Employee referral, social proximity and worker discipline

Amrita Dhillon, Vegard Iversen and Gaute Torsvik
Employee referral, social proximity and worker discipline*

June 18, 2012

Amrita Dhillon
University of Warwick

Vegard Iversen
University of Manchester

Gaute Torsvik
University of Bergen

Abstract

We study ex-post hiring risks in low income countries with limited legal and regulatory frameworks. In our theory of employee referral, the new recruit internalises the rewards and punishments of the in-house referee meted out by the hiring firm. This social mechanism makes it cheaper for the firm to induce worker discipline. The degree of internalization depends on the unobserved strength of the endogenous social tie between the referee and the recruit. When the referee’s utility is increasing in the strength of ties, referee workplace incentives do not matter and referee and employer incentives are aligned: in this case industries and jobs with high costs of opportunism and where dense kinship networks can match the skill requirements of employers will have clusters of close family and friends, they will show a high incidence of referrals rather than anonymous hiring and will show a wage premium to referred workers matched by their higher productivity. This no longer applies if the referee’s utility is decreasing in the strength of ties: referrals are then more costly for firms, they will be used less frequently by employers and will require higher referee wages (or status). We illustrate how these insights add to our understanding of South-Asian labour markets.

JEL: D21, D85,D86, J41, J6,O12, O17.

Keywords: Efficiency wage Contracts, Moral hazard, Referee incentives, Referrals, Networks, Strength of ties, Spot market.

*We thank Chris Woodruff, Rocco Machiavello, Daniel Sgroi, Andres Carvajal, Francois Maniquet, Paul Beaudry, Matt Jackson, Parikshit Ghosh, Eric Maskin, Stelios Michalopoulos, E.Somanathan, the IAS Economics Workshop, Dilip Mookherjee, Mukesh Esaran, the PET development workshop, Peter Vida and especially Rohini Somanathan for comments. Special thanks to the Indian Statistical Institute, Delhi and the IAS Princeton for their hospitality. Corresponding author: Amrita Dhillon. Email: a.dhillon@warwick.ac.uk
1 Introduction

Much recent evidence points towards networks and referrals as a preferred mode of recruitment. Starting with Montgomery (1991), a large theoretical and empirical literature has developed trying to understand why referrals are used by employers rather than anonymous hiring. However, there are some important differences in the prevalence of referrals in white and blue collar jobs as well as in the "types" of referrals that are used by them.

For example, Munshi and Rosenzweig (2006), confirming similar findings by Gore (1970), report that 70% of blue collar jobs in Dadar, Mumbai were found through referral (with a corresponding figure for white collar jobs of around 44%), pointing to some differences in the use of referrals vs other types of recruitment methods for blue collar jobs and white collar jobs. Second, there is also some evidence on differences in the type of referrals across these jobs. In the seminal work by Granovetter (1973) in a Massachusetts town "weak ties" (bridging social capital) was more important in generating non-redundant information about vacancies for white collar jobs, while for blue collar jobs "strong ties" (bonding social capital) were crucial. Table 1 in Montgomery (1991) gives some information on job variation and the use of referrals: Rees and Schultz (1970) shows that less skilled workers rely more heavily on referrals by close friends and family compared to other methods like employment agencies. In a developing country parallel of a weak tie advantage, Wahba and Zenou (2005) propose that search through friends and acquaintances is more efficient for illiterate and poor workers who are unable to read vacancy posters and advertisements. Iversen and Torsvik (2011) report findings using data from an in-depth survey of migrants from Western Uttar Pradesh, North India. 176 or 61.3 % of these migrants entered their first migrant job through a workplace insider. Among these workplace referral cases about 30 % of the new employees were recruited by a member of their own household (usually fathers recruiting sons) while another 49.5% were recruited by a (typically close) relative.

1In Montgomery’s (1991) theory, employers invite the more talented in the workforce to recruit new staff: homophily or assortative matching by talent in employee networks makes employee referral a screening device. Marsden and Gorman (2001) propose an alternative screening mechanism where referral, by virtue of an employee’s informational advantage, reduces uncertainty about a known candidate’s labour productivity. Simon and Warner (1992) and Fernandez et al (2000) highlight, instead, gains in the quality of the match between workplace and employee. Kugler (2003), studying high skill industries in the United States, suggests that employers, through peer pressure from referee to the new recruit, gain from a reduction in monitoring costs. Heath’s (2010) research on garment factories in Bangladesh, finds that referrals serve as a disciplining device by employers.
Social networks and connections thus appear to be vitally for recruitment into low- and unskilled jobs, and the use of close friends and family also seems to be a recurring pattern in these jobs. So why are referrals so popular among employers hiring blue collar workers and why do we see this dependence on family for referrals? Existing explanations of the use of referrals do not distinguish between these differences in job requirements. These explanations rely either on job search cost reduction (e.g. Topa (2003), Calvo-Armengol and Jackson (2004, 2007)) or improving the quality of match between employers and employees (Saloner, (1985) and Simon and Warner (1992), Mortenson and Vishwanath (1994)) or on screening of workers by referees (Montgomery (1991), Kono (2006)) and finally on worker disciplining effects when there is moral hazard (e.g. Kugler 2003, Heath (2010)). None of these explanations however differentiate between different jobs, although implicitly it seems that the explanations are indeed focused on particular jobs. It seems quite natural to think that referrals might be used for completely different reasons in different jobs. Trying to isolate one mechanism for all jobs might be an exercise in futility! Indeed, recent evidence on wage differentials between referred and non-referred workers in developed country settings suggests that such differentials are highly job, industry and country specific (e.g. Ioannides and Loury (2004); Pellizari (2010)). Moreover, within an industry some jobs have higher incidence of referrals than others. These empirical findings suggest that wage premia, referral incidence and productivity outcomes should be expected to be highly job specific. Our paper is firmly centred around this belief.

