Referral Networks and Inequality*

Manolis Galenianos†

April 2016

Abstract

I develop a theoretical model to study the welfare effects of using referrals in the labor market. In the model, firms use referrals to hire, workers are heterogeneous and the social network endogenous. Consistent with empirical evidence, referred workers are more likely to be hired, to receive a higher wage and to be more productive. The use of referrals exacerbates inequality among workers. Higher inequality is efficient if heterogeneity is driven by productivity differentials but is detrimental to efficiency if the probability of forming a match is weakly correlated with productivity, which is likely in the presence of discrimination.

*Parts of this paper were circulated in an early version of Galenianos (2014). I thank seminar and conference participants for useful comments.

†Royal Holloway, University of London. Email: manolis.galenianos@gmail.com
1 Introduction

It is well-known that social networks play an important role in labor markets: a majority of workers report finding their jobs through friends, acquaintances or relatives.\textsuperscript{1} But is this a good thing? On the one hand, some authors have argued that the use of referral networks increases inequality because it benefits the networked at the expense of qualified but less well-connected workers.\textsuperscript{2} On the other hand, several empirical studies find that workers hired through a referral have better labor market outcomes than those hired using different (formal) channels, suggesting that referrals alleviate informational frictions.\textsuperscript{3} These two observations are not, of course, mutually exclusive but they suggest that the welfare implications of referral use in the labor market are not immediate.

In this paper I build a theoretical model to study the welfare effects of referral use in the labor market. I consider a setting where firms use both formal and informal channels to hire, workers are heterogeneous and the social network is endogenous. The model’s predictions about the characteristics of referred workers are consistent with the empirical evidence and, furthermore, it turns out that the use of referrals exacerbates inequality. Higher inequality improves efficiency if heterogeneity is driven by productivity differentials but is detrimental if productivity and the probability of forming a match are weakly correlated, as in the case of discrimination.

In the model there are two types of worker and a worker’s type (high or low) determines his productivity and the probability of forming a match when meeting a firm. The network is formed at an initial stage and mutual consent is required to create a link between two workers. Subsequently, workers and firms interact in a frictional labor market.

In the labor market vacancies are created both through the free entry of new firms and

\textsuperscript{1}See the surveys by Ioannides and Loury (2004) and Topa (2011).
\textsuperscript{2}Calvo-Armengol and Jackson (2004) make this point in a theoretical model where agents learn about jobs from their neighbors. Topa (2001) finds strong local spillovers in unemployment rates across geographical locations. Beaman, Keleher and Magruder (2015) find evidence in a field experiment that the use of referrals reinforces unequal access to jobs between men and women.
\textsuperscript{3}See Dustmann, Glitz, Schoenberg and Bruecker (2015) and Hensvik and Skans (2015) among others.
through the expansion of producing firms. A firm and a worker meet either through search in the frictional market or through a referral, which occurs when a producing firm expands and asks its current employee to refer a link. When a firm and a worker meet all information becomes public, including the worker’s type and network, and the probability the match is formed depends on the worker’s type. The flow surplus of a worker-firm match is equal to productivity plus the value of the referrals and the wage is determined by Nash bargaining.

The equilibrium admits a sharp characterization despite the model’s complexity. The benefit of forming a link is that a referral may be provided at some future date. High-type workers are more likely to be employed, and hence provide a referral, which makes them more desirable as links. In equilibrium, every worker regardless of type has a majority of links with high-type workers. The equilibrium network is hierarchical rather than homophilous, which is the more common assumption in the literature (Montgomery, 1991).

The equilibrium structure of the network means that referrals are mostly received by high type workers which has two implications. First, the use of referrals exacerbates inequality. Second, referred workers’ labor market outcomes are consistent with the a wealth of empirical findings: conditional on observable worker characteristics, referred candidates are more likely to be hired (Fernandez and Weinberg, 1997; Castilla, 2005; Brown, Setren and Topa, 2015; Burks, Cowgill, Hoffman, Housman, 2015), to receive higher wages (Simon and Warner, 1992; Bayer Ross and Topa, 2008; Brown, Setren and Topa, 2015; Dustmann, Glitz, Shoenberg and Bruecker, 2015) and to be more productive (Castilla, 2005; Pinkston, 2011; Burks, Cowgill, Hoffman, Housman, 2015).

The optimal network is different, generically, from the equilibrium network. Ceteris paribus, the optimal network increases inequality if differentials in productivity are higher. Ceteris paribus, the optimal network reduces inequality if differentials in the probability of forming a match are higher. Therefore, the use of referrals improves efficiency when produc-

---

4However, the distribution of firm sizes is degenerate: each firm hires one worker and vacancies created through an expansion are immediately sold off.
tivity and match probability are strongly correlated but it is detrimental for efficiency when they are weakly correlated. The latter case might occur when discrimination is present.

This paper is related to two broad literatures. The first is theoretical and models the interaction of labor markets with social networks. Montgomery (1991) is one of the first papers to model social networks of heterogeneous workers in the labor market and assumes that the network is exogenous and homophilous. Calvo-Armengol and Jackson (2004) use graph theory to model the exogenous network among homogeneous (in terms of productivity) workers and show how network structure correlates to inequality. In models similar to the present one, Galenianos (2014) studies the effect of referrals on matching efficiency and Igarashi (2016) examines the welfare effect of banning referrals when only a subset of the workforce has access to the network. Calvo-Armengol (2004) and Galeotti and Merlino (2014) endogenize network formation among homogeneous workers and examine the effect of referrals on hiring intensity.

The second literature is empirical and estimates how referrals affect labor market outcomes. Dustmann, Glitz, Schoenberg and Bruecker (2015) use German data and identify referrals by focusing on the immigrant background of new hires and incumbent workers. This is the first paper to use firm fixed effects and decompose the wage effect of a referral from selection on the firm side. Hensvik and Skans (2015) use Swedish data with access to workers’ military draft ability tests and they identify co-worker networks. They find that referred workers have higher ability controlling for observables. The firm-level studies of Fernandez and Weinberg (1997), Castilla (2005), Brown, Setren and Topa (2015) and Burks, Cowgill, Hoffman and Housman (2015) provide information on all applicants for a job, through both formal and informal channels.

Section 2 presents the model, Section 3 characterizes the equilibrium, Section 4 shows that the model’s predictions are consistent with the data and Section 5 studies the efficiency properties of equilibrium. Proofs are collected in the Appendix.

---

\(^5\)See Galenianos (2013) for a theoretical study of this selection.
2 The Model

The economy is populated by workers and firms. There is measure 1 of high type workers (H-type) and measure 1 of low type workers (L-type) who can be thought of as above-median and below-median according to some attribute such as employability or productivity.\(^6\) The measure of firms is endogenously determined through a free entry condition.

There are two distinct stages. In the first stage workers form a network. In the second stage workers and firms interact in a frictional labor market. Network formation is costly but creates the opportunity of finding a job through a referral in the frictional labor market. This structure is motivated by the idea that forming a network is a long-term investment which is not affected by short-term fluctuations in one’s employment status.\(^7\)

Section 2.1 describes the network formation process. Section 2.2 describes the labor market and the interaction between referrals and labor market outcomes.

2.1 Network Formation

Network formation is modeled as a non-cooperative game with non-transferable utility and attention is restricted to equilibria in symmetric strategies by workers. The guiding principle behind the assumptions on network formation (detailed below) is that mutual consent is necessary for the formation of a link between two workers.

Denote the network of worker \(j\) of type \(i\) \(n_i^j\), the cost of forming the network by \(C_i(n_i^j)\)

\(^6\)Having arbitrary measures of the two types yields qualitatively identical results regarding the labor market interaction (proof available upon request). Network formation is affected by the proportions of each type but in a somewhat uninteresting way: if there are very few low type workers then the network of a high-type worker will mechanically include very few low types. For this reason, the focus of this paper is on the case of equal measures for each type.

\(^7\)The model can be reformulated so that the network-formation decisions are made, and associated costs are borne, during the labor market stage. Such a reformulation complicates the analysis but does not qualitatively affect the allocation so long as a worker cannot adjust his network in response to his temporary employment status.
and the steady state utility in the labor market by $\Lambda_i(n^j_i)$. Worker $j$ chooses $n^j_i$ to maximize:

$$\mathcal{L}_i(n^j_i) = \Lambda_i(n^j_i) - C_i(n^j_i)$$

(1)

In more detail, the network of a worker is fully described by the measure of links that he has with type-$H$ and type-$L$ types workers. The network of worker $j$ of type $i$ is denoted by $n^j_i = (n^j_{iH}, n^j_{iL})$ where $n^j_{ik}$ is the measure of links that he has with workers of type $k$.

