Referral Networks and Inequality*

Manolis Galenianos†

June 2020

Abstract

This is the first paper to study the interaction between labour markets and endogenous referral networks in the context of worker heterogeneity. In equilibrium the structure of the referral network is hierarchical, which is different from the usual assumption of homophily but is consistent with the evidence. Hierarchy exacerbates inequality. The welfare effects of the use of referrals are subtle and depend on the nature of heterogeneity. If heterogeneity is due to productivity differences, referrals improve welfare. If workers face different probability of forming a match despite having the same productivity, as in the case of discrimination, referrals reduce welfare.

Keywords: referral networks; endogenous networks; inequality; search; matching

JEL codes: J01; J64; J71

*I would like to thank Rachel Kranton, the editor, and four anonymous referees for comments that improved the paper. I also thank Philipp Kircher, Giuseppe Moscarini, Oskar Skans and seminar and conference participants for useful discussions.

†Royal Holloway, University of London. Egham Hill, Egham TW20 0EX, UK. Email: manolis.galenianos@gmail.com

1 Introduction

It is well-known that personal contacts play an important role in labour markets: approximately half of all workers report finding their jobs through referrals (Granovetter, 1995).[1] Perhaps surprisingly, however, there is no consensus about referrals' effect on welfare. On the one hand some authors have argued that the use of referrals may increase inequality by benefitting the networked at the expense of qualified but less well-connected workers.[2] On the other hand several empirical studies find that hiring through referrals yields more productive workers, suggesting that referrals alleviate some of the informational frictions that pervade labour markets.[3]

These views are not necessarily incompatible with each other; rather, they suggest that determining the welfare consequences of referrals requires taking into account several factors which could individually point in different directions. The first such factor is the importance of worker heterogeneity and the difficulties firms face in finding desirable workers. The second factor is that workers and firms conduct their search through both formal and informal channels. The third, and most crucial, factor is that referral networks are the outcome of choices made by workers, rather than some exogenous and immutable feature of the economic or social environment. Taking into account the endogenous nature of referral networks is the key for understanding the potential of referrals to alleviate frictions, to exacerbate inequality, and to affect overall welfare.

In this paper I develop the first model that studies the interaction between worker heterogeneity and endogenous referral networks in labour markets. In the model there are two

[1] The surveys of Ioannides and Loury (2004) and Topa (2011) provide a wealth of additional evidence.

[2] Calvo-Armengol and Jackson (2004) make this point theoretically in a model where people hear about job opportunities from their network neighbours. Empirically, Topa (2001) finds strong local spillovers in unemployment rates across geographical locations in Chicago, Kugler (2003) finds that referrals are associated with higher wage inequality among equally productive workers and Beaman et al. (2018) present evidence that the use of referrals reinforces unequal access to jobs between men and women.

[3] Castilla (2005), Burks et al. (2015) and Barr et al. (2019) examine direct measures of productivity on the job and report that referred workers are more productive than non-referred workers. Hensvik and Skans (2016) show that referred workers have higher measures of cognitive ability than non-referred workers, conditional on observable characteristics.

worker types, A and B. The workers form a referral network among them at an initial stage and, subsequently, firms and workers interact in a frictional labour market. In the labour market vacancies are created in two ways: a new firm enters or a producing firm expands. Firms and workers meet either through a referral, which occurs when a producing firm expands and asks its current employee to refer a member of his network, or through search in the market. At a meeting, the probability that a match is formed depends on the worker's type. Type-A workers dominate type-Bs in the probability of forming a match when meeting a firm and, potentially, in productivity on the job.

Despite the model's complexity, the equilibrium admits a sharp characterization. In equilibrium workers *of both types* have most of their links with type-A workers and, furthermore, type-A workers have more links and enjoy a higher arrival rate of referrals. Therefore, the equilibrium network has a hierarchical structure where type-A workers reap most of the benefits from the use of referrals.[4] The hierarchical structure is a novel result and contrasts with the literature's usual assumption of (exogenously-given) homophilous structure for referral networks (Montgomery, 1991). An important positive implication of the network's hierarchical structure is that it exacerbates inequality because, under such a structure, most referrals are received by type-A workers who have better labour market prospects to begin with.

The effects of referrals on welfare are subtle and depend on the nature of worker heterogeneity. In a labour market where worker heterogeneity is due to productivity differences, which I will call a *meritocratic* labour market, the optimal network favours the highly-productive type-As and, therefore, the equilibrium structure of the network enhances welfare despite increasing inequality. When, however, workers of similar productivity face consistently different probabilities of being hired, which I will call a *discriminatory* labour market, the optimal network favours type-B workers and *reduces* inequality, which is the exact opposite outcome from that of the equilibrium network. In that setting the use of referrals exacerbates the effects of discrimination and is strictly detrimental to welfare. Policies that

[4] I use the term "hierarchical structure" in its literal meaning rather than as a graph-theoretic term.

promote hiring through formal channels, rather than referrals, can alleviate (though not resolve) some of the consequences of discrimination. Overall, the optimal network is generically different from the equilibrium network.

The sociology literature has long documented that social networks are homophilous (McPherson et al., 2001) which appears to be at odds with this model's theoretical prediction of a hierarchical structure. However, there is no contradiction in principle: the sociology findings refer to social networks which need not be identical to referral networks. To evaluate the empirical relevance of the hierarchy prediction, I characterise the labour market equilibrium under a homophilous and a hierarchical network structure, derive an empirical test that can distinguish between the two referral network structures and use the test to re-evaluate the evidence published in two very detailed empirical studies on referrals. I find that the evidence in Hensvik and Skans (2016) and Beaman et al. (2018) is consistent with a hierarchical referral network structure in a meritocratic and a discriminatory labour market, respectively.

This evidence is particularly striking in Beaman et al. (2018) who run a field experiment in Malawi to explore the effect of referrals on gender inequality. They find that referrals from women are mostly sent to men despite the gender-homophily of social networks and similar ability levels between men and women, thereby exacerbating, rather than alleviating, gender imbalances in the labour market. This finding is hard to understand if one equates social and referral networks but is easily interpretable through the lens of the present model: the referral network is hierarchical with men at the top of the hierarchy.

The theoretical literature that studies the interaction between social networks and economic activity is vast and I will focus on the subset that relates to labour markets.[5] The

[5] Some strands of the literature study the effect of referrals on labour market outcomes without explicitly modelling the underlying network. One line of work posits that referrals help firms screen for better match quality among ex ante homogeneous workers, e.g. Simon and Warner (1992), Galenianos (2013), Brown et al. (2016) and Dustmann et al. (2016). Bentolila et al. (2010) examine the possibility that referrals lead to mismatch by distorting workers' occupational choice decisions. Some recent work examined whether referrals alleviate ex post moral hazard problems but the evidence is mixed: Heath (2018) finds supportive evidence for this channel but Pallais and Sands (2016) find that referrals do not generate a positive effort effect. The

interaction between heterogeneous workers and referrals was first studied by Montgomery (1991) who assumes an exogenous and homophilous network. In that paper using referrals helps firms hire high-productivity workers but there is no exploration of the welfare properties of equilibrium or the determinants of the network's structure.

In the following papers workers are assumed to be homogeneous in their productivity and the role of referrals is to facilitate the search process. Calvo-Armengol and Jackson (2004) examine the role of (exogenous) network structure in sustaining inequality. Calvo-Armengol (2004) and Galeotti and Merlino (2014) endogenise the network and characterise equilibrium outcomes such as network density and architecture. Calvo-Armengol and Zenou (2005) and Cahuc and Fontaine (2009) study the welfare consequences of using informal search methods and find that network use is inefficient. Mortensen and Vishwanath (1994) and Arbex et al. (2019) examine the effect of heterogeneous access to referrals on the wage distribution and Igarashi (2016) on the welfare consequences of banning referrals. Galenianos (2014) studies the effect of heterogeneous intensity of referrals on matching efficiency.

The paper is organised as follows. Section 2 presents the model, Section 3 characterises the equilibrium and derives its positive and normative predictions and Section 4 discusses the evidence on hierarchy and homophily. Section 5 concludes. Proofs are collected in the appendix.

2 The Model

Workers and firms populate the economy. A free entry condition determines the measure of firms. There is measure 1 of type-A workers and measure 1 of type-B workers. Each worker has a *market* characteristic and a *social* characteristic. The market characteristic determines the worker's interaction with firms, it is exogenous and it depends on the worker's type. The social characteristic is the worker's referral network and it is determined endogenously within

present work focuses on the structure of the referral network and will not examine these channels.

the model.[6] There are two distinct stages in the model. In the first stage workers form a referral network. In the second stage workers and firms search and produce in a frictional labour market.

2.1 Network Formation

The formation of the referral network is modelled as a large, non-cooperative game with non-transferable utility. A worker's referral network consists of the measure of links that he has with type-A and type-B workers. Let n^j_{ik} denote the measure of links that worker j of type i has with type-k workers; worker j's referral network is $(n^j_{iA}, n^j_{iB}) \in [0,1] \times [0,1]$. The focus is on equilibria with symmetric referral networks and a referral network is symmetric if $n^j_{ik} = n_{ik}$ for all j, i, k.[7]

Define the *homophily rate* of a type-i worker, ϕ_i, to be the proportion of links that he has with workers of his own type. Therefore:

$$\phi_i = \frac{n_{ii}}{n_{ii} + n_{ik}}. \tag{1}$$

A referral network is homophilous when $\phi_i \geq 1/2$ for $i \in \{A, B\}$ and hierarchical when $\phi_i \geq 1/2 \geq \phi_k$ for $i \neq k$.[8]

Modelling a worker's network as a continuum is consistent with the (spirit of the) sociology literature's finding that job-seekers receive referrals mostly from their (relatively numerous) acquaintances rather than their (relatively fewer) close friends (Granovetter, 1995).[9] Furthermore, as a modelling device it is crucial for tractability: a worker's employment opportunities

[6] A worker's referral network represents the set of workers that he might refer, and be referred by, if a job opportunity arises, rather than the full set of people with whom he interacts socially. Therefore a worker's referral network might differ from his social network in substantive ways. Section 4 discusses in more detail the distinction between referral and social networks and its empirical implications.

[7] I will restrict attention to equilibria in symmetric network-formation effort, defined below, which lead to symmetric referral networks.

[8] For completeness, a network is heterophilous when $\phi_i \leq 1/2$ for $i \in \{A, B\}$.

[9] In his classic study Granovetter (1995) finds that more than 80% of respondents who found their job through a personal contact interacted with their referrer "occasionally" or "rarely".

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, due to the law of large numbers, the aggregate (un)employment rate of a worker's network is deterministic (and constant, in steady state), thereby greatly simplifying the analysis.

Of course, a continuum model cannot replicate some of the richness in network architecture that can be found in graph-theoretic models of networks. From the perspective of the present paper, the benefits of explicitly considering such richness are limited because there is currently little, if any, evidence about how network architecture affects labour market outcomes. Therefore, while potentially significant in some settings, the limitations inherent in a continuum model are not particularly salient for this study.

The measure of links that a worker acquires depends on the effort he exerts in network formation and the aggregate effort of all other workers. Workers simultaneously choose their effort levels. Denote the effort of worker j of type i by (e^j_{iA}, e^j_{iB}) where $e^j_{ik} \geq 0$ is the effort he exerts in linking with type-k workers. The aggregate effort of all workers is denoted by $(E_{AA}, E_{AB}, E_{BA}, E_{BB})$ where E_{ik} denotes the effort of type-i workers towards linking with type-k workers and will be labelled the i-types' "demand" for such links (equivalently, E_{ki} is the "supply" of such links by k-type workers). I restrict my analysis on equilibria where $e^j_{ik} = E_{ik}$ for all i, j, k, which leads to referral networks that are symmetric.

Effort is mapped to links according to two link-formation equations. 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. Therefore:

$$n^j_{ii}(e^j_{ii}) \quad = \quad e^j_{ii} \; I[E_{ii} > 0], \tag{2}$$

where the indicator function takes a value of 1 if $E_{ii} > 0$ and 0 if $E_{ii} = 0$.