In this paper, we focus on low skilled jobs and differentiate between jobs on the basis of the scope for "opportunistic behaviour" (moral hazard): in this setting referrals may be used by employers to reduce moral hazard. We model social preferences in the relationship between an employee referee and a potential worker as one of the key features that employers can exploit. In the low- and unskilled labour markets we focus on, labour relations are typically governed by informal, unwritten agreements. A new work relation exposes an employer to a pre-hiring screening challenge and post-hiring behavioural risk. We focus on this context of low skilled jobs (where pre-hire screening is often easy to conduct), in low income countries (where enforcement of explicit contracts is costly for low wage jobs). Of course

2The anthropological literature on India is replete with examples of how kinship morality, social preferences or felt obligations among members of the workforce may be actively used by employers to discipline the workforce (e.g. de Neve 2007), thus exemplifying what Bandiera et al (2009) describe as social incentives and Karlan et al (2009) call social collateral.
when referees are being used to discipline workers, one may well ask why a referee would want to refer a worker in the first place especially if he then bears the costs of being held responsible for the worker’s good behaviour in the firm. Our second contribution is to introduce referee incentives explicitly in the modelling. When both the referee and employer want workers who are closely related to referees (i.e. referee and employer incentives are *aligned*), we show that not only will referrals *always* be preferred to anonymous hiring in this situation, but that referrals will feature close relatives and friends, subject to their availability. In addition, there will be higher wages for referred workers, no difference in wages (or status) between referee-employees and other employees and better outcomes for the firm (e.g. higher labour productivity, lower turnover and absenteeism and lower propensity to take part in industrial action (e.g. Holmstrom (1984)) of employees who are referred, less theft of valuable items etc). Implicitly, this is the case studied by most of the theoretical literature. The empirical evidence on referral and employer outcomes is sparse in the setting of less developed countries except for the numerous examples in Holmstrom (1984), Kajisa’s (2007) evidence from the Philippines and Iversen et al (2009) for India using NSSO Employment Survey data suggesting employer preferences for referral and a link between referrals and lower turnovers. Barr and Oduro (2002) study Ghanaian manufacturing and find strong evidence that relatives are favoured in pay and in the allocation of jobs, and find that such recruits perform well on the job.

However, when referee and employer incentives are *not* aligned, we may not see such a high incidence of referrals and the implications for wages and social ties will be different: conditional on referrals being used, we should observe higher wages (or status) of employee-referees compared to other employees in similar jobs and weaker ties between workers and referees. This is confirmed by the findings of Iversen and Torvik (2011) who show that the job profile of workplace referees and the new recruits is very different. Figure 5 in the Appendix illustrates the job profile of the new recruits with category 1 representing owners, category 2-4 the higher skill levels, 5 are vendors, 6 apprentices and 7-9 are less skilled jobs where 8 and 9 represent hard manual labour. Figure 6 reports the corresponding job profiles for those acting as referees on behalf of the firm. These findings are also indirectly confirmed (when screening is the main reason for referrals) by Bandiera et al (2009) and Beaman and Magruder (2011) who demonstrate in separate experiments, referees or supervisors may not have incentives aligned with employers’ interests and when there
is a mismatch, referrals may not provide a productivity advantage, by Fafchamps and Moradi (2010) who find evidence that recruits referred by army staff who had reached the summit of their careers were more likely to underperform.

2 The Model

Consider a firm who needs to hire 1 worker. Profit per worker in period $t$ is $e_t - w_t$ if the worker does not behave opportunistically and $1 - w_t$ if the worker behaves opportunistically. Hence per capita costs of opportunism to the firm are given by $c_t = (e_t - 1)$. The worker gains $\alpha c_t$ when he behaves opportunistically where $\alpha < 1$. The firm has access to a monitoring technology where a worker can be caught if he behaves opportunistically with an exogenous probability $q \in (0, 1)$. Increasing the level of $q$. is assumed to have infinite costs. The firm can recruit workers either anonymously or using referrals by an existing employee. Each firm comes with 1 referee (existing employee) who comes with a network. $\rho \in [0, \bar{\rho}]$ denotes the social proximity between referee and worker and $\rho$ is interpreted as the strength of this social tie. The referee’s network is a distribution on $\rho$, so $\tau(\rho)$ is the number of workers with whom he has a connection with strength at least $\rho$. $\rho$ and $\tau(\rho)$ is not observed by the firm, nor is opportunistic behaviour ex-ante. We assume $\tau'(\rho) < 0$, i.e. the density of the referee’s network is declining with the strength of the social tie. Put differently, a referee has few very close relatives, but many distant relatives and acquaintances. The strength of the social tie is chosen by the referee: what this means is that the referee introduces a worker with strength of tie $\rho$ to the firm: since the referee’s network is not observed by the firm, the referee can bring any $\rho$ that he wants.

The timeline of the stage game is as follows:

1. In stage 0 new workers $N_t$ enter (free entry), at a cost of entry $y$. Workers exit the industry at an exogenous rate $\gamma$. For simplicity we assume that the distribution $\tau(\rho)$ is simply replicated every period that workers enter.

2. In stage 1 the firm chooses how to hire a worker: $\rho \in [0, \bar{\rho}]$, where $\rho = 0$ corresponds to anonymous hiring.

\footnote{While we focus on a single worker problem, the underlying assumption is that there are many such workers and it is not possible to pinpoint who is responsible for the opportunistic behaviour.}

\footnote{We discuss the implications for the capacity of the network to supply the type of skills employers require in a separate section below.}
3. In stage 2, the firm chooses its wage contracts for the worker and the referee. Contracts are of 2 types: Efficiency wage

- \((\rho, \bar{w}_N(\rho), \tilde{w}_N(\rho))\), where \(\bar{w}_N\) denotes worker wage and \(\tilde{w}_N\) denotes referee wage. If the worker is caught in the next period with probability \(q\) then the worker is fired and the referee gets \(V_R^P(\rho)\) the PV of the implicit "punishment".
- Market wage: If cheating is detected, the worker gets his outside option and there is no punishment for opportunistic behaviour. The outside option is normalised to \(w = 0\)

4. In stage 3, workers choose whether to accept or reject the contract and if they accept they decide whether to follow the contract. Fired workers can only be hired at wage \(\bar{w}\).