Modeling a worker’s network as a continuum of links is consistent with the (spirit of the) sociology literature’s finding that it is a person’s more numerous weak ties that help most with finding a job (Granovetter, 1973; 1995) and is crucial for the model’s tractability (Galenianos, 2014). A worker’s employment opportunities will in general depend on how many of his links are employed which necessitates keeping track of each link’s time-varying employment status. Having a continuum of links means that the aggregate (un)employment rate of a worker’s social contacts is determinist due to the law of large numbers (and constant, in steady state), thereby greatly simplifying the analysis.

The measure of links that a worker acquires depends on the effort he exerts in network creation and on the aggregate effort of all other workers. The effort of worker $j$ of type $i$ is denoted by $(e^j_{iH}, e^j_{iL})$ where $e^j_{ik}$ is the effort he exerts in linking with type-$k$ workers. Workers simultaneously choose their effort levels. The aggregate effort of all workers is denoted by $(E_{HH}, E_{HL}, E_{LH}, E_{LL})$ where $E_{ik}$ denotes the effort of type-$i$ workers towards linking with type-$k$ workers, and will be referred to as the “demand” for such links. I will focus attention on equilibria in symmetric strategies where $e^j_{ik} = E_{ik}$ for all $i, j, k$.

Links are formed as follows. If there is positive demand for within-$i$ links ($E_{ii} > 0$), then the measure of links that worker $j$ of type $i$ forms is equal to his effort: $n^j_{ii} = e^j_{ii}$; if there is no demand ($E_{ii} = 0$), then he forms no such links ($n^j_{ii} = 0$) regardless of his effort level. 

---

8In the labor market, the worker transits between employment and unemployment. The steady state utility is calculated using the proportion of time that he spends at each labor market state which, in general, depends on his network.
Therefore:

\[ n_{ji}^j(e_{ij}^j) = e_{ij}^j I[E_{ji} > 0] \]

where the indicator function takes a value of 1 if \( E_{ji} > 0 \) and 0 if \( E_{ji} = 0 \).

The measure of links between workers of different types needs to respect aggregate consistency, i.e. the high types’ measure of links with low-type workers is equal to the measure of links that low types have with high-type workers: \( n_{HL} = n_{LH} \). The following structure delivers this requirement:

\[ n_{ik}^j(e_{ik}^j) = e_{ik}^j \frac{E_{ki}}{E_{ik} + E_{ki}} \]

Under this formulation, the probability of forming across-type links depends on the relative demand across types in a natural way. Notice that mapping effort to links is symmetric for within- and across-type links.\(^9\)

Cost is quadratic in effort: \( \tilde{C}_i(e_{ii}^j, e_{ik}^j) = \frac{c}{2} \left( (e_{ii}^j)^2 + (e_{ik}^j)^2 \right) \). Therefore the cost of network formation is given by:

\[ C_i(n_{ii}^j, n_{ik}^j) = \frac{c}{2} \left( (n_{ii}^j(e_{ii}^j))^2 + \left( \frac{n_{ik}^j(e_{ik}^j)(E_{ki} + E_{ik})}{E_{ki}} \right)^2 \right) \]

for \( E_{ii} > 0 \) and \( E_{ik} > 0 \). The cost is infinite in the case where \( E_{ii} = 0 \) and \( n_{ii}^j > 0 \) and in the case where \( E_{ki} = 0 \) and \( n_{ik}^j > 0 \). I assume throughout that \( c \) is high enough for the solution to be interior.

The costs of network formation are separable across links with the two worker types which

---

\(^9\)To see this recall that in equilibrium \( e_{ik}^j = E_{ik} \) and assume, for simplicity, that \( E_{HL} = E_{LH} = E_{ii} = E > 0 \). The measure of links of a type-\( H \) worker with type-\( L \)s is \( n_{HL}^j = \frac{E}{2E} = \frac{E}{2} \) and, furthermore, \( n_{HL}^j = n_{LH}^j \). Therefore the total measure of links across types is \( (n_{HL} + n_{LH}) = E \). The measure of links of a type-\( i \) worker with workers of his own type is \( n_{ii}^j = e_{ii}^j = E \) and the total measure of links within type is \( n_{ii}^j = E \). This can be easily extended to the general case where \( E_{ik} \neq E_{ii} \) and shows that the mapping from effort to links is symmetric between within- and across-type link formation.
means that the marginal cost of forming a link with a worker starts at zero for both types. This assumption is motivated by the observation that there are many reasons for forming links with other workers, some of which are non-pecuniary (friendship, etc.). To the extent that the distribution of non-pecuniary benefits is not strongly type-dependent (so that an $H$-type worker values the non-pecuniary benefits he receives from $L$-type workers and vice versa) this effect is captured, in a reduced-form way, by the separability in the cost of network formation: the marginal cost, net of non-pecuniary benefits, of linking with workers of a type starts at zero because the initial link is formed with the worker who provides the highest level of non-pecuniary benefits. Variations of this assumption which might affect the network structure are discussed in the Conclusions.

2.2 The Labor Market

I first describe a labor market where all workers have the same network, conditional on type. I then derive the value functions for a worker with an off-equilibrium network, which will be useful for solving the network formation stage.

Time runs continuously, the horizon is infinite, the discount rate is $r > 0$ and the labor market is at steady state. Firms are homogeneous, risk-neutral and maximize expected discounted profits. Each firm hires one worker and a firm is either filled and producing or vacant and searching, where $k$ denotes the flow cost of a vacancy.

Workers are heterogeneous, risk-neutral and maximize expected discounted utility. A worker is either employed or unemployed and the flow utility of unemployment is $b > 0$. The heterogeneity captured by the two worker types does not refer to education or some other easily contractible attribute. In other words firms cannot post type-specific vacancies and all workers search for jobs in the same market. This is precisely the setting where the information embedded in a worker’s social network becomes relevant.

Vacancy creation occurs in two ways: (1) a new firm enters and starts searching through
the market; (2) an existing firm expands, which occurs at exogenous rate $\rho$. Following an expansion, the new position is immediately sold off which keeps firms’ employment at one worker. A firm and a worker meet either through search in the market or through a referral, which occurs when a firm expands and asks its current employee to refer a link. The rate of meeting through the market is determined by a matching function and the rate of meeting through referrals is determined by the rate at which firms expand.

When a firm and a type-\(i\) worker meet, a match is formed with probability $p_i$ and produces flow output $y_i$. With probability $1 - p_i$ there is no match and search continues. All payoff-relevant variables (e.g. worker’s type and network) are common knowledge when the match is formed. The probability of a match and productivity are weakly higher for high types ($p_H \geq p_L$ and $y_H \geq y_L$) and low types are hireable ($y_L > b$). The correlation between differentials in productivity and probability of forming a match across the two types depends on the specific interpretation of the nature of heterogeneity. I provide two examples which will turn out to have very different implications in the efficiency analysis of Section 5.

**Example 1:** When a firm and a type-\(i\) worker meet they draw the match specific productivity from some continuous distribution $F_i$ where the distribution of $H$-type workers first order stochastically dominates that of $L$-types. In this case, the match is formed if the draw is above an endogenous cutoff, a high type is more likely to make a draw above his cutoff than a low type and, conditional on forming a match, his productivity is higher than that of a low type.\(^{10}\) This example is consistent with a labor market where heterogeneity is due to productivity differences and differentials in $y_i$s and $p_i$s are strongly positively correlated.

**Example 2:** The match productivity distribution is binary, with productivity levels $\{\bar{y}, y\}$ which occur with probabilities $\{\bar{p}, p\}$, respectively, for both types. A match is profitable only if match productivity is $\bar{y}$ and the firm draws a binary signal about match productivity which is always accurate for $H$-type workers and is inaccurate with probability $\zeta$ for $L$-type workers.

\(^{10}\)Standard assumptions, such as log-concavity of the match productivity distribution, are sufficient for this result.
worker in highly productive matches. A match with an $L$-worker is only formed after a signal of high match productivity, and we have $y_H = y_L = \bar{y}$ and $p_L = \bar{p}(1 - \zeta) < \bar{p} = p_H$. The second example is consistent with a labor market where some form of (taste or statistical) discrimination is prevalent and the correlation between $y_i$s and $p_i$s is weak.

When a firm that employs a type-$i$ worker expands, one of the links of the incumbent worker is referred at random. If the referred worker is unemployed then he meets with the firm and a match if formed with the type-specific probability. If the match is not formed or the referred worker is employed then the referral opportunity is lost and search in the market begins.

Denote the proportion of links that a worker type $i$ has with his own type by $\phi_i$:

$$\phi_i = \frac{n_{ii}}{n_{ii} + n_{ik}}$$

The assumption of symmetric networks means that we can dispense with the superscript $j$ which denotes the individual worker. A referral from a type-$i$ worker arrives at an unemployed type-$i$ worker with probability $\phi_i u_i$ and an unemployed type-$k$ ($\neq i$) worker with probability $(1 - \phi_i) u_k$, where $u_l$ is the unemployment rate of type-$l$ workers. Therefore, the referral pool is in general different from the unemployment pool in terms of worker types.