The measure of links between workers of different types needs to respect aggregate consistency, i.e. the A-types' measure of links with B-type workers is equal to the measure of links that B-types have with A-type workers: $n_{AB} = n_{BA}$ (recall that the economy is populated by equal measures of A- and B-workers). The measure of links as a function of effort is:

$$n^j_{ik}(e^j_{ik}) \quad = \quad e^j_{ik} \frac{E_{ki}}{E_{ik} + E_{ki}}, \tag{3}$$

if $E_{ik} > 0$ or $E_{ki} > 0$ and zero otherwise.

The key features of the link-formation equations (2) and (3) are the following. The number of links that worker j of type-i acquires with i-type and k-type workers is increasing linearly in j's own effort (e^j_{ii} and e^j_{ik}), conditional on positive demand for such links ($E_{ii} > 0$ and $E_{ki} > 0$, respectively). The slope of the within-type link-formation equation is one, while the slope of the across-type link-formation equation is always less than one; therefore, less effort is required to acquire a link with a worker of the same type than a worker of a different type, a natural property in the context of heterogeneity.[10] The slope of the across-type link equation decreases in the relative congestion faced by worker j's type (E_{ik}/E_{ki}), which is an intuitive feature. Changes in E_{ii} do not affect congestion in the within-type link-formation process because higher aggregate effort increases both the demand for and the supply of within-type links.

I assume that the network formation cost is quadratic in effort:

$$C_i\big(e^j_{ii}, e^j_{ik}\big) \quad = \quad \frac{c}{2}\Big((e^j_{ii})^2 + (e^j_{ik})^2\Big),$$

where c is high enough to guarantee an interior solution. The substantive implication of this cost structure is that the net returns of exerting effort to link with each type is positive for

[10]Note that this feature does not dominate equilibrium outcomes: if it did, then the resulting referral network would be homophilous, which turns out not to be the case in equilibrium. Therefore, it is the benefits of forming links with particular types that derive from the labour market (and is analysed below), rather than the details of the link-formation process, which determine the equilibrium structure of the referral network.

low effort and is diminishing in effort, which will lead to interior effort choice in equilibrium. The qualitative results of the model do not depend on the details of how diminishing net returns are modelled, though my specific assumptions are analytically convenient.

Denote the steady state utility in the labour market by $\Lambda_i(n^j_{iA}, n^j_{iB})$.[11] Worker j chooses effort to maximise:

$$\mathcal{L}_i(e^j_{iA}, e^j_{iB}) = \Lambda_i\big(n_{iA}(e^j_{iA}), n_{iB}(e^j_{iB})\big) - \frac{c}{2}\Big((e^j_{ii})^2 + (e^j_{ik})^2\Big). \tag{4}$$

2.2 The Labour Market

Time runs continuously, the horizon is infinite, the discount rate is $r > 0$ and the labour market is in steady state. Firms are homogeneous, risk-neutral and maximise expected discounted profits. Each firm hires one worker and a firm is either filled and producing or vacant and searching for a worker, where k denotes the flow cost of a vacancy. Workers are heterogeneous in their market and, potentially, social characteristics and maximise expected discounted utility. A worker is either employed and producing or unemployed and searching for a firm. The flow utility of employment equals the wage w and the flow utility of unemployment is $b > 0$. Matches are destroyed at rate δ.

2.2.1 Match formation and production

A worker's market characteristic has two dimensions: his probability of forming a match when meeting with a firm (p) and his productivity when employed (y). Type-A workers are weakly more productive ($y_A \geq y_B$) and have strictly higher probability of forming a match ($p_A > p_B$) than type-B workers.[12] B-types generate positive surplus when employed ($y_B > b$). All payoff-relevant variables (e.g. the worker's type and network) are common

[11]In the labour market a worker transits between employment and unemployment. His steady state utility is calculated using the proportion of time that he spends at each labour market state which, in general, depends on his network.

[12]The case where $p_A = p_B$ will be considered as a point of comparison for some of the characterization results but the applications of the model focus on $p_A > p_B$.

knowledge when the match is formed and the wage is determined through Nash bargaining.

The two dimensions of the market characteristic provide a parsimonious way of capturing a variety of economic environments. To fix ideas, I provide examples of two very different environments, corresponding to a meritocratic and a discriminatory labour market, which can be represented in this framework by an appropriate choice of parameters. These examples correspond to the two empirical studies that I will discuss in section 4 and will turn out to have very different implications for welfare in section 3.3.

Example 1: Meritocratic Labour Market. When a firm and a type-i worker meet they draw the match-specific productivity from some distribution $F_i(\cdot)$, where the distribution of A-type workers first order stochastically dominates that of B-types. The worker and firm observe match quality and the match is formed if the productivity draw is above an endogenous cutoff. Under standard restrictions on $F_i(\cdot)$, such as log-concavity, a type-A worker is more likely to make a draw above his cutoff than a B-type ($p_A > p_B$) and, conditional on forming a match, his productivity is higher on average than that of an employed B-type ($y_A > y_B$). In this example the market characteristic captures worker ability (possibly, productivity differences beyond observable characteristics) and it is consistent with the labour market for high-skill Swedish men, as described in Hensvik and Skans (2016).

Example 2: Discriminatory Labour Market. When a firm and a worker meet they draw match quality which is good with probability $\bar{p}$ and bad with probability $1-\bar{p}$ for both worker types. A good match produces $\bar{y}$ and a bad match produces $\underline{y} < b$ for both types. Bad matches are never formed (as their surplus is negative), good matches with A-types are always formed and good matches with B-types are formed with probability $\zeta < 1$. These assumptions lead to $y_A = y_B = \bar{y}$ and $p_A = \bar{p} > \bar{p}\zeta = p_B$. In this example the market characteristic captures the effects of discrimination since the differences in the probability of forming a match is not driven by productivity differentials,[13] and it is consistent with the

[13] I also abstract away from within-type productivity differences.

labour market for men and women in Malawi, as described in Beaman et al. (2018).[14]

I provide a rough outline of two ways of rationalizing $\zeta < 1$ that are consistent, respectively, with statistical and taste discrimination (a more detailed analysis is beyond the scope of this paper). First, suppose that match quality is uncertain when the firm and worker meet and the firm draws a binary signal about match quality. For A-type workers the signal is always accurate and for B-type workers it is always accurate if match quality is bad but it is accurate with probability ζ if match quality is good.[15] In equilibrium a match is only formed after a good signal, so long as ζ is not too low. Second, suppose that a share $1-\zeta$ of recruiters incur high disutility from hiring B-workers regardless of match quality and, therefore, B-workers are hired only if match quality is good *and* they meet a non-discriminatory recruiter.[16]

2.2.2 Search through the market and referrals

Unemployed workers and vacancies meet through the market or through a referral. The meeting rate through the market is determined by a standard meeting function. The meeting rate through a referral depends on firms' expansion rate, as described below.

Vacancies are created in two ways: (1) a new firm enters (free entry); (2) a producing firm expands at exogenous rate ρ, where $\rho < \delta$. When a producing firm expands, the incumbent employee refers one of his links at random to fill the new vacancy. If the referred worker is unemployed, he meets with the firm and they form a match with the type-specific probability. If the match is not formed or the referred worker is employed, the new vacancy loses the referral opportunity and begins searching in the market, together with the vacancies that were created through free entry. Following an expansion, the new position is sold off,

[14]Since the two worker types are in equal shares in the population the model needs to be extended in order to study discrimination against ethnic or racial minorities.

[15]Statistical discrimination has often been modelled to arise due to less accurate productivity signals regarding the discriminated group, e.g. Aigner and Cain (1977).

[16]In this example discrimination might take place at any level of the recruiting process and, therefore, does not need to be identified with the firm's owner or the firm's policy.

keeping firms' employment at one worker.[17]

Three remarks are in order regarding the referral process. First, the firm benefits from the referral in (potentially) two ways: it samples a worker immediately without having to go through time-consuming search in the market, which can only add to the firm's value; furthermore, the worker is drawn from a different type distribution than the unemployment pool which might, and in equilibrium will, turn out to be better than the type distribution among the unemployed (see equation (5) below). Therefore the firm always prefers to use a referral rather than forgo the opportunity and there is no loss in generality in assuming that a referral always takes place after an expansion.

Second, although the referring worker does not select a particular worker type from his network, he does choose the type composition of his network. In other words, the worker's strategic considerations about who to refer are incorporated in the network formation stage rather than the referral stage and this is, essentially, an assumption about the timing rather than the substance of decisions.[18]

Third, the assumption that the referring worker does not distinguish between employed and unemployed workers when sending a referral is consistent with the evidence that the referral network mostly consists of one's acquaintances, rather than one's close friends, and hence the referring employee might not know which of his links is currently looking for a job.[19] The practical implication of this assumption is that the referral channel is subject to frictions that are qualitatively similar to those of the market channel in that, for example, the probability that the firm meets with a worker through that channel is increasing in the

[17]The principal motivation for this assumption is to keep the model tractable. In the data, small firms use referrals at a higher rate than large firms which means that the relation between firm size and referrals is not mechanistic. See Galenianos (2013) for a model that rationalizes size-dependent referral use.

[18]The model can be reformulated so that the network-formation and referral decisions are made, and associated costs are borne, during the labour 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 time-varying employment status.

[19]Granovetter (1995) finds that in almost 60% of cases where a job was found through a referral, contact was initiated by the job-searcher's link and in half of those cases the link was not aware whether the job-searcher was actually interested in a job.

unemployment rate.

A referral from a type-i worker is received by 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_k is the unemployment rate of type-k workers and ϕ_i is the homophily rate of type-i workers, defined in equation (1). The composition of worker types in the referral pool will generally differ from that in the unemployment pool.

Denoting the value of employing a type-i worker by J_i and the value of a vacancy by V, the surplus generated by an expansion when employing a type-i worker is equal to:

$$X_i = V + \phi_i u_i p_i (J_i - V) + (1-\phi_i) u_k p_k (J_k - V). \tag{5}$$

The flow value to a match between a firm and a type-i worker is $y_i + \rho\gamma X_i$, where $\gamma \in [0,1]$ is a parameter that determines the share of the expansion's surplus that is received by the incumbent match. A firm that enters the labour market through free entry has value V (the value of a vacancy) and the second and third term of equation (5) reflect the additional value generated by a referral.

When unemployed, worker j of type i receives a referral if the employer of one of his links expands and worker j is chosen among the referrer's links. Worker j has n_{ii}^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 ρ. A referrer of type i has $n_{ii}+n_{ik}$ links and each of them is equally likely to receive the referral. And similarly for type-k links. Therefore, worker j is referred to a job at rate

$$\alpha_{Ri}^j = \frac{\rho n_{ii}^j (1-u_i)}{n_{ii}+n_{ik}} + \frac{\rho n_{ik}^j (1-u_k)}{n_{kk}+n_{ki}}. \tag{6}$$

Using network symmetry ($n_{ii}^j = n_{ii}, n_{ik}^j = n_{ik}$), consistency ($n_{ik} = n_{ki}$) and the definition of the homophily rate from equation (1), equation (6) becomes:

$$\alpha_{Ri} \;=\; \rho\phi_i(1-u_i)+\rho(1-\phi_k)(1-u_k). \tag{7}$$

Two features of equation (7) are noteworthy. First, the referral rate does not directly depend on the size of the network. Instead, it only depends on the homophily rates of the two worker types, ϕ_A and ϕ_B, which greatly simplifies the analysis.[20]

Second, a worker's referral rate increases in the homophily rate of his own type and decreases in the homophily rate of the other type: type-A unemployed workers receive share ϕ_A of referrals generated by type-A employed workers and share $1-\phi_B$ of referrals generated by type-B employed workers; and similarly for type-B unemployed workers. As a result the referral rates of the two types are more unequal under a hierarchical network rather than a homophilous network and it is crucial to have an equilibrium model of the network structure in order to study the impact of referral use on inequality.

The key friction in searching through the market is that all workers search in the same pool and, in other words, a vacancy cannot target its search towards, say, type-A workers. The reason is that workers' market characteristics capture features which are either not contractible (e.g. productivity beyond observable characteristics in the meritocratic case) or cannot be legally used for recruiting purposes (e.g. gender, in the discrimination case). The information embedded in a worker's network is valued precisely in settings where information that is harder to advertise for is more relevant.

