NETWORK-BASED HIRING: LOCAL BENEFITS; GLOBAL COSTS

ARUN G. CHANDRASEKHAR[‡], MELANIE MORTEN[§], AND ALESSANDRA PETER[⋆]

Abstract. Entrepreneurs, particularly in the developing world, often hire from their networks: friends, family, and resulting referrals. Network hiring has two benefits, documented extensively in the empirical literature: entrepreneurs know more about the ability of their network (and indeed they are often positively selected), and network members may be less likely to engage in moral hazard. We study theoretically how network hiring affects the size and composition (i.e., whether to hire friends or strangers) of the firm. Our primary result is that network hiring, while *locally beneficial*, can be *globally inefficient*. Because of the existence of a network, entrepreneurs set inefficiently low wages, firms are weakly too small, rely too much on networks for hiring, and resulting welfare losses increase in the quality of the network. Further, if entrepreneurs are uncertain about the true quality of the external labor market, the economy may become stuck in an *information poverty trap* where forward-looking entrepreneurs or even entrepreneurs in a market with social learning never learn the correct distribution of stranger ability, exacerbating welfare losses. We show that the poverty trap can worsen when network referrals are of higher quality.

Date: This Version: February 21, 2020.
We thank Ran Abramitzky, Abhijit Banerjee, Nick Bloom, Emily Breza, Gabriel Carroll, Ben Golub, Rachel Heath, Nate Hilger, Matt Jackson, Victor Lavy, Kaivan Munshi, Ben Olken, Isaac Sorkin, Meredith Startz, Adam Szeidl, and Xiao Yu Wang for helpful discussions. Financial support from the NSF under grant SES-1530791 is gratefully acknowledged. Chandrasekhar and Morten thank the Alfred P. Sloan Foundation for financial support.
[‡] Department of Economics, Stanford University; NBER; JPAL.
[§] Department of Economics, Stanford University; NBER.
[⋆] Department of Economics, Princeton University.

1. Introduction

"The best way to find out if you can trust somebody is to trust them."

—Ernest Hemingway

Entrepreneurs often leverage networks in order to hire workers, particularly so in the developing world. The literature has identified two primary benefits of network hiring. First, there is *referral value*. Entrepreneurs know the ability of network members, and, further, screening in referrals may mean these individuals are better than average. Second, as entrepreneurs are often in a setting where effort is not observable, there is a *relational value* to network-based hiring. While there may not be strong external incentives for strangers to exert effort, with network members they can leverage repeated interactions, supported links, implicit enforcement by within-firm referral provision, or even altruism to provide incentives. These ideas have been studied at length in a large empirical and theoretical literature (e.g., Munshi (2003); Karlan, Mobius, Rosenblat, and Szeidl (2009); Beaman and Magruder (2012); Pallais and Sands (2016); Beaman (2016); Heath (2018)).[1] However, network-based hiring is not without its limits. A key practical constraint is that networks tend to be sparse and clustered: an entrepreneur who relies exclusively on hiring from their own network will not have many potential employees.[2] This leads to two natural questions: first, whether the reliance on network hiring distorts an entrepreneur's organizational choices of firm size and composition, and second, whether each entrepreneur's decision has broader implications for other firms in the market through experimentation, learning from other firms, or general equilibrium effects.

The empirical literature points out three positive *local* benefits of hiring from networks owed to the referral and relational value. The entrepreneur knows the quality of the worker, does not doubt worker effort, and understands that these workers are positively selected.[3] But we show that exactly these local features can be *globally* detrimental. Referral and relational value can result in inefficient firm organization, where entrepreneurs run firms that are too small and too network reliant, generating welfare losses. Further, in dynamic settings where firms

[1]Other key references (although incomplete), include Montgomery (1991); Calvó-Armengol (2006); Bandiera, Barankay, and Rasul (2009); Bloom, Sadun, and Van Reenen (2012); Fainmesser (2013); Fafchamps and Moradi (2015); Burks, Cowgill, Hoffman, and Housman (2015).

[2]For a literature on sparse networks see, e.g., Ioannides and Datcher Loury (2004); Cross and Borgatti (2004); Jackson (2008); Leskovec et al. (2009); Ugander et al. (2011); Schweinberger and Handcock (2015); Chandrasekhar (2016); Graham (2015). In the context of this paper, by an entrepreneur's network we mean the collection of her friends, kin, or referrals thereof for whom information frictions are alleviated, rather than the broader network or graph of links between all individuals. In other contexts, one might call this an (extended) "network neighborhood".

[3]Another natural consideration with network hiring could be if entrepreneurs derive utility (nepotism) or expect reciprocity for hiring friends. We briefly describe in Section 2 how our model could be adapted to accommodate this. We show the extent to which such an extension does not change our results and therefore omit it for parsimony.

are uncertain about the quality of the general labor market and learn through experimenting with organizational structure or observing others, these pathologies can be amplified and lead to an *information poverty trap* whereby the market fails to learn that firm expansion is indeed profitable.

Specifically, in this paper, we provide a theoretical characterization of an entrepreneur's organizational decision and the tension they face between exclusively relying on network-based hiring, and therefore running a smaller firm, or additionally hiring strangers from a general labor market pool and building a larger firm. Throughout the paper we use the term friend synonymously with the entrepreneur's network members. The general pool can consist of workers from other entrepreneurs' networks and workers, such as recent migrants, that are not connected to any entrepreneurs. Friends differ from strangers in two ways. First, the entrepreneur knows the friend quality (modeled as the cost of providing high effort), and so can set a wage exactly sufficient to induce the friend to work. In contrast, the entrepreneur knows only the distribution of worker quality in the pool, so the wage offered by the entrepreneur implies a threshold such that a stranger with effort below than this threshold will accept the wage and one with effort above it will not. The former is costly as the stranger extracts an information rent from the entrepreneur; the latter is costly as the entrepreneur forfeits their fixed capital cost if they do not successfully hire. Our baseline analysis considers wages that are linked (*tethered*) between friends and strangers. As we discuss in Section 2, this assumption is the empirically-relevant one to consider for wage setting in firms in low-income countries, motivated by the empirical literature.[4] An entrepreneur with a high-quality friend will need to offer a stranger a higher wage than they would pay their friend; because of tethering, they will need to pay their friend this wage as well. This change in incentives when a stranger is employed discourages the entrepreneur from expanding and makes it more likely they will run a small networked firm.[5]

The second difference between a friend and a stranger is that the friend may not be morally hazardous, while a stranger will be. As a result, the stranger wage contract needs to incentivize high effort, but the friend contract can be flat. Again, in the base case of tethered wages, the entrepreneur trades off expanding and employing a stranger at the cost of needing to now pay the friend the incentivized wage. This factor also leads the employer, for the same quality friend, to be more likely to run a small networked firm than to expand to a large firm.

[4]Studying tethered wages allows us to nest the case of concerns of fairness and equal treatment or even–quite distinctly–imperfect capacity to execute side payments which all have empirical support. See, among others, Breza, Kaur, and Shamdasani (2017) and references therein. Nonetheless, we provide all results for the case of independent wages for completeness.

[5]We have written the problem such that better-quality friends (those with a lower cost of effort) will earn a lower wage. Our setup is isomorphic to one where the entrepreneur pays a piece-rate wage per unit of output, and a worker with a lower cost of effort will produce more units per hour and so will indeed earn more income. We present it this way for simplicity.

The firm choice is therefore between running a small, networked firm (i.e., employing only the friend); running a large, networked firm (i.e., employing a friend and a stranger); running a large, non-networked firm (employing two strangers); or choosing not to operate.[6] In the case of tethered wages, the entrepreneur chooses to run a small, networked firm when the cost of capital is high and the friend is of sufficiently high quality; chooses to run a large, networked firm when the cost of capital is low and the friend is of sufficiently high quality; chooses to run a large, non-networked firm when the cost of capital is low and the friend is not high quality; and chooses to shut down when the cost of capital is high and the friend is not high quality. The referral and relational value of the network affects the entrepreneur's organizational choice. The effect of the referral value on hiring decisions is non-monotonic. The stronger the referral value, the more likely the entrepreneur will start a business. However, since the referral value is only present when hiring network members, it discourages firm expansion. The relational value tends to strengthen this effect. When hiring only a friend, the wage contract includes insuring the worker, since effort is observable, which generates an additional disincentive to firm expansion. In the case of independent wages the intuition is similar, but the nonmonotonicity between friend quality and organizational choice is not present. Instead, the decision to run a small or large firm depends on whether capital is above a threshold value; the choice of whether or not to hire a friend depends on the threshold value of the friend.

We then study whether network hiring leads to welfare losses for the economy. We define welfare as the sum of profits and workers' surplus; welfare is thus broader than simply firm's profits (which is what the entrepreneur maximizes). Our economy features three sources of inefficiency: (i) the quality of the stranger is not observed before the wage is set, so the entrepreneur is trading off the probability of the worker accepting the wage against the higher wage payments; (ii) effort is unobservable, so the wage needs to incentivize effort, which inefficiently exposes the workers to risk; (iii) in the baseline case, wages are equal for all workers in the same firm, which discourages firms from hiring strangers when they have a particularly good network connection. While it may seem like the network correctly solves the first two sources of inefficiency, and so hence welfare may be higher, we show that welfare may be lower in the economy when entrepreneurs are networked than when they are not. This striking result occurs because a network inefficiently encourages firms to operate at a small scale and hire too many friends relative to strangers.

We prove three results. First, being connected to a better friend can reduce welfare. If the entrepreneur has a very high-quality friend and the capital cost is low enough, there is a quality threshold where the entrepreneur switches between running a small, friend-only firm and running a large, networked (friend/stranger) one. Welfare in the economy discontinuously jumps at this threshold. Second, welfare may be higher if the entrepreneur had no friend at

[6] In our setup, revenues are linear in the number of workers and therefore hiring only one stranger is never optimal.

all. If the entrepreneur has a low-quality friend and the capital cost is high enough, there is a threshold where the entrepreneur switches from running a small, friend-only firm to running a large, stranger-only one. Again, welfare discontinuously jumps at this threshold. Third, having a friend who solves the moral hazard problem (one with both referral and relational value) may be worse than having a friend with only referral value. Similar to the previous thresholds, the negative welfare effect is larger if the network provides more benefits to the entrepreneur as these extra benefits lead the entrepreneur to continue to hire the network member instead of expanding.[7]

Finally, we consider two model extensions – general equilibrium effects and a social learning problem – to characterize how network hiring interacts with the broader economy. In terms of general equilibrium effects, we first consider whether endogenizing the outside option of workers changes the equilibrium. Our model assumes that an entrepreneur who has a high-quality friend can pay this friend a low wage. If the friend had the option to work in the labor market, would they still accept this low wage? Focusing on only the referral effect, we show that general equilibrium considerations could actually exacerbate the key tension that entrepreneurs are too likely to hire from their network. The mechanism for this operates via the outside option. As more entrepreneurs hire their network members, fewer jobs are available, and thus it becomes harder for other workers to find a job outside their own network. As a result, workers' outside options are lower, and entrepreneurs can pay their high-quality friend lower wages and are therefore disincentivized from hiring an additional stranger.

Second, in terms of the social learning problem, all the previous results rest on the fact that entrepreneurs know the distribution of worker quality in the economy, and based on this, they choose whether to operate large or small firms. We relax this assumption and study the arguably more realistic case where entrepreneurs have only beliefs over this distribution. Then, if an entrepreneur hires from the general market pool based on these beliefs, this choice constitutes an experiment. The outcome of the production process generates a signal about the worker's type, and since the worker was drawn randomly from the distribution, this is a signal about the underlying quality distribution itself. We study both a single forward-looking Bayesian entrepreneur and a social learning problem where, as entrepreneurs learn about the quality of strangers, they pass this information on to other firms. The main result is that, because of the incentives to run small networked firms, if non-networked individuals are truly high quality, the economy can become stuck in an information poverty trap. Entrepreneurs then stop hiring strangers and so, therefore, maintain their beliefs that strangers are low quality, leading them to

[7]In the case of independent results we still find scope for network hiring to cause welfare losses. In this case, the result is not that firms are too small, but that firms inefficiently hire friends over strangers, i.e., there is a welfare discontinuity at the friend quality such that the entrepreneur switches between running a large, networked (friend/stranger) firm and a large, non-networked (two stranger) one. The information rent effect drives this discontinuity.

run a small, networked firm with the highest possible welfare losses.[8] Further, the information poverty trap can increase with the strength of both the referral and relational values and lead to even more firms organizing suboptimally as small and networked. We show that there is an important asymmetry: if strangers are truly low quality, then entrepreneurs will always learn that it is optimal to rely on network-based hiring only.

Our paper makes three core contributions. First, we contribute to an understanding of the organizational structure of a firm. We take the applied literature on referral networks, particularly in development, as a starting point. Experiments have shown that entrepreneurs (a) tend to know the quality of their network members (Beaman and Magruder, 2012; Pallais and Sands, 2016; Chandrasekhar et al., 2018), and (b) face fewer moral hazard concerns when employing their referred members (Heath, 2018; Pallais and Sands, 2016). Our model takes these results at face value, assuming that the friend quality is known to the entrepreneur (and we focus on the case where they are typically of high quality) and also that friends intrinsically have no moral hazard concerns.[9] We simply add these assumptions to the usual moral hazard contracting setup (see, e.g., (Salanie, 2005)). Despite these unambiguously positive, local effects of networks that serve as informal substitutes for poor formal contracting environments with limited information, there is a potential for significant adverse consequences globally because the presence of such networks discourages entrepreneurs from hiring non-networked members. These unintended consequences already arise when considering an entrepreneur in isolation and are exacerbated at the market level.

