frac{m(v,u)}{u}$$

Under this specification, search effort is a “rat-race” among firms in that it does not affect the total number of meetings. However, it does affect the proportion of meetings that occurs through the two channels which has real effects because of the accuracy differential between the signals.

The labor market is in steady state. A worker can be in one of four states: unemployed, employed at a match whose quality is uncertain and was created through the market or through a referral and employed at a certain, good match (recall that a match is endogenously dissolved once it is certain that match quality is bad). The measure of workers at each state is denoted by $u, n_M, n_R, n_G$ respectively, where $u + n_M + n_R + n_G = n$. The steady state is described by the following conditions which equate the flows in and out of each employment state:

$$u \alpha_{Wm} \pi_M = (\delta + \lambda)n_M \quad (1)$$

$$u \alpha_{Wr} \pi_R = (\delta + \lambda)n_R \quad (2)$$

$$\lambda(n_M p_M + n_R p_R) = \delta n_G \quad (3)$$

The value functions can now be defined. A worker-firm pair can be in one of three states.
They can be in a certain match which in equilibrium is a good match and is denoted by a subscript $G$. They can be in an uncertain match which is denoted by a subscript $M$ or $R$ depending on whether it was created through the market or referrals, respectively. The channel through which a match is created affects the probability that the match will turn out to be good.

The value functions of the firm and the worker in a good match are given by:

$$r J_G = y - w_G + \delta(V - J_G)$$
$$r W_G = w_G + \delta(U - W_G)$$

The value functions of the firm and the worker in state $i \in \{M, R\}$ are given by:

$$r J_i = yp_i w_i + \lambda p_i (J_G - J_i) + (\delta + \lambda(1 - p_i))(V - J_i)$$
$$r W_i = w_i + \lambda p_i (W_G - W_i) + (\delta + \lambda(1 - p_i))(U - W_i)$$

The value of an unemployed worker is:

$$r U = z + \alpha_{WM} \pi_M \max[W_M - U, 0] + \alpha_{WR} \pi_R \max[W_R - U, 0]$$

Recall that we focus on equilibria where a worker is never hired after a bad signal.

The value of a vacancy $j$ that chooses effort levels $e_M$ and $e_R$ is

$$r \tilde{V}^j(e_M, e_R) = -K + \alpha_{FM}^j \pi_M \max[J_M - V, 0] + \alpha_{FR}^j \pi_R \max[J_R - V, 0] - \frac{c_M e_M^2 + c_R e_R^2}{2}$$

and the value given an optimal choice of effort is:

$$V = \max_{e_M, e_R} \tilde{V}^j(e_M, e_R)$$
Finally, wages are determined by Nash bargaining:

\[ w_k = \arg\max_w (W_k - U)^\beta (J_k - V)^{1-\beta} \]  \hspace{1cm} (7)

where \( k \in \{G, M, R\} \).

The Equilibrium can now be defined.

**Definition 2.1** An Equilibrium is the steady state level of unemployment \( u \), the number of vacancies \( v \), the matching decision rules \( \{d_M, d_R\} \) and the effort levels \( E_M \) and \( E_R \) such that:

- The labor market is in steady state as described in equations (1), (2) and (3).
- The surplus is split according to (7).
- The choice of effort maximizes (6).
- Forming a match after a bad signal is suboptimal.
- There is free entry of firms: \( V = 0 \).

### 2.2 Equilibrium

This section proves the following result.

**Proposition 2.1** An equilibrium exists.

To prove the existence of equilibrium, we will describe the equilibrium as a function of the value of unemployment.

First, we use firms’ optimal search effort to uniquely express the value of unemployment as a function of the number of vacancies, the number of unemployed workers and the match surplus.
Second, we use the steady state condition for the flows to solve for the number of unemployed workers and express the value of unemployment as a function of the number of vacancies and the match surplus.

Third, we use the value functions to express the value of unemployment as a function of the number of vacancies alone.

Finally, we show that there is a number of vacancies that satisfies free entry condition.

We begin by simplifying notation. It will prove useful to define:

\[
\alpha_W = \frac{m(v, u)}{u}, \\
\alpha_F = \frac{m(v, u)}{v},
\]

so that

\[
\alpha_{Wi} = \frac{\alpha_WE_i}{E_M + E_R}, \\
\alpha_{Fi} = \frac{\alpha_{FE_i}}{E_M + E_R}.
\]

The surplus of a match between a firm and a worker who are at state \(k \in \{M, R, G\}\) is denoted by \(S_k = \max[\tilde{S}_k, 0]\) where \(\tilde{S}_k = W_k - U + J_k - V\). Our assumptions mean that \(S_G > 0\) but for \(i \in \{M, R\}\) it is possible that \(\tilde{S}_i < 0\). In that case, a meeting through channel \(i\) does not lead to a match even after a good signal. Our assumptions on the model’s primitives will rule out the trivial case where \(S_M = S_R = 0\).