5. Stage 4: payoffs

The stage game is infinitely repeated with discrete time. We look for the stationary sub game perfect equilibrium of this game.

We focus on low and unskilled informal labour markets where workers are paid wages, contract enforcement is costly and where output contracts (piece rates) or bonds are not feasible.\(^5\) The way the referral works is that the referee brings a worker of the required \(\rho\) to the firm. Workers are altruistic towards referees. This means that the worker cares about the present value of utility the referee gets, \(V_R(\rho)\), conditional on his being recruited by the firm and what the referee gets if the worker behaves opportunistically, \(V_R^P(\rho)\). \(\beta(\rho)\) measures the degree of altruism with \(\beta' > 0, \beta'' > 0\), i.e. the closer the relationship between worker and referee, the stronger is this altruism.

Worker utility per period, if hired, is \(w_l + \beta(\rho)V_R(\rho)(1 - \delta)\), if he is referred, \(w_l\) if not referred. If a worker cheats when the contract calls for non-opportunistic behaviour, he will be fired if caught. Referee motivations for referring workers may vary: e.g. an important motive is a strong altruistic relationship measured by \(\beta_R(\rho)\): the referee and worker may be close relatives (e.g. father-son, brother-brother etc), or the referee and worker may be in another relationship e.g. a risk

\(^5\)While piece rates have attractive incentive properties and are common in garment factories (e.g. Heath 2010), low and unskilled workers in shops, bakeries, all types of domestic work etc are typically paid wages.
sharing network in the village where the referee benefits from a higher status or gains through reciprocating favours. When jobs are scarce, referees may have particularly strong incentives to refer their own relatives. We assume that $\beta(0) = \beta_R(0) = 0$. Referee gets per period utility $U_R(\rho)$ from getting a worker of type $\rho$. A special case is where $U_R = \beta_R(\rho)w$: the referee’s utility increases with the sum $\tilde{w}_{N,t} + w_t$.

We model these alternative possibilities in Section 4. Firms’ per period profits are $e_t - (w_t + (\tilde{w}_{N,t}(\rho) - \tilde{w}_{N,t}(0)))$. Firms maximise the Present Value of profits.

We can now compute the steady state levels of the stock of workers of each type in the economy. Workers of type $\rho$ enter in period $t$ until expected gains equal cost of entry: $\frac{1}{S_{\tau}(\rho)}w_t = y$. In a stationary equilibrium this leads to $S$ as the total number of workers who enter every period. The stock of workers in the steady state, $N$, is then given by $N(\rho) = \gamma S_{\tau}(\rho) + f(\rho)S$ where $f(\rho)$ is the equilibrium probability of a type $\rho$ worker being fired. We assume the underlying market structure is perfect competition both among firms and workers so that workers take wages as given, there is no bargaining and firms pay workers the minimum needed for them to accept the contract. Although firms make positive profits in the equilibrium, this is consistent with profits being competed away in the long run with free entry (which we do not model for simplicity).

A key ingredient in our model is that the firm is able to punish the referee if the new recruit underperforms or behaves opportunistically. We denote the present value of the punishment utility stream to the referee when a worker he has referred is caught cheating by $V^P_R$. In a series of examples Holmstrom (1984) illustrates how strong ties and felt obligations within the workforce are used strategically by Indian employers to curb absenteeism and other unwanted behaviour and stimulate effort. Other anthropologists and sociologists have argued that the ability to do this improves the incentives of the worker to work hard (Greico, 1987, Kim 1987, de Neve 2007). In this paper, we focus on the incentives confronting referees in such a situation: given that referees can be punished for worker underperformance, what motivates a referee to recruit a worker and what kind of workers (strength of the social tie) will they recruit given the incentive structure? There can be advantages accruing in the network and material and non-material advantages within the workplace (higher status and goodwill with the employer, prospects for promotion etc). We specify a very general punishment that operates through $V^P_R(\rho)$. This allows for contracts where the referee can be fired ($V^P_R = 0$) or where the referee suffers a loss of reputation so that $V^P_R(\rho) < V_R(\rho)$ so that $V^P_R(\rho)$ is below what the referee would
get if the worker did not underperform.

As discussed above, since workers and referees take $w_t, \tilde{w}_{N,t}$ as given and firms do not internalize the effects of $w_t$ on future entry, we can use the steady state stationary levels of $w_t, \tilde{w}_{N,t}, N_t, S_t$ in the following computations.

3 Efficiency wage contracts

Workers always accept the contract if they get at least $\bar{w} = 0$. Hence now consider stage 2 of the game: Under what conditions do workers choose to accept the efficiency wage contract. We show this in the following sections, first for the spot market and then for the referrals market:

**Standard Efficiency wage contracts: Spot market, $\rho = 0$**

Suppose workers behave opportunistically, then in a stationary equilibrium they get:

$$V^S_S = \frac{\bar{w}_S + \alpha c}{1 - \delta(1 - q)}$$  \hspace{1cm} (1)

The subscript here denotes $S$ for spot market (anonymous hiring) while the superscript $S$ denotes "shirking" or opportunistic behaviour. If they do not behave opportunistically, then in a stationary equilibrium they get:

$$V^S_S = \frac{\bar{w}_S}{1 - \delta}$$  \hspace{1cm} (2)

Hence efficiency wage given by $V^S_S \geq V^S_{\tilde{w}}$ is:

$$\bar{w}_S = \frac{\alpha c(1 - \delta)}{\delta q}$$  \hspace{1cm} (3)

where $\delta$ is the discount factor. This expression is the standard efficiency wage result, that as $q, \delta$ increase, the efficiency wage decreases and as $\alpha, c$ increase so do the efficiency wages.