Second, we study market-level firm organization considering two cases: dynamic experimentation and social learning. In the dynamic experimentation case, we consider a single, forward-looking Bayesian entrepreneur over time. In the social learning case, we consider the classical environment of an individual sequence of Bayesian entrepreneurs, each living a single period (e.g., Banerjee (1992); Smith and Sørensen (2000); Eyster and Rabin (2010)). We depart somewhat from this literature in that each entrepreneur observes the entire information set generated by all prior agents. In both cases, we show that there may be an information poverty trap at the market level. When entrepreneurs are uncertain about the quality of a general stranger market pool, network quality discourages hiring from the pool, which would be necessary for self or social learning to occur. As a result, even if entrepreneurs have incentives to experiment and, in the social learning case, even if they have full information set about all other firms' previous organizational decisions and outcomes, there is a positive probability that they will hire only their friends and thus a share of firms will remain small. This is so

[8]In our setup draws from the referral distribution are not informative about the general stranger distribution, which seems a natural representation given that strangers — for example, a recent migrant in the city — are often strangers precisely as they are different from the network and so plausibly have a different distribution of abilities. Thus, knowledge of a friend's idiosyncratic type, therefore, does not help to learn the aggregate state — the distribution of types in the population — but instead inhibits learning by discouraging experimentation with the pool.

[9]We deliberately abstract from nepotism or other ex ante detrimental effects of network-based hiring.

even when a social planner (that maximizes total welfare in the economy) would want them to hire strangers as well and create large firms. Again, better-performing referral networks may exacerbate this effect.[10]

Lastly, our paper highlights how the quality of institutions can affect patterns of firm structure, organization, and growth and thereby generate systemic welfare losses. In countries where there is either more scope for moral hazard or less information about workers, we would expect entrepreneurs to rely more on their networks. Given that networks are sparse, this reliance not only affects static choices of organizational structure but also impedes learning about the general market and hence entrepreneurs' willingness to expand their firms. Our paper proposes a novel way of thinking about the slow growth of firms in developing countries (Hsieh and Klenow, 2014; Bento and Restuccia, 2017; Hsieh and Olken, 2014), and the seeming presence of "subsistence entrepreneurs," typically interpreted as lacking the will or ability to expand their firm (Schoar, 2010; Akcigit, Alp, and Peters, 2016; McKenzie and Woodruff, 2016).[11]

Organization. The remainder of the paper is organized as follows. Section 2 presents our main environment. We also characterize the optimal contracts and expected profits for each organizational structure as a function of fundamental parameters (e.g., capital costs, worker quality distribution, and friend quality). In Section 3, we study the optimal firm size and composition between networked and pool workers as well as the associated welfare losses. Section 6 turns to learning dynamics. We study the case where firms' organizational decisions serve as experiments and look at the limiting distribution of firm organization in the market. Section 7 is a conclusion.

2. Model of Entrepreneurs and Workers in and out of Network

2.1. **Overview.** We study an environment in which an entrepreneur (she) hires a worker (he) or a set of workers either from her network (also called friends) or from a pool of strangers. The entrepreneur chooses the scale of the firm (here the number of workers), the composition of workers (share of friends and strangers), and the wage structure. For simplicity we study the case where the firm could have either one or two workers, capturing small or large firms.[12] And because networks are sparse, so an individual has only few links, in the context of our model we will say that the entrepreneur has a single friend. To expand, the entrepreneur must therefore be willing to hire from the general stranger pool.

[10]This may be hard to overcome; indeed, the few labor drop experiments find that short-term subsidies do not shift entrepreneurs to hire strangers post-experiment, suggesting strong priors (see, e.g., Groh et al. (2016)). This finding is also consistent with our own piloting in the field where entrepreneurs exhibit considerable skepticism over stranger quality.

[11]Our paper may also offer another perspective on some of the patterns documented in the family firm literature, which primarily focus on different mechanisms (see Bertrand and Schoar (2006) and references therein).

[12]The hiring and expansion incentives we study in this simple setup would be equally at play in a more general setting where firm size depends on demand and productivity.

Motivated by the applied literature we focus on two information frictions – asymmetric information and nonobservable effort – that occur with general labor market hiring but that might be overcome by network hiring. First, asymmetric information exists in that the entrepreneur does not know the worker's cost of exerting effort. Second, she might not be able to observe whether the worker exerts effort, generating the potential for moral hazard.[13] There are, of course, other reasons for why entrepreneurs hire their friends and family, chief among them altruism. We deliberately focus on asymmetric information and moral hazard, since these have been documented to be locally beneficial.[14]

The first friction, asymmetric information, exists only with strangers. We assume that the entrepreneur knows her friend, and therefore she knows the private cost for this friend to exert effort. We call this the *referral value* of the network. This positive value of hiring network members is consistent with the empirical literature, which documents that individuals know more about their network members' abilities (Beaman et al., 2014; Pallais and Sands, 2016). However, the entrepreneur does not know the stranger and therefore has only beliefs about the stranger's cost of effort.

We study the second friction—no observable effort— both where is equally present with friends and strangers and where it is absent with friends. In the latter case, we say there is *relational value* to network-based hiring. The relational value captures a reduced form for any force, such as a repeated game between the entrepreneur and friend or altruism between the entrepreneur and friend. Again, this positive value of network hires is documented by the empirical literature (Heath, 2018) and, more generally, by a literature showing that interactions with network members overcomes contracting frictions (Pallais and Sands, 2016; Heath, 2018; Chandrasekhar et al., 2018). The asymmetric incentive constraint makes the optimal wage the entrepreneur offers to a worker depend on whether the worker is a friend or stranger.

In our primary analysis, we maintain a *tethered wage* assumption—that the wages of any two workers hired by the same firm are tethered. That is, the entrepreneur must offer the same wage to every worker in her firm, whether friend or stranger. The focus on tethered wages is natural, in line with empirical and theoretical literatures that document how considerations of fairness prevent entrepreneurs from paying workers in the same job different wages and how contracting frictions make it difficult to overcome this by executing side-payments.[15] Qualitatively, all of

---

[13]Our framework nests two additional frictions that could be alleviated by hiring friends: first, strangers might be able to steal the installed capital, and second, hiring in the general labor market might take more time than hiring through referrals. We discuss these in more detail below.

[14]If entrepreneurs derive utility, or also expect reciprocity, for hiring their friends, this would act as a shift in profits in our analysis. While shifting the relevant parameter space, this extension would not affect our analysis of who gets hired at the margin.

[15]In fact, in our own survey of 85 small firms (such as small convenience stores, restaurants, medical shops, clothing stores, and stationery shops) in six neighborhoods in Bangalore, Delhi, and Patna, we found that 76% of entrepreneurs tether wages and pay workers identical piece rates. Qualitative evidence behind this survey is provided in Appendix D. Though beyond the scope of our paper, prior research has delved into the motivations behind this tethering, perhaps chief among them coming from morale effects of pay inequality. See Breza, Kaur,

our results for the tethered wage case go through as long as there is some — even minimal — cost of paying workers differently. For completeness, all propositions also include results for the knife-edge case where wages can be entirely independent between workers.

Let us briefly discuss the role of each of the three key components — referral value, relational value, and tethered wages — in our theory. First, the referral value on its own gives entrepreneurs a strong incentive to hire out of their network and hence affects the *composition* of firms. Friends are sure to accept the wage contract and are not paid the usual information rent that arises with private information. Since part of the benefit of network-based hiring to the entrepreneur is that she avoids paying rents to workers, this force tends toward hiring fewer strangers relative to a social planner, as we will discuss in detail in Section 4. Second, the relational value gives entrepreneurs an even greater incentive to hire out of their network, since workers need not be exposed to income risk in order to incentivize high effort. In contrast to the referral value, the relational value on its own is always welfare improving. However, when the other two frictions are present, it can amplify welfare losses from network-based hiring. Third, the tethered wage generates an interaction between different workers in the firm, one that increases with worker heterogeneity. This assumption is what generates most of the effects on firm *size*. When an entrepreneur hires an additional worker, the larger the difference between the two workers, the lower the profits. In our setup, the difference between workers arises naturally as a consequence of information frictions being solved for a subset of workers – the network connections.

Taken together, all three core elements of our model allow us to study an environment where network-based hiring is useful to overcome contracting frictions, as has been emphasized by an applied literature. As such, welfare is increasing *locally* in the extent to which networks solve information frictions. However, due to the sparseness of networks, firm expansion requires hiring from the general labor market. The decision whether or not to expand is intuitively interlinked with the firm's network-based hires. As such, *globally*, i.e., when taking firm size and composition into account, the welfare effects of network-based hiring are ambiguous.

2.2. **Baseline Model.**

2.2.1. *Environment.* A single entrepreneur hires from two potential sources of workers: the network (friends) and strangers. The entrepreneur picks the number of friends n_f and the number of strangers n_s and therefore selects a firm size of $n = n_f + n_s \leq 2$. As networks are sparse, we model this by letting the entrepreneur only have one friend, so to run a large firm they must hire a stranger as well. Hiring each worker involves a fixed cost of K. This can be thought of as a required capital investment made by the entrepreneur before extending an offer to a potential worker.

and Shamdasani (2017) for both an experiment on this theme and an extensive discussion of the empirical literature.

Output by a given worker is determined by worker effort, $e \in \{0, 1\}$. If $e = 0$ then output is 0, whereas if $e = 1$ then output is y with probability p, and 0 otherwise. Without loss of generality and for simplicity, all workers have outside option zero.[16]

Workers from the pool have common strictly increasing, concave, differentiable utility functions $u(\cdot)$ with $u(0) = 0$. Workers are heterogeneous in their cost of effort c drawn from an absolutely continuous CDF $F_c(\cdot)$ on $\mathbb{R}$. Entrepreneurs cannot observe an individual worker's cost of effort c. They offer a wage contract to a randomly chosen worker from the pool. If the worker does not accept, the firm cannot make another offer. This is a reduced form for the idea that search can be slow in practice. We assume here and in Section 3 that the population distribution F_c is known to entrepreneurs. In Section 6 we make the natural assumption that nascent businesses may not know the distribution itself and may need to learn it.

In what follows, we assume that the entrepreneur faces the problem of potentially morally hazardous behavior by her employees. Moral hazard is the natural situation in our analysis. For instance, effort on typical jobs held by the poor in the developing world is not observable. In Appendix C, for completeness, we characterize the case with no moral hazard as well.

The entrepreneur's friend has a cost of providing effort c_f, which is known to the entrepreneur. The fact of this knowledge simply encodes the findings from the empirical literature (e.g., Beaman et al. (2014); Pallais and Sands (2016)). In the economy as a whole, c_f is distributed according to some distribution, $F_f(\cdot)$ not necessarily equivalent to $F_c(\cdot)$.[17] While we study both cases, we will often focus on the case motivated by the findings of the empirical literature where network referrals are of "good types" corresponding to known low c_f (e.g., Beaman et al. (2014); Pallais and Sands (2016)).

The entrepreneur's friend has a cost of lying to them, parametrized by δ. This is a reduced form for repeated game dynamics, trust, altruism, or any such force that would make a friend easier to incentivize than a stranger. If $\delta = 0$, there is no trust advantage of hiring a friend, whereas $\delta = \infty$ maps into the case where there is no moral hazard when hiring a friend. While we study both cases, the predominantly empirically relevant case, motivated by the findings of the empirical literature, is that δ is very high and so friends cooperate with entrepreneurs even

[16]Here we assume that the outside option of workers is fixed and, in particular, does not depend on the other parameters of the market environment. This will be true, for instance, in a newly developing industry, where the outside option can be thought of as the payoff from working in more established industries. Section 5 analyzes the case where the outside is endogenous to hiring decisions of entrepreneurs.

[17]There are a number of reasons these distributions do not need to coincide. First, we think that network referrals in part have screening value as noted by the empirical literature. So friends will be positively selected. Second, a general worker pool need not simply consist of friends of other entrepreneurs, but furthermore can also include other sources of potential workers such as recent migrants. Third, by the friendship paradox, those who are linked to entrepreneurs must have more links than a random person and therefore do not constitute a random draw. For further information on the friendship paradox see Feld (1991) and Jackson (2019). Fourth, there is homophily. If entrepreneurs are of a "higher type," then their connections will tend to be of higher type as well (Montgomery (1991)). We do not formally explore these interpretations and view our assumptions on c_f as a reduced form.

without formal contracts (e.g., Pallais and Sands (2016); Heath (2018); Chandrasekhar et al. (2018)).

Our setup is chosen to nest the primary justifications in the literature for referrals and usage of network hiring. By allowing a friend's effort cost to be observable and drawn from a different distribution, F_c, we capture the *referral value*, which directs higher-quality individuals to the entrepreneurs. δ, the friend's cost of lying, captures the *relational value*, which generates better behavior workers who are friends. The difference between F_f and F_c could also capture, in a reduced-form way, the fact that labor search is slow in practice. We would expect that it takes longer to hire from the general labor market than through a network, reflected here as a difference between the "effective" effort cost distributions F_c and F_f. Our setup also nests the notion that strangers might not be trustworthy in the sense that they could steal, for instance, the installed capital from the entrepreneur. Any mass of F_c above the threshold where a worker is employable can be interpreted as the worker stealing K. If no production happens, K is wasted and from the point of view of the entrepreneur this is analogous to the worker running away with her capital.