Three types of agents search in the market: measure v vacancies, measure u_A type-A unemployed workers and measure u_B type-B unemployed workers. A Cobb-Douglas function

[20] This feature is clearly the result of this model's maintained assumptions. Extending the model to a more general setting where network size plays a non-trivial role, though interesting and worthwhile, is beyond the scope of the present paper.

determines the flow of meetings between vacancies and workers of either type:

$$m(v, u_A, u_B) = \mu v^\eta (u_A + u_B)^{1-\eta},$$

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

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

$$\alpha_{Fi} = \frac{m(v, u_A, u_B)}{v} \frac{u_i}{u_A + u_B} = \mu(\frac{u_A + u_B}{v})^{1-\eta} \frac{u_i}{u_A + u_B}.$$

The rate at which a type i worker meets a firm through the market is

$$\alpha_{Mi} = \frac{m(v, u_A, u_B)}{u_A + u_B} = \mu(\frac{v}{u_A + u_B})^\eta.$$

Since this rate does not depend on the worker's type, the i-subscript is henceforth dropped.

The steady state conditions equate each type's flows in and out of unemployment:

$$u_A(\alpha_M + \alpha_{RA})p_A = (1 - u_A)\delta, \tag{8}$$

$$u_B(\alpha_M + \alpha_{RB})p_B = (1 - u_B)\delta. \tag{9}$$

I assume that the vacancy cost is low enough for $u_A \leq 1/2$ and $u_B \leq 1/2$, which is the empirically relevant case.

2.2.3 Value functions

I now describe the agents' value functions. First, consider a firm. When vacant, it searches in the market, meets with a type-i worker at rate α_{Fi} and forms a match with probability p_i. When producing with a type-i worker, 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 δ. The firm's values of a vacancy (V)

and employing a type-i worker (J_i) satisfy

$$\begin{aligned} rV &= -k + \alpha_{FA} p_A (J_A - V) + \alpha_{FB} p_B (J_B - V), \\ rJ_i &= y_i + \rho\gamma X_i - w_i + \delta(V - J_i). \end{aligned}$$

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 equals his wage and the match is destroyed at rate δ. The worker's value of unemployment (U_i) and employment (W_i) satisfy

$$\begin{aligned} rU_i &= b + (\alpha_M + \alpha_{Ri}) p_i (W_i - U_i), \\ rW_i &= w_i + \delta(U_i - W_i). \end{aligned}$$

The wage solves the Nash bargaining problem

$$w_i = \text{argmax}_w (W_i - U_i)^\beta (J_i - V)^{1-\beta}. \tag{10}$$

where β denotes the worker's bargaining power.

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

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

2.2.4 Labour market outcomes with off-equilibrium network

To determine network-formation incentives, I describe the payoffs to worker j of type i whose network is (n_{ii}^j, n_{ik}^j) and might differ from that of the other type i workers. Since worker j is measure zero, his network does not affect any aggregate quantity such as the values of other agents, the unemployment rates or vacancy creation.

Worker j's network affects the referrals that he receives and the referrals that he generates.

His arrival rate of job opportunities through a referral (α_{Ri}^j) is given by equation (6), before introducing the restriction to symmetric networks, and the proportion of time that he spends unemployed is determined by $u_i^j(\alpha_M + \alpha_{Ri}^j)p_i = (1 - u_i^j)\delta$.

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 due to the referrals that he generates, i.e. the worker assumes $X_i^j = X_i$ regardless of his network. 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 steady state utility of worker j of type i is:

$$\Lambda_i^j = u_i^j U_i^j + (1 - u_i^j) W_i^j, \tag{11}$$

where U_i^j and W_i^j are determined as before but with the worker-specific referral rate α_{Ri}^j.

2.3 Equilibrium definition

The equilibrium in the labour market for a given symmetric referral network is defined as follows.

Definition 1 *A Labour Market Equilibrium given referral network* (n_{AA}, n_{AB}) *and* (n_{BB}, n_{BA}) *is the steady state measures of unemployed workers* $\{u_A, u_B\}$ *and the measure of vacancies* v *such that:*

- *The labour market is in steady state as described by* (8) *and* (9).
- *The surplus is split between worker and firm according to* (10).
- *There is free entry of firms:* $V = 0$.

The equilibrium is defined as follows.

Definition 2 *An Equilibrium is* (e_{AA}, e_{AB}) *and* (e_{BB}, e_{BA}) *which solve* (4) *subject to the symmetry restriction, where the labour market payoffs are given by equation* (11).

2.4 The optimal network

I describe the problem of a social planner whose objective is to maximise steady state output subject to the search and informational frictions underlying the economic environment. In other words, the planner takes as given the market meeting function, the expansion/referral rate and the probability that a meeting turns into a match with each worker type, p_A and p_B. One interpretation is that this is a utilitarian social planner who internalises firms' preference for hiring each type of worker, as described by p_A and p_B, regardless of the origin of these preferences.

The planner has two instruments at his disposal: the network structure and the entry of vacancies. In Galenianos (2014) I studied a similar referral model (but without heterogeneity) and showed 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 surplus-sharing parameters β and γ take exactly the "right" (non-generic) values. That logic transfers to the present model and, since the current focus is on inequality between the two types of workers rather than vacancy entry, I will focus the analysis on network structure and assume that vacancy creation is constrained efficient. The results on the optimal network do not qualitatively depend on vacancy entry and, therefore, this assumption does not affect the rest of the welfare analysis.

The planner solves the following problem:

$$\max_{\phi_A,\ \phi_B} \mathcal{W}(\phi_A, \phi_B) \quad = \quad \big(1 - u_A(\phi_A, \phi_B)\big)y_A + \big(1 - u_B(\phi_A, \phi_B)\big)y_B + \big(u_A(\phi_A, \phi_B) + u_B(\phi_A, \phi_B)\big)b, \tag{12}$$

where $u_A(\phi_A, \phi_B)$ and $u_B(\phi_A, \phi_B)$ are determined by the referral rate (7) and the steady state conditions (8) and (9).

3 Equilibrium Characterization and Welfare

This section characterises the equilibrium and derives its positive and normative implications.

3.1 Equilibrium characterization

The surplus of a match between a firm and a type-i worker equals:

$$S_i = W_i - U_i + J_i - V.$$

Nash bargaining implies:

$$\begin{aligned} W_i - U_i &= \beta S_i, \\ J_i - V &= (1-\beta) S_i. \end{aligned}$$

The value of a vacancy is, therefore, given by

$$rV = -k + \alpha_{FA} p_A (1-\beta) S_A + \alpha_{FB} p_B (1-\beta) S_B. \tag{13}$$

The first Proposition proves that, given an arbitrary network structure (ϕ_A, ϕ_B), an equilibrium exists and a worker's unemployment rate does not depend on his productivity.[21]

Proposition 1 *A Labour Market Equilibrium exists. The steady state unemployment rate of a type-i worker depends on his probability of matching when meeting a firm (p_i) and does not depend on his productivity on the job (y_i).*

[21] Recall the maintained assumption that the firm entry is low enough for the unemployment rate to be below 50% for both worker types.

To characterise the equilibrium network, I first show that a worker's choice of network formation effort is unique and then aggregate the optimal choice of all workers to determine the equilibrium structure.

Consider worker j of type i. Differentiating his value with respect to networking effort with his own type (e^j_{ii}) and the other type (e^j_{ik}) yields:[22]

$$\begin{aligned}\frac{d\mathcal{L}^j_i}{de^j_{ii}} &= \frac{\partial \Lambda^j_i}{\partial \alpha^j_{Ri}} \frac{d\alpha^j_{Ri}}{de^j_{ii}} - ce^j_{ii} \\ &= \frac{\partial \Lambda^j_i}{\partial \alpha^j_{Ri}} \frac{\rho(1-u_i)}{n_{ii}+n_{ik}} - ce^j_{ii} \qquad (14) \\ \frac{d\mathcal{L}^j_i}{de^j_{ik}} &= \frac{\partial \Lambda^j_i}{\partial \alpha^j_{Ri}} \frac{\rho(1-u_k)}{n_{kk}+n_{ki}} \frac{E_{ki}}{E_{ki}+E_{ik}} - ce^j_{ik}. \qquad (15)\end{aligned}$$

The benefit of additional effort is given by the effect of effort on the referral rate times the effect of the referral rate on steady utility. Following from the assumption of quadratic costs, the marginal cost is linear in effort. As a result:

Proposition 2 *In the network formation stage, the optimal effort of worker j $(e^{j*}_{ii}, e^{j*}_{ik})$ is unique.*

The proposition proves that steady state utility Λ^j_i is strictly increasing and strictly concave in a worker's referral rate α^j_{Ri} and, therefore, the worker's optimal effort level is unique. This is an intuitive result: a worker's steady state utility increases in the rate at which he contacts job opportunities; furthermore, a higher referral rate reduces time spent in unemployment and therefore reduces the benefit from additional rises in α^j_{Ri}.

I now state the main result regarding the existence and characterization of equilibrium.

Proposition 3 *An equilibrium exists. Furthermore:*

[22] I only consider the case where $E_{ii} > 0$ and $E_{ik} > 0$. Equilibria without networks exist but are of no particular interest to this study.

1. *In equilibrium* $\phi_A = 1 - \phi_B$ *and the structure of the referral network is characterised by a single variable:* $\phi^* = \phi_A$.

2. *The equilibrium homophily rates relate to the relative employment rate according to:*

   - $\frac{\phi^*}{1-\phi^*} = \sqrt{\frac{1-u_A}{1-u_B}}$

3. *The referral network exhibits hierarchy under type A:*

   - *If* $p_A > p_B$ *then* $u_A < u_B,\ \phi^* > \frac{1}{2}$, $n_{AA} + n_{AB} > n_{BB} + n_{BA}$ *and* $\alpha_{RA} > \alpha_{RB}$
   - *If* $p_A = p_B$ *then* $u_A = u_B,\ \phi^* = \frac{1}{2}$, $n_{AA} + n_{AB} = n_{BB} + n_{BA}$ *and* $\alpha_{RA} = \alpha_{RB}$

To understand these characterization results it is useful to first reiterate the agents' incentives for engaging in costly network-building: forming more links generates additional referrals, the magnitude of which depends on the new link's employment rate (since it is employed workers who can make referrals). As a result all workers, independently of their own type, prefer to link with high-employment A-types.[23]

Consequently, type-A workers spend less effort "linking-down" than do B-types "linking-up", which leads to greater rationing of across-type links for B-workers. Furthermore, type-B workers do not respond to this rationing by creating more links with other B-types because such links are less valuable. This explains why B-types have more links with A- than with B-types ($n_{BB} < n_{AA}$ or, equivalently, $\phi_B < \phi_A$) which leads to a hierarchical network and a lower referral rate for B-types than for A-types.

Notice that the hierarchical structure is an outcome of many-to-many linking in the referral network. If, instead, each worker could only create a single connection, then the A types would only link with other A-types and the B-types would, by necessity, link with each other, leading to a homophilous (or positive-assortative) network. Burdett and Coles (1997) study a model of marriage, where it is natural to restrict to a single connection, and show

[23] More precisely, the result states that there is no equilibrium with $u_B < u_A$ due to, say, a self-fulfilling favourable network structure for B-workers.

that a homophilous (positive-assortative) matching structure emerges. This observation suggests that the assumption on whether linking (matching) is 1-1 or many-to-many is crucially important for the resulting structure of connections.

3.2 Positive implications

I now explore the effect of the referral network's equilibrium structure on inequality and the determinants of the strength of network hierarchy.

The use of referrals affects workers' arrival rate of job opportunities, employment rates and bargained wages. To explore how the referral network affects inequality, I compare the relative labour market outcomes of the two worker types under two referral network structures: the equilibrium network structure of the baseline model and a *homogeneous* network where all workers, regardless of type, have the same network ($\hat{\phi}_A = \hat{\phi}_B = 1/2$) and, therefore, face the same arrival rate of job opportunities. The model with a homogeneous network provides a natural benchmark because it exhibits the same search and informational frictions as the baseline model but without the effect of the equilibrium referral network structure.