Denoting the value of employing a type-$i$ worker by $J_i$ and the value of a vacancy by $V$, the value of expanding when employing a type-$i$ worker is equal to:

$$X_i = V + \phi_i u_i p_i (J_i - V) + (1 - \phi_i) u_k p_k (J_k - V).$$

The flow value to a match between a firm and a type-$i$ worker is $y_i + \rho \gamma X_i$, where $\gamma \in [0, 1]$

---

11This holds so long as the $\zeta$ is not too high.

12For example, Bertrand and Mullainathan (2004) document that the call-back rates after sending CVs to help-wanted ads are with white-sounding names are 50% higher than those with black-sounding names. More broadly, the literature on statistical discrimination posits higher variance in the ability signals of minority workers, e.g. Aigner and Cain (1977).
is the share of the expansion’s surplus that is received by the incumbent.

When unemployed, a worker is referred to a firm when the employer of one of his links expands and this worker is chosen among the referrer’s links. Consider worker \( j \) of type \( i \) who has \( n_{ji}^j \) links with type-\( i \) workers and \( n_{ik}^j \) links with type-\( k \) workers. Each link of type \( i \) is employed with probability \( 1 - u_i \) and gets the opportunity to refer at rate \( \rho \). A referrer of type \( i \) has \( n_{ii} + n_{ik} \) links and each of them is equally likely to receive the referral. Therefore, our worker is referred to a job at rate

\[
\alpha_{Ri}^j = \frac{\rho n_{ii}^j (1 - u_i) + \rho n_{ik}^j (1 - u_k)}{n_{ii} + n_{ik} + n_{ki}}
\]

\[
\Rightarrow \alpha_{Ri} = \rho \phi_i (1 - u_i) + \rho (1 - \phi_k)(1 - u_k),
\]

where I used network symmetry was imposed and the consistency condition \( n_{ik} = n_{ki} \). Notice that the referral rate only depends on the proportion of links with each type and is not directly affected by the size of the network.

Three types of agents search in the market: measure \( v \) vacancies, measure \( u_H \) high-type unemployed workers and measure \( u_L \) low-type unemployed workers. The flow of meetings in the market between vacancies and workers of either type is given by a Cobb-Douglas function

\[
m(v, u_H, u_L) = \mu v^\eta (u_H + u_L)^{1-\eta},
\]

where \( \mu > 0 \) and \( \eta \in (0, 1) \).

The rate at which a vacancy meets with a type \( i \) worker through the market is

\[
\alpha_{Fi} = \frac{m(v, u_H, u_L)}{v} \frac{u_i}{u_H + u_L} = \mu (\frac{u_H + u_L}{v})^{1-\eta} \frac{u_i}{u_H + u_L}.
\]
The rate at which a type \( i \) worker meets a firm through the market is

\[
\alpha_{Mi} = \frac{m(v, u_H, u_L)}{u_H + u_L} = \mu\left(\frac{v}{u_H + u_L}\right)^n.
\]

Since this rate does not depend on the worker’s type, the \( i \)-subscript is henceforth dropped.

The steady state conditions are that each type's flows in and out of unemployment are equal:

\[
\begin{align*}
    u_H(\alpha_M + \alpha_{RH})p_H &= (1 - u_H)\delta, \quad (3) \\
    u_L(\alpha_M + \alpha_{RL})p_L &= (1 - u_L)\delta. \quad (4)
\end{align*}
\]

I now describe the agents' value functions. Consider a firm. When vacant, it searches in the market, meets with a type-\( i \) worker at rate \( \alpha_{Fi} \) and forms a match with probability \( p_i \). When producing, the firm’s flow payoffs are \( y_i + \rho\gamma X_i - w_i \) where \( w_i \) denotes the wage. The match is destroyed at rate \( \delta \). The firm’s values of a vacancy (\( V \)) and employing a type-\( i \) worker (\( J_i \)) are given by:

\[
\begin{align*}
    rV &= -k + \alpha_{FHPH}(J_H - V) + \alpha_{FLPL}(J_L - V), \\
    rJ_i &= y_i + \rho\gamma X_i - w_i + \delta(V - J_i).
\end{align*}
\]

Consider a worker of type \( i \). When unemployed, his flow utility is \( b \). Job opportunities appear at rate \( \alpha_M + \alpha_{Ri} \) and a match is formed with probability \( p_i \). When employed, the worker’s flow utility is equal to the wage and the match is destroyed at rate \( \delta \). The worker’s value of unemployment (\( U_i \)) and employment (\( W_i \)) are given by:

\[
\begin{align*}
    rU_i &= b + (\alpha_M + \alpha_{Ri})p_i(W_i - U_i), \\
    rW_i &= w_i + \delta(U_i - W_i).
\end{align*}
\]
The wage solves the Nash bargaining problem

\[ w_i = \arg\max_w (W_i - U_i)^\beta (J_i - V)^{1-\beta}. \] (5)

Finally, the steady state utility of a type \( i \) agent is:

\[ \Lambda_i = u_i U_i + (1 - u_i)W_i. \]

I now determine the payoffs to worker \( j \) of type \( i \) whose network is \((n^j_{ii}, n^j_{ik})\) and might differ from that of the other type \( i \) workers. Worker \( j \) is measure zero and therefore his network does not affect any aggregate quantity such as the values of other agents, unemployment rates or vacancy creation.

Worker \( j \)’s network size and composition affects the referrals that he receives and the referrals that he generates. His arrival rate of job opportunities through a referral is given by:

\[ \alpha^j_{Ri} = \frac{n^j_{ii}(1 - u_i)\rho}{n_{ii} + n_{ik}} + \frac{n^j_{ik}(1 - u_k)\rho}{n_{kk} + n_{ki}}. \]

The proportion of time that he spends unemployed is determined by:

\[ u^j_i (\alpha_M + \alpha^j_{Ri})p_i = (1 - u^j_i)\delta \Rightarrow u^j_i = \frac{\delta}{\delta + (\alpha_M + \alpha^j_{Ri})p_i}. \]

I will assume that, when forming his network, a worker does not take into account the differential value that he might have on his employer: \( X^j_i = X_i \). The assumption has a second-order effect on the equilibrium to the extent that the benefits from forming a network arise mostly from the opportunity of receiving referrals rather than from the increased wage that the firm is willing to pay to someone who can generate referrals.
The value of unemployment for worker \( j \) and of the firm employing him are given by:

\[
\begin{align*}
    rU_i^j &= b + (\alpha_M + \alpha_R)\pi_i(W_i^j - U_i^j) \\
rJ_i^j &= y + \rho\gamma X_i - w_i^j + \delta(V - J_i^j)
\end{align*}
\]

Notice that the worker’s referral rate directly affects his value of being unemployed and therefore his outside option when bargaining with his employer.

The steady state utility of worker \( j \) of type \( i \) is:

\[
\Lambda_i^j = u_i^j U_i^j + (1 - u_i^j)W_i^j
\]

### 2.3 Definition of Equilibrium

The equilibrium in the labor market for a given network is defined as follows.

**Definition 2.1** A Labor Market Equilibrium given network \((n_{HH}, n_{HL})\) and \((n_{LL}, n_{LH})\) is the steady state measures of unemployed workers \(\{u_H, u_L\}\) and the measure of vacancies \(v\) such that:

- The labor market is in steady state as described by (3) and (4).
- The surplus is split according to (5).
- There is free entry of firms: \(V = 0\).

The equilibrium is defined as follows.

**Definition 2.2** An Equilibrium is \((e_{HH}, e_{HL})\) and \((e_{LH}, e_{LH})\) which solve (1) subject to the symmetry restriction where the labor market payoffs are given by equation (6).
3 Equilibrium Characterization

I begin the analysis of the labor market for any symmetric network. I then solve for the equilibrium network.

3.1 The labor market

The surplus of a match between a firm and a type-$i$ worker is given by $S_i = W_i - U_i + J_i - V$.

Nash bargaining implies that

\[
W_i - U_i = \beta S_i, \\
J_i - V = (1 - \beta)S_i,
\]

and the value functions can be rearranged to yield

\[
(r + \delta)S_i = y_i + \rho \gamma X_i - b - (\alpha_M + \alpha_R)\beta S_i.
\]

Combine the above with equation (2) and the free entry condition to arrive at:

\[
S_i = \frac{y_i - b + \rho \gamma (1 - \beta)(1 - \phi_i)u_k p_k S_k}{r + \delta + (\alpha_M + \alpha_R) p_i \beta - \rho \gamma (1 - \beta) \phi_i u_i p_i}.
\] (7)

Equation (7) illustrates that the dependence between $S_i$ and $S_k$ is due to the fact that a type-$i$ worker may refer a type-$k$ in the case of an expansion. If $\phi_i = 1$ then $i$ types only refer workers of the same type and the term multiplying $S_k$ drops out.