The timing is as follows:

(1) The entrepreneur decides n_s, n_f, and a wage contract s_y^i for each worker i as a function of output $y \in \{0, 1\}$, paying $K \cdot n$ to start a firm of size n.
(2) Each of the n workers decides, independently, whether to accept or decline the contract.
(3) A set of $\hat{n} = \hat{n}_f + \hat{n}_s$ workers remain with the firm and each selects an effort level e_i for $i = 1, ..., \hat{n}$.
(4) Outputs y_i are realized.
(5) Payoffs are received.

The entrepreneur decides the firm size $(n_f + n_s)$, the composition of workers in terms of friends (n_f) and strangers from a general market pool (n_s), and the wages $\{s_0^i, s_1^i\}_{i=1}^n$ for each worker i to maximize expected profits:

$$\max_{\{s_0^i, s_1^i\}_{i=1}^n, n_f, n_s} n_f \left[r_f(s_0^i, s_1^i; c_f, \delta) \right] + n_s \, \mathbb{E}_c \left[r_s(s_0^i, s_1^i; c) \right] - (n_f + n_s) \cdot K$$

$$\text{s.t. } n_f + n_s \leq 2 \qquad \text{(a) size constraint}$$

$$\varphi \left(\sum_{i,j} (s_0^i - s_0^j)^2 + \sum_{i,j} (s_1^i - s_1^j)^2 \right) \qquad \text{(b) tethered wages constraint}$$

$$r_f(s_0^i, s_1^i; c_f, \delta) := \begin{cases} p(y - s_1^i) + (1-p)(0 - s_0^i) & \text{if } \ pu(s_1^i) + (1-p)u(s_0^i) - c_f \geq 0 \\ & \text{and } \ pu(s_1^i) + (1-p)u(s_0^i) - c_f \geq u(s_0^i) - \delta \\ (0 - s_0^i) & \text{if } \ u(s_0^i) \geq 0 \\ & \text{and } \ pu(s_1^i) + (1-p)u(s_0^i) - c_f < u(s_0^i) - \delta \\ 0 & \text{o/w} \end{cases} \qquad \text{(c) IC: friend}$$

$$r_s(s_0^i, s_1^i, c, \delta) := \begin{cases} p(y - s_1^i) + (1-p)(0 - s_0^i) & \text{if } \ pu(s_1^i) + (1-p)u(s_0^i) - c \geq 0 \\ & \text{and } \ pu(s_1^i) + (1-p)u(s_0^i) - c \geq u(s_0^i) \\ (0 - s_0^i) & \text{if } \ u(s_0^i) \geq 0 \\ & \text{and } \ pu(s_1^i) + (1-p)u(s_0^i) - c < u(s_0^i) \\ 0 & \text{o/w.} \end{cases} \qquad \text{(d) IC: stranger}$$

That is, entrepreneurs consider the fact that friends want to participate and apply effort (incorporating the cost of lying, δ); strangers who participate find it incentive-compatible to apply effort; both friends and strangers receive at least their outside option; and firms need to choose a feasible scale. The constraint (b) nests both the case we focus on, tethered wages ($\varphi = \infty$), and the knife-edge case with fully independent wages ($\varphi = 0$) which we include for completeness.[18]

2.2.2. *Optimal contracts.* In this section, we describe the trade-offs entrepreneurs face when choosing wages for each of the four possible organizational forms: friend only, stranger only, friend and stranger, and two strangers. Appendix C contains a full characterization.

When hiring only a friend, the entrepreneur chooses a wage contract that exactly satisfies the friend's participation and incentive constraints. The optimal wage depends on the cost of lying, δ. If $\delta = \infty$, the incentive constraint is always satisfied, and the entrepreneur pays her friend a flat wage. If $\delta = 0$ on the other hand, the wage profile is chosen to be as steep as possible by setting the wage in case of failure equal to zero to ensure that friends exert high effort.

When hiring a stranger, the optimal wage also features maximal incentives, so zero payments in the case of low output. The wage offered in case of high output (s_1) solves the following first-order condition:

$$F(\bar{c}) = f(\bar{c})\, u'(s_1)\, p(y - s_1) \tag{2.1}$$

where $\bar{c} = pu(s_1)$ is the effort cost corresponding to the worker who is just indifferent between accepting the contract or not. The optimal wage trades off the benefits of a marginal increase in the probability of the worker accepting — the right-hand side of equation (2.1) — against extra payments that have to be made to all workers who accept the contract – the left-hand side. The

[18]Qualitatively, all results we show for $\varphi = \infty$ hold for any $\varphi > 0$. For exposition, we focus on the two polar cases only.

equation also illustrates how asymmetric information affects who gets hired. If the entrepreneur were able to observe c, she would hire any worker for whom $c \leq pu(y)$. That worker would be paid the entire surplus $s_1 = y$. This cannot be the optimal contract with unobserved type and $\bar{c}$ is lower than $pu(y)$. Because the model has linear payoffs, an entrepreneur who hires two strangers offers the same wage, and payoffs simply scale linearly.

As long as firm wages are tethered (i.e. $\varphi > 0$), interesting interactions occur between friends and strangers in a large, mixed firm. Consider first the case where $\delta = 0$, which corresponds to a world in which both friends and strangers are offered steep incentives. The wage offered in case of high output solves the following first-order condition

$$1 + F(\bar{c}) = f(\bar{c})\, u'(s_1)\, p(y - s_1) + pu'(s_1)\,(\lambda + \mu) \tag{2.2}$$

where λ and μ are the Lagrange multipliers on the friend's participation and incentive constraints. If the friend is of low quality, so c_f is sufficiently high, then these constraints are binding, and the wage is simply set to satisfy the friend's constraints. If the friend is of high enough quality, $\lambda = \mu = 0$, and the optimal wage is the solution to (2.2). Compared to the case where only a stranger is hired, an additional cost of offering a higher wage occurs. This higher wage must be paid not only to all strangers who accept the contract, but also has to be paid to the friend. The optimal wage is, therefore, lower than in a stranger-only firm, and fewer workers from the pool are hired. In other words, the presence of a high-quality friend who must be paid the same wage offered to the pool makes the entrepreneur more selective in terms of hiring strangers.

Next consider the case where $\delta = \infty$. Now there is not only a potential level difference between the optimal friend and stranger wages but also a difference in the slopes. For sufficiently good friends, the optimal wage contract is the same as in the $\delta = 0$ case, as it is dictated by the incentives of workers from the pool. For high enough values of friend cost, the optimal wage might now feature a positive compensation in the case of firm failure. The trade-off here is between cheaply satisfying the friend's participation constraint (by providing insurance through the wage contract) against discouraging effort by pool workers.

If instead, wages are fully independent within the firm, the contracts offered to a friend and stranger in a large, mixed firm are simply identical to the ones offered by small firms.

Figure 1 illustrates how the firm's profits (gross of the capital cost K) depends on both the friend's effort cost and the presence of the relational value. When the firm hires only strangers, profits are, of course, independent of friend type. When only a friend is hired, profits decrease in c_f and are higher when the friend has a cost of lying. The grey lines in Figure 1 correspond to the additional profit from hiring a stranger if the entrepreneur already employs her friend. This illustrates three important properties of the model in the baseline case of tethered wages.

The first two properties hold independently of the value of δ, that is, both where the friend solves moral hazard concerns and where he does not. First, hiring from the general labor market

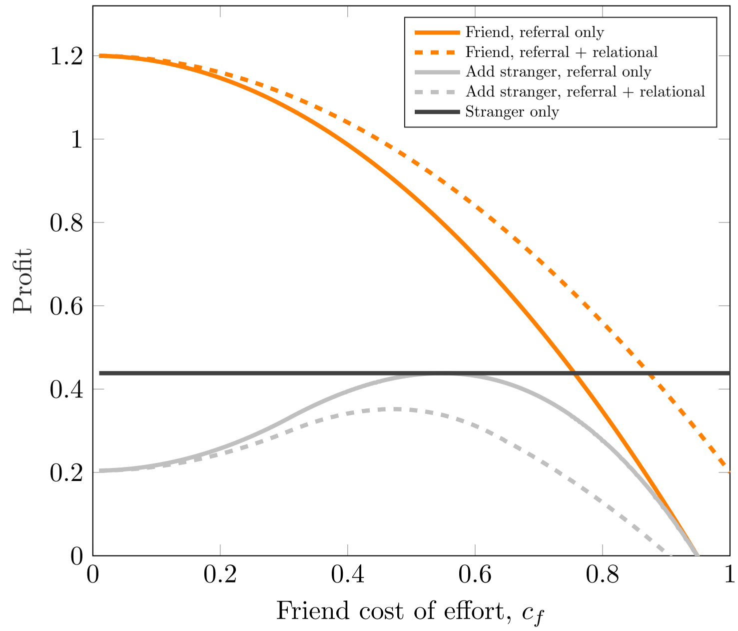


FIGURE 1. This figure shows firm profit, gross of capital cost, as a function of the known friend's cost of effort. The figure shows the profit from hiring a friend only as well as the marginal profit from expanding the firm and hiring a stranger. The profit from hiring a stranger only (which does not depend on the friend cost of effort) is given by the horizontal line.

is always less profitable when the entrepreneur also hires her friend. Since all workers must be paid the same wage, except in the knife-edge case where optimal payments to the friend and the pool worker happen to coincide, payments to one of the types of agents are distorted, and profits from expanding the firm are lower.

Second, the additional profit from hiring from the pool is non-monotonic in friend quality. For very high-quality friends, the optimal wage is lower if they are the only employee. When a friend is added to the firm, the entrepreneur must pay her friend a higher wage in order to have a reasonable chance of attracting a pool worker. The higher is the friend's effort cost, the smaller is the difference in optimal wages, and hence the higher the additional profit from hiring a stranger. At some point, the optimal wage paid to the pool worker is lower than what the entrepreneur must pay her friend. When running a mixed firm, therefore, she overpays workers from the general labor market, and in this region, additional profits are declining in the friend's effort cost.

Third, when the friend has a high cost of lying to the entrepreneur, profits from hiring the friend only increase, but additional profits from hiring a stranger decrease. In this case, there is not only a difference in the optimal *level* of wages between the friend and the stranger but also in the optimal *slope* of the wage contract. This additional difference in optimal compensation reduces the profitability of a large, mixed firm relative to a small, networked one.

In the knife-edge case where wages are fully independent within the firm, the additional profit from hiring a stranger is independent of who else is hired and always equal to the horizontal solid line in Figure 1. This is true irrespective of whether or not friends solve the moral hazard problem.

3. Optimal Firm Size and Workforce Composition

We now analyze the optimal firm organizational structure—size and composition—as a function of c_f, the quality of the entrepreneur's network connection, and K, the cost of capital.[19] We denote the choice by $x(c_f, \delta, K) : \mathbb{R}^0_+ \times \mathbb{R}^0_+ \times \mathbb{R}^0_+ \to \{\text{shutdown}, F, S, FS, SS\}$, where shutdown means the firm does not operate; F (or S) means the entrepreneur operates a small, friend-only (stranger-only) firm; FS means the entrepreneur operates a large, networked (friend/stranger) firm; and SS means the entrepreneur operates a large firm consisting of two strangers. We use $\pi(i, c_f, \delta)),\ i \in \{\text{shutdown}, F, S, SS, FS\}$ to denote expected profits *gross* of capital cost.

We present the referral value effect separately first and then add the relational effect. This allows us to see the role that each force plays.

3.1. Referral Value Effect.

Let us begin with the referral effect. To think about this, we imagine a case with moral hazard for both friends and strangers, so being a network member has no value other than the fact that their quality is known. Friends—or individuals who are referred—tend to be better than average, as discussed in the empirical and theoretical literature (owing to either to screening or the friendship paradox). So let us begin with the idea of a "good friend," formally defined below.

The wage to incentivize a good friend to join the firm does not have to be particularly high. After all, their intrinsic cost of effort is rather low when compared to the distribution of strangers from the general labor market. So expanding the firm and hiring a second worker who is a stranger necessitates incentivizing this unknown. A higher wage is required, but since firms offer tethered wages, the network friend will also receive the higher wage as a premium, which is costly for the entrepreneur. All told, for high enough capital cost to hire the second worker, it may not be worth it. The better the quality of the friend, the lower the tolerance in terms of capital cost there is to expand to a large, mixed firm. That is, entrepreneurs with better friends tend to be more conservative in that, as the capital cost increases, they rather quickly choose to keep only their friend as a worker and remain small.

It is worth also commenting on the case of a so-called "good enough" friend (one with a high cost of effort). We think a priori this case is less empirically relevant, i.e., inconsistent with the aforementioned empirical literature and counter to the typical intuition. Nonetheless, we describe it for completeness. In this case, a premium is paid to any stranger who is hired since the friend needs higher wages. But the friend offers another relevant advantage. The

[19] The cost of capital K and the value of output y have similar effects on the main outcomes of interest. We therefore interpret productivity of the firm as the former, but varying K can be interpreted as (inversely) varying y.