When \(S_k > 0\), the solution to the Nash bargaining problem implies:

\[
W_k - U = \beta S_k \\
J_k - V = (1 - \beta)S_k
\]
Denote the expected surplus of a meeting between a worker and a firm through channel \(i \in \{M, R\}\) by

\[
\bar{S}_i = \pi_i S_i
\]

The value functions of the unemployed worker and vacancy \(j\) can be rewritten as:

\[
r_U = z + \alpha_W \beta \left[ \frac{E_M \bar{S}_M}{E_M + E_R} + \frac{E_R \bar{S}_R}{E_M + E_R} \right]
\]

\[
r_{\tilde{V}}^{j}(e_M, e_R) = -K + \alpha_F (1 - \beta) \left[ \frac{e_M \bar{S}_M}{E_M + E_R} + \frac{e_R \bar{S}_R}{E_M + E_R} \right] - \frac{c_M e_M^2 + c_R e_R^2}{2}
\]

Set the derivative of equation (9) with respect to \(e_i\) to zero to arrive at:

\[
c_i e_i = \frac{\alpha_F (1 - \beta) \bar{S}_i}{E_M + E_R}
\]

Evaluating equation (10) at \(e_i = E_i\) for both \(M\) and \(R\) yields:

\[
\frac{E_i}{E_M + E_R} = \frac{\bar{S}_i}{c_i} \frac{\alpha_F (1 - \beta)}{E_M + E_R} = \frac{\bar{S}_i}{c_i} \frac{\alpha_F (1 - \beta)}{E_M + E_R} = \frac{\bar{S}_i}{c_M} + \frac{\bar{S}_i}{c_R}
\]

where \(\bar{S}_M > 0\) and/or \(\bar{S}_R > 0\). If \(\bar{S}_M = \bar{S}_R = 0\), we have \(E_M = E_R = 0\).

The value of unemployment can therefore be expressed as:

\[
r_U = z + \alpha_W \beta \left( \frac{\bar{S}_M^2}{c_M} + \frac{\bar{S}_R^2}{c_R} \right)
\]

which defines \(U = T_1(\bar{S}_M, \bar{S}_R, u, v)\) and completes the first part of the proof.
We now consider the flows. Equations (1), (2) and (3) can be rearranged as follows:

\[ n_i = \frac{u \alpha_i E_{\pi_i}}{(\delta + \lambda)(E_M + E_R)}, \quad i \in \{M, R\} \]

\[ n_G = \frac{u \alpha_W \lambda}{\delta(\delta + \lambda)} \left( \frac{E_M \gamma q_{GM} + E_R \gamma q_{GR}}{E_M + E_R} \right) \]

Therefore

\[ n_M + n_R + n_G = u \alpha_W \left[ \frac{E_M}{E_M + E_R} \left( \frac{\gamma q_{GM}}{\delta} + \frac{(1 - \gamma)(1 - q_{BM})}{\delta + \lambda} \right) + \frac{E_R}{E_M + E_R} \left( \frac{\gamma q_{GR}}{\delta} + \frac{(1 - \gamma)(1 - q_{BR})}{\delta + \lambda} \right) \right] \]

\[ = \frac{u \alpha_W E_{\Gamma_{1M}} + E_{\Gamma_{1R}}}{E_M + E_R} \]

where

\[ \Gamma_{1i} = \frac{\gamma q_{Gi}}{\delta} + \frac{(1 - \gamma)(1 - q_{Bi})}{\delta + \lambda} \]

Using \( n = u + n_M + n_R + n_G \) and the optimal choice of search effort we have:

\[ n = u + \mu v^\eta u^{1-\eta} \frac{\Gamma_{1M} \bar{s}_M + \Gamma_{1R} \bar{s}_R}{\bar{s}_M + \bar{s}_R} \]  \hspace{1cm} (12)

The right-hand side of equation (12) is strictly increasing in \( u \), is equal to zero when \( u = 0 \) and is greater than \( n \) when \( u = n \). Therefore the equality defines the level of unemployment uniquely given the number of vacancies and match surplus: \( u = u(\bar{s}_M, \bar{s}_R, v) \).

Equation (11) defines \( U = T_1(\bar{s}_M, \bar{s}_R, u, v) \) and equation (12) defines \( u = u(\bar{s}_M, \bar{s}_R, v) \). We combine these two equations to end up with \( U = T_1(\bar{s}_M, \bar{s}_R, u(\bar{s}_M, \bar{s}_R, v), v) = T_2(\bar{s}_M, \bar{s}_R, v) \).