The value of a vacancy is given by

\[
rV = \alpha_{FH} p_H (1 - \beta) S_H + \alpha_{FL} p_L (1 - \beta) S_L.
\] (8)

The following proposition states the Section’s main result.
Proposition 3.1  A Labor Market Equilibrium exists.

3.2 Network formation and equilibrium

This section characterizes the equilibrium in two steps. First, I show that a worker’s choice of effort is unique. Second, the optimal choice of all workers is aggregated to determine the equilibrium.

Consider the problem of worker $j$ of type $i$. The first order conditions of the worker’s problem with respect to effort on networking with his own type are:

$$\frac{dL^j_i}{de^j_{ii}} = \frac{\partial \Lambda^j_i}{\partial \alpha^j_{Ri}} \frac{d\alpha^j_{Ri}}{de^j_{ii}} - ce^j_{ii}$$

$$= \frac{\partial \Lambda^j_i}{\partial \alpha^j_{Ri}} \rho(1-u_i)n_{ii} + n_{ik} - ce^j_{ii}$$

(9)

The first ratio of equation (9) describes how the steady state utility changes with the referral rate while the second ratio describes how effort affects the referral rate.

Similarly:

$$\frac{dL^j_i}{de^j_{ik}} = \frac{\partial \Lambda^j_i}{\partial \alpha^j_{Ri}} \rho(1-u_k)E_{ki} + E_{ik} - ce^j_{ik}$$

(10)

The next Proposition proves that steady state utility $\Lambda^j_i$ is strictly increasing and strictly concave in a worker’s referral rate $\alpha^j_{Ri}$ and, therefore, the worker’s optimal effort level is unique. The concavity result is intuitive: a worker’s steady state utility increases in the rate at which he meets with job opportunities; a high referral rate leads to less time spent unemployed and therefore lower benefit from additional increases in $\alpha^j_{Ri}$, yielding concavity.

Proposition 3.2  Worker $j$’s optimal effort $(e^j_{ii}, e^j_{ik})$ in the network formation stage is unique.

$^{13}$I only consider the case where $E_{ii} > 0$ and $E_{ik} > 0$. Equilibria without networks exist but are of no particular interest for this study.
I now state the main result regarding the existence and characterization of Equilibrium. I focus on the case where $u_H \leq \frac{1}{2}$ and $u_L \leq \frac{1}{2}$, as this is most relevant.

**Proposition 3.3** An Equilibrium exists. In equilibrium:

1. $p_H > p_L \Rightarrow u_H < u_L$ and $p_H = p_L \Rightarrow u_H = u_L$.

2. $\left(\frac{\phi_H}{\phi_L}\right)^2 = \frac{1-u_H}{1-u_L}$.

3. $\phi_H = 1 - \phi_L$.

In equilibrium the network is characterized by one variable: $\phi^* \equiv \phi_H = 1 - \phi_L$.

Therefore, when $p_H > p_L$ we have $1-u_H > 1-u_L$ and $\phi^* > \frac{1}{2}$.

**3.3 Discussion**

Proposition 3.3 provides a precise characterization of equilibrium and shows that a potentially very complicated network structure can be reduced to a single variable.

To interpret these results, it is useful to reiterate the agents’ incentives for forming a network. When contemplating a new link, a worker cares about receiving a referral which only depends on his potential link’s employment status. This observation has two important implications. First, all workers, independent of their own type, have a preference for linking with high-employment $H$-types. This explains why $L$-types have the same proportion of links with high-types as the $H$-types themselves ($1 - \phi_L = \phi_H$): a low-type worker does not respond to the rationing that he faces when linking with $H$-types by creating more links with other $L$-types because these links are less valuable. Since $H$-types are less interested in “linking down”, this results in fewer total links for $L$-types.

The outcome is a network structure where high-type workers benefit most from the use of referrals and any initial inequality is exacerbated by the network’s endogeneity. That
high types benefit most is easily visible in the equilibrium referral rates which are higher for $H$-types and also increasing in the degree of $H$-type homophily:

\[
\alpha_{RH} = \rho \phi_H (1 - u_H) + \rho (1 - \phi_L) (1 - u_L) = \rho \phi^* (2 - u_H - u_L)
\]

\[
\alpha_{RL} = \rho (1 - \phi_H) (1 - u_H) + \rho \phi_L (1 - u_L) = \rho (1 - \phi^*) (2 - u_H - u_L)
\]

The second implication is that the differential desirability of the two types is parametrized by the probability of forming a match when meeting a firm, $p_i$. Therefore $p_i$ directly affect the network formation decision while productivity on the job, $y_i$, does not, a distinction which becomes relevant when examining the efficiency properties of equilibrium (Section 5). Furthermore, increasing the differential between types leads to a more skewed network, further exacerbating inequality. Assuming for simplicity that $p_H = 1 - p_L = p$, the following proposition illustrates this point:

**Proposition 3.4** An increase in $p$ leads increases $\phi^*$, in equilibrium.

These implications mean that the network that emerges in equilibrium is a hierarchical network, which differs from the common assumption of homophily, according to which workers are more likely to link within their type. The assumption of homophily in social networks, originally made in Montgomery (1991) in a labor market context, was used to generate the prediction that a referred worker might dominate a randomly chosen unemployed worker along payoff-relevant characteristics, which is identical to the predictions of the present model, as discussed in Section 4.

Montgomery (1991) was motivated by a wealth of sociological evidence that social networks tend to homophilous (see McPherson, Smith-Lovin and Cook 2001 for a detailed survey). This evidence typically concerns homophily along observable characteristics (after all, it is measured) rather than unobservable characteristics, which might be a more relevant dimension in labor markets where observables such as education can be readily searched for.
In the case of observable heterogeneity, there is some evidence that homophily affects the choice of who is referred, although it is far from conclusive. The experimental study of Beaman, Keleher and Madruger (2015) examines the referrer-referred identity across gender. They find some support for homophily, namely that women are almost twice as likely to refer other women than men are (43% vs. 23%) although women are still more likely to refer a man than a woman. They also find that men have a higher probability of being hired in the job which suggests that likelihood of success of the referred candidate is an important factor in deciding who to refer for the job (the mechanism in which such benefits might be exchanged, however, is probably quite different in their setting). The possibility of extending the present framework to allow for homophily in equilibrium is discussed in the Conclusions.

4 Testable Predictions and Evidence

In equilibrium, when a firm meets a worker it is more likely that the worker is of a high type if the meeting occurs through a referral rather than the market. The intuition is quite straightforward: regardless of own type, each worker is linked with more high-type workers than low-type workers. Therefore, the recipient of a referral is more likely to be a high-type worker regardless of the referrer’s type. The following proposition formalizes this intuition.

**Proposition 4.1** When a firm and a worker meet, it is more likely that the worker is of high type if the meeting is through a referral rather than through the market.

Proposition 4.1 leads to the following predictions:

**Prediction 1:** When a worker and a firm meet, the match is more likely to be formed if they meet through a referral.

**Prediction 2:** When a worker and a firm meet, the match is more productive in expectation if they meet through a referral.
**Prediction 3:** When a worker and a firm meet, the wage is higher in expectation if they meet through a referral.

These predictions are supported by ample empirical evidence. In their firm-level studies, Fernandez and Weinberg (1997), Castilla (2005), Brown, Setren and Topa (2015) and Burks, Cowgill, Hoffman and Housman (2015) examine all the applicants for a job (both successful and unsuccessful) and find that referred applicants are more likely to be hired after controlling for their observable characteristics, consistent with Prediction 1. Consistent with Prediction 2, Castilla (2005), Pinkston (2012) and Burks, Cowgill, Hoffman and Housman (2015) find that the referred are more productive after controlling for worker observables (Castilla and Burks et al. have direct measures of output by worker while Pinkston has subjective measures, reported by the employer). Hensvik and Skans (2015) use a very detailed Swedish data set to document that referred workers have higher scores in the armed forces ability test, conditional on observables.

Finally, consistent with prediction 3, Simon and Warner (1992), Bayer, Ross and Topa (2008), Brown, Setren and Topa (2015), Hensvik and Skans (2015) and Dustmann, Glitz, Schoen-berg and Bruecker (2015) find that referred workers receive higher wages than non-referred workers after controlling for worker observables and firm/place of employment fixed effects.\(^{14}\)\(^{15}\)

**Prediction 4:** The wage premium of a referred worker depends on the extent of unobservable heterogeneity.

\(^{14}\)These wage differentials decline with tenure but a significant part persists over time. Simon and Warner (1992), Galenianos (2013), Brown, Setren and Topa (2015) and Dustmann, Glitz, Schoenberg and Bruecker (2015) develop models with gradual learning about match quality that can rationalize this pattern. The present paper’s interpretation of ex ante worker heterogeneity refers to the persistent part of the wage differential and is therefore complementary to the learning interpretation.