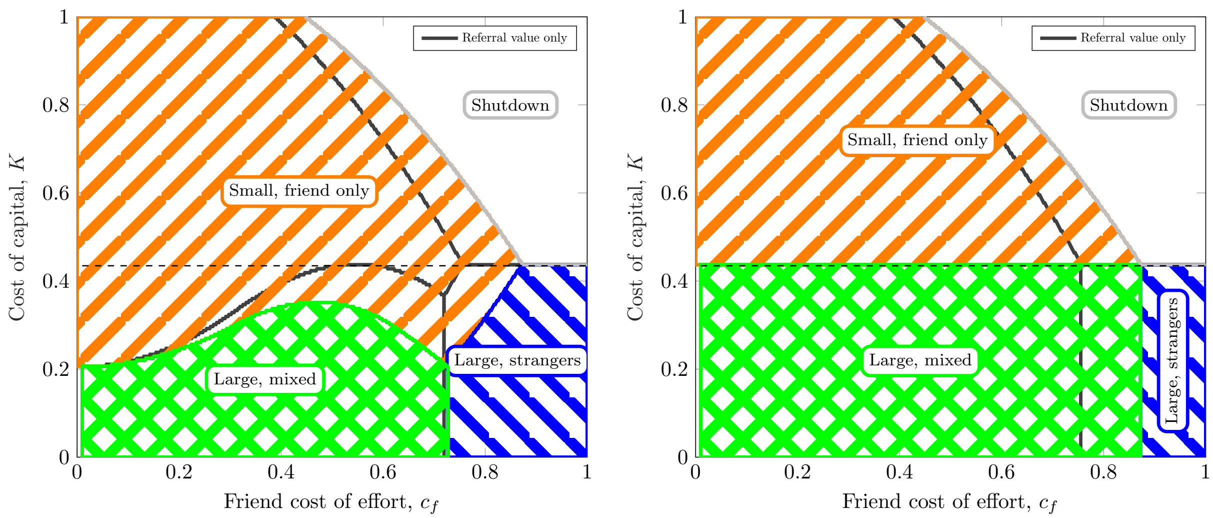
entrepreneur knows the friend's type and so can guarantee that he does not quit. So there is a trade-off here between the certainty that the friend will not quit (as opposed to expanding to a second stranger who may leave) versus the premium cost that must be provided to this "good enough" friend. When the latter dominates, a large firm comprised of strangers is preferable, whereas when the former dominates, hiring the good enough friend is preferable.

The presence of a friend, good or good enough, in fact, always reduces the tolerance in terms of capital cost for scaling up. When the entrepreneur has a good friend, expanding and hiring a stranger comes at the cost of overpaying her friend, and she is less likely to expand than a non-networked individual. That is, the threshold capital cost to run a big firm $\bar{K}(c_f, 0)$ which depends on two arguments—network quality and scope of moral hazard by friends ($\delta = 0$)—actually declines with the quality of the friend. With a good enough friend, expanding and hiring a stranger comes at the cost of overpaying the stranger, and again the entrepreneur is less willing to expand. In summary, unless the friend is precisely such that his wage coincides with the optimal stranger wage, entrepreneurs who can hire through referrals are more reluctant to run large firms.[20]

We also briefly mention the case of fully independent wages. In this case, the threshold capital cost to run a big firm $\bar{K}(c_f, 0)$ is independent of the network referral quality. This comes from the separability in profits in hiring the network referral and hiring the stranger.

Figure 2 provides a visualization of the four quadrants for type of firm. The case where the network has both referral and relational value is shown in the hatched portion of each figure; for comparison, the allocation when the network has only referral value is shown in the background. Proposition 3.1 formalizes the relationship between an entrepreneur's willingness to scale up and the quality of her friend. Appendix C extends Proposition 3.1 to the entire range of friend qualities for completeness.

[20]The point where the wages coincide is precisely the cutoff between good and good enough.



(A) Tethered wages (B) Independent wages

FIGURE 2. This figure shows the organizational choice of the firm. Panel (a) considers the case where wages of strangers and friends are tethered to one another. Panel (b) considers the case where wages of strangers and friends are independent. For each panel, the figure shows the organizational choice for the case where the friend has both referral and relational value. The black line shows how the organizational choice changes if the friend only has referral, but not relational, value. The dashed horizontal line shows the maximum cost of capital before the stranger-only firm would shut down.

PROPOSITION **3.1.** *Suppose $\delta = 0$. Let $\hat{c_f}$ be the quality level of friend s.t. the optimal wage coincides with the one offered to the pool: $u(s_1(SS)) = \hat{c_f}$. Define a "good friend" to be any friend with $c_f \leq \hat{c_f}$. Then*

(1) $\forall c_f \leq \hat{c_f},\ x(c_f, 0, K) = \begin{cases} FS & \forall K \leq \bar{K}(c_f, 0) \\ F & \textit{otherwise} \end{cases}$

(2) *If $\varphi = \infty$ (wages are tethered)*
   (a) $\frac{\partial \bar{K}(c_f,0)}{\partial c_f} > 0$.
   (b) $\bar{K}(\hat{c_f}, 0) = \max_{c_f} \bar{K}(c_f, 0) = \frac{1}{2}\pi(SS)$.

(3) *If $\varphi = 0$ (wages are fully independent)*
   (a) $\forall c_f :\ \bar{K}(c_f, 0) = \frac{1}{2}\pi(SS)$ *and therefore independent of c_f.*

Proof. All proofs are in Appendix A. ■

3.2. **Relational Value Effect.** With any friend, the relational value effect means that there is no need to provide a slope for incentive because the friend automatically applies high effort. Meanwhile, hiring a worker from the pool involves providing incentives. Because wages are tethered, the friend in a mixed firm also receives this higher wage. Therein lies the issue: a premium is given to the friend when the stranger must be incentivized to work. Among good friends, the premium can be high, and the better the friend, the higher the premium. That is, the relative cost to the firm of merely paying a wage to satisfy the participation constraint of a good friend versus giving them a pooled wage written up to induce effort by a stranger in a larger firm is higher the better the friend.

Owing to this fact, an entrepreneur with a higher-quality friend requires lower levels of capital cost per worker to simply decide to maintain a small firm with the network member. The worse the friend, the higher the capital cost the entrepreneur is willing to withstand before deciding she wants a small networked firm instead of a large mixed firm. So $\bar{K}(c_f, \infty)$, the threshold capital cost to run a big firm as a function of friend quality and lack of moral hazard for the friend, declines in the friend quality.

Compared to the previous case, when there is only a referral effect, the effective premium that must be paid to a friend in a mixed firm increases. While previously there was a level of friend quality such that a networked entrepreneur was just as willing as an unnetworked entrepreneur (shown by the black dashed line) to expand her firm, this is now no longer the case. As illustrated by Figure 2, for all levels of friend quality, the entrepreneur is willing to tolerate only a lower cost of expansion before deciding to run a small, networked firm. Notice that this is not true for non-networked individuals (the right side of both panels of Figure 2), as the relational value only affects choices in the presence of a potentially qualified friend.

In the case of fully independent wages, again the threshold capital cost is independent of the friend quality, and indeed for high enough quality friends, the threshold capital cost is the same whether or not friends also avoid moral hazard.

Proposition **3.2.** *Suppose $\delta = \infty$. Let $\tilde{c_f}$ be the quality level of friend s.t. the friend's participation constraint starts to bind at the optimal wage offered to the mixed firm: $pu(s_1(FS, \infty)) = \tilde{c_f}$. Define a 'very good friend' to be any friend with $c_f \leq \tilde{c_f}$. Then*

(1) $\forall c_f \leq \tilde{c_f}$, $x(c_f, \infty, K) = \begin{cases} FS & \forall K \leq \bar{K}(c_f, \infty) \\ F & \text{otherwise} \end{cases}$

(2) *If* $\varphi = \infty$
   (a) $\frac{\partial \bar{K}(c_f,\infty)}{\partial c_f} > 0$.
   (b) $\forall\, c_f \leq \tilde{c_f}$, $\bar{K}(c_f, \infty) < \bar{K}(c_f, 0)$
   (c) $\max_{c_f} \bar{K}(c_f, \infty) \leq \frac{1}{2}\pi(SS)$

(3) *If* $\varphi = 0$
   (a) $\forall c_f : \bar{K}(c_f, \infty) = \frac{1}{2}\pi(SS)$ *and therefore independent of* c_f.
   (b) $\forall\, c_f \leq \tilde{c_f}$, $\bar{K}(c_f, \infty) = \bar{K}(c_f, 0)$

To further clarify the different trade-offs faced by entrepreneurs with tethered versus independent wages, Proposition 3.3 compares firm composition in the two cases. We show that with fully independent wages, there is a larger effect on firm composition. This contrasts with the previous two propositions, which highlighted the effect of referral and relational value on size. This effect was only present with tethered wages. When wages instead are fully independent within the firm, entrepreneurs choose to hire their friend and a stranger for a larger range of friends, and hire purely strangers only for very high-cost friends. The reason for this is that the value of the good enough friend is higher when the entrepreneur does not need to overpay the pool worker.

Proposition **3.3.** *Let $c_f^*(\delta)$, $\delta \in \{0, \infty\}$ be the quality level of friend s.t. the entrepreneur is indifferent between choosing FS and SS when $\varphi = 0$ (no fairness constraint): $\pi(FS, c_f^*(\delta), \delta) = \pi(SS, c_f^*(\delta), \delta)$. Let $c_f^{**}(\delta), \delta \in \{0, \infty\}$ be the same quality level of friend when $\varphi = \infty$: $\pi(FS, c_f^{**}(\delta), \delta) = \pi(SS, c_f^{**}(\delta), \delta)$.*

(1) *$\forall\delta$, $c_f^*(\delta)$ and $c_f^{**}(\delta)$ are unique;*

$$x(c_f, \delta, \varphi = \infty) = \begin{cases} FS & \forall c \leq c_f^*(\delta) \\ SS & otherwise \end{cases}; \quad x(c_f, \delta, \varphi = 0) = \begin{cases} FS & \forall c \leq c_f^{**}(\delta) \\ SS & otherwise \end{cases}$$

(2) *$\forall\delta$, $c_f^{**}(\delta) < c_f^*(\delta)$.*

The fact that with independent wages, avoiding the information rent leads to more network-based hiring lies at the core of welfare losses resulting from referral and relational value, which we turn to next.

4. Efficiency and Welfare Properties

We now discuss efficiency and welfare properties of the market allocation described above. There are three sources of inefficiency in this setup. First, entrepreneurs cannot observe a worker's cost of effort before setting the wage. They set a wage that optimally trades off a higher probability of a worker accepting an offer against higher wage payments. The resulting wage is inefficiently low in that a set of workers exists that would produce in the first-best, but they are priced out in the market allocation. Second, effort is unobservable. Wages set by the entrepreneur are steep in order to provide incentives, inefficiently exposing the worker to risk. Third, in the baseline case of tethered wages, all workers in a firm must be offered the same wage contract. As discussed in Propositions 3.1 and 3.2, this discourages entrepreneurs from hiring strangers when their network connection is of sufficiently high (or low) quality. In the case of fully independent wages, as discussed in Proposition 3.3, entrepreneurs are more likely to hire their friend to avoid paying an information rent to strangers.

While it may seem like the network exactly solves the first two sources of inefficiency, and hence welfare may be higher, we show a striking result that welfare may actually be lower in the economy when entrepreneurs are networked than when they are not, because a network inefficiently encourages firms to operate on a small or excessively networked scale.

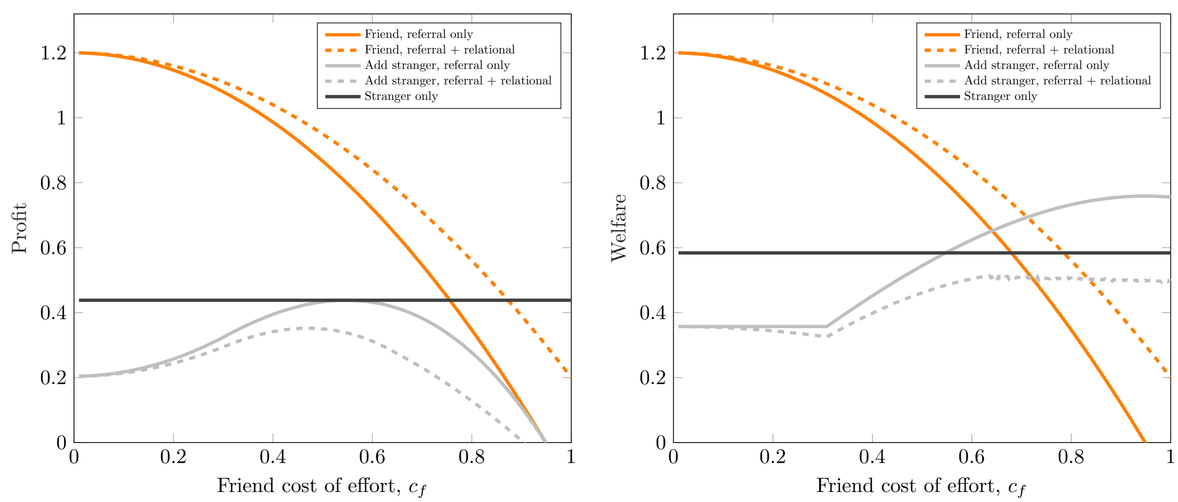
We consider a social planner that maximizes total surplus in the economy. To avoid distributional concerns, we evaluate the total surplus as the sum of entrepreneur's profits and the maximum monetary transfers workers could pay while still respecting their participation constraint. The social planner faces the exact economic environment as the entrepreneur, taking the wage contract set by entrepreneurs as given and choosing the organizational form. For instance, if an entrepreneur offers a wage contract $\{s_0 = 0, s_1 = w\}$ to the pool and also hires her friend, welfare is given by

$$WF(\text{FS}, c_f) = \Pi(FS, c_f) + \int_0^{p\,u(w)} p\,\tau(c)\,f(c)\,dc + p\,\tau(c_f)$$

where $p\,u(w - \tau(c)) = c$.

To be explicit, the only difference between entrepreneurs and the social planner is that the latter chooses an organizational structure to maximize welfare instead of profits. We refer to this welfare-maximizing allocation as the second-best. We focus on this approach for pedagogical reasons: it allows us to concentrate purely on the welfare consequences of organizational choice, i.e., hiring a friend or a stranger.[21]

Figure 3b illustrates how welfare in each organizational form depends on the friend's cost of effort. Welfare is computed assuming that the entrepreneur chooses the level of wages, so this represents welfare in the market allocation. For comparison, Figure 3a repeats the picture for profits from Section 2.