Rearrange equation (11) as follows:

\[ rU = z + \frac{\mu v^\eta \beta \left( \frac{s^2_M}{c_M} + \frac{s^2_R}{c_R} \right)}{u^\eta \left( \frac{s_M}{c_M} + \frac{s_R}{c_R} \right)} \Rightarrow u = \left( \frac{\mu v^\eta \beta \left( \frac{s^2_M}{c_M} + \frac{s^2_R}{c_R} \right)}{(rU - z) \left( \frac{s_M}{c_M} + \frac{s_R}{c_R} \right)} \right)^{\frac{1}{\eta}} \]
Introduce this expression into equation (12):

\[
n = \left( \frac{\mu \nu \beta (\frac{S^2_M}{c_M} + \frac{S^2_R}{c_R})}{(rU - z)(\frac{S^2_M}{c_M} + \frac{S^2_R}{c_R})} \right)^{\frac{1}{\beta}} + \left( \frac{\mu \nu \beta (\frac{S^2_M}{c_M} + \frac{S^2_R}{c_R})}{(rU - z)(\frac{S^2_M}{c_M} + \frac{S^2_R}{c_R})} \right)^{\frac{1}{\gamma}} \frac{\mu \nu \beta (\frac{\Gamma_M S_M}{c_M} + \frac{\Gamma_R S_R}{c_R})}{\frac{S^2_M}{c_M} + \frac{S^2_R}{c_R}} \tag{13}\]

The right-hand side of equation (13) approaches infinity as \(U \to z/r\), is strictly decreasing in \(U\) and it approaches zero as \(U \to +\infty\). Therefore, equation (13) uniquely defines \(U = T_2(S_M, S_R, v)\) which completes the second part of the proof.

We now express the surplus levels as a function of \(U\). The surplus from a good match is:

\[
(r + \delta)S_G = y - rU - rV \tag{14}
\]

To get the surplus of a match of type \(i \in \{M, R\}\) we combine equations (4) and (5) with the Nash bargaining solution, the free entry condition \(V = 0\) and equation (14) to get:

\[
\tilde{S}_i = \frac{yp_i(r + \delta + \lambda) - rU(r + \delta + \lambda p_i)}{(r + \delta)(r + \delta + \lambda)} \tag{15}
\]

\[
\Rightarrow \pi_i\tilde{S}_i = y\Gamma_{2i} - rU(\Gamma_{2i} + \Gamma_{3i}) \tag{16}
\]

where

\[
\Gamma_{2i} = \frac{\gamma q Gi}{r + \delta} \quad \Gamma_{3i} = \frac{(1 - \gamma)(1 - q Bi)}{r + \delta + \lambda}
\]

Observe that \(\tilde{S}_i\) might be negative, in which case \(\tilde{S}_i = 0\) and a meeting through channel \(i\) does not lead to a match even after a good signal. The condition for this to occur is

\[
\tilde{S}_i \leq 0 \iff U \geq \bar{U}_i
\]
where

\[ \bar{U}_i = \frac{y \Gamma_{2i}}{r(\Gamma_{2i} + \Gamma_{3i})} \]

In equilibrium \( U \geq z/r \) because an unemployed worker always has access to the flow value of unemployment. At this point we can confirm that the surplus of a new match will be positive for at least one of the two channels. If not, then \( rU = z \). Therefore, we need that:

\[ y\Gamma_{2i} - z(\Gamma_{2i} + \Gamma_{3i}) > 0 \]

\[ \Rightarrow y > z(1 + \frac{r + \delta}{r + \delta + \lambda} \frac{1-\gamma}{\gamma} \frac{1-q_{Bi}}{q_{Gi}}) \]

Assumption 2 means that the term in the parenthesis is less than 2 and therefore Assumption 1 \((y > 2z)\) suffices to prove that \( \bar{S}_i > 0 \) for some \( i \).