\(^{15}\)Pistaferri (1999), Pellizzari (2010) and Bentolila, Michelacci and Suarez (2010) find zero or even negative effect of referrals on wages. These studies, however, do not control for firm fixed effects, unlike the studies cited above, which suggests that selection is important on the firm side. See Galenianos (2013) for a model where low-productivity firms use referrals more intensely.
Prediction 4 provides a direction for future research. The literature documents that different types of jobs or different industries use referrals at different rates, as documented in Topa (2011). Galenianos (2014) studies the implications on matching efficiency from heterogeneous referral use. A detailed analysis of the source of this heterogeneity is still missing.

5 Efficiency Analysis

As noted in Section 3, the use of referrals exacerbates inequalities in labor markets as it mostly benefits high-type workers. This feature, however, is not a priori detrimental from an efficiency viewpoint if it facilitates the employment of more productive workers. This Section examines under what conditions the use of referrals improves efficiency.

The social planner’s objective is to maximize steady state output, subject to search and informational frictions. The planner has two instruments at his disposal: the network structure and the entry of vacancies.

In Galenianos (2014) I study a similar referral model (but with homogeneous workers) and show that vacancy entry is generically inefficient in equilibrium. The search process is subject to externalities typical of random search models and the efficient level of vacancy entry obtains only when the bargaining parameters (\(\beta\) and \(\gamma\)) take exactly the “right” (non-generic) values. This logic transfers to the present model. However, the current focus is on inequality between the two types of workers rather than vacancy entry and so I will focus the efficiency analysis on network structure.

5.1 Optimal network

The network structure is described by \(\phi_H\) and \(\phi_L\) where, in equilibrium, \(\phi_H = 1 - \phi_L\). I restrict attention to networks with the feature \(\phi_H = 1 - \phi_L\). The planner solves the following
max \( \phi \) \( W(\phi) = (1 - u_H)y_H + bu_H + (1 - u_L)y_L + u_Lb \)

s.t. \( (1 - u_H)\delta = u_H p_H (\alpha_M + \alpha_{RH}) \)

and \( (1 - u_L)\delta = u_L p_L (\alpha_M + \alpha_{RL}) \)

where \( \phi = \phi_H = 1 - \phi_L \). Denote the planner’s solution by \( \phi^P \).

The proposition provides the main efficiency result.

**Proposition 5.1** *The equilibrium network is generically inefficient: \( \phi^P \neq \phi^* \). If \( y_H = y_L \) and \( p_H = p_L \) then the equilibrium network is efficient: \( \phi^P = \phi^* = \frac{1}{2} \).

To explore the source of the inefficiency, it is useful to compare the planner’s objectives with the agents’ incentives. Recall that an agent only cares about the employment rate of his potential links, which is determined by the probability of forming a match when meeting a firm \( p_i \). The planner, by contrast, weighs the different employment rates by the productivity of each type of worker. This is a crucial difference which is responsible for the divergence of objectives between the workers and the planner.

To illustrate this difference, I perform two comparative statics exercises, which are summarized in the proposition below. First, fix the probability of forming a match for the two types, \( p_H \) and \( p_L \), and increase the productivity difference across the two types, i.e. increase \( \frac{y_H}{y_L} \). This change will not affect the workers’ employment rates and therefore the equilibrium network structure, \( \phi^* \), will remain unchanged.\(^{16}\) The planner, however, will now prefer to increase the employment rate of high-types since they produce more output, even at the cost of higher unemployment for the low types. As a result the optimal network structure will

\(^{16}\)Of course, a high-type worker whose productivity has increased will want to create more links in order to increase his employment rate. In equilibrium, however, this will not affect the proportion of links that he creates with other high-types.
feature an increase in $\phi^P$ which increases the referral rate of high types and reduces the referral rate of low types. Ceteris paribus, the optimal network structure exacerbates inequality when productivity differentials are high.

Second, fix the productivity levels, $y_H$ and $y_L$, and consider an increase in the probability of forming a match for the high types relative to that of the low types. For simplicity, let $p = p_H = 1 - p_L$ and consider an increase in $p$. In equilibrium, this will lead to an increase in the relative employment rate of high-types and, hence, higher $\phi^*$ (Proposition 3.4). From the planner’s viewpoint, however, the relative merits of the employment of a high- vs. a low-type worker are the same as before and now low-type workers have lower chances of becoming employed. It turns out that the planner prefers to mitigate the increase in $p$ by reducing $\phi^P$ which will increase the referral rate of low types, at the expense of the high types’ referral rate. Ceteris paribus, the optimal network structure reduces inequality when the differentials in match probability are high. Notice that in this case the optimal network moves in the exactly opposite direction than the equilibrium network.

These results are stated formally in the proposition.

**Proposition 5.2** Assume that $\phi^P \in (0, 1)$. Then:

1. An increase in $\frac{y_H}{y_L}$ leads to an increase in $\phi^P$ and no change in $\phi^*$.  

2. Let $p = p_H = 1 - p_L$. An increase in $p$ reduces $\phi^P$ and increases $\phi^*$.  

**5.2 Discussion and policy implications**

These results provide some insights into the settings where the inefficiencies are likely to be more severe. Qualitatively, the worst scenario from an efficiency viewpoint occurs when the planner’s optimal network structure reduces inequality ($\phi^P < \frac{1}{2}$) while the equilibrium network structure increases inequality ($\phi^* > \frac{1}{2}$). Example 2 from Section 2.2 presents a model where this scenario occurs in equilibrium. This case is consistent with discrimination in the
labor market. By contrast, when productivity differentials are significant then the planner’s optimal network structure need not differ much from the equilibrium network structure (although, of course, they will not generically coincide). This occurs in Example 1 from Section 2.2 where heterogeneity is driven by productivity differentials.

The policy implications are nuanced. In a well-functioning market, referral networks do exacerbate pre-existing inequality but this need not reduce the efficiency of equilibrium, if the inequality across types is driven by productivity differentials. In such a case, the use of referrals has mostly benign effects as it reduces search and informational frictions.

However, if the distinction between the two types is not about productivity but, rather, about other features which nevertheless affect a worker’s probability of passing an interview, then social networks can have a very detrimental effect in magnifying that inequality, at the time when efficiency requires the exact opposite action. In such cases, discouraging the use of referrals or taking actions that promote the mixing across types might be beneficial. Furthermore, policies that increase the employment rate of the disadvantaged group might have high payoffs because they also lead to an adjustment in the network structure.

Using the model to determine which of the two polar cases best describes the actual data is much harder to do. The evidence that referred workers have higher probability of being hired and receiving higher wages are consistent with both examples (lower wages for the disadvantaged group are due to lower employment prospects, rather than lower productivity). Therefore, actual productivity measures are needed. Castilla (2005) and Burks, Cowgill, Hoffman and Housman (2015) provide such data for their firm-level studies and show that referrals are mostly associated with higher productivity. In a more detailed data-set, Hensvik and Skans (2015) have objective ability measures and suggests that referrals are, on average, given on merit at least for the population under study: men who completed the draft in Sweden. Whether this generalizes to women or immigrant populations is very much an open question.
6 Conclusions

This paper studies the implication of referral use for inequality and efficiency. I develop a model with heterogeneous workers where referral networks are endogenously formed and respond to labor market incentives. Consistent with empirical evidence, referred workers are more likely to be hired, to receive higher wages and to be more productive. The use of referrals benefits high-type workers most and exacerbates pre-existing inequality among workers.

The network is generically inefficient. If worker heterogeneity is due to productivity differentials, then the increased inequality improves efficiency. If, however, the correlation between the probability of forming a match and productivity is weak, then the optimal network delivers qualitatively opposite structure to equilibrium and inefficiencies are severe. The latter case is likely to occur in the presence of discrimination in the labor market.

This model can be readily extended in several directions. The network size does not currently play a crucial role but one might expect that a denser network might lead to more information transmission, thereby extending the advantage of well-networked high type workers. Furthermore, a model where high type workers refer more might create a force for sorting of types across firms, which has not been explored in the literature.

To study in greater depth the case where heterogeneity represents different social groups, one possibility is to allow for heterogeneous costs for linking within and across groups. If linking within one’s group is easier, then the ensuing network will balance the pecuniary with non-pecuniary incentives and will provide richer implications for network structure. This extension can replicate the homophily that is observed in sociological studies and examine the interaction with labor markets. A further topic of interest is the occupational specialization that is often observed across ethnic groups.
APPENDIX

A Proofs

Proposition 3.1.