(A) Profits (B) Welfare

FIGURE 3. Panel (a) shows firm profit as a function of the known friend's cost of effort. The figure shows the profit from hiring a friend only as well as the marginal profit from expanding the firm and hiring a stranger. The profit from hiring a stranger only (which does not depend on the friend cost of effort) is given by the horizontal line. Panel (b) shows welfare as a function of the known friend's cost of effort. The figure shows welfare from hiring a friend only as well as the marginal welfare from expanding the firm and hiring a stranger. The welfare from hiring a stranger only (which does not depend on the friend cost of effort) is given by the horizontal line.

[21] If the social planner could also choose wages, they would offer a higher wage to the pool compared to the entrepreneur, which would increase the probability that a match occurs and production happens. They would also hire strangers at a higher value of K, further increasing production.

When hiring only a friend, welfare and profits coincide. Since the friend's effort cost is observable, there are no inefficiencies. When hiring only a stranger, welfare is independent of the friend's effort cost, but importantly higher than pure profits. The difference between welfare and profits corresponds exactly to the workers' information rent. Since the entrepreneur must set a wage before observing the effort cost, workers who accept the wage contract in general enjoy a surplus.

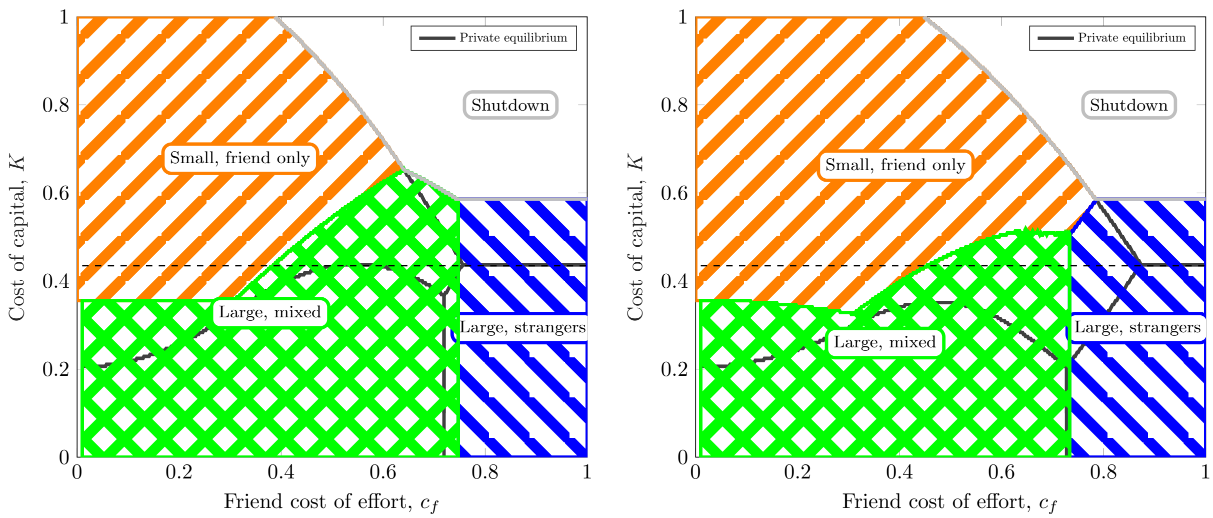
In the case of fully independent wages, the gap between profits and welfare when hiring a stranger is exactly why the social planner and the entrepreneur make different choices. The entrepreneur does not internalize the surplus to the worker. For a friend with a relatively high cost of effort, the entrepreneur would keep hiring her friend, while the social planner would have already switched to hiring from the general labor market.

With tethered wages, welfare also depends on the quality of the entrepreneur's friend, since this affects all wages offered. In the absence of the referral effect, the welfare of hiring an additional stranger is weakly increasing everywhere.[22] For low values of the friend's effort cost, the difference between welfare and profits comes from the fact that hiring an additional stranger means paying the friend a higher wage. While this reduces profits, it is a pure transfer and as such has no effect on welfare. As soon as the friend's constraint starts binding, a higher effort cost has the positive effect that entrepreneurs now offer higher wages to the pool, counteracting some of the inefficient wage shading that results from observable effort cost. Taken together, the additional welfare from hiring a stranger is strictly increasing whenever the friend's constraint is binding.

In the presence of the referral effect, there is a welfare cost of hiring an additional stranger, since friends must now be paid inefficiently steep wages. While welfare is still in general higher than profits, welfare from hiring an additional stranger is lower when friends also solve the moral hazard concern.

Figure 4 illustrates the second-best organizational choice, that is, the welfare-maximizing choice of firm size and composition conditional on the entrepreneur setting wages (suboptimally). The left panel corresponds to the case with only the referral effect ($\delta = 0$). The dashed line marks the maximum level of capital the entrepreneur would be willing to pay to run a large firm in the market allocation. For all levels of referral quality, the second-best features a weakly larger firm. When both the market allocation and the second-best involve hiring only strangers (this corresponds to the far right part of panel 4a), the discrepancy between the planner and market choices comes from the fact that workers' cost of effort is not observable. All workers who accept a contract get a surplus equal to the difference between their reservation wage and the market wage. The entrepreneur does not take this into account when deciding whether to operate. This is precisely the sense in which these frictions lead entrepreneurs to run inefficiently small firms.

[22] Depending on the assumption about F_c, it could start declining for very large values of c_f.



(A) Referral value only

(B) Referral and relational value

FIGURE 4. This figure shows the social planner's organization choice, taking the wage contract as given. Panel (A) shows the case where the network has referral value only. Panel (B) shows the case where the network has referral and relational value. In each panel, the private equilibrium is shown as a black line. The dashed horizontal line shows the maximum cost of capital before the stranger-only firm would shut down in the private equilibrium.

Through the referral effect, friends solve the inefficiency of unobserved effort cost. One might therefore expect that the presence of a friend removes the difference between the second-best and the market allocation. In fact this is true whenever the capital cost is high enough that the second-best organizational form is also a small, networked firm. However, as Proposition 3.1 shows, the presence of a friend—especially a "good one"—reduces the entrepreneur's incentive to scale up and hire outside of her network. Figure 4a shows that the second-best organizational choice is (weakly) larger for all values of network referral quality than the private equilibrium (shown as a solid black line in the figure), and the disagreement in size is larger in the presence of a good friend. Even though the referral value solves the first inefficiency for network connections, entrepreneurs still choose too small—and now also too networked—a firm. Proposition 4.1 formalizes this.

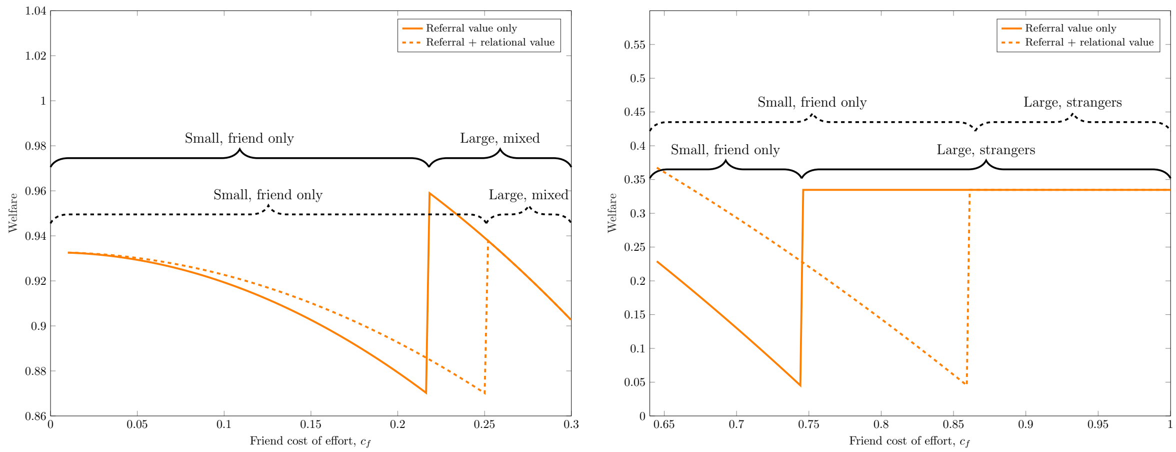
Figure 4b illustrates how the presence of the relational effect changes the second-best allocation. When the network connection also solves the inefficiency of unobservable effort, the welfare-maximizing choice of organizational form is (weakly) smaller and more networked. In this case there is a real cost of having to pay all workers the same wage, as friends are given inefficiently steep wage contracts. However, as Figure 4b illustrates and Proposition 4.1 formalizes, the entrepreneur's choice of organizational form is also too small and too networked in this case.

In the case of fully independent wages, the market and planner allocation differ mostly in terms of *composition*.[23] The set of friend effort cost c_f for which the planner would like to run a networked firm (FS versus SS) is strictly smaller. Part of why the entrepreneur chooses to hire her friend is because this avoids paying an information rent to workers. The planner does not value this and would therefore switch to hiring only from the general labor market much earlier. Appendix B plots the equivalent of Figure 4 for the case of fully independent wages.

Having discussed efficiency properties and the planner's choice, we now turn to analyzing how welfare in the market allocation depends on the network quality. On the one hand, through the referral and relational value effects, network connections solve some of the underlying inefficiencies in the market. On the other hand, when entrepreneurs are connected to high-quality workers, organizational choices are further away from the welfare-maximizing ones. This begs the question of whether it is possible that the referral and relational value effects reduce welfare on net. We look at three aspects of this overall question: (1) does being connected to a better friend reduce welfare, (2) is welfare higher when the entrepreneur has no network connection at all, and (3) does having a network connection with relational in additional to referral value reduce welfare. Note that without considering organizational choice, the answer to all three questions is no.

When taking into account how the referral and relational values alter the choice of firm composition, however, the effects on welfare can indeed flip. Consider first the case of a "very good friend", as defined in Proposition 3.2 and a level of capital cost such that the entrepreneur chooses to run a mixed firm as long as the friend's quality is sufficiently low. Figure 5a plots welfare as function of friend quality in this region. The jump in welfare occurs exactly where the market allocation switches from running a small, network-reliant firm to also employing a stranger. Both in the presence and in the absence of the relational value effect (dashed and solid lines), welfare is declining in the friend's effort cost as long as the organizational form does not change. This is simply a consequence of surplus being lower when workers have a higher cost of production. At the point where the entrepreneur has such a high-quality friend that she chooses to hire only the friend, welfare drops discontinuously. The drop in welfare is exactly equal to the expected surplus of workers hired from the pool, or their information rent from unobservable cost of effort.

[23]The same result of too-low tolerance of K when strangers are hired of course also applies here.



(A) Very good friend (B) Good enough friend

Figure 5. This figure shows the welfare achieved by the private organizational structure. Panel (A) shows the case for a low cost of capital and the case of tethered wages, for the jump where the market allocation switches from running a small networked firm to running a large mixed firm (the very good friend case). Panel (B) shows the case for a higher cost of capital and the case of tethered wages, for the jump where the market allocation switches from running a small networked firm to a large stranger-only firm (the good enough friend case).

Figure 5b illustrates that the welfare losses from inefficient organizational choice can be large enough that welfare is higher in the absence of a friend (represented here as a friend with prohibitively high effort cost). This is most likely to happen when the entrepreneur has a friend of relatively low quality, a good enough friend. Due to the referral (and possibly relational) value effect, she values her friend enough to be willing to pay her a relatively high wage. This wage is too high to justify offering the contract to the pool, however, so she chooses to hire her friend only. Facing the same cost of installing capital, an entrepreneur with no network connections would have run a large firm, generating much higher welfare gains.

Again taking into account the effect on organizational choice, it is possible that the relational value effect reduces welfare. The dotted orange line in Figure 5a plots welfare with both the referral and relational value effects, whereas the solid orange line corresponds to the case with only the referral effect. When only the friend is hired in both cases, welfare is higher in the presence of the relational value, since contracts can be flat and provide insurance to the friend. When the organizational choice is a mixed firm in both cases, welfare coincides. For low enough values of c_f, the friend's participation and incentive constraints are not binding. Wages and hence welfare are identical in this region. However, because hiring the friend only is more profitable in the presence of the relational value effect, the entrepreneur is even more discouraged from hiring an additional worker from the pool. It is precisely for this range of friend quality where the relational value effect deters outside hiring that welfare is reduced by having a friend for whom effort is effectively contractible.

In the case of fully independent wages, there is a similar possibility for welfare declines with better-quality networks. Now this effect will only happen in the case of "good enough" friends. Since the planner and entrepreneur disagree on the highest cost friend that should still be hired, welfare drops discontinuously at the level of c_f such that the entrepreneur starts hiring her friend. Again, even in the knife-edge case where wages are fully independent within the firm, since the relational value effect encourages even more network-based hiring, it is entirely possible that welfare is lower than with the referral value only. Appendix Figure B.2a shows the welfare discontinuity for the case of independent wages.

Proposition 4.1 formalizes the above discussion.