**Efficiency wage contracts with referrals, $\rho > 0$**

Recall that an efficiency wage contract specifies $\rho, \tilde{w}(\rho), \tilde{w}_N(\rho)$. Of course $\rho, \tilde{w}_N(\rho)$ and $\tilde{w}_N(\rho)$ are determined simultaneously.

However, we first compute the efficiency wage contract of the worker taking as given the choice of $\rho$ and $\tilde{w}_N(\rho)$. If a worker behaves opportunistically then in a stationary equilibrium, he gets:
\[ V^S_N = \bar{w}_N + \alpha c + \beta(\rho)V_R(\rho)(1 - \delta) + \delta (q\beta(\rho)V^P_R (\rho) + (1 - q)(V^S_n)) \] (4)

The subscript \( N \) is used to denote referrals ("network") here while the superscript \( S \) denotes opportunistic behaviour (shirking). If the worker does not behave opportunistically then in a stationary equilibrium he gets:

\[ V_N = \bar{w}_N + \beta(\rho)V_R(\rho)(1 - \delta) + \delta(V_n) \] (5)

If opportunistic he gets:

\[ V^S_N = \bar{w}_N + \alpha c + \beta(\rho)V_R(\rho)(1 - \delta) + \delta (q\beta(\rho)V^P_R (\rho) + (1 - q)(V^S_n)) \] (6)

Hence the efficiency wage with referrals is:

\[ \bar{w}_N = \frac{\alpha c(1 - \delta)}{\delta q} - \beta(\rho) \left[ V_R(\rho) - V^P_R (\rho) \right] (1 - \delta) \] (7)

where \( V^P_R \) is the "punishment" PV of the referee if the worker behaves opportunistically and is caught. It is easy to see that if \( X(\rho) \equiv [V_R(\rho) - V^P_R (\rho)](1 - \delta) > 0 \) then referral efficiency wages are lower than anonymous spot market wages. We capture some important implications in the following lemma:

**Lemma 1** Referral efficiency wages are given by: \( \bar{w}_N = \bar{w}_S - \beta(\rho)X(\rho) \). \( \bar{w}_S - \bar{w}_N(\rho) \geq 0 \) iff \( X(\rho) \geq 0 \). Moreover \( \bar{w}_N \) is increasing in \( \alpha, c \) and in \( V^P_R (\rho) \) and decreasing in \( \delta, q \) and in \( \rho \) and \( V_R(\rho) \). \( \bar{w}_S - \bar{w}_N \) is therefore increasing in \( \rho \) and \( X(\rho) \).

This lemma says that when \( \rho, V_R, V^P_R \) are fixed exogenously, the firm benefits from referrals by paying lower efficiency wages to workers than if they were to pay efficiency wages to workers hired through the spot market. This is because workers internalize the impact of their behaviour on the referee - the punishment for opportunistic behaviour is amplified by the impact on the referee. This is similar to the mechanism in Heath (2011) (although we do not have any skill heterogeneity between workers) and resemble group liability schemes which work through monitoring advantages between members of the group relative to the firm. Firms use referrals to reduce the costs of incentivizing workers who have low productivity. They do this
by increasing the (explicit) punishment on referrers for poor performance of referred workers. They assume that referees will be repaid in other ways by the referred worker. Referee incentives and the strength of ties are not explicitly modelled and piece rate contracts are used instead of efficiency wages.

This advantage to the firm from using referrals $\bar{w}_S - \bar{w}_N$ is increasing in the strength of the tie $\rho$, $\bar{w}_N$ and decreasing in $V_P$: put differently, the more a new recruit cares about the referee’s utility and the larger the difference between $V_R$ and $V_P$ (the penalty imposed on the referee by the firm), the stronger is the punishment of the recruit from opportunism and the lower is the premium on wages necessary to induce good behaviour. Kugler (2003), in contrast to us, relies on advantages in monitoring between a referred worker and a referee to obtain lower efficiency wages from referrals that we achieve with altruism and the resulting felt social obligations. In Kugler (2003), as in most of the literature on referrals however, the question of referee incentives is not addressed.

A key innovation in our paper is the way we address the referee’s problem, specifically the decision about whom to recruit, given the various incentive structures that may confront the referee. Intuitively it may seem that when there are no ability differences, and no screening incentives then "favouritism" does not affect the firm: indeed it benefits the firm if productivity is unaffected while wages can be kept lower. This may or may not hold and depends on how the referee benefits from referring a worker. We discuss these different possibilities in the next section.

4 The Referee’s problem

In this part we take $\rho$ and $\bar{w}_N(\rho)$ as given and compute the minimum referee wages $\bar{w}_N(\rho)$ necessary to induce him to choose the optimal $\rho$ when he refers a worker. This is because $\rho$ is unobserved. We compute the total costs of offering the joint contracts to the referee and worker efficiency wages for the worker along with rewards for the referee to make the optimal referral. Given efficiency wage $\bar{w}_N(\rho^*)$ all workers with $\rho < \rho^*$ will cheat for sure, while those above will not cheat. Let $U_R(\rho)$ be the per-period gain to the referee from hiring the worker (so in this case, $U_R(\rho) = V_R(\rho)(1 - \delta)$).