Proof. Define:

\[ H(u_H, u_L) \equiv u_H \mu \left( \frac{v}{u_H + u_L} \right)^\eta + u_H \rho \left( \phi_H (1 - u_H) + (1 - \phi_L)(1 - u_L) \right) - \frac{\delta}{p_H} (1 - u_H) \]

\[ L(u_H, u_L) \equiv u_L \mu \left( \frac{v}{u_H + u_L} \right)^\eta + u_L \rho \left( \phi_L (1 - u_L) + (1 - \phi_H)(1 - u_H) \right) - \frac{\delta}{p_L} (1 - u_L) \]

and note that in a steady state \( H(u_H, u_L) = L(u_H, u_L) = 0 \) holds. Define \( h(u_L) \) and \( l(u_H) \) such that \( H(h(u_L), u_L) = 0 \) and \( L(u_H, l(u_H)) = 0 \).

Let \( H_x(u_H, u_L) \equiv \partial H(u_H, u_L) / \partial x \) and similarly for \( L(u_H, u_L) \) and observe that

\[
\begin{align*}
H(0, u_L) &= -\frac{\delta}{p_H} < 0 \\
H(1, u_L) &= \mu \left( \frac{v}{1 + u_L} \right)^\eta + \rho (1 - \phi_L)(1 - u_L) > 0 \\
H_{uH}(u_H, u_L) &= \mu \left( \frac{v}{u_H + u_L} \right)^\eta \left( 1 - \frac{\eta u_H}{u_H + u_L} \right) + \rho \left( \phi_H (1 - u_H) + (1 - \phi_L)(1 - u_L) \right) \\
&\quad + \frac{\delta}{p_H} - u_H \rho \phi_H > 0 \\
H_{uL}(v, u_H, u_L) &= -\eta u_H \mu \left( \frac{v}{u_H + u_L} \right)^\eta - u_H \rho (1 - \phi_L) < 0 \\
h'(u_L) &= -\frac{H_{uL}}{H_{uH}} > 0
\end{align*}
\]

The first two equations mean that \( h(u_L) \) exists and is in \((0, 1)\) for any \( u_L \in [0, 1] \). The third equation proves that \( h(u_L) \) is a function (not a correspondence) and the fourth equation proves that \( h'(u_L) > 0 \). And similarly for \( l(u_H) \).

Define \( T_1(u_H) = h(l(u_H)) \). A steady state is a fixed point of \( T_1(u_H) \). The results above
prove that $T_1(0) > 0$ and $T_1(1) < 0$ and therefore a fixed point exists. To prove the fixed point is unique it suffices to show that $T_1'(u_H) < 1$. We have:

$$T_1'(u_H) = h'(l(u_H))l'(u_H) = \frac{H_H(u_H, l(u_H))L_H(u_H, l(u_H))}{H_H(u_H, l(u_H))L_H(u_H, l(u_H))}$$

Notice that:

$$H_H(u_H, u_L) + L_H(u_H, u_L) = (1 - \eta)\alpha_M + \alpha_{RH} + \frac{\delta}{\rho_H} - \rho(\phi_H u_H + (1 - \phi_H)u_L) > 0$$

$$L_L(u_H, u_L) + H_L(u_H, u_L) = (1 - \eta)\alpha_M + \alpha_{RL} + \frac{\delta}{\rho_L} - \rho(\phi_L u_L + (1 - \phi_L)u_H) > 0$$

and therefore $T_1'(u_H) < 1$ and the steady state is uniquely defined.

Note that:

$$H_v = \frac{\eta u_H}{v} \mu\left(\frac{v}{u_H + u_L}\right)^\eta > 0$$

$$L_v = \frac{\eta u_L}{v} \mu\left(\frac{v}{u_H + u_L}\right)^\eta > 0$$

which imply that $\frac{dh(u_L)}{dv} < 0$ and $\frac{dl(u_H)}{dv} < 0$. Therefore $u_H$ and $u_L$ are decreasing in $v$.

The value of a vacancy is given by:

$$rV = \alpha_{FH}(1 - \beta)S_H + \alpha_{FL}(1 - \beta)S_L$$

(11)

where the $S_i$’s can be written as follows:

$$S_H = \frac{(y_H - b)D_{L1} + (y_L - b)D_{H2}}{D_{H1}D_{L1} - D_{H2}D_{L2}}$$

$$S_L = \frac{(y_L - b)D_{H1} + (y_H - b)D_{L2}}{D_{H1}D_{L1} - D_{L2}D_{H2}}$$
where

\[ D_{i1} = r + \delta + (\alpha_M + \alpha_{R_i})p_i \beta - \rho \gamma (1 - \beta) \phi_i u_i p_i, \]

\[ D_{i2} = \rho \gamma (1 - \beta)(1 - \phi_i) u_k p_k. \]

Note that \( D_{i1} \) is increasing in \( v \) and \( D_{i2} \) is decreasing in \( v \) and therefore \( S_i \) is decreasing in \( v \). Furthermore the steady state equations imply that

\[ v \to 0 \Rightarrow (u_H, u_L) \to (1, 1) \Rightarrow \alpha_{F_i} \to \infty, \]

\[ v \to \infty \Rightarrow (u_H, u_L) \to (0, 0) \Rightarrow \alpha_{F_i} \to 0. \]

These observations mean that a vacancy’s value is above \( k \) for \( v \) near zero and below \( k \) for \( v \) very large and, therefore, a labor market equilibrium exists. ■

**Proposition 3.2.**

**Proof.** The second and cross-derivatives of the worker’s objective function are:

\[ \frac{\partial^2 L_j^i}{\partial (e_{ii})^2} = \frac{\partial^2 \Lambda_j^i}{\partial (\alpha_{R_i})^2} (\frac{\rho (1 - u_i)}{n_i + n_{ik}})^2 - c \]

\[ \frac{\partial^2 L_j^i}{\partial (e_{ik})^2} = \frac{\partial^2 \Lambda_j^i}{\partial (\alpha_{R_i})^2} (\frac{\rho (1 - u_k)}{n_{kk} + n_{ki}} E_{ki} + E_{ik})^2 - c \]

\[ \frac{\partial^2 L_j^i}{\partial (e_{ii} e_{ik})} = \frac{\partial^2 \Lambda_j^i}{\partial (\alpha_{R_i})^2} \frac{\rho (1 - u_i) \rho (1 - u_k)}{n_i + n_{ik} n_{kk} + n_{ki}} \frac{E_{ki}}{E_{ki} + E_{ik}} \]

Proving that \( (\partial^2 \Lambda_j^i)/(\partial (\alpha_{R_i})^2) \) is negative suffices to show that the maximization problem has a unique equilibrium.

Recall that worker \( j \)’s unemployment rate and match surplus are defined by:

\[ u_j^i = \frac{\delta}{\delta + (\alpha_M + \alpha_{R_i})p_i} \]

\[ S_j^i = \frac{y_i - b + \rho \gamma X_i}{r + \delta + (\alpha_M + \alpha_{R_i})p_i \beta} \]
Rewrite worker $j$’s steady state utility as follows:

$$
\Lambda_i^j = u_i^j U_i^j + (1 - u_i^j) W_i^j \\
= (1 - u_i^j)(W_i^j - U_i^j) + \frac{1}{r} \left(b + (\alpha_M + \alpha_{Ri}^j) p_i (W_i^j - U_i^j)\right) \\
= \frac{b}{r} + \left((1 - u_i^j) + \frac{(\alpha_M + \alpha_{Ri}^j) p_i}{r}\right) \beta S_i^j
$$

Letting

$$
\alpha_i^j = (\alpha_M + \alpha_{Ri}^j) p_i
$$

and going through some algebra leads to:

$$
\Lambda_i^j = \frac{b}{r} + \beta \left(y_i - b + \rho \gamma X_i\right) \frac{\alpha_i^j (r + \delta) + (\alpha_i^j)^2}{\delta (r + \delta) + \alpha_i^j (r + \delta + \delta \beta) + (\alpha_i^j)^2 \beta}
$$

Differentiating $\Lambda_i^j$ with respect to $\alpha_{Ri}^j$ we have:

$$
\frac{\partial \Lambda_i^j}{\partial \alpha_{Ri}^j} = \frac{\beta(y_i - b + \rho \gamma X_i) p_i}{r (\delta (r + \delta) + \alpha_i^j (r + \delta + \delta \beta) + (\alpha_i^j)^2 \beta)^2} \left[(r + \delta + 2 \alpha_i^j)(\delta (r + \delta) + \alpha_i^j (r + \delta + \delta \beta) + (\alpha_i^j)^2 \beta) - (\alpha_i^j (r + \delta) + (\alpha_i^j)^2)(r + \delta + \beta \delta + 2 \alpha_i^j \beta)\right] \\
= \frac{\beta(y_i - b + \rho \gamma X_i) p_i}{r (\delta (r + \delta) + \alpha_i^j (r + \delta + \delta \beta) + (\alpha_i^j)^2 \beta)^2} \left[(r + \delta)^2 \delta + 2 \alpha_i^j \delta (r + \delta) + (\alpha_i^j)^2 (r(1 - \beta) + \delta)\right] > 0
$$