Proposition **4.1.**

(1) *For all values of referral quality c_f:*
   (a) *If $\varphi = \infty$, the entrepreneur chooses an organizational form that is smaller than in the welfare-maximizing allocation.*
   (b) *If $\varphi = 0$, the entrepreneur chooses an organizational form that is more networked than in the welfare-maximizing allocation.*

(2) *Define welfare as the sum of entrepreneur's profits and workers' surplus. For all values of φ:*
   (a) $\exists\ (c'_f, c_f)$ *with* $c'_f < c_f$ *s.t.* $WF(c'_f, \delta) > WF(c_f, \delta)\ \forall\ \delta \in \{0, \infty\}$
   (b) $\exists\ c_f$ *s.t.* $WF(c_f, \infty) > WF(c_f, 0)$
   (c) $\exists\ c_f$ *s.t.* $WF(\infty, \delta) > WF(c_f, \delta)\ \forall\ \delta \in \{0, \infty\}$

To sum up, in our setup (a) entrepreneurs are choosing firms that are too small and too network-reliant and (b) the referral and relational value effects can lower welfare precisely if they induce the entrepreneur to run a small and networked firm. These results also motivate an analysis of overall market organization, where we question whether a market learns to organize itself in a manner of efficient scale and composition, and how all of this depends on fundamentals such as network quality in the market.

5. General Equilibrium

Our paper primarily focuses on the case of small industries, which is appropriate for burgeoning markets. This is a setting where workers' outside options naturally consist of the market at large in a variety of industries and, therefore, can be thought of as fixed—i.e., a partial equilibrium setup. However, we also briefly examine a setting where this phenomenon applies to a large and mature industry—i.e., a general equilibrium setup — to explore the interaction between hiring decisions of entrepreneurs in a common labor market. We leave a more general characterization of the market-wide effects of network-based hiring for future work, as this is beyond the scope of this paper. For simplicity, we focus on the case where there is only referral value; we conjecture that the same mechanism carries over to the case where there is also relational value.

We find that general equilibrium interactions have the potential to amplify the negative effect of network quality on firm expansion. As before, being connected to a better friend makes an entrepreneur less likely to hire from the pool. In equilibrium, this can change workers' outside option sufficiently to induce other entrepreneurs also to stop hiring from the pool. We also show that the intuition of high-quality friends deterring outside hiring is still present when these friends have the option to join the general labor market.

5.1. **Setup.** Consider a market with a measure N of entrepreneurs, each of whom is connected to one friend. There are also $M > N$ workers who do not have a connection to an entrepreneur; these could, for instance, be recent migrants to the city. As before, migrants' cost of providing effort is unknown to entrepreneurs and drawn from $F_c(.)$.

We denote by $z(x)$ the share of entrepreneurs who choose the organizational form $x \in \{0, F, S, FS, SS\}$. The stranger pool then consists of M migrants in addition to the $N \cdot (z(0) + z(SS))$ workers that were not hired by their friend. The number of jobs offered to pool workers is $N \cdot (z(FS) + 2z(SS))$.

The timing is as follows. Entrepreneurs decide how many workers to hire and whether these are friends or strangers. Workers observe the distribution of wage offers in the population; the ratio of vacancies to job applicants given by $\theta := \frac{N\cdot[z(FS)+2z(SS)]}{M+N\cdot[z(0)+z(SS)]}$; and in the case of networked workers, their wage offer from the connected entrepreneur. Connected workers then decide whether to accept the wage offer from their known entrepreneur, join the general labor market pool, or exit the labor market altogether. The M unconnected workers decide whether to enter the labor market. We normalize the value of exiting the labor market to 0 for all workers.

Workers from the labor market pool are randomly matched to firms; the number of matches that form is equal to the minimum of vacancies and jobs offered. Given $M > N$, there are always more applicants than jobs, and the ex ante job-finding probability for workers, θ, is less than 1. Once matched, workers cannot redraw a new wage offer this period. This assumption captures the fact that a job search takes time in practice. Therefore, workers accept a contract as long as their expected utility exceeds zero. Workers who are connected to an entrepreneur compare the wage contract offered to the expected benefits from joining the labor market pool. Let v^* denote the expected wage payments received conditional on being matched in the general labor market. A worker with effort cost c_f has an outside option of $\theta\,(v^* - c_f)$.

5.2. **Equilibrium.** An equilibrium is a set of $\{v^*, \theta\}$, a distribution over organizational forms $z(x; v^*, \theta)$, and wage contracts $\{s_0(x, c_f, v^*, \theta), s_1(x, c_f, v^*, \theta)\}$ such that

(1) Given $\{v^*, \theta\}$, organizational choice x and wage contracts $\{s_0(x, c_f, v^*, \theta), s_1(x, c_f, v^*, \theta)\}$ solve the entrepreneur's problem as laid out in Section 2, with the friend's participation constraint replaced by $pu(s_1) + (1-p)u(s_0) - c_f \geq \theta\,(v^* - c_f)$

(2) $\theta = \dfrac{N \cdot [z(FS; v^*, \theta) + 2z(SS; v^*, \theta)]}{M + N \cdot [z(0; v^*, \theta) + z(SS; v^*, \theta)]}$

(3)
$$v^* = \frac{2z(SS)}{2z(SS) + z(FS)} \left[p\, u(s_1(SS, v^*, \theta) + (1-p)\, u(s_0(SS, v^*, \theta)\right] + \frac{z(FS)}{2z(SS) + z(FS)} \mathbb{E}_{c_f} \left[p\, u(s_1(c_f, FS, v^*, \theta) + (1-p)\, u(s_0(c_f, FS, v^*, \theta)\right]$$

From the new participation constraint, it is clear that, as long as $\theta < 1$, good friends still require a lower wage than higher-cost friends. Even though both types of friend could join the general labor market, they might not be matched to any employer. This possibility is more costly to a higher-quality worker, who has a higher expected surplus from being matched, than a lower-quality worker. As a result, the main mechanism driving Propositions 3.1 and 3.2 is still present in general equilibrium. Being connected to a very high quality friend reduces the marginal profits of hiring from the pool and increases the incentives to run a small and networked firm.

5.3. **Quality of Friends and Labor Market Equilibrium.** For simplicity, we consider a distribution of friend types that has two mass points. A fraction χ of connected workers are of type c_f^H, and the remainder has $c_f^L < c_f^H$. We start from a baseline equilibrium with tethered wages in which entrepreneurs connected to the low-cost friends choose FS, corresponding to the good friends region defined in Section 3. Entrepreneurs with high-cost friends choose SS. We then analyze the general equilibrium effects of reducing c_f^H just enough to induce entrepreneurs to hire their friends.

This change affects the other entrepreneurs — those connected to low-cost friends — through a change in their friend's outside option.[24] When a fraction χ of entrepreneurs switch from hiring two strangers to hiring their friend only, the probability of being matched, θ, decreases for all pool workers. In addition, the only entrepreneurs who still hire strangers are those connected to low-cost friends, who chose FS in the baseline. The wage they offer is lower since we assumed that they were operating in the good friend region where the friend's constraints are not binding. Taken together, not only does the job-finding probability decrease, but the expected wage when finding a job is also lower.

Since their outside option decreased, low-cost friends are now even more valuable to the entrepreneur. However, the reduction in the outside option can only be fully capitalized on when no strangers are hired. For a large enough change in the outside option, these entrepreneurs could now be induced to stop hiring from the pool altogether. In the new equilibrium, no entrepreneurs hire from the general labor market, and all friends have an outside option of zero.

[24]Throughout the analysis, we abstract from changes in the composition of the labor pool. We conjecture that for M large enough, the effect through the outside option dominates.

This setup shows how an increase in the friend quality in one small part of the distribution can not only induce directly affected entrepreneurs to stop hiring from the labor market but also, through general equilibrium interactions, result in other entrepreneurs choosing to run small, networked firms. Proposition 5.1 formalizes.

Proposition 5.1. *Suppose $\varphi = \infty$, $x(c_f^L; v^*, \theta) = FS$, and $x(c_f^H; v^*, \theta) = SS$. Consider a shift of c_f^H by ε such that $x(c_f^H - \varepsilon; v^*, \theta) = F$. Then, in the new equilibrium, $x(c_f^H - \varepsilon; v^{*'}, \theta') = F$, $x(c_f^L; v^{*'}, \theta') \in \{F, FS\}$, and*

1. *$\theta' < \theta$; the job-finding probability is lower.*
2. *$v^{*'} < v^*$; expected payments are lower.*
3. *For small enough values of c_f^L, $x(c_f^L; v^{*'}, \theta') = F$; entrepreneurs with low cost friends also hire their friend only.*

In the case of independent wages, we assume that friend quality is low enough such that both types of entrepreneurs hire only from the general labor market in the baseline. We then consider an improvement in the friend quality of the lower cost friend, such that it induces entrepreneurs to start hiring their friends.

Since a fraction $1-\chi$ of entrepreneurs now hire only one worker from the pool, the job-finding probability decreases. In addition, the average wage offered in the labor market decreases. This is a result of the fact that without the c_f^L workers, the pool is now of higher expected quality. After all, the set of friends now hired directly by connected entrepreneurs is of relatively low quality. With a higher expected quality, wages offered to unknown workers are lower. Both a lower job-finding probability and a lower wage reduce the workers' outside option.

With a lower outside option, workers become cheaper to employ; for a large enough shift, entrepreneurs connected to the low-cost workers also switch from hiring only strangers to hiring their friend in the new equilibrium. Similar to the case with tethered wages, general equilibrium effects imply that an increase in friend quality in some part of the distribution can lead to more network-based hiring everywhere.

Proposition 5.2. *Suppose $\varphi = 0$ and $x(c_f^L; v^*, \theta) = x(c_f^H; v^*, \theta) = SS$. Consider a shift of c_f^L by ε such that $x(c_f^L - \varepsilon; v^*, \theta) = FS$. Then, in the new equilibrium, $x(c_f^L - \varepsilon; v^{*'}, \theta') = FS$, $x(c_f^H; v^{*'}, \theta') \in \{FS, SS\}$, and*

1. *$\theta' < \theta$; the job finding probability is lower.*
2. *$v^{*'} < v^*$; expected payments are lower.*
3. *For small enough c_f^H, $x(c_f^H; v^{*'}, \theta') = FS$; entrepreneurs with high-cost friends also hire their friend.*

In Section 4, we discussed that, through the effect on organizational choice, welfare might be decreasing in friend quality. The general equilibrium interactions described here reinforce this effect. The overall welfare effects of an increase in c_f depend on the relative strength of the direct effect (welfare is lower when "technology" worsens) and the indirect effect (a worse friend might induce hiring from the pool). As established in Propositions 5.1 and 5.2, a change in the distribution of friend quality has the potential to change hiring behavior in an unaffected part of the distribution. We know from the previous section that, in partial equilibrium, a reduction in c_f can reduce welfare. In general equilibrium, this change also induces other entrepreneurs to also rely more on network-based hiring, lowering welfare further. Therefore, the general equilibrium effects of network-based hiring, through their effect on outside options, exacerbate welfare losses from high-quality network connections.

6. Learning and Market-Level Organization

We now turn to learning in this environment. As in much of the paper, we focus here on the case with tethered wages. In the previous analysis we assumed that, while a specific worker's type is not observable, entrepreneurs know the distribution F_c of types in the population. In this section, we relax this arguably strong assumption and suppose instead that entrepreneurs have initial beliefs and update these based on their experience of hiring strangers. We study the distribution of firm organizational structures when — as before — entrepreneurs face contracting frictions and have access to referral networks and — in addition — face uncertainty about the quality of the general labor market pool.

Crucially, the nature and quality of social learning will depend on the firm organizational structure. Hiring a stranger, for instance, is necessary to deliver a signal to the entrepreneur about the worker pool distribution. Clearly, operating a small firm with only one friend as a worker never allows the entrepreneur to update her beliefs. In contrast, each stranger hired delivers a signal about worker quality.

There are two natural learning environments to investigate, but because of the technical similarity in the arguments and results, we look at them in parallel. For tractability, we study an environment in which the state of the world could be either "good," in which case the optimal organizational choice involves hiring strangers, or "bad," in which case entrepreneurs optimally hire only their friend or exit the market.

The first environment looks at a single forward-looking Bayesian entrepreneur over time, who is infinitely lived and who can choose how to organize each period. Therefore the entrepreneur chooses an organizational structure balancing current-period profits with the value of experimenting to learn about the quality of the stranger pool, which affects expected future profits. Here the market at large is characterized by the limiting distribution of organization choices by a continuum of entrepreneurs in a market.

The second environment studies social learning. In this case, we consider an individual sequence of Bayesian entrepreneurs, each living a single period. Each entrepreneur observes information generated by all prior agents. This is therefore along the lines of the classical sequential model of social learning (Banerjee, 1992; Smith and Sørensen, 2000; Eyster and Rabin, 2010).[25] In the typical sequential model, agents do not see the full signal received by prior agents but instead see coarsenings, such as the outcome of a discrete decision, which are not sufficient statistics for beliefs. Our setup departs from this literature because in our case the agent at time t observes the entire history of all organizational structures, wages, profits — the full information set — of all previous firms.

For both environments, we characterize the limiting distribution of firm organization and demonstrate the existence of an *information poverty trap.* When the worker pool is actually good, we show that in the limit, a positive fraction of entrepreneurs — or sequences of entrepreneurs — do not learn the state of the world and instead run small, networked firms. Being pessimistic about the general labor market pool is an absorbing state — a trap — since changing these beliefs would require overcoming the pessimism enough to start hiring from that labor market. We show that, relative to a social planner defined as in Section 4, both learning setups lead to an inefficiently large fraction of entrepreneurs stuck in the information poverty trap. Furthermore, a simple but important asymmetry exists. If the worker pool is actually bad, the market must learn it, and there is no such trap. Finally, we show how it can be the case that an improvement in network quality exacerbates the information poverty trap.