Denote the arrival rate of job opportunities for a type-i worker in the baseline network model and the homogeneous network model by α_i and $\hat{\alpha}_i$, respectively. Therefore:

$$\begin{aligned} \hat{\alpha}_A &= \hat{\alpha}_B = \alpha_M + \rho\frac{1}{2}(2 - u_A - u_B), \\ \alpha_A &= \alpha_M + \rho\phi^*(2 - u_A - u_B), \\ \alpha_B &= \alpha_M + \rho(1 - \phi^*)(2 - u_A - u_B). \end{aligned}$$

Proposition 3 proves that if $p_A > p_B$ then $\alpha_A > \hat{\alpha}_A = \hat{\alpha}_B > \alpha_B$ and, therefore, the employment rate across types is more unequal in the equilibrium network model than in the homogeneous network model. The same holds for wages since they depend on workers' outside options.

Proposition 4 *If $p_A > p_B$ then the use of referrals exacerbates inequality.*

Though intuitive, this result is not necessarily immediate from an ex ante point of view. Since type-B workers face worse employment prospects in the labour market, investing in network formation could, in principle, provide an opportunity to reduce inequality in labour market outcomes. It turns out that the exact opposite outcome arises in equilibrium: low(er) employment rates make B-types less attractive as links, leading to reduced access to referrals.

The extent of the referral-induced inequality is parametrised by the equilibrium variable ϕ^* which determines the strength of the network's hierarchy. To see this, notice that the two worker types' relative referral rate is:

$$\frac{\alpha_{RA}}{\alpha_{RB}} = \frac{\phi^*}{1-\phi^*}.$$

If ϕ^* is near 1/2, then the two worker types have similar referral rates and the network is not very hierarchical. If ϕ^* is close to 1, then the A-types enjoy a much higher referral rate and hierarchy is much stronger. I now consider three comparative statics exercises to explore the circumstances that lead to stronger or weaker network hierarchy and, by extension, greater or lesser referral-induced inequality.

Proposition 5 *The strength of referral network hierarchy (ϕ^*) depends on the following variables:*

1. *Workers' productivity does not affect the strength of referral network hierarchy (ϕ^*).*

2. *Increasing workers' contact rate through the market α_M leads to a less hierarchical referral network in equilibrium (lower ϕ^*).*

3. *Assume $p_A = 1 - p_B = p$. Increasing p leads to a more hierarchical referral network in equilibrium (higher ϕ^*).*

The interpretation of the proposition is as follows. First, recalling that a potential link's employment rate depends only on p_i and does not depend on y_i (Proposition 1) it follows that a link's attractiveness does not depend on productivity which, therefore, does not affect network structure.

Second, the strength of the network's hierarchical structure depends on the relative importance of the referral channel in generating job opportunities. If the arrival rate of job opportunities through the market is high, say because of greater vacancy entry or greater efficiency in the market meeting process, then referrals play a less important role in job-finding which reduces the incentives for incurring the additional costs of a hierarchical network.

Third, as shown in Proposition 3, the strength of the network's hierarchy depends on the relative employment rates of the two types which, in turn, depends on the relative probability that a match is formed when a worker and firm meet. As a result, a greater differential in the probability of forming a match increases the employment rate differential, thereby leading to a more hierarchical referral network.

3.3 Normative implications

I now derive the optimal network structure. I restrict attention to networks where $\phi_A = 1-\phi_B$, as this network structure obtains in equilibrium.[24] The planner chooses the homophily rate of type-A workers and the planner's solution is ϕ^P. The next proposition provides this section's main result.

Proposition 6 *The equilibrium network is generically inefficient: $\phi^P \neq \phi^*$. The planner's solution has the following features:*

1. *If the market characteristics of the two worker types are identical ($y_A = y_B$, $p_A = p_B$) then the equilibrium referral network is optimal: $\phi^P = \phi^* = 1/2$.*

[24] Alternatively, the planner could separately choose ϕ_A and ϕ_B. This will make it harder to compare the constrained optimum with the equilibrium outcome and draw conclusions about potential sources of inefficiency.

2. *If the two worker types are equally productive ($y_A = y_B$) and type-A workers form a match with greater probability ($p_A > p_B$) then the optimal referral network is hierarchical in favour of type-B workers, while the equilibrium referral network is hierarchical in favour of type-A workers: $\phi^P < 1/2 < \phi^*$.*

3. *Increasing the productivity differential between the two worker types leads to an optimal referral network that is more hierarchical in favour of type-A workers and does not affect the equilibrium referral network.*

To explore the source of the inefficiency, note that the planner chooses ϕ^P to maximise productivity-weighted employment. The optimal network, therefore, depends on workers' productivity while the equilibrium network does not (Proposition 5), which creates a wedge between the optimal and the equilibrium solution and is responsible for the inefficiency of equilibrium.

To see this in more detail, Parts 1 and 2 of the Proposition consider the case where the two worker types are equally productive. In this case the planner's objective becomes to maximise aggregate employment. If the two types also enjoy the same probability of forming a match, this objective is achieved through a network where all workers, regardless of type, have the same number and composition of links ($\phi^P = 1/2$), which I call a homogeneous network. If all workers have the same probability of forming a match, this coincides with the equilibrium solution.

If, however, $p_A > p_B$ then the planner compensates for the reduced employment opportunities of type-B workers by increasing their referral rate ($\phi^P < 1/2$) which points in the opposite direction from the equilibrium solution: the optimal network reduces inequality while the equilibrium network increases inequality, in comparison to the homogeneous network. Underlying this result is a tension between private incentives in link formation and the optimal recipient of referrals. Private incentives inevitably lead to a network that "favours the favoured": everyone wants to link with high-employment workers, who end up dominating the network. However, the optimal referral recipient need not be the type-As.

Part 3 demonstrates that in some settings the planner prefers a network that increases inequality, over the homogeneous benchmark. When productivity differences are substantial, then the planner prefers to increase the employment rate of the A-types even at the expense of the B-types' employment. Therefore, as productivity differentials increase, the optimal network features greater hierarchy in favour of the type-A workers.

This analysis makes clear that, when discrimination is present in the labour market, the use of referrals creates additional inefficiencies because the referral network favours type As even though they have no productivity advantage. Sections 3.1 and 3.2 present some ways of alleviating this problem by reducing the strength of hierarchy. First, a policy that promotes formal hiring reduces the relative importance of referrals, which leads to a decrease in the strength of the hierarchy (Proposition 5, Part 2) and the additional inequality created by the referral hiring channel. Second, a policy that subsidises the employment of type-B workers increases their relative employment and gives them more opportunities to become the source of referrals which, in turn, reduces the strength of hierarchy (Proposition 3, Part 2). These results demonstrate that there are potentially important indirect benefits to subsidizing the employment of discriminated groups.[25] These policies, of course, do not directly tackle the sources of discrimination, which is of course important but beyond the scope of this paper.

4 On Hierarchy and Homophily

The model delivers the sharp theoretical prediction that referral networks are hierarchical: all workers, regardless of type, have a majority of links with type-A workers and, furthermore, type-A workers have more links in total than type-B workers and enjoy a higher arrival rate

[25] However, promoting referrals without improving the employment rates of discriminated groups is unlikely to be successful since, according to the model, these referrals will be received by the advantaged group. The study of Beaman et al. (2018) shows evidence of this behaviour and is described in detail in section 4.3.

of referrals. At first sight, this prediction appears to contradict the sociological work that documents the pervasive homophily of social networks.

In this section I, first, briefly discuss why the sociological evidence does not *in principle* contradict the model's prediction, which implies that the structure of the referral network a feature to be determined through the empirical study of *referring behaviour*. Second, I use the theoretical model to derive an empirical test that can distinguish between a hierarchical and a homophilous referral network. Third, I use this test to evaluate the evidence in two published empirical studies and show that the evidence is consistent with referral networks that have a hierarchical, rather than homophilous, structure. While this section does not provide any new data, it demonstrates the usefulness of evaluating the evidence through the lens of the theoretical model.

4.1 A discussion of the evidence from sociology

The sociology literature documents that social networks are homophilous along several dimensions (see McPherson et al. (2001) for a survey). These findings refer to aspects of social interaction that differ in, at least, two important ways from the behaviour that the present study focuses on.

First, it is important to distinguish between a person's social network and his referral network: the former encompasses the set of people that he is linked with through a great variety of interactions,[26] while the latter consists only of the set of people that he refers to a job (in general, the referral network might be considered to be a subset of the social network). This distinction is important because strategic considerations are likely to be more salient when deciding whether to provide someone with information about a job and, so, the structure of the referral network need not simply mirror that of the social network.

Second, social networks are documented to be homophilous along observable character-

[26] For example, the type of interactions that the sociology literature studies include marriage, close friendships and confiding, acquaintances etc.

istics such as religion, race, ethnicity, gender or education (McPherson et al., 2001). In the application of the model to a meritocratic labour market (example 1 in Section 2.2) worker heterogeneity explicitly refers to differences in productivity after conditioning on observable characteristics. Therefore, there is no a priori contradiction in having a network which is, say, homophilous across observable educational attainment and simultaneously hierarchical across the separate dimension of unobservable ability *within* each level of education.

Therefore, to determine the structure of referral networks in a satisfactory way requires analysing evidence about actual referring behaviour, which is the subject of the remainder of this section.

4.2 From theory to empirical tests

I use the characterization of the labour market equilibrium for an arbitrary referral network (Proposition 1) to derive implications that can empirically distinguish a homophilous from a hierarchical referral network.

Proposition 7 *Consider the equilibria of two separate labour markets: a labour market with a hierarchical referral network ($\phi_A > 1/2 > \phi_B$) and a labour market with a homophilous referral network ($\phi_A \geq \phi_B > 1/2$), where ϕ_i is the proportion of links that a type-i worker has with workers of his own type. In equilibrium:*

1. *A worker hired through a referral is more likely to be type-A than a worker hired through the market under both referral network structures.*

2. *A type-A worker is more likely to refer another type-A worker under both referral network structures.*

3. *A type-B worker is more likely to refer a type-A worker under a hierarchical referral network and a type-B worker under a homophilous referral network.*

Part 1 follows from the observation that type-A workers receive more referrals than type-B workers. This observation clearly holds under a hierarchical referral network since all referrers (regardless of own type) are more likely to refer a type-A worker. It also holds under a homophilous referral network because A-type workers generate more referrals, since they are employed more often, and their referrals are more likely to be received by other A-type workers, due to homophily. Parts 2 and 3 follow directly from the definition of hierarchy and homophily.

Using Proposition 7, I derive a two-step empirical test that distinguishes the referral network structure.

Step 1: Determine the empirical counterpart to type-A and type-B workers. According to Part 1 of Proposition 7, type-A workers are referred more frequently than type-B workers. Given a dataset with information on worker characteristics and method of hire (referral or market), workers with frequently-referred characteristics correspond to type-A workers. Notice that the nature of these characteristics (e.g. high ability) is not in itself informative about the structure of the referral network.

Step 2: Determine the referral network structure. Part 3 of Proposition 7 shows that type-B workers refer their own type under a homophilous network structure but refer A-types under a hierarchical network structure. Given a dataset that, additionally, has information on referrer-referee pairs, I can determine whether type-B workers refer their own types or the type-As and, therefore, determine the structure of the referral network.

Operationalizing the second step depends on the particular worker characteristics under consideration. If the characteristic that distinguishes workers is a binary variable, such as gender, then it suffices to estimate how often men refer other men etc.[27] If the characteristic is a continuous variable, e.g. some continuous ability measure, then one can examine the coefficient of a regression of referred type on referrer type: if positive, then an increase in the ability measure of the referrer is associated with an increase in referred workers' ability,

[27] It is straightforward to extend this to discrete variables, more generally.

thereby indicating the presence of homophily; if zero, then the ability measure of referred workers does not depend on referrers' ability which indicates a hierarchical network structure (assuming that referred workers have higher ability, on average).

The data requirements for distinguishing between the alternative network structures are quite high. It is not only necessary to determine the potentially unobservable type of a referred worker (which is already challenging) but it is also required to determine the identity and type of the referrer.