**Case 1:** $U_R' > 0$.

---

6This can also be interpreted as saying that the employer delegates the choice of the worker to the referee, so even if he can observe $\rho$, he does not know $\tau(\rho)$ and hence he cannot control who the referee chooses.
This case captures a commonly observed situation where the referee gains in status in a network where both referee and the new recruit belong such as an extended family network. For example, \( U_R(\rho) \) is the expected utility from a favour (getting the potential worker a job) being returned in some other network that the referee and the worker are part of. The probability that, conditional on successful referral, the favour will be returned is increasing in \( \rho \). Preliminary Evidence from experiments (Bapna, Gupta, Rice and Sundarajan (2011)) shows that indeed trust and reciprocity increase with the strength of ties. Buchan et al (2006) show that social distance is negatively correlated with the degree of trust and reciprocity. In this case, referee incentives are aligned with the firm: both want higher \( \rho \). In this case it is clear that the referee will always choose the highest \( \rho \) that is feasible, as long as \( U_R(\rho) - V^F_R(\rho)(1 - \delta) \) is increasing in \( \rho \). This condition also guarantees that the participation constraint of referees is satisfied.

**Proposition 1** Suppose that \( U_R - V^F_R(1 - \delta) \) is increasing in \( \rho \). Then the firms optimal \( \rho^* = \min(\bar{\rho}, \hat{\rho}) \), \( \bar{w}_N = \check{w}_N(0) \) and \( \check{w}_N = \bar{w}_S - \beta(\rho^*) [V_R(\rho) - V^F_R(1 - \delta)] \).

So if \( U'_R > 0 \) then referrals are always better as long as \( V_R(\rho) > V^F_R(\rho) \): stronger ties are chosen subject to availability of connected people and there is no need for extra incentive provision through referee wages. The spot market will be used only if there are no connected workers available. In this case referee and employer interests are closely aligned and the referee will not behave opportunistically vis-a-vis the employer.

**Case 2:** \( U'_R(\rho) < 0 \)

The opposite case is when \( U'_R(\rho) < 0 \). Some examples of this are when the referee maximises bribes taken from workers for getting them a job: assume that the closer the worker to the referee the lower the bargaining power of the referee: in this case \( U_R \) is decreasing in \( \rho \). This could possibly correspond to situations where an employee of the firm starts acting as a middleman in referring workers. Another compelling objective function that is decreasing in \( \rho \) is when the referee cares about the workers wage in the firm as well as his own. \( U_R = \beta_R(\rho)w \) and \( \beta_R(\rho) = \beta_R > 0 \) a constant (just to ensure that \( U_R \) is decreasing in \( \rho \) when the firm uses the efficiency wage contract with referrals). In this case the incentive constraint of the referee is binding. Assume that \( V^F_R \) is decreasing in \( \rho \): lower \( \rho \) corresponds to less responsibility attributed to the referee for the worker. The incentive constraints are then given by the following expression:
\[
\frac{\tilde{w}_N(\rho) + U_R(\rho)}{1 - \delta} \geq \frac{\tilde{w}_N(\tilde{\rho}) + U_R(\rho) + \delta q V^P_R(\tilde{\rho})}{1 - \delta(1 - q)} \quad \forall \tilde{\rho} < \rho
\] (8)

The participation constraint is given by:

\[
\tilde{w}_N(\rho) + U_R(\rho) \geq \tilde{w}_N(0) + U_R(0) \quad \forall \rho > 0
\] (9)

This implies that \(\tilde{w}_N(\rho)\) is increasing in \(\rho\): the referee must be incentivised to announce a higher \(\rho\). This effect is exacerbated when \(V^P_R\) is decreasing in \(\rho\) as well.

5 A special case I

We focus on the following simple case: let \(U_R = \beta_R w\), with \(\beta_R > 0\). Let there be only 3 possible values of \(\rho = \rho_H, \rho_L, 0\), chosen to capture both constraints. The two constraints on the referee can then be written as:

\[
\frac{\tilde{w}_N(\rho_H) + \beta_R \tilde{w}_N(\rho_H)}{1 - \delta} \geq \frac{\tilde{w}_N(\rho_L) + \beta_R \tilde{w}_N(\rho_L) + \delta q V^P_R(\rho_L)}{1 - \delta(1 - q)}
\] (10)

Inequality 10 is the incentive constraint. The (binding) participation constraint is given by:

\[
\frac{\tilde{w}_N(\rho_L) + \beta_R \tilde{w}_N(\rho_L)}{1 - \delta} \geq \frac{\tilde{w}_N(0) + \beta_R \frac{1}{H} \bar{w}_S}{1 - \delta}
\] (11)

while the worker’s efficiency wages (for \(i = H, L\), as before, are given by:

\[
\bar{w}_N(\rho_i) = \bar{w}_S - \beta(\rho_i) \left[ \bar{w}_N(\rho_i) + \beta_R \bar{w}_N(\rho_i) - V^P_R(\rho_i)(1 - \delta) \right]
\] (12)

The participation constraint 11 captures the fact that when the worker is not hired through a referral process he must go to the spot market where he has a \(\frac{1}{H}\) chance of getting the higher efficiency wage (recall that \(H\) is the steady state stock of workers in every period).