6.1. **Limit Distribution of Firm Organization.** We index by $t \in \mathbb{N}$. When we study one forward-looking entrepreneur, t indexes time, and when we study social learning, t indexes entrepreneurs. Forward-looking entrepreneurs are infinitely lived and discount the future at rate β. Assume that the distribution of worker quality is given by $F_{c,\theta}(\cdot)$. θ denotes the binary state of the world, $\theta \in \{1, 0\}$. $F_{c,0}(\cdot)$ first-order stochastically dominates $F_{c,1}(\cdot)$. $F_{c,1}(\cdot)$ is the low-cost state of the world under which the optimal decision would be for the entrepreneur to choose to organize as a large firm. $F_{c,0}(\cdot)$ is the high-cost state of the world under which the optimal decision would be to either remain small and networked or exit. Until otherwise noted, the state is 1, unknown to the agents.

An entrepreneur at t has belief $q_t := \mathrm{P}(\theta = 1|Z_{t-1})$ where the probability is taken with respect to their information set: here this is a Bayesian update using their prior and all past signals $Z_{t-1} := (z_1, ..., z_{t-1})$ described below. An entrepreneur at t takes a decision $x_t \in \mathcal{X}$ where $\mathcal{X} = \{\text{shutdown}, F, S, FS, SS\}$ is a finite set of firm organizational decisions.

Given x_t the entrepreneur generates a binary signal $z_t \sim G(\cdot|x_t, s_t, \theta)$ if $x \in \mathcal{X} \setminus \mathcal{X}_0$. The distribution of the signal depends on the state of the world θ, the organizational choice x_t, and the wages offered s_t. z_t is either 1, if output is high, or 0, if no output is produced. Intuitively, when output is high, the entrepreneur knows that the worker put in high effort, an outcome more likely in the "good" state of the world. But if the decision $x_t \in \mathcal{X}_0 \subset \mathcal{X}$ then no signal

[25]We conjecture that, with some effort, one could extend our results to a setting that nests our own, such as stochastic neighborhoods of observation (Acemoglu et al., 2011) but we leave that to future work and content ourselves with the simpler setup.

is generated. In our case, the set $\mathcal{X}_0 = \{F, \text{shutdown}\}$ since for these organizational structures, there can be no signal delivered about the quality distribution of the general worker pool.

To select an organization in $\mathcal{X}_0$, the entrepreneur must be sufficiently pessimistic about the distribution of overall worker quality, so we say a decision in $\mathcal{X}_0$ is optimal if $q_t \leq q^*$ for some q^*.

An entrepreneur at $t+1$ forms beliefs according to Bayes' rule:

$$\frac{\mathrm{P}(\theta = 0 | Z_t, x_t, s_t)}{\mathrm{P}(\theta = 1 | Z_t, x_t, s_t))} = \frac{\mathrm{P}(\theta = 0 | Z_{t-1}, x_{t-1}, s_{t-1}))}{\mathrm{P}(\theta = 1 | Z_{t-1}, x_{t-1}, s_{t-1}))} \cdot \frac{\mathrm{P}(z_t | \theta = 0, Z_{t-1}, x_{t-1}, s_{t-1}))}{\mathrm{P}(z_t | \theta = 1, Z_{t-1}, x_{t-1}, s_{t-1}))}.$$

Firm organization as well as the wage are chosen to maximize flow profits and, in the case of one forward-looking entrepreneur, the continuation value.

$$\{x_t, s_t\} := \underset{\{x \in \mathcal{X}, s_t\}}{\operatorname{argmax}} \; \underbrace{\mathbb{E}[\Pi(x, s) | q_t]}_{\text{flow profits}} + \underbrace{\beta \mathbb{E}[V(q_{t+1}(q_t, x, s) | x, s)]}_{\substack{\text{continuation value for} \\ \text{infinitely lived entrepreneur}}}, \text{ with}$$

$$V(q_t) = \max_{\{x \in \mathcal{X}, s_t\}} \; \mathbb{E}[\Pi(x, s) | q_t] + \beta \mathbb{E}[V(q_{t+1}(q_t, s, x) | x, s)].$$

Recall that in the social learning case, t indexes individual entrepreneurs, each living one period and therefore behaving with $\beta = 0$.

This generates a stochastic process of beliefs $\{q_t\}$, which depends on the equilibrium firm organization decisions in previous periods. This also generates a stochastic process over the organization of the firm $\{x_t\}$.

We are interested in the distribution of this stochastic process. When we are looking at a single forward-looking entrepreneur making decisions over time, the limit distribution tells us the share of firms under each organizational structure in the limit in the market. To the extent that the share of firms making suboptimal decisions, i.e., decisions different from the ones they would make if they knew the state of the world, is larger than what the social planner would choose, there is an inefficiency.

When we are considering the social learning case, we analogously consider the fraction of markets in which all but a finite number of firms choose the optimal firm organizational structure. If that is smaller than what a social planner would choose, there is more mass on suboptimally small and networked organizational decisions, and we say there is inefficient social learning and a failure of information aggregation.

Recall if $x_t \in \mathcal{X}_0 = \{F, \text{shutdown}\}$ then no signal z_t is generated. So at any belief, $q < q^*$, that justifies such a decision learning stops: this is absorbing. Practically this means that if an entrepreneur at time t ever becomes convinced that it is not worthwhile to try hiring from the pool, then they may exit or hire only from the network, but in that case, no further information on the worker pool is generated. Therefore, in subsequent periods the exact same decisions are made.

This begs the question as to the behavior of this stochastic process. Does it converge to the truth? Does it converge to this absorbing region of exit and conservative organization?

Proposition **6.1.**

(1) *Consider a continuum of forward-looking, infinitely lived entrepreneurs in a market with no social learning. Under the above assumptions, for any initial priors that are interior and justify experimentation, i.e.,* $q^* < q_0 < 1$*,*
 (a) *the share of firms that either exit or organize as small firms with only a friend as a worker is positive and*
 (b) *the share of firms that learn the state of the world and optimally organize is positive.*

(2) *Consider an infinite number of sequences of one-period-lived Bayesian entrepreneurs. Under the above assumptions, for any priors that are interior and justify experimentation, i.e.,* $q^* < q_0 < 1$*,*
 (a) *the share of markets where all but a finite number of agents either exit or organize as a small firm with only a friend as a worker is positive and*
 (b) *the share of markets where all but a finite number of firms learn the state of the world and optimally organize is positive.*

This means that in both setups, for any initial priors that are interior such that entrepreneurs are willing to experiment at least in the first period, there is always a positive probability that entrepreneurs or the market never learn the true distribution even in the limit and remain suboptimally small and networked. In the social learning setup, this information aggregation failure happens despite the fact that full information sets are transmitted and entrepreneurs see all choices and outcomes of all previous entrepreneurs. This does not always happen, though, and depends on the equilibrium quality of signals that depend on fundamentals, which also determines the cutoff q^*.

6.1.1. *Asymmetry in learning process.* So far, we have assumed that the state of the world is 1, unknown to agents. In this state of the world, there is a positive share of entrepreneurs or markets where firms fail to learn the state of the world even in the long run and are stuck in an information poverty trap. Would a similar failure of learning imply that a positive share of entrepreneurs or markets fail to learn that the state of the world is 0 and instead keep organizing as large firms even in the long run? Proposition 6.2 shows that this is not the case. If the state of the world is truly 0, all forward-looking entrepreneurs will eventually learn that hiring from the general labor market pool is not optimal. Further, in the social learning setup, in all markets, all but a finite number of firms are sufficiently pessimistic and behave as if they know the state of the world is 0.

Proposition **6.2.** *Consider a continuum of forward-looking, infinitely lived entrepreneurs or an infinite number of sequences of one-period-lived Bayesian entrepreneurs. Under the above assumptions, when* $\theta = 0$*, then* $\lim_{t\to\infty} \mathrm{P}(x_t \in \mathcal{X}_0) = 1$.

We omit the proof of the straightforward result. While simple, it is important to notice the asymmetry. If the priors are optimistic, since the negative state of the world is absorbing, then, in any case, the optimal choice of scale and composition will be correct in the long run. In contrast, if the world is actually good, then we get the result that not all entrepreneurs or markets optimally organize and a nontrivial share suboptimally gets absorbed.

6.2. **Network Quality and Learning.** Next we look at how the relational and referral value of network-based hiring affect learning. Specifically, we ask how forward-looking entrepreneurs and the market organize as a function of whether or not the referred network members tend to be of higher or lower quality. Here, as in much of the paper, we focus on what we consider to be the empirically relevant case: network members are "good" in the sense that they have a low cost of effort. This is consistent with the idea that referrals have some screening value.

In this setting, the threshold belief q^* such that below which the individual selects F instead of FS actually declines in friend quality. That is, when the agent has a better friend, they are more conservative in terms of taking a chance on a stranger as well and paying a premium to their friend, which is required to incentivize the stranger. As a consequence, they simply remain small, relying only on their network. The result is that with a better friend, individual entrepreneurs or the market are more likely to be stuck in an information poverty trap, in which firms are small and network-reliant.[26]

PROPOSITION **6.3.** *Consider a value of the capital cost K such that $\forall c \in [0,\ \hat{c}]$, $x(1, c_f, K, \delta) = FS$ and $x(0, c_f, K, \delta) = F$. Let $\hat{c_f}$ and $\tilde{c_f}$ be as defined in Propositions 3.1 and 3.2.*

(1) $\forall c \in [0,\ \hat{c}]$, $\exists q^* \in [0,\ 1]$ *s.t. the entrepreneur is indifferent between F and FS .*
(2) $\forall c \in [0,\ \hat{c}]$, $\frac{\partial q*(c_f)}{\partial c_f} < 0$
(3) *In the case of a continuum of infinitely lived forward-looking entrepreneurs, $\exists\ \{c_f, c_f'\}$ with $c_f' > c_f$, such that the share of entrepreneurs who fail to learn the state of the world is larger under c_f than under c_f'.*
(4) *In the case of infinite number of sequences of one-period-lived Bayesian entrepreneurs, $\forall c \in [0,\ \tilde{c}]$, the share of markets where all but a finite number of firms organize as a small firm with only a friend as a worker is decreasing in c_f.*

6.3. **Welfare and Learning.** Last, we turn to the efficiency and welfare properties of the information poverty trap. To this end, we consider a social planner in the spirit of Section 4. The social planner takes wages set by entrepreneurs as given but chooses the organizational form to maximize welfare. Welfare is defined as previously, with the addition that the social

[26]In the case of a fully-forward-looking entrepreneur, we show existence by simulation. In our example, the proposition holds not only for two points but for the entire range of friend quality considered.

planner now maximizes the expected discounted value of all future welfare, using the same discount factor as entrepreneurs.[27]

In either setup, there is too little experimentation in the private equilibrium, and consequently, when the state of the world is truly "good," too many entrepreneurs or too many markets fail to learn and so are stuck in the information poverty trap. The social planner has a lower cutoff level of beliefs q^* and would like to experiment more. In general, the fact that the absorbing threshold is lower will also lead to a lower share of entrepreneurs or markets who suboptimally organize small, networked firms. Compared to an infinitely-lived entrepreneur, the social planner is more willing to hire strangers since the welfare of doing so exceeds pure profits. In addition, since there is no difference between welfare and profits when only a friend is hired, there are more returns to experimentation from the social planner's perspective. In the social learning environment, the planner also corrects an additional learning externality. Since entrepreneurs only live for one period, they do not internalize the value of signals generated by their choice of organizational form or wages.

Proposition **6.4.**

(1) *Consider a continuum of forward-looking, infinitely lived entrepreneurs in a market with no social learning. Under the above assumptions*
   (a) $\exists \{c_f, \delta\}$, $q^*_{SP}(c_f, \delta) < q^*(c_f, \delta)$*; the social planner stops experimenting for lower values of beliefs.*
   (b) *There exist parameters such that the share of entrepreneurs who fail to learn the state of the world in the limit is lower under social planner experimentation.*
(2) *Consider an infinite number of sequences of one-period-lived Bayesian entrepreneurs. Under the above assumptions*
   (a) $\forall \{c_f, \delta\}$, $q^*_{SP}(c_f, \delta) < q^*(c_f, \delta)$*; the social planner stops experimenting for lower values of beliefs.*
   (b) *The share of markets in which all but a finite number of firms either exit or organize as a small firm with only a friend as a worker is lower under social planner experimentation.*

In summary, Proposition 6.4 shows that in the presence of uncertainty, there is an additional source of welfare losses from the relational and referral value. Similar to the previous case in which the state of the world is known, welfare losses come from the interaction between friend quality and the decision to scale up. When incorporating learning dynamics, higher-quality friends can increase welfare losses not only because they discourage outside hiring conditional on information, but also because they discourage information generation and increase the likelihood that an information poverty trap occurs.

[27]In the case of forward-looking entrepreneurs, they do not internalize, when setting wages, that the planner might force them to choose a different organizational form in the future. This is the most natural extension of the one-period setup in the previous sections.

7. Conclusion

In developing countries, particularly in new markets, entrepreneurs often hire from their networks. We explore why this may be the case and study the welfare and market organizational consequences. We set up a model that incorporates the primary reasons for network-based hiring established in the empirical literature. We call these *local benefits* because the studies primarily look at variation in outcomes conditional on a network-based worker being hired, or the likelihood thereof. The literature argues that networks provide (a) referral value, including screening benefits, and (b) relational value. Our study demonstrates that exactly these local features can generate *globally detrimental* welfare effects, which stem from distorting the choice of firm size and the composition of the workforce. Further, negative welfare consequences can worsen with the strength of referral and relational value. In markets where entrepreneurs must learn about the quality of the general labor market, such as when it consists of migrants or general workers at large, information poverty traps will arise and again these traps can be exacerbated as network value increases.