Define:

$$
D = \delta (r + \delta) + \alpha_i (r + \delta + \delta \beta) + (\alpha_i^j)^2 \beta
$$
Taking the second derivative of $\Lambda^j_i$ with respect to $\alpha^j_{Ri}$ leads to:

$$\frac{\partial^2 \Lambda^j_i}{\partial (\alpha^j_{Ri})^2} = \frac{\beta(y - b + \rho\gamma X_i)p_i}{rD^3} \left[ (2\delta(r + \delta) + 2\alpha_i(r(1 - \beta) + \delta)) (\delta(r + \delta) + \alpha_i(r + \delta + \delta\beta) + \alpha_i^2 \beta)^2 - (r + \delta)^2 \delta + \alpha_i 2\delta(r + \delta) + \alpha_i^2 (r(1 - \beta) + \delta)) 2D_i ((r + \delta + \delta\beta) + \alpha_i 2\beta) \right]$$

$$= \frac{2\beta(y - b + \rho\gamma X_i)p_i}{rD^2} \left[ -\delta(r + \delta)^2 (r + \beta\delta) - \alpha_i^j \delta(r + \delta) (\beta(r + \delta) + (r + \delta) 2\beta) - (\alpha_i^j)^2 3\delta\beta(r + \delta) - (\alpha_i^j)^3 (r(1 - \beta) + \delta) \beta \right] < 0$$

Therefore, an agent’s steady state utility is a strictly increasing and strictly concave function of his job finding rate.

As a result, the worker’s optimal effort level is given by the first order conditions of his optimization problem.

Proposition 3.3.

Proof. Using the definition for how effort leads into link creation:

$$\frac{e^j_{ii}}{e^j_{ik}} = \frac{E_{ik} + E_{ki}}{E_{ki}} n^j_{ii} \frac{n^j_{ik}}{n^j_{ki}}$$  \hfill (12)

Equate the first order conditions with respect to $e^j_{ii}$ and $e^j_{ik}$ with zero, rearrange and take their ratio:

$$\frac{e^j_{ii}}{e^j_{ik}} = \frac{1 - u_i n_{ki} + n_{kk} E_{ik} + E_{ki}}{1 - u_k n_{ik} + n_{ii} E_{ki}}$$  \hfill (13)

Combining equations (12) and (13):

$$\frac{E_{ik} + E_{ki}}{E_{ki}} n^j_{ii} n^j_{ik} = \frac{1 - u_i n_{ki} + n_{kk} E_{ik} + E_{ki}}{1 - u_k n_{ik} + n_{ii} E_{ki}}$$

$$\Rightarrow \frac{n^j_{ii}}{n^j_{ik}} = \frac{1 - u_i n_{ki} + \frac{n_{kk} E_{ik}}{1 - u_k n_{ik} + n_{ii}}}{1 - u_k n_{ik} + n_{ii}}$$

30
Recall that:

\[
\begin{align*}
n_{ji}^j &= \phi_i^j (n_{ii}^j + n_{ik}^j) \\
n_{ik}^j &= (1 - \phi_i^j)(n_{ii}^j + n_{ik}^j) \\
\Rightarrow \frac{n_{ii}^j}{n_{ik}^j} &= \frac{\phi_i^j}{1 - \phi_i^j}
\end{align*}
\]

Combining the above together with the equilibrium symmetry condition \(\phi_i^j = \phi_j^i\) for all \(j\) we have:

\[
\begin{align*}
\frac{\phi_H}{1 - \phi_H} &= \frac{1 - u_H n_{LL} + n_{LH}}{1 - u_L n_{HH} + n_{HL}} \\
\frac{1 - \phi_L}{\phi_L} &= \frac{1 - u_H n_{LL} + n_{LH}}{1 - u_L n_{HH} + n_{HL}}
\end{align*}
\]

which proves part 2.

Consistency requires \(n_{HL} = n_{LH}\) which implies that:

\[
\begin{align*}
(n_{HH} + n_{HL})(1 - \phi_H) &= (n_{LL} + n_{LH})(1 - \phi_L) \\
\Rightarrow \frac{n_{LL} + n_{LH}}{n_{HH} + n_{HL}} &= \frac{1 - \phi_H}{1 - \phi_L} = \frac{1 - \phi_H}{\phi_H}
\end{align*}
\]

Combining equations (14) and (15) proves part 3.

The equilibrium is characterized by \((v, u_H, u_L, \phi)\), where \(\phi = \phi_H = 1 - \phi_L\), which satisfy part 3, the steady state and free entry conditions. Given \(v\) (e.g. at its equilibrium value) this is determined by the root of the following equation:

\[
T_2(\phi) = \phi^2 (1 - u_L(\phi)) - (1 - \phi)^2 (1 - u_H(\phi))
\]

where \(u_H(\phi)\) and \(u_L(\phi)\) are defined by the steady state conditions \(H(u_H, u_L, \phi) = L(u_H, u_L, \phi) = 0\) (with a slight abuse of notation).
I show that $T_2(\phi) = 0$ has a unique solution at $\phi^*$. Note that:

\[
T_2(0) = -(1 - u_H(0)) < 0 \\
T_2(1) = 1 - u_L(1) > 0 \\
T_2'(\phi) = 2\phi(1 - u_L(\phi)) + 2(1 - \phi)(1 - u_H(\phi)) - \phi^2 u_L'(\phi) + (1 - \phi)^2 u_H'(\phi)
\]

\[
= \frac{(1 - \phi)^2}{\phi} \left( 2\phi^2(1 - u_L(\phi)) + \phi u_H'(\phi) \right) + \frac{\phi^2}{1 - \phi} \left( 2(1 - \phi)^2(1 - u_H(\phi)) - (1 - \phi) u_L'(\phi) \right)
\]

It suffices to show that $T_2'(\phi) > 0$ when $T(\phi) = 0$. Using $T(\phi) = 0$ we can rewrite:

\[
T_2'(\phi) = \frac{(1 - \phi)^2}{\phi} \left[ 2(1 - u_H(\phi)) + \phi u_H'(\phi) \right] + \frac{\phi^2}{1 - \phi} \left[ 2(1 - u_L(\phi)) - (1 - \phi) u_L'(\phi) \right]
\]

Using implicit differentiation we have:

\[
\begin{align*}
u_H'(\phi) &= -\frac{L_{uL}H_{\phi} - H_{uL}L_{\phi}}{H_{uH}L_{uL} - H_{uL}L_{uH}} \\
u'_L(\phi) &= -\frac{H_{uH}L_{\phi} - L_{uH}H_{\phi}}{H_{uH}L_{uL} - H_{uL}L_{uH}}
\end{align*}
\]

We examine the square brackets in equation (16) separately:

\[
2(1 - u_H) + \phi u_H' = \frac{1}{\Delta} \left( L_{uL} \left( 2(1 - u_H)H_{uH} - u_H\alpha_{RH} \right) - H_{uL} \left( 2(1 - u_H)L_{uH} + \alpha_{RH} u_L \right) \right)
\]

where $\Delta = H_{uH}L_{uL} - H_{uL}L_{uH} > 0$.

Recalling that $L_{uL} + H_{uL} > 0$ it suffices to show that:

\[
2(1 - u_H)H_{uH} - u_H\alpha_{RH} + 2(1 - u_H)L_{uH} + \alpha_{RH} u_L > 0 \\
\Rightarrow \alpha_M 2(1 - u_H)(1 - \eta) + \alpha_{RH} (2(1 - u_H) - u_H + u_L) + 2(1 - u_H) \left( \frac{\delta}{p_H} - \rho \phi u_H - \rho(1 - \phi) u_L \right) > 0
\]

which is positive when $u_H \leq \frac{1}{2}$. 

32
Similarly:

\[ 2(1 - u_L) - (1 - \phi)u_L' = \frac{1}{\Delta} \left( H_{uH} (2(1 - u_L)L_{uL} - u_L\alpha_{RL}) - L_{uH} (2(1 - u_L)H_{uL} + \alpha_{RL}u_H) \right) \]

which is positive because \( H_{uH} + L_{uH} > 0 \) and

\[ \alpha_M 2(1 - u_L)(1 - \eta) + \alpha_{RL} (2(1 - u_L) - u_L + u_H) + 2(1 - u_L)\left( \frac{\delta}{p_L} - u_L\rho(1 - \phi) - u_H\rho\phi \right) > 0 \]

which is positive if \( u_L \leq \frac{1}{2} \). Therefore, given \( v, (\phi, u_H, u_L) \) are unique.