We can now define the expected surplus of a meeting as a function of the value of unemployment. Suppose without loss of generality that \( \bar{U}_k \leq \bar{U}_i \). Then:

\[ U \in [\frac{z}{r}, \bar{U}_k) \ \Rightarrow \ \bar{S}_i(U) = y\Gamma_{2l} - rU(\Gamma_{2l} + \Gamma_{3l}) \quad l \in \{i, k\} \] (17)

\[ U \in [\bar{U}_k, \bar{U}_i) \ \Rightarrow \ \bar{S}_i(U) = y\Gamma_{2i} - rU(\Gamma_{2i} + \Gamma_{3i}) \quad \text{and} \quad \bar{S}_k(U) = 0 \] (18)

\[ U \in [\bar{U}_i, \infty) \ \Rightarrow \ \bar{S}_i(U) = \bar{S}_k(U) = 0 \] (19)

We now determine the value of unemployment as function of the number of vacancies alone, \( U = T_3(v) \). Using equation (13) we define

\[ \hat{T}(U, v) = (\mu \beta)^{\frac{1}{n}} \left( \frac{Q_1(U)}{rU - z} \right)^{\frac{1}{n}} + \mu \beta^{\frac{1-n}{n}} \left( \frac{Q_1(U)}{rU - z} \right)^{\frac{1-n}{n}} Q_2(U) - \frac{n}{v} \] (20)

---

\(^5\)A bit of algebra shows that \( \bar{U}_i \geq \bar{U}_k \iff \frac{q_{Gi}}{1-q_{Bi}} \geq \frac{q_{Gk}}{1-q_{Bk}} \).
where

\[ Q_1(U) = \frac{S_M(U)^2}{c_M} + \frac{S_B(U)^2}{c_R} \]

\[ Q_2(U) = \frac{\Gamma_{1M}S_M(U)}{c_M} + \frac{\Gamma_{1R}S_B(U)}{c_R} \]

where \( \bar{S}_i(U) \) is given by equations (17), (18) and (19). Note that \( \hat{T}(U, v) = 0 \) if and only if equation (13) holds.

Note that

\[
\lim_{U \to z=r} \hat{T}(U, v) = +\infty
\]  

(21)

When \( U \in [\bar{U}_k, \bar{U}_i) \) we have simplify equation (20) as follows:

\[ Q_1(U) = \bar{S}_i(U) \]

\[ Q_2(U) = \Gamma_{1i} \]

\[ \Rightarrow \hat{T}(U, v) = (\mu^2 \beta)^{\frac{1}{\gamma}} \left( \frac{\bar{S}_i(U)}{rU - z} \right)^{\frac{1}{\gamma}} + \mu^{\frac{1}{\gamma}} \beta^{\frac{1}{\gamma} - \frac{n}{\gamma}} \left( \frac{\bar{S}_i(U)}{rU - z} \right)^{\frac{1}{\gamma}} \Gamma_{1i} - \frac{n}{v} \]  

(22)

Finally, for \( U \geq \bar{U}_i \) we have

\[ Q_1(U) = 0 \]

\[ \Rightarrow \hat{T}(U, v) = -\frac{n}{v} < 0 \]  

(23)

Equations (21) and (23) together with the continuity of \( \hat{T}(U, v) \) prove that there exists \( T_3(v) \) such that \( \hat{T}(T_3(v), v) = 0 \). Since equation (13) holds whenever \( \hat{T}(U, v) = 0 \), we have defined the equilibrium value of unemployment as a function of the number of vacancies: \( U^*(v) = T_3(v) \).
One equilibrium feature is that $\hat{T}(U, v)$ depends only on the ratio of the labor force over the number of vacancies. Therefore the size of the labor force ($n$) can be scaled without affecting $U^*(v)$.

A second feature is that if $\hat{T}(\bar{U}_k, v) < 0$ then firms use both the market and referrals to hire in equilibrium. It is clear from equation (17) that $\bar{S}_i(U)$ is decreasing in $U$ and therefore $\hat{T}(U, v)$ is strictly decreasing in $U$ (observe equation (22)). This implies that that $\hat{T}(U, v)$ is strictly decreasing in $U$ if $U \in [\bar{U}_k, \bar{U}_i)$ and $\hat{T}(\bar{U}_k, v) < 0 \Rightarrow \hat{T}(U, v) < 0 \forall U \geq \bar{U}_k$. Therefore $U^*(v) < \hat{U}_k$.

This completes the third part of the proof.

We now consider firm entry and show that given flow vacancy cost $K$ there is a unique number of vacancies $v$ such that $V = 0$.

When search effort is chosen optimally, the value of a vacancy is given by

$$rV = -K + (1 - \beta)\alpha_F \frac{E_M}{E_M + E_R} \bar{S}_M - \frac{c_M E_M^2}{2} + (1 - \beta)\alpha_F \frac{E_R}{E_M + E_R} \bar{S}_R - \frac{c_R E_R^2}{2}$$

$$= -K + \frac{1}{2}(1 - \beta)\mu \frac{\bar{V}}{\bar{U}}^{-\frac{1}{\eta}} Q_1$$

Rearranging equation (11) we have

$$\frac{v}{u} = \left(\frac{rU - z}{\mu Q_1}\right)^{\frac{1}{\eta}}$$

and therefore:

$$rV = -K + \frac{1}{2}(1 - \beta)\mu^{\frac{1}{\eta}} Q_1 (U^*(v))^{\frac{1}{\eta}} (rU^*(v) - z)^{-\frac{1 - \eta}{\eta}}$$
Notice that:

\[
\lim_{v \to 0} rV = +\infty, \\
\lim_{v \to \infty} rV = -K
\]

which proves the existence of an equilibrium.