We assume first that both constraints are binding and solve for the lowest wages that satisfy the referee constraints. This is w.l.o.g in the equilibrium. Solving first, equations (11) and (12) for \(\rho_H\) we get the following:

\[
\bar{w}_N(\rho_H) = \frac{1}{1 + \beta(\rho_H)\beta_R} \left[ \bar{w}_S - \beta(\rho_H)\bar{w}_N(\rho_H) + \beta(\rho_H)V^P_R(\rho_H)(1 - \delta) \right]
\] (13)
\[ \bar{w}_N(\rho_H) = \frac{(1 - \delta)(1 + \beta(\rho_H)\beta_R)}{1 - \delta(1 - q)} \bar{w}_N(0) \]
\[ + \bar{w}_S \frac{\beta_R}{N(1 - \delta(1 - q))} \left((1 - \delta)(1 + \beta(\rho_H)\beta_R) - H(1 - \delta(1 - q))\right) \]
\[ + (1 - \delta) \left[ \delta q V^P_R(\rho_L) \frac{(1 + \beta(\rho_H)\beta_R)}{1 - \delta(1 - q)} - \beta(\rho_H)\beta_R V^P_R(\rho_H) \right] \]
\[ (14) \]

Hence

\[ \bar{w}_N(\rho_H) = \bar{w}_S \left[ 1 - \frac{\beta(\rho_H)\beta_R(1 - \delta)}{H(1 - \delta(1 - q))} \right] - \frac{\beta(\rho_H)(1 - \delta)}{1 - \delta(1 - q)} \bar{w}_N(0) \]
\[ + \beta(\rho_H)V^P_R(\rho_H)(1 - \delta) - \frac{\beta(\rho_H)(1 - \delta)\delta q V^P_R(\rho_L)}{1 - \delta(1 - q)} \]
\[ (15) \]

\[ \bar{w}_N(\rho_L) = \bar{w}_N(0)(1 + \beta(\rho_L)\beta_R) + \bar{w}_S \frac{\beta_R}{N} ((1 + \beta(\rho_L)\beta_R) - H) \]
\[ - \beta(\rho_L)\beta_R V^P_R(\rho_L)(1 - \delta) \]
\[ (16) \]

\[ \bar{w}_N(\rho_L) = \frac{1}{1 + \beta(\rho_L)\beta_R} \left[ \bar{w}_S - \beta(\rho_L)\bar{w}_N(\rho_L) + \beta(\rho_L)V^P_R(\rho_L)(1 - \delta) \right] \]
\[ (17) \]

Hence

\[ \bar{w}_N(\rho_L) = \bar{w}_S \left[ \frac{(1 - \beta(\rho_L)\beta_R)}{1 + \beta(\rho_L)\beta_R} - \frac{\beta(\rho_L)\beta_R(1 + \beta(\rho_L)\beta_R)}{H} \right] + \beta(\rho_L)V^P_R(\rho_L)(1 - \delta) - \beta(\rho_L)\bar{w}_N(0) \]
\[ (18) \]

\[ \bar{w}_N(\rho_L) > 0 \text{ if (a) } H > \frac{\beta(\rho_L)(1 + \beta(\rho_L))^2}{1 - \beta(\rho_L)} \text{ and (b) } \bar{w}_S \left[ \frac{(1 - \beta(\rho_L)\beta_R)}{1 + \beta(\rho_L)\beta_R} - \frac{\beta(\rho_L)\beta_R(1 + \beta(\rho_L)\beta_R)}{H} \right] > \beta(\rho_L)\bar{w}_N(0) - \beta(\rho_L)V^P_R(\rho_L)(1 - \delta). \]
Total costs when $\rho = \rho_H$ are given by: $\tilde{w}_N(\rho_H) - \tilde{w}_N(0) + \bar{w}_N(\rho_H)$ i.e.

$$
\tilde{w}_N(0) \left[ \frac{1 - \delta}{1 - \delta(1 - q)} (1 - \beta(\rho_H)(1 - \beta_R)) - 1 \right] + \bar{w}_S \left[ \frac{1 - \delta}{H(1 - \delta(1 - q))} \left( \beta_R(1 - \beta(\rho_H)) + \beta(\rho_H)\beta_R^2 \right) + 1 - \beta_R \right] + V_P^P(\rho_L) \frac{(1 - \delta)q}{1 - \delta(1 - q)} [1 - \beta(\rho_H)(1 - \beta_R)]
$$

(19)

Total costs when $\rho = \rho_L$

$$
- \tilde{w}_N(0) \left[ \beta(\rho_L)(1 - \beta_R) \right] + \bar{w}_S \left[ \beta_R \left( \frac{1 + \beta(\rho_L)\beta_R(1 - \beta(\rho_L))}{H} - 1 \right) + \frac{1 - \beta(\rho_L)\beta_R}{1 + \beta(\rho_L)\beta_R} \right] + V_P^P(\rho_L)(1 - \delta)\beta(\rho_L)(1 - \beta_R)
$$

(20)

And total costs when referrals are not used are just given by $\bar{w}_S$.

**Proposition 2** Assume that $H \geq \max \left( 2, \frac{\beta(\rho_L)\beta_R(1 + \beta(\rho_L) + \beta_R^2)}{1 - \beta(\rho_L)\beta_R} \right)$. Then referee wages $\tilde{w}_N(\rho_H)$ are decreasing in $\bar{w}_S$, and in $V_P^P(\rho_H)$ and increasing in $\tilde{w}_N(0)$ and in $V_P^P(\rho_L)$. Referee wages $\tilde{w}_N(\rho_L)$ are decreasing in $\bar{w}_S$ and in $V_P^P(\rho_L)$ and increasing in $\tilde{w}_N(0)$. Worker wages, $\tilde{w}_N(\rho_H)$ are increasing in $\bar{w}_S$ and in $V_P^P(\rho_H)$ and decreasing in $\tilde{w}_N(0)$ and in $V_P^P(\rho_L)$. Worker wages $\tilde{w}_N(\rho_L)$ are increasing in $\bar{w}_S$, and in $V_P^P(\rho_L)$ and decreasing in $\tilde{w}_N(0)$

Roughly speaking therefore, the wages of the referee are positively related to his "stakes" from the referral measured by $\tilde{w}_N(0) - V_P^P(\rho_i)$, $i = H, L$ while the worker wages are negatively linked to the stakes of the referee in the firm. Referee wages are negatively linked to $\bar{w}_S$, for large enough $H$, while worker wages are positively related to $\bar{w}_S$. It is not clear therefore whether the total costs of referrals is actually lower or higher than anonymous hiring for the firm. We turn to this question in the next section.