Note that the substance of this empirical test does not depend on the exact modelling details of the present study and, for instance, holds in the model of Montgomery (1991).[28]

4.3 Using the test to evaluate published empirical evidence

I use the empirical test to evaluate the evidence presented in the studies of Hensvik and Skans (2016) and Beaman et al. (2018). The labour markets analysed in these studies correspond to the meritocratic and discriminatory labour markets in the present paper, respectively, and are based on very detailed data which includes information on referrer-referred links.

4.3.1 A meritocratic labour market: high-skill Swedish men

Hensvik and Skans (2016) study the role of referrals in the labour market for high-skill Swedish men by combining two very detailed administrative datasets. The first dataset is a matched employer-employee dataset which contains every worker's employment history and *coworker network*, defined as the set of coworkers that a worker encountered at each establishment

[28]In Montgomery (1991) type-A workers are more productive and the firm chooses whether to ask for a referral from its incumbent worker based on his type. To introduce a similar feature in the present model, assume that firms choose the rate at which they request a referral ρ, subject to some cost. Firms that expect to receive a high benefit from the referral will choose a high ρ and vice versa. Under a hierarchical network, all workers have similar networks and, therefore, employed workers of both types will be asked for a referral at similar rates. Under a homophilous network, which is the maintained assumption in Montgomery (1991), the types of referrer and referred are positively correlated and, therefore, a firm is more likely to ask a referral from a type-A worker than from a type-B worker. As a result, the substance of Proposition 7 holds in the Montgomery model, with the addition that B-workers are expected to refer less frequently under a homophilous network.

where he was employed in the past. The second dataset contains scores from ability tests run by the Swedish army (similar to the US army's AFQT test) that all men in Sweden take for the military draft. The authors focus on workers employed at occupations with high skill requirements (top third of the occupation distribution) and use test scores on cognitive ability as a measure of worker ability. Furthermore, they use the score component which is orthogonal to the worker's observable characteristics, such as education, as a measure of the worker's unobserved ability.

The authors assume that a referral has occurred when a newly-hired worker is linked with one of his new employer's incumbent workers through his coworker network (i.e. when the new hire and the incumbent employee were employed in the same establishment at the same time in the past). The authors confirm that this link captures a genuine personal connection by performing a number of statistical tests.[29] Connecting these two datasets allows the authors to associate a referral hiring with the ability of both referrer and referred.

Hensvik and Skans (2016) find that a newly-hired worker's cognitive ability is on average 13.6% of a standard deviation higher and his entry wage is 3.6% higher if he is referred to a job, controlling for the worker's observable characteristics (Hensvik and Skans, 2016, tables 4 and 7). These estimates are statistically and economically significant and they suggest that the market for high-skilled Swedish men is consistent with the present paper's meritocratic labour market: referred workers (type-As according to the first step of the empirical test) have higher ability than non-referred workers.[30]

[29]To confirm that this connection captures a genuine link, the authors also define two types of placebo links for newly-hired workers. These are new hires who worked at the incumbent's previous employer but at a different time from the incumbent and new hires who worked at the incumbent's firm at the same time as the incumbent but at a different establishment of that firm. Workers with placebo links have a similar mobility pattern to workers with genuine links but lack a direct connection with the incumbent worker. It turns out that new hires with placebo links are mostly indistinguishable in terms of their characteristics and outcomes to new hires unconnected to incumbent workers; by contrast, new hires with genuine links to incumbent workers are statistically and economically different to other hires (Hensvik and Skans, 2016, table 9). This finding suggests that coworker networks do transmit information about workers.

[30]Several other studies from advanced countries (mostly the US) find supportive evidence that the meritocratic model is a good approximation for the use of referrals. Using direct, personal productivity measures Castilla (2005), Burks et al. (2015) and Barr et al. (2019) show that referred workers are more productive.

The next set of findings relates to network structure. Hensvik and Skans (2016) find that increasing the ability of an incumbent worker by one standard deviation leads to a rise of 0.1 percentage point in the probability of becoming a referrer and a rise of 2% of a standard deviation in the cognitive ability of the referred worker (Hensvik and Skans, 2016, tables 5 and 6). Both estimates are statistically significant under the authors' preferred specification.

The first finding means that an incumbent worker's ability has little effect on whether he becomes a referrer.[31] The second finding shows a positive but quantitatively weak relationship between the ability of a referred worker and the ability of his referrer: a referral by a high-ability referrer (one standard deviation above average) is associated with an increase in the newly-hired worker's ability by 15.7% of a standard deviation; however, only one-eight of that gain (2 percentage points out of 15.7) is associated to the referrer's ability. Through the lens of the model, therefore, the evidence of a strong positive overall effect of a referral and weak effect of referrer ability points more towards a hierarchical, rather than a homophilous, structure for the referral network, according to the second step of the empirical test.[32]

These findings are consistent with circumstantial evidence from firms' personnel policies. It is common practice among large firms to provide monetary incentives to all their employees who refer candidates to apply for a job at the firm (indeed, many of the firm-level studies on referrals are based on data collected by human resources departments for purposes of remunerating referrers). These incentives are typically available to all workers: none of the studies report that better-than-average workers are somehow favoured in such schemes. This observation suggests that firms believe that their lower-ability employees are able to refer higher-ability applicants to the job.

Simon and Warner (1992), Bayer et al. (2008), Brown et al. (2016), and Dustmann et al. (2016) show that referred workers receive higher wages and Fernandez et al. (2000) and Brown et al. (2016) show that referred applicants are more likely to be hired when they apply for a job.

[31] The Montgomery (1991) model predicts this effect to be strongly positive.

[32] Hensvik and Skans (2016) write that their results support the Montgomery (1991) model, in part because they interpret the positive effect of a referral on ability as consistent with that model. As discussed in section 4.2 this effect is also consistent with a hierarchical network structure. Furthermore, their emphasis is on identifying whether referrals lead to high-ability hires rather than to distinguish between different potential network structures.

4.3.2 A discriminatory labour market: Malawian women

Beaman et al. (2018) design and run a field experiment to examine the referrer-referee relationship across gender in a developing country, Malawi. The aim of their study is to explore whether encouraging referrals by women improves labour-market opportunities for other women, which is what the gender-homophilous structure of the social network predicts.

The field experiment uses the recruitment efforts of an NGO (IPA-Malawi) to hire workers in a job for which it is common to employ both men and women. The NGO advertises a job opening, tests applicants to measure their ability and, after employing qualified applicants for a day, asks them to refer someone to do the same job in the following week. Referrers are randomly split in three groups regarding the gender of who they might refer: constrained to refer a man, constrained to refer a woman or unconstrained. All referring and referred workers are tested and their ability levels recorded, thereby providing information about the ability levels of referred workers and referrer-referred pairs.

The authors find that men and women refer men for the job 77% and 57% of the time, respectively (Beaman et al., 2018, table 1). Since men are more likely to be referred, they are type-As and women are type-Bs, according to the first step of the empirical test.[33] Furthermore, the authors find that both men and women referrers are able to refer equally qualified women when constrained to do so: the average ability of referred workers is statistically identical when restricted to refer a woman, both for men and for women referrers (tables 2 and 4). Therefore the network available to referrers includes men and women of similar ability levels, which suggests that the high referral rates enjoyed by men are not due to superior ability. As a result, the Malawian labour market is consistent with this paper's discriminatory labour market.[34]

[33] Understanding the difference in the intensity of referring men is certainly interesting but beyond the scope of the present paper.

[34] The interpretation is not, of course, that the NGO is a discriminatory employer. In fact, a crucial part of the analysis depends on the objective nature of the ability test given by the NGO to the job candidates. Rather, the interpretation is that the referral network, i.e. the preference for referring men, reflects workers' overall experience in the Malawian labour market.

The finding that women are more likely to refer men rather than women is inconsistent with a homophilous and consistent with a hierarchical referral network, according to the second step of the empirical test. Indeed, the authors conclude that referrals tend to reinforce, rather than alleviate, gender inequalities even when the referrers are women, which is the opposite from what they were expecting. This observation is consistent with the present paper's theoretical prediction that the use of referrals exacerbates inequality.

Beaman et al. (2018) offer a striking illustration of the distinction between a social and a referral network. The sociology literature finds that social networks are gender homophilous (McPherson et al., 2001) which, if anything, is likely stronger in a developing country like Malawi. The finding that women are more likely to refer a man rather than a woman is inconsistent with a homophilous referral network which, at the very least, suggests that the choice of who to refer is based on different criteria from the choice of who to socialise with.

The present paper offers an interpretation of this finding which is based on individual incentives that lead to hierarchical referral network. The incentives should be broadly interpreted. This paper's substantive prediction is that a worker with a referral opportunity will refer someone who is able to provide a favour and, in the model, a favour is identified with a future referral. It is possible, however, that there are alternative social arrangements where favours take different forms which are not directly related to the labour market or are directed more broadly (e.g. gift exchange, long-run family or kinship relations etc.). Further study is required to determine whether the incentives analysed here are indeed the drivers of behaviour and, particularly, to account for other possible explanations for the aforementioned findings.

5 Conclusions

This is the first paper to study the interaction between labour markets and endogenous referral networks in the context of worker heterogeneity. The analysis yields an important novel result: referral networks have a hierarchical structure. The network's hierarchical structure exacerbates inequality in the labour market and drives the model's normative implications, which are subtle. When worker heterogeneity is due to productivity differentials, the structure of referral networks improves welfare, despite the rise in inequality. When workers have similar productivity and, nevertheless, some workers form matches with higher probability, then the equilibrium network structure leads to severe inefficiencies. The latter case is likely to occur in the presence of discrimination in the labour market.

Using the theoretical model, I provide an empirical test for distinguishing whether the structure of the referral network is hierarchical, as predicted by my theory, or homophilous, as typically assumed based to the evidence on the pervasive homophily of social networks. I then show that the evidence from two detailed studies is consistent with a hierarchical, and inconsistent with a homophilous, referral network which suggests that referral networks do not necessarily mirror social networks. Since the empirical evidence that is detailed enough to distinguish between different referral network structures is still rather limited, these finding should be interpreted with caution and more empirical work is needed to confirm whether they are robust. One promising way forward is to use matched employer-employee data to define coworker networks and study referrer-referred links in more detail. Field experiments can also provide more detailed information about the particular incentives related to referrals.

There are several directions for theoretical work to pursue. First, to study the interaction between referral networks and minorities. Second, to study the interaction between referral networks and workers' job-to-job transitions, where the work of Arbex et al. (2019) is a promising initial step. Finally, to examine the role of network size, which I have abstracted from in the present model.

APPENDIX

A Proofs

Proposition 1.
Proof. The unemployment rate for each worker type is determined by the steady state equations. These equations include a worker's probability of forming a match when meeting a firm (p_i) and do not include the worker's productivity (y_i).

Define:

$$\begin{aligned} H(u_A, u_B) &\equiv u_A \mu\left(\frac{v}{u_A + u_B}\right)^\eta + u_A \rho\big(\phi_A(1-u_A) + (1-\phi_B)(1-u_B)\big) - \frac{\delta}{p_A}(1-u_A), \\ L(u_A, u_B) &\equiv u_B \mu\left(\frac{v}{u_A + u_B}\right)^\eta + u_B \rho\big(\phi_B(1-u_B) + (1-\phi_A)(1-u_A)\big) - \frac{\delta}{p_B}(1-u_B). \end{aligned}$$

and note that in a steady state $H(u_A, u_B) = L(u_A, u_B) = 0$ holds. Define $h(u_B)$ and $l(u_A)$ such that $H(h(u_B), u_B) = 0$ and $L(u_A, l(u_A)) = 0$.