We now find conditions so that *first* matches occur through both the market and referral channel and *second* it is not optimal to form a match after a bad signal. We keep the convention that \( \bar{U}_i \geq \bar{U}_k \).

For the first result we need that \( U \leq \bar{U}_k \) where

\[
r\bar{U}_k = \frac{y\Gamma_{2i}}{\Gamma_{2i} + \Gamma_{3i}}
\]

For the second result, denote the probability the match quality is good after a bad signal through channel \( i \) by \( p_{bi} \). Then

\[
p_{bi} = \frac{\gamma(1-q_{Gi})}{\gamma(1-q_{Gi}) + (1-\gamma)q_{Bi}}
\]

The expected surplus of hiring a worker with a bad signal from channel \( i \) can be found by adjusting equation (15):

\[
\tilde{S}_{bi} = yp_{bi}(r + \delta + \lambda) - rU(r + \delta + \lambda p_{bi}) \\
= \frac{(r + \delta + \lambda)(r + \delta)}{\gamma(1-q_{Gi}) + (1-\gamma)q_{Bi}} \left( \frac{y\gamma(1-q_{Gi})}{r + \delta} - rU\left( \frac{\gamma(1-q_{Gi})}{r + \delta} + \frac{(1-\gamma)q_{Bi}}{r + \delta + \lambda} \right) \right) \\
= \frac{(r + \delta + \lambda)(r + \delta)}{\gamma(1-q_{Gi}) + (1-\gamma)q_{Bi}} \left( y\Gamma_{2i} - rU(\Gamma_{2i} + \Gamma_{3i}) \right)
\]
where

\[ \Gamma_{b2i} = \frac{\gamma(1-q_{Gi})}{r+\delta} \]

\[ \Gamma_{b3i} = \frac{(1-\gamma)q_{Bi}}{r+\delta+\lambda} \]

Note that \( \bar{U}_i \geq \bar{U}_k \Rightarrow \tilde{S}_{bi} \geq \tilde{S}_{bk} \). Therefore we want a condition for \( \tilde{S}_{bi} \leq 0 \) which corresponds to

\[ rU \geq r\bar{U} = \frac{y\Gamma_{b2i}}{\Gamma_{b2i} + \Gamma_{b3i}} \]

Note that:

\[ r\bar{U}_k > r\bar{U}_i \]

\[ \frac{y\Gamma_{2k}}{\Gamma_{2k} + \Gamma_{3k}} > \frac{y\Gamma_{b2i}}{\Gamma_{b2i} + \Gamma_{b3i}} \]

\[ \frac{\gamma q_{Gk}}{r+\delta} \left( \frac{\gamma(1-q_{Gi})}{r+\delta} + \frac{(1-\gamma)q_{Bi}}{r+\delta+\lambda} \right) > \frac{\gamma q_{Gk}}{r+\delta} + \frac{(1-\gamma)(1-q_{Bi})}{r+\delta+\lambda} \]

\[ q_{Gk}q_{Bi} > (1-q_{Gi})(1-q_{GBk}) \]

which holds by assumption. Therefore we need \( U \in (\bar{U}_i, \bar{U}_k) \). Using equation (22), this corresponds with having \( v \in (\bar{v}, \bar{v}) \) where \( \bar{T}(U, \bar{v}) = \bar{T}(\bar{U}_k, \bar{v}) = 0 \). This corresponds to a condition on vacancy costs: \( K \in (\bar{K}, \bar{K}) \).

2.3 Predictions

We now present the predictions of the homogeneous firm version of the model.

**Proposition 2.2** If \( \frac{q_{GR}}{1-q_{BR}} \geq \frac{q_{GM}}{1-q_{BM}} \) then:

1. A worker that is hired through a referral receives a higher wage than a worker hired through the market.
2. A worker that is hired through a referral has lower separation rate than a worker hired through the market.