### 5.1 Optimal choice of $\rho$

Now we are ready to check the the profit maximising level of $\rho$ for the firm among these three choices: $\rho_H, \rho_L, 0$. One general result we get is that as $\beta_R \to 1, H \to \infty$, the optimal $\rho$ is $\rho_L > 0$. 

14
Proposition 3 Let \( \beta_R \to 1 \). Then for \( H > \max \left( \frac{(1+\beta(\rho_L)(1-\beta(\rho_L))^{1/2})}{\beta(\rho_L)}, \frac{\beta(\rho_L)(1+\beta(\rho_L))^{1/2}}{1-\beta(\rho_L)} \right) \) and \( \bar{w}_S \left[ \frac{(1-\beta(\rho_L))}{1+\beta(\rho_L)} - \frac{\beta(\rho_L)(1+\beta(\rho_L))}{H} \right] > \beta(\rho_L) \left( \bar{w}_N(0) - V^P_R(\rho_L)(1-\delta) \right) \). Then the optimal choice of \( \rho \) for the employer is \( \rho_L \).

Proof. The employers problem is:

Maximize \( \pi_N(\rho) \) w.r.t \( \rho \in \{\rho_H, \rho_L, 0\} \), subject to

\[
\tau(\rho) \geq 1, \\
\bar{w}_N(\rho) \geq \bar{w} \\
\dot{w}_N(\rho) = \frac{\alpha c(1-\delta)}{\delta q} - \beta(\rho) (V_R(\rho) - V^P_R(\rho)) (1-\delta) \\
\frac{\dot{w}_N(\rho) + \beta_R\bar{w}_N(\rho)}{1-\delta} \geq \frac{\dot{w}_N(\tilde{\rho}) + B_R\bar{w}_N(\tilde{\rho}) + \delta qV^P_R(\rho)}{1-\delta(1-q)} \text{ for } \tilde{\rho} < \rho
\]

and

\[
V_R(\rho) \geq V_R(0)
\]

We ignore the feasibility constraint \( \tau(\rho) \geq 1 \), which is discussed in the next section. The constraint \( \bar{w}_N \geq 0 \) is non-binding under the conditions of the proposition. Hence it is clear that both the incentive and participation constraints are binding in equilibrium. Hence as \( \beta_R \to 1 \), we can re-write the two constraints (10) and (11) as:

\[
\frac{\dot{w}_N(\rho_H) + \dot{w}_N(\rho_H)}{1-\delta} = \frac{\dot{w}_N(\rho_L) + \beta_R\bar{w}_N(\rho_L) + \delta qV^P_R(\rho_L)}{1-\delta(1-q)}
\]

and:

\[
\frac{\dot{w}_N(\rho_L) + \dot{w}_N(\rho_L)}{1-\delta} = \frac{\dot{w}_N(0) + \beta_R\bar{w}_S}{1-\delta}
\]

Putting these together, the total costs with \( \rho_H \) are given by

\[
\frac{\bar{w}_N(0) + \beta_R\bar{w}_S + \delta qV^P_R(\rho_L)}{1-\delta(1-q)}
\]

and total costs with \( \rho_L \) are given by \( \beta_R\bar{w}_S \). Hence clearly total costs are lower with \( \rho_L \) than with \( \rho_H \). To see that costs are lower with \( \rho_L \) than with anonymous hiring we compare total costs when \( \rho_L \) for the limit as \( \beta_R \to 1 \):

\[
\bar{w}_S \left[ \frac{1}{H} + \frac{\beta(\rho_L)(1-\beta(\rho_L))}{H} - 1 + \frac{1-\beta(\rho_L)}{1+\beta(\rho_L)} \right]
\]

15
with total costs of anonymous hiring at $\bar{w}_S$. Hence for $H > \frac{(1+\beta(\rho)(1-\beta(\rho)(L))(1+\beta(\rho))}{2\beta(\rho)L}$, total costs with $\rho_L$ turn out to be lower than with anonymous hiring. ■

6 A special case II

To check the intuition from the previous section, an even simpler case is considered in this section. Suppose that there are only two possible announcements for the referee $\rho = 1$ or $\rho = 0$, i.e. he can either say that he refers a relative/friend or he does not refer. Moreover we assume that when the referee is not related to the worker, the firm cannot hold him responsible for bad behaviour. However when $\rho = 0$ then the worker can only be hired through the spot market so his chances of being hired decrease to $\frac{1}{H}$. Solving these two equations leads to:

$$\tilde{w}_N(1) = \tilde{w}_N(0) \frac{1-\beta_R}{1-2\beta_R} - \tilde{w}_S \left( \frac{\beta_R(1-H)}{H(1-2\beta_R)} \right)$$

and

$$\tilde{w}_N(1) = \tilde{w}_S \frac{H}{H(1-2\beta_R)} - \tilde{w}_N(0) \frac{\beta}{1-2\beta_R}$$

$$+ V_R^P (1-\delta) \frac{\beta}{(1-\beta_R)(1-2\beta_R)}$$

(25)

(26)

$\tilde{w}_N(1) + \beta_R(1)\tilde{w}_N(1) = \tilde{w}_N(0) + \beta_R(1)\frac{1}{H} \tilde{w}_S$ (27)

where $\tilde{w}_N(1) = \tilde{w}_S - \beta(1) \left[ \tilde{w}_N(1) + \beta_R \tilde{w}_N(1) - V_R^P (1-\delta) \right]$. Assume that $1/2 > \beta\beta_R$. Then referee wages are positively related to $\tilde{w}_N(0)$, negatively related to $\tilde{w}_S$ and to $V_R^P (1-\delta)$. Worker wages are negatively related to $\tilde{w}_N(0)$, positively related to $\tilde{w}_S$ and to $V_R^P (1)$. When $V_R^P$ is low, then worker wages can be lower but referee wages need to be higher to satisfy the participation constraint of the referee.