Let $H_x(u_A, u_B) \equiv (\partial H(u_A, u_B)/\partial x)$ and similarly for $L(u_A, u_B)$ and observe that

$$\begin{aligned} H(0, u_B) &= -\frac{\delta}{p_A} < 0, \\ H(1, u_B) &= \mu\left(\frac{v}{1 + u_B}\right)^\eta + \rho(1-\phi_B)(1-u_B) > 0, \\ H_{u_A}(u_A, u_B) &= \mu\left(\frac{v}{u_A + u_B}\right)^\eta\left(1 - \frac{\eta u_A}{u_A + u_B}\right) + \rho(\phi_A(1-u_A) + (1-\phi_B)(1-u_B)) \\ &\quad + \frac{\delta}{p_A} - u_A \rho \phi_A > 0. \end{aligned}$$

The first two equations mean that $h(u_B)$ exists and is in $(0,1)$ for any $u_B \in [0,1]$ and the third equation proves that $h(u_B)$ is a function (not a correspondence). Furthermore:

$$h'(u_B) = -\frac{H_{u_B}}{H_{u_A}} = -\frac{-\frac{\eta u_A}{u_A+u_B}\mu\left(\frac{v}{u_A+u_B}\right)^\eta - u_A\rho(1-\phi_B)}{H_{u_A}} > 0.$$

And similarly for $l(u_A)$.

Define $T_1(u_A) = h(l(u_A))$. A steady state is a fixed point of $T_1(u_A)$. 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 $T_1'(u_A) < 1$. We have:

$$T_1'(u_A) = h'(l(u_A))l'(u_A) = \frac{H_{u_B}(u_A, l(u_A))L_{u_A}(u_A, l(u_A))}{H_{u_A}(u_A, l(u_A))L_{u_B}(u_A, l(u_A))}.$$

Notice that:

$$
\begin{aligned}
H_{u_A}(u_A,u_B)-|L_{u_A}(u_A,u_B)| &= (1-\eta)\alpha_M+\alpha_{RA}+\frac{\delta}{p_A}-\rho\big(\phi_A u_A+(1-\phi_A)u_B\big)>0,\\
L_{u_B}(u_A,u_B)-|H_{u_B}(u_A,u_B)| &= (1-\eta)\alpha_M+\alpha_{RB}+\frac{\delta}{p_B}-\rho\big(\phi_B u_B+(1-\phi_B)u_A\big)>0,
\end{aligned}
$$

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

Note that $H_v=(\eta u_A/v)\mu\big(v/(u_A+u_B)\big)^\eta>0$ and $L_v=(\eta u_B/v)\mu\big(v/(u_A+u_B)\big)^\eta>0$ which imply that $(dh(u_B)/dv)<0$ and $(dl(u_A)/dv)<0$. Therefore u_A and u_B are decreasing in v.

The value of a vacancy is given by $rV=\alpha_{FA}(1-\beta)S_A+\alpha_{FB}(1-\beta)S_B$ where the S_i's can be written as $S_A=((y_A-b)D_{B1}+(y_B-b)D_{A2})/(D_{A1}D_{B1}-D_{A2}D_{B2})$ and $S_B=((y_B-b)D_{A1}+(y_A-b)D_{B2})/(D_{A1}D_{B1}-D_{B2}D_{A2})$, where $D_{i1}=r+\delta+(\alpha_M+\alpha_{Ri})p_i\beta-\rho\gamma(1-\beta)\phi_i u_i p_i$ and $D_{i2}=\rho\gamma(1-\beta)(1-\phi_i)u_k p_k$ for $i\in\{A,B\}$.

Note that D_{i1} is increasing in v and D_{i2} is decreasing in v and therefore S_i is decreasing in v. Furthermore, recalling $\delta>\rho$, the steady state equations imply that $v\to 0\Rightarrow(u_A,u_B)\to(1,1)\Rightarrow\alpha_{Fi}\to\infty$ and $v\to\infty\Rightarrow(u_A,u_B)\to(0,0)\Rightarrow\alpha_{Fi}\to 0$. 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 2.
Proof. The second and cross-derivatives of the worker's objective function are:

$$
\begin{aligned}
\frac{\partial^2\mathcal{L}_i^j}{\partial(e_{ii}^j)^2} &= \frac{\partial^2\Lambda_i^j}{\partial(\alpha_{Ri}^j)^2}\Big(\frac{\rho(1-u_i)}{n_{ii}+n_{ik}}\Big)^2-c,\\
\frac{\partial^2\mathcal{L}_i^j}{\partial(e_{ik}^j)^2} &= \frac{\partial^2\Lambda_i^j}{\partial(\alpha_{Ri}^j)^2}\Big(\frac{\rho(1-u_k)}{n_{kk}+n_{ki}}\frac{E_{ki}}{E_{ki}+E_{ik}}\Big)^2-c,\\
\frac{\partial^2\mathcal{L}_i^j}{\partial(e_{ii}^j e_{ik}^j)} &= \frac{\partial^2\Lambda_i^j}{\partial(\alpha_{Ri}^j)^2}\frac{\rho(1-u_i)}{n_{ii}+n_{ik}}\frac{\rho(1-u_k)}{n_{kk}+n_{ki}}\frac{E_{ki}}{E_{ki}+E_{ik}}.
\end{aligned}
$$

Proving that $(\partial^2\Lambda_i^j)/(\partial(\alpha_{Ri}^j)^2)$ 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 $u_i^j=\delta/(\delta+(\alpha_M+\alpha_{Ri}^j)p_i)$ and $S_i^j=(y_i-b+\rho\gamma X_i)/(r+\delta+(\alpha_M+\alpha_{Ri}^j)p_i\beta)$, respectively. Rewrite worker j's steady state utility as follows:

$$
\begin{aligned}
\Lambda_i^j &= u_i^j U_i^j+(1-u_i^j)W_i^j\\
&= \frac{b}{r}+\Big((1-u_i^j)+\frac{(\alpha_M+\alpha_{Ri}^j)p_i}{r}\Big)\beta S_i^j.
\end{aligned}
$$

Letting $\alpha_i^j=(\alpha_M+\alpha_{Ri}^j)p_i$ and going through some algebra leads to:

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

Differentiating Λ_i^j with respect to α_{Ri}^j we have:

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

Define $D = \delta(r+\delta) + \alpha_i(r+\delta+\delta\beta) + (\alpha_i^j)^2\beta$. Taking the second derivative of Λ_i^j with respect to α_{Ri}^j leads to:

$$\begin{aligned}
\frac{\partial^2 \Lambda_i^j}{\partial (\alpha_{Ri}^j)^2} &= \frac{\beta(y_i - b + \rho\gamma X_i)p_i}{rD^3}\Big[\big(2\delta(r+\delta) + 2\alpha_i\big(r(1-\beta)+\delta\big)\big)\big(\delta(r+\delta) + \alpha_i(r+\delta+\delta\beta) + \alpha_i^2\beta\big)^2 \\
&\quad -\big((r+\delta)^2\delta + \alpha_i 2\delta(r+\delta) + \alpha_i^2\big(r(1-\beta)+\delta\big)\big)2D_i\big((r+\delta+\delta\beta) + \alpha_i 2\beta\big)\Big] \\
&= \frac{2\beta(y_i - b + \rho\gamma X_i)p_i}{rD^2}\Big[-\delta(r+\delta)^2(r+\beta\delta) - \alpha_i^j\delta(r+\delta)\big(\beta(r+\delta) + (r+\delta)2\beta\big) \\
&\quad -(\alpha_i^j)^2 3\delta\beta(r+\delta) - (\alpha_i^j)^3\big(r(1-\beta)+\delta\big)\beta\Big] < 0.
\end{aligned}$$

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.
Proof. To prove the first two parts of the proposition, combine the mapping from effort to links, the first order conditions, the definition of the homophily rate and the network consistency requirement. The definition of how effort maps to links implies:

$$\frac{e_{ii}^j}{e_{ik}^j} = \frac{E_{ik} + E_{ki}}{E_{ki}}\frac{n_{ii}^j}{n_{ik}^j}. \tag{16}$$

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

$$\frac{e_{ii}^j}{e_{ik}^j} = \frac{1-u_i}{1-u_k}\frac{n_{ki}+n_{kk}}{n_{ik}+n_{ii}}\frac{E_{ik}+E_{ki}}{E_{ki}}. \tag{17}$$

Combining equations (16) and (17) yields:

$$\frac{n_{ii}^j}{n_{ik}^j} = \frac{1-u_i}{1-u_k}\frac{n_{ki}+n_{kk}}{n_{ik}+n_{ii}}. \tag{18}$$

The definition of the homophily rate implies:

$$\begin{aligned}
n_{ii}^j &= \phi_i^j(n_{ii}^j + n_{ik}^j) \\
n_{ik}^j &= (1-\phi_i^j)(n_{ii}^j + n_{ik}^j) \qquad (19)\\
\Rightarrow \frac{n_{ii}^j}{n_{ik}^j} &= \frac{\phi_i^j}{1-\phi_i^j} \qquad (20)
\end{aligned}$$

Combining equations (18) and (20) with the equilibrium symmetry condition $\phi_i^j = \phi_i$ for all j

we have:

$$\begin{aligned}
\frac{\phi_A}{1-\phi_A} &= \frac{1-u_A}{1-u_B}\frac{n_{BB}+n_{BA}}{n_{AA}+n_{AB}}, \qquad (21)\\
\frac{1-\phi_B}{\phi_B} &= \frac{1-u_A}{1-u_B}\frac{n_{BB}+n_{BA}}{n_{AA}+n_{AB}},\\
\Rightarrow \frac{\phi_A}{1-\phi_A} &= \frac{1-\phi_B}{\phi_B},\\
\Rightarrow \phi_A &= 1-\phi_B. \qquad (22)
\end{aligned}$$

which proves part 1.

The consistency requirement $n_{AB}=n_{BA}$ and equation (19) lead to:

$$\begin{aligned}
(n_{AA}+n_{AB})(1-\phi_A) &= (n_{BB}+n_{BA})(1-\phi_B),\\
\Rightarrow \frac{n_{BB}+n_{BA}}{n_{AA}+n_{AB}} &= \frac{1-\phi_A}{1-\phi_B} = \frac{1-\phi_A}{\phi_A}. \qquad (23)
\end{aligned}$$

Combining equations (21) and (23) proves part 2: $\big(\phi_A/(1-\phi_A)\big)^2=(1-u_A)/(1-u_B)$.

The equilibrium is characterized by (v,u_A,u_B,ϕ), where $\phi=\phi_A=1-\phi_B$, which satisfy $\big(\phi/(1-\phi)\big)^2=(1-u_A)/(1-u_B)$, the steady state and free entry conditions. Given v (e.g. at its equilibrium value) the equilibrium ϕ is determined by the root of $T_2(\phi)=\phi^2\big(1-u_B(\phi)\big)-(1-\phi)^2\big(1-u_A(\phi)\big)$ where $u_A(\phi)$ and $u_B(\phi)$ are defined by the steady state conditions $H(u_A,u_B,\phi)=L(u_A,u_B,\phi)=0$ (with a slight abuse of notation).

I show that $T_2(\phi)=0$ has a unique solution at ϕ^*. Note that $T_2(0)=-\big(1-u_A(0)\big)<0$, $T_2(1)=1-u_B(1)>0$ and

$$\begin{aligned}
T_2'(\phi) &= 2\phi\big(1-u_B(\phi)\big)+2(1-\phi)\big(1-u_A(\phi)\big)-\phi^2u_B'(\phi)+(1-\phi)^2u_A'(\phi)\\
&= \frac{(1-\phi)^2}{\phi}\Big(2\frac{\phi^2\big(1-u_B(\phi)\big)}{(1-\phi)^2}+\phi u_A'(\phi)\Big)+\frac{\phi^2}{1-\phi}\Big(2\frac{(1-\phi)^2\big(1-u_A(\phi)\big)}{\phi^2}-(1-\phi)u_B'(\phi)\Big).
\end{aligned}$$

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

$$T_2'(\phi) = \frac{(1-\phi)^2}{\phi}\Big[2\big(1-u_A(\phi)\big)+\phi u_A'(\phi)\Big]+\frac{\phi^2}{1-\phi}\Big[2\big(1-u_B(\phi)\big)-(1-\phi)u_B'(\phi)\Big]. \qquad (24)$$

Implicit differentiation leads to: $u_A'(\phi)=-\Big(L_{u_B}H_\phi-H_{u_B}L_\phi\Big)/\Delta$ and $u_B'(\phi)=-\Big(H_{u_A}L_\phi-L_{u_A}H_\phi\Big)/\Delta$ where $\Delta=H_{u_A}L_{u_B}-H_{u_B}L_{u_A}$.