3. The differences in wages and separation rates across hiring channels decline with the workers’ tenure on the job.

Proof. The condition corresponds to \( p_R \geq p_M \) which states that the signals of the referral channel are more accurate than those of the market channel: conditional on a good signal, a referred worker is more likely to be in a good match than a non-referred worker.

\[
\frac{p_R}{\gamma q_{GR}} \geq \frac{p_M}{\gamma q_{GM}}
\]

\[
\frac{\gamma q_{GR}}{(1-\gamma)(1-q_{BR}) \geq \frac{\gamma q_{GM}}{(1-\gamma)(1-q_{BM})}}
\]

\[
\frac{\gamma^2 q_{GM} q_{GR} + \gamma (1-\gamma) q_{GR} (1-q_{BM}) \geq \gamma^2 q_{GM} q_{GR} + \gamma (1-\gamma) q_{GM} (1-q_{BR})}{\frac{q_{GR}}{1-q_{BR}} \geq \frac{q_{GM}}{1-q_{BM}}}
\]

A higher posterior probability of a good match leads to higher wage \( w_R \geq w_M \) and also lower separation rate: \( \lambda (1-p_R) + \delta \leq \lambda (1-p_M) + \delta \). These differentials disappear after learning has occurred and the posterior of surviving workers is equal to 1 regardless of the channel through which they were hired. ■

Under the proposition’s condition, the model gives predictions that are consistent with findings (1), (2) and (4) as reported in the introduction.

Proposition 2.3 If \( \frac{q_{GM} + q_{BM}}{c_M} \geq \frac{q_{GR} + q_{BR}}{c_R} \) and \( U \leq \hat{U} \) then the firm exerts more effort in hiring through the market.

Proof. The effort exerted by firms depends on the cost-adjusted accuracy.
First, we describe that firms exert if \( rU = z \). Then we derive the maximum \( U \) for which they exert more effort for \( M \).

\[
\frac{S_M(z)}{c_M} \geq \frac{S_R(z)}{c_R}
\]

\[
\frac{y\Gamma_{2M}}{c_M} - z\left(\frac{\Gamma_{2M}}{c_M} + \frac{\Gamma_{3M}}{c_M}\right) \geq \frac{y\Gamma_{2R}}{c_R} - z\left(\frac{\Gamma_{2R}}{c_R} + \frac{\Gamma_{3R}}{c_R}\right)
\]

\[
\frac{y\gamma}{z(r + \delta)}\left(\frac{q_{GM}}{c_M} - \frac{q_{GR}}{c_R}\right) \geq \frac{\gamma}{r + \delta}\left(\frac{q_{GM}}{c_M} - \frac{q_{GR}}{c_R}\right) + \frac{1 - \gamma}{r + \delta + \lambda}\left(\frac{1 - q_{BM}}{c_M} - \frac{1 - q_{BR}}{c_R}\right)
\]

The condition above suffices to prove that this expression is positive.

Since \( S_M(\hat{U}_M) = \), the above inequality will eventually reverse. It will hold for \( U \leq \hat{U} \) where:

\[
S_M(\hat{U}) = S_R(\hat{U})
\]

\[
\Rightarrow r\hat{U} = \frac{y\gamma}{r + \delta}\left(\frac{q_{GM}}{c_M} - \frac{q_{GR}}{c_R}\right) + \frac{1 - \gamma}{r + \delta + \lambda}\left(\frac{1 - q_{BM}}{c_M} - \frac{1 - q_{BR}}{c_R}\right)
\]

If the conditions of the two proposition hold simultaneously, then we would observe that referred workers appear to be “better” and still most hires occur through the market, as is the case in the data where in most data sets 50% or less of workers find their jobs through referrals.

**Proposition 2.4** If \( \gamma(q_{GR} - q_{GM}) \geq (1 - \gamma)(q_{BR} - q_{BM}) \) then a worker that the firm meets through a referral is more likely to be hired.
Proof. This condition corresponds to:

\[ \pi_i \geq \pi_k \]

\[ \gamma q_G i + (1 - \gamma)(1 - q_{Bi}) \geq \gamma q_G k + (1 - \gamma)(1 - q_{Bk}) \]

\[ \gamma(q_G i - q_G k) \geq (1 - \gamma)(q_{Bi} - q_{Bk}) \]

Under the proposition’s condition, the model gives predictions that are consistent with findings (3), as reported in the introduction.

3 Firm heterogeneity and endogenous signal accuracy

This Section introduces two additional features to the model: endogenous choice of signal accuracy and firm heterogeneity. The goal is to explore how the use of referrals differs by firm productivity.

3.1 The Extended Model

There are two types of firm which differ in productivity \((y^H > y^L)\) and there are two markets, one for each type of firm denoted by \(t \in \{H, L\}\). An entrepreneur can post a vacancy at either market subject to a flow cost which is given by \(K(v^t)\) where \(v^t\) is the number of vacancies in market of type \(t\) and \(K(0) = K'(0) = 0, K''(v) > 0\). The assumption that vacancy costs depend on the number of vacancies in the market will guarantee that both markets are open.