Furthermore notice that:

\[ T_2 \left( \frac{1}{2} \right) = \frac{1}{4}(u_H - u_L) \]

When \( \phi = \frac{1}{2} \) we have \( \alpha_{RH} = \alpha_{RL} \) and therefore \( u_H < u_L \Leftrightarrow p_H > p_L \). Therefore:

\[ p_H = p_L \Rightarrow u_H = u_L \Rightarrow T_2 \left( \frac{1}{2} \right) = 0 \Rightarrow \phi^* = \frac{1}{2} \]

\[ p_H > p_L \Rightarrow u_H < u_L \Rightarrow T_2 \left( \frac{1}{2} \right) < 0 \Rightarrow \phi^* > \frac{1}{2} \]

This proves part 1 and completes the proof. ■

**Proposition 3.4.**

**Proof.** Note that:

\[ \frac{dT_2(\phi)}{dp} = -\phi^2 \frac{du_L(\phi)}{dp} + (1 - \phi)^2 \frac{du_H(\phi)}{dp} \]

Using implicit differentiation we have:

\[ \frac{du_H(\phi)}{dp} = \frac{L_{uL} H_p - H_{uL} L_p}{H_{uH} L_{uL} - H_{uL} L_{uH}} \]

\[ \frac{du_L(\phi)}{dp} = \frac{H_{uH} L_p - L_{uH} H_p}{H_{uH} L_{uL} - H_{uL} L_{uH}} \]

33
where:

\[ H_p = (\alpha_M + \alpha_{RH})u_H \]
\[ L_p = -(\alpha_M + \alpha_{RL})u_L \]

Therefore \( \frac{du_H}{dp} < 0 < \frac{u_L}{dp} \) and \( \frac{dT_2(\phi)}{dp} < 0 \) which implies \( \frac{d\phi^*}{dp} > 0. \) □

**Proposition 4.1.**

**Proof.** In a meeting through the market the probability that the worker is of high type is:

\[ P[H|\text{market}] = \frac{u_H}{u_H + u_L} \]

Denote the rate that referral are generated from high type workers and are received by high and low type unemployed workers by \( R_{HH} \) and \( R_{HL} \), respectively. We have:

\[ R_{HH} = \rho(1 - u_H)\phi Hu_H \]
\[ R_{HL} = \rho(1 - u_H)(1 - \phi_H)u_L \]

The rate at which referrals are generated from low type workers is defined by \( R_{LH} \) and \( R_{LL} \) in an equivalent way.

In a meeting through referrals, the probability that the referred worker is of a high type is:

\[ P[H|\text{referral}] = \frac{R_{HH} + R_{LH}}{R_{HH} + R_{HL} + R_{LH} + R_{LL}} \]
\[ = \frac{[(1 - u_H)\phi_H + (1 - u_L)(1 - \phi_L)]u_H}{[(1 - u_H)\phi_H + (1 - u_L)(1 - \phi_L)]u_H + [(1 - u_H)(1 - \phi_H) + (1 - u_L)\phi_L]u_L} \]
Noting that
\[(1 - u_H)\phi_H + (1 - u_L)(1 - \phi_L) \geq (1 - u_H)(1 - \phi_H) + (1 - u_L)\phi_L\]
proves that \(P[H|\text{referral}] > P[H|\text{market}]\). ■

**Proposition 5.1.**

**Proof.** The problem becomes:

\[
\max_{\phi} W(\phi) = y_H + y_L - u_H(\phi)(y_H - b) - u_L(\phi)(y_L - b)
\]

where \(u_H(\phi)\) and \(u_L(\phi)\) are defined by the steady state conditions.

Differentiating and going through some algebra (\(\Delta = H_u L_u - H_u L_h\)):

\[
W'(\phi) = -u'_H(\phi)\bar{y}_H - u'_L(\phi)\bar{y}_L
\]

\[
= -\frac{\rho(2 - u_H - u_L)u_H u_L}{\Delta} \left[ \frac{(y_H - b)\delta}{p_L u_L^2} - \frac{(y_L - b)\delta}{p_H u_H^2} - (\bar{y}_H - \bar{y}_L) \left( \frac{2\eta \alpha_M}{u_H + u_L + \rho} \right) \right]
\]

Let \(T_3(\phi)\) denote the term inside the square brackets and note that the planner’s solution is given by \(T_3(\phi^*) = 0\). I show that \(T_3'(\phi) < 0\) when \(T_3(\phi)\), which suffices to show that \(T_3(\phi)\) has at most one root.

Differentiating with respect to \(\phi\):

\[
T'_3(\phi) = -\frac{2(y_H - b)\delta}{p_L u_L^2} \frac{du_L}{d\phi} + \frac{2(y_L - b)\delta}{p_H u_H^2} \frac{du_H}{d\phi} + (y_H - y_L) \frac{2\eta \alpha_M(1 + \eta)}{(u_H + u_L)^2} \left( \frac{du_H}{d\phi} + \frac{du_L}{d\phi} \right)
\]

\[
= \frac{2}{\rho} \frac{du_H}{d\phi} \left[ \frac{(y_L - b)\delta}{p_H u_H^2} + (y_H - y_L) \frac{\eta \alpha_M(1 + \eta)}{(u_H + u_L)^2} \right] - 2 \frac{du_L}{d\phi} \left[ \frac{(y_L - b)\delta}{p_H u_H^2 u_L} + (y_H - y_L) \frac{\rho}{u_L} \right]
\]

\[
+ (y_H - y_L) \frac{\eta \alpha_M}{u_L(u_H + u_L)} \left( 2 - \frac{(1 + \eta)u_L}{u_H + u_L} \right)
\]

where we used the condition for \(T_3(\phi) = 0\). Notice that both terms in the square brackets
are positive. Furthermore:

\[
\frac{du_H}{d\phi} = -\frac{u_H \rho (2 - u_H - u_L)}{\Delta} \left[ \alpha_M \left(1 - \frac{2\eta u_L}{u_H + u_L}\right) + \alpha_{RL} + \frac{\delta}{p_L} - u_L \rho \right] < 0
\]

\[
\frac{du_L}{d\phi} = \frac{u_L \rho (2 - u_H - u_L)}{\Delta} \left[ \alpha_M \left(1 - \frac{2\eta u_H}{u_H + u_L}\right) + \alpha_{RH} + \frac{\delta}{p_H} - u_H \rho \right] > 0
\]

and we have proved that \( T'_3(\phi) < 0 \) when \( T_3(\phi) = 0 \).

As a result there is a unique \( \phi^P \) that solves the planner’s problem. We have:

\[
T_3(0) \leq 0 \quad \Rightarrow \quad \phi^P = 0
\]

\[
T_3(0) > 0 \quad \text{and} \quad T_3(1) \geq 0 \quad \Rightarrow \quad \phi^P = 1
\]

\[
T_3(0) > 0 \quad \text{and} \quad T_3(1) < 0 \quad \Rightarrow \quad \phi^P \in (0, 1) \quad \text{and} \quad T(\phi^P) = 0
\]

If \( y_H = y_L \) and \( p_H = p_L \) then \( T_3(\phi) = 0 \leftrightarrow u_H = u_L \). Therefore the planner’s solution is \( \phi^P = \frac{1}{2} = \phi^* \). For general values of \( y_i \) and \( p_i \), note that \( T_3(\phi) \neq T_2(\phi) \) and therefore \( \phi^P \neq \phi^* \).

\[\blacksquare\]

**Proposition 5.2.**

**Proof.** It is easy to see that:

\[
\frac{dT_3(\phi)}{d\bar{y}_H} > 0 > \frac{dT_3(\phi)}{d\bar{y}_L}
\]

and therefore when \( \phi^P \) is interior we have:

\[
\frac{d\phi^P}{d\bar{y}_H} > 0 > \frac{d\phi^P}{d\bar{y}_L}
\]

Suppose that \( p = p_H = 1 - p_L \). To see how \( T'_3(\phi) \) is affected by a change in \( p \), simply note
that we should replace $\frac{du_i}{d\phi}$ with $\frac{du_i}{dp}$. Note that:

$$\frac{du_H}{dp} = -\frac{1}{\Delta}(L_{u_L}H_p - H_{u_L}L_p)$$

$$\frac{du_L}{dp} = -\frac{1}{\Delta}(H_{u_H}L_p - L_{u_H}H_p)$$

where:

$$H_p = (\alpha_M + \alpha_{RH})u_H$$

$$L_p = -(\alpha_M + \alpha_{RL})u_L$$

Therefore:

$$\frac{du_H}{dp} < 0 < \frac{du_L}{dp} \Rightarrow \frac{\partial T_3(\phi)}{\partial p} < 0 \Rightarrow \frac{d\phi^p}{dp} < 0$$

This completes the proof. ■

References