Each term in the square brackets of equation (24) is examined separately:

$$2(1-u_A)+\phi u_A'(\phi) = \frac{1}{\Delta}\Big(L_{u_B}\big(2(1-u_A)H_{u_A}-u_A\alpha_{RA}\big)-H_{u_B}\big(2(1-u_A)L_{u_A}+\alpha_{RA}u_B\big)\Big).$$

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

$$\begin{aligned}
2(1-u_A)H_{u_A} &- u_A\alpha_{RA}+2(1-u_A)L_{u_A}+\alpha_{RA}u_B>0\\
\Rightarrow \alpha_M 2(1-u_A)\big(1 &- \eta\big)+\alpha_{RA}\big(2(1-u_A)-u_A+u_B\big)+2(1-u_A)\big(\frac{\delta}{p_A}-\rho\phi u_A-\rho(1-\phi)u_B\big)>0,
\end{aligned}$$

which is positive when $u_A \leq 1/2$.

Similarly:

$$2(1-u_B)-(1-\phi)u_B'(\phi) = \frac{1}{\Delta}\Big(H_{u_A}\big(2(1-u_B)L_{u_B}-u_B\alpha_{RB}\big)-L_{u_A}\big(2(1-u_B)H_{u_B}+\alpha_{RB}u_A\big)\Big),$$

which is positive because $H_{u_A}+L_{u_A}>0$ and

$$\alpha_M 2(1-u_B)\big(1-\eta\big)+\alpha_{RB}\big(2(1-u_B)-u_B+u_A\big)+2(1-u_B)\big(\frac{\delta}{p_B}-u_B\rho(1-\phi)-u_A\rho\phi\big)>0,$$

which is positive if $u_B \leq 1/2$. Therefore, given v, (ϕ, u_A, u_B) are unique.

Furthermore notice that $T_2(1/2)=(u_A-u_B)/4$. When $\phi=1/2$ we have $\alpha_{RA}=\alpha_{RB}$ and therefore $u_A<u_B \Leftrightarrow p_A>p_B$. Therefore $p_A=p_B \Rightarrow u_A=u_B \Rightarrow T_2(1/2)=0 \Rightarrow \phi^*=1/2$ and $p_A>p_B \Rightarrow u_A<u_B \Rightarrow T_2(1/2)<0 \Rightarrow \phi^*>1/2$. This completes the proof. ■

Proposition 5.

Proof. Part 1. Obvious from proposition 1.

For part 2 notice that $(\partial T_2(\phi)/\partial\alpha_M)=-\phi^2(du_B/d\alpha_M)+(1-\phi)^2(du_A/d\alpha_M)$. If the sign is positive then $(d\phi^*/d\alpha_M)<0$ and vice versa.

Recall that $\phi^2/((1-\phi)^2)=(1-u_A)/(1-u_B)$ when $T_2(\phi)=0$ and therefore:

$$\frac{\partial T_2(\phi)}{\partial\alpha_M} = \frac{(1-\phi)^2}{1-u_B}\Big((1-u_B)\frac{du_A}{d\alpha_M}-(1-u_A)\frac{du_B}{d\alpha_M}\Big).$$

We have $(du_A/d\alpha_M)=-\Big(L_{u_B}H_{\alpha_M}-H_{u_B}L_{\alpha_M}\Big)/\Delta$ and $(du_B/d\alpha_M)=-\Big(H_{u_A}L_{\alpha_M}-L_{u_A}H_{\alpha_M}\Big)/\Delta$ where $\Delta=H_{u_A}L_{u_B}-H_{u_B}L_{u_A}$, $H_{\alpha_M}=u_A$ and $L_{\alpha_M}=u_B$. Therefore $(du_A/d\alpha_M)=-\Big(\delta u_A/(p_B u_B)+u_B u_A\rho(2\phi-1)\Big)/\Delta$ and $(du_B/d\alpha_M)=-\Big(\delta u_B/(p_A u_A)-u_B u_A\rho(2\phi-1)\Big)/\Delta$.

Hence

$$\begin{aligned}\frac{\partial T_2(\phi)}{\partial\alpha_M} &= -\frac{(1-\phi)^2}{\Delta(1-u_B)}\Big[(1-u_B)\Big(\frac{\delta u_A}{p_B u_B}+u_B u_A\rho(2\phi-1)\Big)-(1-u_A)\Big(\frac{\delta u_B}{p_A u_A}-u_B u_A\rho(2\phi-1)\Big)\Big]\\ &= -\frac{(1-\phi)^2}{\Delta(1-u_B)}\Big[\alpha_M(u_A-u_B)+\rho(2-u_A-u_B)\big(u_A(1-\phi)(1-u_B)-u_B\phi(1-u_A)\big)\Big]>0.\end{aligned}$$

and we have $(d\phi^*/d\alpha_M)<0$.

For Part 3 notice that $(\partial T_2(\phi)/\partial p)=\big((1-\phi)^2/(1-u_B)\big)\Big((1-u_B)(du_A/dp)-(1-u_A)(du_B/dp)\Big)$ and also $(\partial T_2(\phi)/\partial p)<0 \Leftrightarrow (d\phi^*/dp)>0$. We have $(du_A/dp)=-\Big(L_{u_B}H_p-H_{u_B}L_p\Big)/\Delta$ and $(du_B/dp)=-\Big(H_{u_A}L_p-L_{u_A}H_p\Big)/\Delta$ where $\Delta=H_{u_A}L_{u_B}-H_{u_B}L_{u_A}$ and where $H_p=\delta(1-u_A)/p^2$ and $L_p=-\delta(1-u_B)/(1-p)^2$.

Therefore:

$$
\begin{aligned}
\frac{\partial T(\phi)}{\partial p} &= -\frac{(1-\phi)^2}{\Delta(1-u_B)}\Big[(1-u_B)\big(L_{u_B}H_p - H_{u_B}L_p\big) - (1-u_A)\big(H_{u_A}L_p - L_{u_A}H_p\big)\Big] \\
&= -\frac{(1-\phi)^2}{\Delta(1-u_B)}\Big[\frac{\delta(1-u_B)}{(1-p)^2}\Big(\alpha_M\big(1-\frac{\eta u_A(2-u_A-u_B)}{u_A+u_B}\big) + \rho\phi(2-u_A-u_B)(1-u_A)\Big) \\
&\quad + \frac{\delta(1-u_A)}{p^2}\Big(\alpha_M\big(1-\frac{\eta u_B(2-u_A-u_B)}{u_A+u_B}\big) + \rho(1-\phi)(2-u_A-u_B)(1-u_B)\Big)\Big].
\end{aligned}
$$

The terms multiplied by ρ are positive. Furthermore:

$$
\begin{aligned}
&\frac{1-u_B}{(1-p)^2}\big(1-\frac{\eta u_A(2-u_A-u_B)}{u_A+u_B}\big) + \frac{1-u_A}{p^2}\big(1-\frac{\eta u_B(2-u_A-u_B)}{u_A+u_B}\big) > \\
&(1-u_B)\big(1-\frac{\eta u_A(2-u_A-u_B)}{u_A+u_B}\big) + (1-u_A)\big(1-\frac{\eta u_B(2-u_A-u_B)}{u_A+u_B}\big) = \\
&(2-u_A-u_B)\Big(1-\frac{\eta u_A(1-u_B)}{u_A+u_B} - \frac{\eta u_B(1-u_A}{u_A+u_B}\Big) > \\
&(2-u_A-u_B)(1-\eta) > 0,
\end{aligned}
$$

and therefore $(\partial T_2(\phi)/\partial p) < 0$ and $(d\phi^*/dp) > 0$. ■

Proposition 6.
Proof. Write the planner's problem as follows:

$$
\max_{\phi} \mathcal{W}(\phi) = y_A + y_B - u_A(\phi)(y_A - b) - u_B(\phi)(y_B - b),
$$

where $u_A(\phi)$ and $u_B(\phi)$ are defined by the steady state conditions.

Implicit differentiation on the steady state conditions yields:

$$
\begin{aligned}
u_A'(\phi) &= -\frac{u_A\rho(2-u_A-u_B)}{\Delta}\Big[\alpha_M\big(1-\frac{2\eta u_B}{u_A+u_B}\big) + \alpha_{RB} + \frac{\delta}{p_B} - u_B\rho\Big] < 0, \\
u_B'(\phi) &= \frac{u_B\rho(2-u_A-u_B)}{\Delta}\Big[\alpha_M\big(1-\frac{2\eta u_A}{u_A+u_B}\big) + \alpha_{RA} + \frac{\delta}{p_A} - u_A\rho\Big] > 0.
\end{aligned}
$$

where $\Delta = H_{u_A}L_{u_B} - H_{u_B}L_{u_A} > 0$. The two derivatives' sign uses the assumption that the entry cost is low enough for $u_H \leq 1/2$ and $u_L \leq 1/2$.

The first order conditions of the planner's problem are:

$$
\begin{aligned}
\mathcal{W}'(\phi) &= -u_A'(\phi)(y_A - b) - u_B'(\phi)(y_B - b) \\
&= -\frac{\rho\big(2-u_A(\phi)-u_B(\phi)\big)u_A(\phi)u_B(\phi)}{\Delta}\Big[\frac{(y_A-b)\delta}{p_B u_B(\phi)^2} - \frac{(y_B-b)\delta}{p_A u_A(\phi)^2} - (y_A-y_B)\big(\frac{2\eta\alpha_M}{u_A+u_B}+\rho\big)\Big],
\end{aligned}
\tag{25}
$$

where I used the steady state conditions to simplify some terms.

Denote the term inside the square brackets of equation (25) by $T_3(\phi)$ and note that the planner's solution is given by $T_3(\phi^P) = 0$. I show that $T_3'(\phi) < 0$ which suffices to show that $T_3(\phi)$ has at most one root:

$$
\begin{aligned}
T_3'(\phi) &= -\frac{2(y_A-b)\delta}{p_B u_B(\phi)^3}u_B'(\phi)+\frac{2(y_B-b)\delta}{p_A u_A(\phi)^3}u_A'(\phi)+(y_A-y_B)\frac{2\eta\alpha_M(1+\eta)}{\big(u_A(\phi)+u_B(\phi)\big)^2}\big(u_A'(\phi)+u_B'(\phi)\big)\\
&= 2u_A'(\phi)\Big[\frac{(y_B-b)\delta}{p_A u_A(\phi)^3}+(y_A-y_B)\frac{\eta\alpha_M(1+\eta)}{\big(u_A(\phi)+u_B(\phi)\big)^2}\Big]-2u_B'(\phi)\Big[\frac{(y_B-b)\delta}{p_A u_A(\phi)^2 u_B(\phi)}\\
&\quad +(y_A-y_B)\frac{\rho}{u_B(\phi)}+(y_A-y_B)\frac{\eta\alpha_M}{u_B(\phi)\big(u_A(\phi)+u_B(\phi)\big)}\big(2-\frac{(1+\eta)u_B(\phi)}{u_A(\phi)+u_B(\phi)}\big)\Big]<0.
\end{aligned}
$$

As a result there is a unique ϕ^P that solves the planner's problem which is given by:

$$
\begin{aligned}
T_3(0) &\leq 0 && \Rightarrow \phi^P=0,\\
T_3(0) &>0 && \text{and } T_3(1)\geq 0 \Rightarrow \phi^P=1,\\
T_3(0) &>0 && \text{and } T_3(1)<0 \Rightarrow \phi^P\in(0,1) \text{ and } T(\phi^P)=0.
\end{aligned}
$$

For general values of y_i and p_i, note that $T_3(\phi)\neq T_2(\phi)$ and therefore $\phi^P\neq\phi^*$.