Every worker chooses which market to enter and we denote the number of workers in market \(t\) by \(n^t\), where \(n^H + n^L = n\).

Each market operates as in the basic model with two simplifications: the cost of exerting search effort is assumed to be the same for both channels and normalized to unity \((c_M = \)
$c_R = 1$) and a signal is correct with the same probability across good and bad match quality ($q_{Gi} = q_{Bi} = q_i$ for $i \in \{M, R\}$). The major change is that the probability that a signal is correct is a choice variable for the firm. Specifically, the firm chooses how much it allocates for human resources, $h$, which determines the accuracy of the signals, $q_M(h)$ and $q_R(h)$. The cost of this choice is equal to $s(h) = (1 - \beta)h$, and it is paid every time the firm meets with a worker (i.e. it is proportional to the amount of interviewing that is done by the firm).

We assume that:

$$\begin{align*}
q_i(0) & \geq \frac{1}{2} \\
q'_i(h) & > 0 \\
\lim_{h \to 0} q''_i(h) & = +\infty \\
q''_i(h) & < 0
\end{align*}$$

Furthermore, to be consistent with the evidence presented in the introduction we assume:

$$\begin{align*}
q_R(h) & \geq q_M(h) \\
q'_R(h) & \leq q'_M(h)
\end{align*}$$

We focus on equilibria where all firms in market $t$ choose the same $h$.

Firms exert effort in searching through the market and referrals and learn about match quality as in Section 2. The value function of firm $j$ of type $t$ is:

$$r\tilde{V}^{jt}(e_M, e_R, h) = -K(v^t) + \alpha_{FM}^{jt}(\pi_M(h)(J_M^t(h) - V^t) - s(h)) + \alpha_{FR}^{jt}(\pi_R(h)(J_R^t(h) - V^t) - s(h)) - \frac{e^2_M + e^2_R}{2}$$

$^6$The term $1 - \beta$ is a normalization that will simplify the algebra shortly.

$^7$One possible micro-foundation of this assumption is that referred candidates emit an additional signal.
where

\[ V^t = \max_{e_M, e_R, h} \hat{V}^{jt}(e_M, e_R, h) \]

The other value functions are identical to the ones in the previous Section.

The Equilibrium is defined as follows:

**Definition 3.1** An Equilibrium is the steady state level of unemployment \( u^t \), the number of vacancies \( v^t \), the matching decision rules \( \{d_M^t, d_R^t\} \), the effort levels \( E_M^t \) and \( E_R^t \), the choice of accuracy \( h^t \) and the value of unemployment \( U^t \) for each market such that in both markets:

- The labor market is in steady state.
- The surplus is split according to the Nash bargaining solution.
- The choice of effort and human resources maximizes the value of a vacancy.
- The matching decision rules are optimal.
- There is free entry of firms: \( V^t = 0 \).

and

- The value of unemployment is the same across markets.

### 3.2 Equilibrium

We prove that

**Proposition 3.1** An equilibrium exists.

We first solve for the market equilibrium within a market with a given number of workers \( n^t \). We then examine the workers' choice of market.
Using the Nash bargaining solution we can rewrite the value function of a vacant firm as follows:

\[
\tilde{V}^{jt}(e_M, e_R, h) = -K(v^t) + \alpha_F \left( \frac{e_M}{E_M^t + E_R^t} - \frac{e_R}{E_M^t + E_R^t} \right) \left( (1 - \beta) \hat{S}_M^t(h) - (1 - \beta)h \right) + \frac{e_R}{E_M^t + E_R^t} \left( (1 - \beta) \hat{S}_R^t(h) - (1 - \beta)h \right) - \frac{e_M^2 + e_R^2}{2} 
\]

where

\[
\hat{S}_i^t(h) = \bar{S}_i^t(h) - h
\]

The first order conditions with respect to \( e_M \) and \( e_R \) yield:

\[
e_i = \frac{\alpha_F(1 - \beta) \hat{S}_i^t(h)}{E_M^t + E_R^t} \tag{24}
\]

Introducing the optimal choice of effort inside the value of a vacancy leads to:

\[
\tilde{V}^{jt}(h) = -K(v^t) + \frac{(1 - \beta) \alpha_F}{E_M^t + E_R^t} \left( (\hat{S}_M^t(h))^2 + (\hat{S}_R^t(h))^2 \right) - \frac{1}{2} \frac{(1 - \beta) \alpha_F}{E_M^t + E_R^t} \left( (\hat{S}_M^t(h))^2 + (\hat{S}_R^t(h))^2 \right) 
\]