Total Costs $\tilde{w}_N(1) - \tilde{w}_N(0) + \tilde{w}_N(1)$ i.e.

$$\tilde{w}_S \frac{\beta(1-2\beta_R)}{H(1-2\beta_R)} - \tilde{w}_N(0) \frac{\beta(1-\beta_R)}{1-2\beta_R}$$

$$+ V_R^P (1-\delta) \frac{\beta(1-\beta_R)(1-\beta\beta_R)}{1-2\beta_R}$$

(28)
Total costs are decreasing with \( \tilde{w}_N(0) - V^P_R(1 - \delta) \) i.e. the stakes of the referee in the firm: what he stands to lose. Total costs increase with \( \tilde{w}_S \) (the costs of opportunism) and decrease with \( H \) i.e. tighter labour market conditions. Comparing total costs with \( \tilde{w}_S \) (total costs with standard efficiency wages), referrals are more attractive when \( \tilde{w}_S \) increases and \( H \) increases. As before, referrals are preferred to anonymous hiring if \( \tilde{w}_N(0) - V^P_R(1 - \delta) \) is sufficiently high.

7 When are efficiency wage contracts used?

In this section we describe the equilibrium of the infinitely repeated game. It is easy to see that efficiency wage contracts will be used whenever the costs of opportunism are sufficiently high. Let \( c(\rho^*) \equiv \tilde{w}_N(\rho^*) + \Delta \tilde{w}_N(\rho^*) - w = \bar{w}_N(\rho^*) + \Delta \tilde{w}_N(\rho^*) \).

**Proposition 4** There exists a subgame perfect equilibrium of the infinitely repeated game where \( H \) workers enter every period, employers choose efficiency wage contracts iff \( c(\rho^*) \). Referees choose a worker with proximity \( \rho^* \), get a wage \( \tilde{w}_N(\rho^*) \) and workers get efficiency wage \( \bar{w}_N(\rho^*) > 0 \). If referees choose \( \tilde{\rho} < \rho^* \) then the worker cheats and is caught with probability \( q \), and the referee gets \( V^P_R(\tilde{\rho}) \). Workers always accept an efficiency wage contract if \( \tilde{w}_N(\rho^*) > 0 \), and if they behave opportunistically, they are fired and never hired again on an efficiency wage contract.

We provide a full description of the equilibrium in the Appendix.

8 Conclusions

In this paper we investigate the mechanisms underlying the use of referrals in low and unskilled labour markets in low income countries. Our main motivation for this comes from India. One important difference between un- and low-skilled labour markets and skilled ones is that screening is less of an issue. What appears to be more important for employers is instead moral hazard issue with referrals often used as a worker disciplining device. Put another way, in many jobs there are substantive costs of opportunism to the employer: unexpected shortages, labour turnover, or even (as in the case of nannies and drivers) theft of valuables is a risk given the lack of knowledge of workers and the ease with which documents can be forged. In this scenario referrals involving a trusted source can be very valuable. We model the special case of employee referrals where the referee is carefully selected often
on the basis of a long term relationship with the employer. Our paper shows how the characteristics of the job, and referee can be important predictors of the use of referrals and how the strength of ties between referees and workers can act as a crucial mechanism to induce worker discipline. Overall our results suggest that whenever a firm pays efficiency wages, it prefers to hire through referrals exploiting the altruism of workers towards referees to pay lower wages. This holds regardless of referee incentives for referral, i.e. whether \( U_R \) is increasing or decreasing in \( \rho \). However, when \( U_R \) is increasing in \( \rho \) the firm prefers to choose the highest feasible \( \rho \) and there are no additional costs via referee wages. This is no longer the case when referees have an incentive to choose a lower \( \rho \) than what the firm wants. Then referee incentives to ensure that the referee complies with firm interests imply that the firm chooses a lower \( \rho \). We can divide testable predictions from our paper into two contingent on the underlying utility function of the referee. When \( U_R(\rho) \) is increasing in \( \rho \) then we claim that referee incentives do not matter, referrals will always be used when feasible and the choice of \( \rho \) will be delegated to the referee. We should therefore observe clusters of close relatives in jobs with high costs of opportunism especially when feasibility is not an issue as in industries and jobs where traditional occupational skills are 'inherited' and passed on through kinship. Comparing workers in different jobs: we should find a higher incidence of referral in jobs where reliability matters and moreover such jobs should come with higher wages for the referred workers. When \( U_R(\rho) \) is decreasing in \( \rho \) then referee incentives and job characteristics, both become important predictors of referrals. Detailed predictions on how the incidence of referrals and strength of ties depends on referee characteristics vary with the exact reasons for why referees may want to refer workers. In contrast to the existing literature an important contribution is to show that referee favouritism is not always bad, that referee incentives matter but not necessarily in the expected ways. We would like to extend the model to explore other plausible \( U_R \) functions and also to screening.
References


de Neve, G. (2008), We are all soudukarar(relatives): kinship and its morality in an urban industry of Tamil Nadu, South India, *Modern Asian Studies*, 42(1), pp. 211-46.


19


20


Appendix

8.1 Proposition 4: Equilibrium strategies:

The steady state stock of workers is $H$ (as discussed in section (2)). Workers always choose efficiency wages as long as $\bar{w}_N(\rho^*) \geq w$. If they choose efficiency wages they never behave opportunistically. If workers behave opportunistically and are caught (which happens with probability $q$) then employers fire them and never hire them again on an efficiency wage contract. This is supported by the beliefs that such a worker will cheat again.

The rest is obvious in the solution of the optimal contract among the set of stationary contracts.

8.2 Figures

0 = High skill levels to 8 = low skills.

![Figure 5: Job classifications first migrant job](image-url)
Figure 6: Job classifications: in-house referees or main contacts