To prove part 1 note that if $y_A=y_B$ and $p_A=p_B$ then $T_3(\phi)=0 \Leftrightarrow u_A(\phi)=u_B(\phi)$. Furthermore, $u_A(\phi)=u_B(\phi)$ only when $\phi=1/2$ (and $p_A=p_B$). In that case, therefore, the planner's solution is $\phi^P=1/2=\phi^*$. For part 2, I show that when $y_A=y_B$ and $p_A>p_B$ we have $T_3(1/2)<0$. This proves that $\phi^P<1/2$ because $T_3(\phi)$ crosses the x-axis from above. When $\phi=1/2$ we have $\alpha_{RA}=\alpha_{RB}$. Let $\alpha=\alpha_M+\alpha_{RA}=\alpha_M+\alpha_{RB}$. The steady state conditions can be rearranged as $u_i=\delta/(\alpha p_i+\delta)$.

When $y_A=y_B$ the planner's solution satisfies:

$$
\begin{aligned}
T_3(\phi) &= (y-b)\delta\Big(\frac{1}{p_B u_B(\phi)^2}-\frac{1}{p_A u_A(\phi)^2}\Big),\\
\Rightarrow T_3(\frac{1}{2}) &= \frac{y-b}{\delta}\Big(\frac{(\alpha p_B+\delta)^2}{p_B}-\frac{(\alpha p_A+\delta)^2}{p_A}\Big).
\end{aligned}
$$

Therefore $T_3(1/2)<0$ if:

$$
\begin{aligned}
p_A\big(\alpha p_B+\delta\big)^2 &< p_B\big(\alpha p_A+\delta\big)^2\\
\Leftrightarrow \delta^2 &< p_A p_B\alpha^2.
\end{aligned}
$$

Multiplying the two steady state conditions yields $\delta^2=p_A\alpha p_B\alpha u_A u_B/((1-u_A)(1-u_B))$ which proves that $T_3(1/2)<0$ since $u_A<u_B\leq 1/2$.

For part 3, it is easy to see that $(dT_3(\phi)/dy_A)>0>(dT_3(\phi)/dy_B)$ and therefore when ϕ^P is interior we have $(d\phi^P/dy_A)>0>(d\phi^P/dy_B)$. This completes the proof. ■

Proposition 7.
Proof. For part 1, note that the probability that a newly-hired worker is type-A when hired through the market with probability $\text{Prob}[A|M]=p_A u_A/(p_A u_A+p_B u_B)$ and when hired through a referral with probability $\text{Prob}[A|R]=p_A u_A(R_{AA}+R_{BA})/(p_A u_A(R_{AA}+R_{BA})+p_B u_B(R_{AB}+R_{BB}))$, where $R_{AA}=\rho(1-u_A)\phi_A$ is the flow of referrals from A workers to other A workers, $R_{BA}=\rho(1-u_B)(1-\phi_B)$ is the flow of referrals from B to A workers and similarly for the other two terms.

Therefore Prob[A|R] > Prob[A|M] if and only if $R_{AA}+R_{BA}>R_{AB}+R_{BB}$. For the hierarchical structure this inequality holds trivially and for the homophilous structure it holds if type-A workers make more referrals than type-B workers, i.e. if $u_A < u_B$, since such referrals are mostly received by other type-As.

To prove $u_A < u_B$ recall that the steady state is defined by $H(u_A, u_B) = L(u_A, u_B) = 0$ (see the proof to Proposition 1). We can define $u_A = A^H(u_B)$ such that $H(A^H(u_B), u_B) = 0$ and $u_A = A^L(u_B)$ such that $L(A^L(u_B), u_B) = 0$. At steady state $A^H(u_B) = A^L(u_B)$ and Proposition 1 shows that it exists and it is unique.

Proposition 1 shows that $A^H(u_B) = h(u_B)$, $A^H(0) > 0$, $A^H(1) < 0$ and $A^{H\prime}(u_B) > 0$ which implies that there is a unique mapping $u_A = A^H(u_B)$ that maps $u_B \in [0,1]$ into $u_A = A^H(u_B)$ such that $H(A^H(u_B), u_B) = 0$.

For $A^L(u_B)$ note that

$$
\begin{aligned}
L(0, u_B) &= u_B\mu(\frac{v}{u_B})^\eta + \rho u_B\big(\phi_B(1-u_B) + 1 - \phi_A\big) - \frac{\delta}{p_B}(1-u_B),\\
L(1, u_B) &= u_B\mu(\frac{v}{1+u_B})^\eta + \rho u_B\phi_B(1-u_B) - \frac{\delta}{p_B}(1-u_B),\\
L_{u_A}(u_A, u_B) &= -\eta u_B\mu v^\eta(u_A+u_B)^{-\eta-1} - u_B\rho(1-\phi_A) < 0.
\end{aligned}
$$

Therefore, if $L(0, u_B) > 0$ and $L(1, u_B) < 0$ there is unique $A^L(u_B)$ such that $L(A^L(u_B), u_B) = 0$.

Furthermore $L(0,0) = -\delta/p_B < 0$ and $L(0,1) = \mu(v)^\eta + \rho(1-\phi_A) > 0$. Proposition 1 shows that $L_{u_B}(u_A, u_B) > 0$ which proves, together with the above, that there exists unique $\underline{u}_B$ such that $L(0, \underline{u}_B) = 0$, $L(0, u_B) < 0$ if $u_B < \underline{u}_B$ and $L(0, u_B) > 0$ if $u_B > \underline{u}_B$. As a result, in steady state $u_B \geq \underline{u}_B$. Moreover $L(1,0) = -\delta/p_B < 0$ and $L(1,1) = \mu(v/2)^\eta > 0$ and, therefore, there exists unique $\overline{u}_B$ such that $L(1, \overline{u}_B) = 0$, $L(1, u_B) < 0$ if $u_B < \overline{u}_B$ and $L(1, u_B) > 0$ if $u_B > \overline{u}_B$. As a result, in steady state $u_B \leq \overline{u}_B$.

The above imply that in steady state we have $u_B \in [\underline{u}_B, \overline{u}_B]$ and $u_A = A^L(u_B)$. To compare the steady state u_A and u_A, define $\tilde{H}(u) = H(u,u)$ and $\tilde{L}(u) = L(u,u)$ and f^A and f^B such that $\tilde{H}(f^A) = 0$ and $\tilde{L}(f^B) = 0$.

Noting that $\tilde{H}(0) = -\delta/p_A < 0$, $\tilde{H}(1) = \mu v^\eta 2^{-\eta} > 0$ and $\tilde{H}'(u) = \mu v^\eta 2^{-\eta}(1-\eta)u^{-\eta} + \rho(\phi_A + 1 - \phi_B)(1-2u) + \delta/p_A > 0$ implies that there is unique $f^A \in (0,1)$, and similarly for f^B. Finally, $\tilde{L}(u) < \tilde{H}(u)$ because $\phi_A \geq \phi_B$, which means that $f^B > f^A$.

Combining the above, we can plot the mappings $A^H(u_B)$ and $A^L(u_B)$ together with the 45-degree line in the $u_B - u_A$ space. The mapping $A^H(u_B)$ starts above the 45-degree line ($A^H(0) > 0$) and cross it at f^A. The mapping $A^L(u_B)$ starts below the 45-degree line at some $\underline{u}_B > 0$ where $A^L(\underline{u}_B) = 0$ and crosses it at f^B. Furthermore the two mappings cross exactly once because there is a unique steady state. Since $f^B > f^A$ the intersection occurs below the 45-degree line and we have $u_A < u_B$ in steady state. ■

References

Aigner, Dennis J. and Glen G. Cain (1977), "Statistical theories of discrimination in labor markets." *Industrial and Labor Relations Review*, 30(2), 175–187.

Arbex, Marcelo, Dennis O'Dea, and David Wiczer (2019), "Network search: climbing the job ladder faster." *International Economic Review*, 60(2), 693–720.

Barr, Tavis, Raicho Bojilov, and Lalith Munasinghe (2019), "Referrals and search efficiency: Who learns what and when?" *Journal of Labor Economics*, 37(4), 1267–1300.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa (2008), "Place of work and place of residence: informal hiring networks and labor market outcomes." *Journal of Political Economy*, 116 (6), 1150–1196.

Beaman, Lori, Niall Keleher, and Jeremy Magruder (2018), "Do job networks disadvantage women? evidence from a recruitment experiment in malawi." *Journal of Labor Economics*, 36 (1), 121–157.

Bentolila, Samuel, Claudio Michelacci, and Javier Suarez (2010), "Social contacts and occupational choice." *Economica*, 77, 20–45.

Brown, Meta, Elizabeth Setren, and Giorgio Topa (2016), "Do referrals lead to better matches? evidence from a firm's employee referral system." *Journal of Labor Economics*, 34 (1), 161–209.

Burdett, Ken and Melvyn G. Coles (1997), "Marriage and class." *Quarterly Journal of Economics*, 112 (1), 141–168.

Burks, Stephen V., Bo Cowgill, Mitchell Hoffman, and Michael Housman (2015), "The value of hiring through employee referrals." *Quarterly Journal of Economics*, 130(2), 805–839.

Cahuc, Pierre and Francois Fontaine (2009), "On the efficiency of job search with social networks." *Journal of Public Economic Theory*, 11(3), 411–439.

Calvo-Armengol, Antoni (2004), "Job contact networks." *Journal of Economic Theory*, 115, 191–206.

Calvo-Armengol, Antoni and Matthew Jackson (2004), "The effects of social networks on employment and inequality." *American Economic Review*, 94, 426–454.

Calvo-Armengol, Antoni and Yves Zenou (2005), "Job matching, social networking and word-of-mouth communication." *Journal of Urban Economics*, 57, 500–522.

Castilla, Emilio J. (2005), "Social networks and employee performance in a call center." *American Journal of Sociology*, 110 (5), 1243–1283.

Dustmann, Christian, Albrecht Glitz, Uta Schoenberg, and Herbert Brucker (2016), "Referral-based job search networks." *Review of Economic Studies*, 83, 514–546.

Fernandez, Roberto M., Emilio J. Castilla, and Paul Moore (2000), "Social capital at work: networks and employment at a phone center." *American Journal of Sociology*, 105(5), 1288–356.

Galenianos, Manolis (2013), "Learning about match quality and the use of referrals." *Review of Economic Dynamics*, 16, 668–690.

Galenianos, Manolis (2014), "Hiring through referrals." *Journal of Economic Theory*, 152, 304–323.

Galeotti, Andrea and Luca Paolo Merlino (2014), "Endogenous job contact networks." *International Economic Review*, 55 (4), 1201–1226.

Granovetter, Mark (1995), *Getting a job: a study of contacts and careers, 2nd edition*. Harvard University Press, Cambridge, Massachusetts.

Heath, Rachel (2018), "Why do firms hire using referrals? evidence from bangladeshi garment factories." *Journal of Political Economy*, 126(4), 1691–1746.

Hensvik, Lena and Oskar Skans (2016), "Social networks, employee selection and labor market outcomes." *Journal of Labor Economics*, 34 (4).

Igarashi, Yoske (2016), "Distributional effects of hiring through networks." *Review of Economic Dynamics*, 20, 90–110.

Ioannides, Yannis and Linda Datcher Loury (2004), "Job information networks, neighborhood effects, and inequality." *Journal of Economic Literature*, XLII, 1056–1093.

Kugler, Adriana (2003), "Employee referrals and efficiency wages." *Labour Economics*, 10, 531–556.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook (2001), "Birds of a feather: Homophily in social networks." *Annual Review of Sociology*, 27, 415–444.

Montgomery, James D. (1991), "Social networks and labor market outcomes." *American Economic Review*, 81, 1408–1418.

Mortensen, Dale T. and Tara Vishwanath (1994), "Personal contacts and earnings: it is who you know!" *Labour Economics*, 1, 187–201.

Pallais, Amanda and Emily Glassberg Sands (2016), "Why the referential treatment? evidence from field experiments on referrals." *Journal of Political Economy*, 124 (6), 1793–1828.

Simon, Curtis and John Warner (1992), "Matchmaker, matchmaker: The effect of old boy networks on job match quality, earning and tenure." *Journal of Labor Economics*, 10 (3), 306–330.

Topa, Giorgio (2001), "Social interactions, local spillovers and unemployment." *Review of Economic Studies*, 68, 261–295.

Topa, Giorgio (2011), "Labor markets and referrals." In *The Handbook of Social Economics* (Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds.), Elsevier North Holland, Amsterdam.