Note that:

\[
\hat{S}_i^t(h) = q_i^t(h) \left( \frac{y^t}{r + \delta} - rU \left( \frac{\gamma}{r + \delta} - \frac{1 - \gamma}{r + \delta + \lambda} \right) \right) - 1 > 0 \\
\hat{S}_i^t(h) = q_i''^t(h) \left( \frac{y^t}{r + \delta} - rU \left( \frac{\gamma}{r + \delta} - \frac{1 - \gamma}{r + \delta + \lambda} \right) \right) < 0
\]
To find the optimal choice of $h$, we solve:

$$\max_h \left( (\hat{S}_M(h))^2 + (\hat{S}_R(h))^2 \right)$$

This is the sum of two strictly concave functions which is therefore also strictly concave and has a unique maximizer.

The optimal choice of $h$ is characterized by:

$$\hat{S}_M(h)\hat{S}_M'(h) + \hat{S}_R(h)\hat{S}_R'(h) = 0 \quad (25)$$

Evaluating this equation at the symmetric equilibrium where $e_i^t = E_i^t$ and performing similar calculations to Section 2 leads to:

$$r\hat{V}^{jt} = -K(v') + \frac{1}{2}(1 - \beta)\alpha_F \left( \frac{(\hat{S}_M)^2 + (\hat{S}_M')^2}{\hat{S}_M^t + \hat{S}_R^t} \right)$$

Note that workers are only affected by the firms’ human resource choice through its effect on $q_M$ and $q_R$. Given these, the equation that determines the equilibrium are identical. When a worker and a firm bargain, the investment in interviewing is already sunk. Furthermore, the firm’s outside option is pinned down by the cost of entry. Therefore the equilibrium within a market is determined in the same way as in Section 2 with the addition of an extra condition (equation (25)).

The Inada conditions on the firms’ entry costs guarantee that if some workers are willing to enter a market, then a number of vacancies will also be created. Furthermore, since the marginal cost of creating the vacancies is lower for small markets, there will be proportionally more vacancies per worker in the case of few workers. This means that it will be worthwhile to enter from the worker’s perspective.
3.3 Predictions

Proposition 3.2 More productive firms exert more effort in hiring through the market than less productive firms.

The optimal choice of search effort implies that:

\[ \frac{E_M^t}{E_R^t} = \frac{\hat{S}_M^t}{\hat{S}_R^t} \]

Note that:

\[
\frac{d(\hat{S}_M^t)/(\hat{S}_R^t)}{dy^t} = \frac{1}{(\hat{S}_R^t)^2}\left[ \left(q_M'(h)\left(\frac{y\gamma}{r + \delta} - rU\left(\frac{\gamma}{r + \delta} - \frac{1 - \gamma}{r + \delta + \lambda}\right)\right) - 1\right)\hat{S}_R^t(h) \\
- \left(q_R'(h)\left(\frac{y\gamma}{r + \delta} - rU\left(\frac{\gamma}{r + \delta} - \frac{1 - \gamma}{r + \delta + \lambda}\right)\right) - 1\right)\hat{S}_M^t(h) \right]
\]

This expression is strictly positive because

\[ \hat{S}_R^t > \hat{S}_M^t \]

\[ q_M'(h) > q_R'(h) \]

Therefore, high productivity firms spend more on human resources, thereby generating more accurate signals and reducing the difference between the market and referral channels of hiring. As a result they hire relatively more from the market than low productivity firms.

This proposition is consistent with the empirical findings of DGS that high productivity firms tend to use referrals less than low productivity firms. To the extent that productivity and firm size are positively correlated, it is also consistent with the fact, documented in Pellizzari (2010), that small firms use referrals to a greater extent than large firms.
4 Conclusions

This paper presents a model where firms have two channels through which to meet with workers: the market or referrals. These channels differ in terms of the informational content that is transmitted regarding the quality of a potential match. In particular, if referrals provide more information, this generates predictions that are consistent with a large number of empirical facts. Specifically, workers who are contacted through referrals appear to have a certain advantage because they are hired more often, they have lower separation rates and they receive higher wages. However, these differentials decline with tenure, which is consistent with learning about match quality.

Furthermore, to the extent that firms can invest in their interviewing technology to improve the accuracy of their signals, high productivity firms will invest to a larger extent and therefore use referrals relatively less than low productivity firms. This is consistent with empirical findings and can also explain the puzzling dissonance between firm-level studies which find positive wage premia for referred workers and wage regressions where the sign is negative.
References