This perspective suggests empirically exploring several policy initiatives. For example, exposing entrepreneurs to outcomes across markets—such as in the experiment studied by Cai et al. (2018)—helps to combat the information poverty trap. In trapped markets, entrepreneurs who observe a large number of firms still only observe a few signals before the absorbing state is reached. Clearly, these signals are predominantly generated from failured FS structures before almost all firms retreat to F structures. In contrast, entrepreneurs from successful markets have seen a large number of firms and their signals are predominantly generated from successful FS structures. Exposing entrepreneurs in failed markets to even a few other entrepreneurs in successful markets would have large peer effects. Other examples of policy initiatives include exploring subsidies to capital, which would also help optimize organizational structure, as well as easy-to-anticipate initiatives such as directly subsidizing external hiring or screening services directly.

Our results also point to a potential and straightforward mechanism as to why we systematically see small, networked firms in the developing world. Complementary to the previous literature (e.g., Akcigit, Alp, and Peters (2016); Bloom, Sadun, and Van Reenen (2012)) we propose a mechanism whereby network-based hiring can exacerbate welfare losses. The information poverty trap we propose also offers a rationale for the presence of subsistence entrepreneurs, who are often interpreted as lacking the will or ability to scale up (Schoar, 2010; Decker, Haltiwanger, Jarmin, and Miranda, 2014; Hurst and Pugsley, 2011). In order to connect to this literature, we plan to analyze, in future work, how network-based hiring shapes the equilibrium in labor markets. In particular, our general equilibrium extension hints at the possibility that network-based hiring might impact (mis-)allocation of workers in the economy, an avenue we plan to explore further in future work.

References

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian learning in social networks," *The Review of Economic Studies*, 78, 1201–1236. 25

Akcigit, U., H. Alp, and M. Peters (2016): "Lack of selection and limits to delegation: firm dynamics in developing countries," Tech. rep., National Bureau of Economic Research. 1, 7

Bandiera, O., I. Barankay, and I. Rasul (2009): "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica*, 77, 1047–1094. 1

Banerjee, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107, 797–817. 1, 6

Beaman, L. (2016): "Social networks and the labor market," in *The Oxford Handbook of the Economics of Networks*. 1

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak (2014): "Can Network Theory based Targeting Increase Technology Adoption?" . 2.1, 2.2.1

Beaman, L. and J. Magruder (2012): "Who Gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, 102, 3574–3593. 1

Bento, P. and D. Restuccia (2017): "Misallocation, Establishment Size, and Productivity," *American Economic Journal: Macroeconomics*, 9. 1

Bertrand, M. and A. Schoar (2006): "The Role of Family in Family Firms," *Journal of Economic Perspectives*, 20, 73–96. 11

Bloom, N., R. Sadun, and J. Van Reenen (2012): "The Organization of Firms Across Countries," *qje*. 1, 7

Breza, E., S. Kaur, and Y. Shamdasani (2017): "The morale effects of pay inequality," *The Quarterly Journal of Economics*, 133, 611–663. 4, 15

Burks, S. V., B. Cowgill, M. Hoffman, and M. Housman (2015): "The value of hiring through employee referrals," *The Quarterly Journal of Economics*, 130, 805–839. 1

Cai, J., A. Szeidl, et al. (2018): "Interfirm Relationships and Business Performance," *The Quarterly Journal of Economics*, 133, 1229–1282. 7

Calvó-Armengol, A. (2006): *Social networks and labour market outcomes*, vol. 17, CREI. 1

Chandrasekhar, A. (2016): "Econometrics of network formation," *The Oxford Handbook of the Economics of Networks*, 303–357. 2

Chandrasekhar, A. G., C. Kinnan, and H. Larreguy (2018): "Social networks as contract enforcement: Evidence from a lab experiment in the field," *American Economic Journal: Applied Economics*, 10, 43–78. 1, 2.1, 2.2.1

Cross, R. and S. P. Borgatti (2004): "The ties that share: Relational characteristics that facilitate information seeking," *Social capital and information technology*, 137–161. 2

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014): "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, 28, 3–24. 7

Eyster, E. and M. Rabin (2010): "Naive Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2, 221–43. 1, 6

Fafchamps, M. and A. Moradi (2015): "Referral and job performance: Evidence from the Ghana colonial army," *Economic development and cultural change*, 63, 715–751. 1

Fainmesser, I. P. (2013): "Social networks and unraveling in labor markets," *Journal of Economic Theory*, 148, 64–103. 1

Feld, S. L. (1991): "Why Your Friends Have More Friends Than You Do," *American Journal of Sociology*, 96:6, 1464–1477. 17

Graham, B. S. (2015): "Methods of identification in social networks," *Annu. Rev. Econ.*, 7, 465–485. 2

Groh, M., N. Krishnan, D. McKenzie, and T. Vishwanath (2016): "Do Wage Subsidies Provide a Stepping-Stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan," *Review of Economics and Statistics*, 98, 488–502. 10

Heath, R. (2018): "Why do firms hire using referrals? evidence from bangladeshi garment factories," *Journal of Political Economy*, 126, 1691–1746. 1, 2.1, 2.2.1

Hsieh, C.-T. and P. J. Klenow (2014): "The Life Cycle of Plants in India and Mexico," *qje*, 129. 1

Hsieh, C.-T. and B. A. Olken (2014): "The Missing "Missing Middle"," *Journal of Economic Perspectives*, 28. 1

Hurst, E. and B. Pugsley (2011): "What Do Small Businesses Do?" *Brookings Papers on Economic Activity*, 2011. 7

Ioannides, Y. M. and L. Datcher Loury (2004): "Job information networks, neighborhood effects, and inequality," *Journal of economic literature*, 42, 1056–1093. 2

Jackson, M. O. (2008): *Social and Economic Networks.*, Princeton University Press. 2

——— (2019): "The friendship paradox and systematic biases in perceptions and social norms," *Journal of Political Economy*, 127, 777–818. 17

Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl (2009): "Trust and social collateral," *The Quarterly Journal of Economics*, 124, 1307–1361. 1

Leskovec, J., K. J. Lang, A. Dasgupta, and M. W. Mahoney (2009): "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," *Internet Mathematics*, 6, 29–123. 2

McKenzie, D. and C. Woodruff (2016): "Business Practices in Small Firms in Developing Countries," *Management Science*, 63, 2967–2981. 1

Montgomery, J. D. (1991): "Social networks and labor-market outcomes: Toward an economic analysis," *American Economic Review*, 81, 1408–1418. 1, 17

Munshi, K. (2003): "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 118, 549–599. 1

Pallais, A. and E. G. Sands (2016): "Why the referential treatment? Evidence from field experiments on referrals," *Journal of Political Economy*, 124, 1793–1828. 1, 2.1, 2.2.1

Salanie, B. (2005): *The Economics of Contracts: A Primer*, The MIT Press. 1

Schoar, A. (2010): "The Divide between Subsistence and Transformational Entrepreneurship," *Innovation Policy and the Economy*, 10. 1, 7

Schweinberger, M. and M. S. Handcock (2015): "Local dependence in random graph models: characterization, properties and statistical inference," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77, 647–676. 2

Smith, L. and P. Sørensen (2000): "Pathological outcomes of observational learning," *Econometrica*, 68, 371–398. 1, 6

Ugander, J., B. Karrer, L. Backstrom, and C. Marlow (2011): "The anatomy of the facebook social graph," *arXiv preprint arXiv:1111.4503*. 2

Appendix A. Proofs

Proof of Proposition 3.1. For $\varphi = \infty$, subsumed in the proof of Propositions C.3 and C.4. ■. For $\varphi = 0$, the result follows from the fact that $\pi(FS, c_f, \delta) = \pi(F, c_f, \delta) + \pi(S)$.

Proof of Proposition 3.2. When $\varphi = 0$, the proof follows again from the fact that $\pi(FS, c_f, \delta) = \pi(F, c_f, \delta) + \pi(S)$. We prove the case of $\varphi = \infty$ below:

Whenever $c_f \leq \tilde{c_f}$, the optimal wage and optimal profits are independent of the parameters of the friend c_f and δ (see Proposition C.1). Therefore $\pi(FS, c_f, \infty) = \pi(FS, c_f, 0)$ in that region.

(1) $\bar{K}(c_f, \infty) = \pi(FS, c_f, \infty) - \pi(F, c_f, \infty)$. Whenever $K > \bar{K}(c_f, \infty)$, hiring a friend only is more profitable. Further, $\pi(FS, c_f, \infty) \geq \pi(F, c_f, \infty)$, since the entrepreneur could always offer the friend's optimal contract to the pool, so for lowe values of K, hiring a friend and a stranger is optimal. Since $\pi(FS, c_f, \infty)$ is independent of c_f and $\pi(F, c_f, \infty)$ second is decreasing in c_f, $\bar{K}(c_f, \infty)$ is increasing in c_f

(2) This follows directly from the fact that $\pi(F, c_f, \infty) \geq \pi(F, c_f, 0)$ (with strict inequality whenever $c_f > 0$) and $\pi(FS, c_f, \infty) = \pi(FS, c_f, 0)$.

(3)
$$\begin{aligned}\bar{K}(c_f, \infty) &= \pi(FS, c_f, \infty) - \pi(F, c_f, \infty)\\ &\leq \pi(F, c_f, \infty) + \pi(S) - \pi(F, c_f, \infty)\\ &= \frac{1}{2}\pi(SS)\end{aligned}$$
where the first inequality follows from the fact that in FS, there is a restriction of common wages.

This completes the proof. ■

Proof of Proposition 3.3.

(1) Proof of Proposition C.3 applies.
(2) Proof of Proposition C.3 applies.

■

Proof of Proposition 4.1.

(1) (a) This follows from the observation that whenever the firm size is either zero or one, welfare is equal to the entrepreneur's profits. If no worker is hired, $WF = \Pi = 0$. If the firm has size one, it must be that the employee is a friend (see Propositions C.4 and C.6). Since the wage is then chosen to exactly satisfy the friend's constraint, $WF(F, c_f) = \Pi(F, c_f)$. In contrast, whenever two workers are hired, welfare is higher than entrepreneurial profits, the difference being given by the surplus of

workers hired from the pool, and any surplus for the friend in case the constraint is slack. It follows that the welfare maximizing choice can never feature a smaller firm.

(b) Whenever a friend is hired, $WF(F, c_f, \delta) = \Pi(F, c_f, \delta)$. But $WF(S) \geq \Pi(S)$, since welfare includes the expected surplus of workers which is strictly positive whenever F_c is non-degenerate. The entrepreneur chooses to hire her friend (and a stranger) whenever $c_f \geq c_f^*$, where c_f^* is defined as $\Pi(F, c_f^*, \delta) = \Pi(S)$. On the other hand, the planner would like the entrepreneur to hire her friend only when $\Pi(F, c_f, \delta) = WF(S) \geq \Pi(S)$.

(2) When $\varphi = 0$, this follows simply from the fact that at the point where the entrepreneur switches from SS to FS, welfare in SS is strictly greater than in FS. We prove the case of $\varphi = \infty$ below.

(a) Suppose K is such that $x(0, \delta) = F$ and $x(c_f, \delta) = FS$ for some $c_f < \tilde{c_f}$, where $\hat{c_f}$ is the level of friend cost for which the friends constraint starts to bind. The argument is structured as follows. We first show that welfare conditional on organizational choice is continuously decreasing in c_f. Then we show that at $\hat{c_f}(\delta)$ (the level of c_f where the entrepreneur is indifferent between F and FS), there are strictly positive welfare losses from choosing F. It follows that there exists an $\epsilon > 0$ s.t. $WF(\hat{c_f}(\delta) - \epsilon) > WF(\hat{c_f}(\delta) + \epsilon)$.

Welfare conditional on hiring a friend is given by

$$WF(F, c_f, 0) = p\left(y - u^{-1}\left(\frac{c_f}{p}\right)\right) - K$$
$$WF(F, c_f, \infty) = py - u^{-1}(c_f) - K$$

Welfare is continuously decreasing in c_f by assumption of concavity and differentiability of $u(.)$.

Conditional on choosing FS, welfare is independent of δ as long as $c_f < \tilde{c_f}$. Whenever the friend's constraint is not binding, the optimal contract is independent of both δ and c_f (see Proposition C.1). Since there is unobservable effort, the contract features $s_0 = 0$ in this region. With $s_0 = 0$, the participation constraint of a friend with $\delta = \infty$ coincides with the incentive constraint of a friend with $\delta = 0$ which also implies that $\tilde{c_f}$ is independent of δ.

$$WF(FS, c_f) = \left(\int_0^{pu(w)} p(y-w)f(c)dc + \int_0^{pu(w)} p\left(w - u^{-1}\left(\frac{c}{p}\right)\right) f(c)dc\right) - K +$$
$$p(y-w) + p(w - u^{-1}\left(\frac{c_f}{p}\right)) - K$$

which is again continuously decreasing in c_f.

Summarizing, welfare is given by

$$WF(c_f,0) = \begin{cases} WF(F, c_f, 0) & \text{if } c_f < \hat{c}_f(0) \\ WF(FS, c_f) & \text{otherwise} \end{cases}$$

and

$$WF(c_f,\infty) = \begin{cases} WF(F, c_f, \infty) & \text{if } c_f < \hat{c}_f(\infty) \\ WF(FS, c_f) & \text{otherwise} \end{cases}$$

It remains to show that $WF(F, \hat{c}_f(\delta), \delta) < WF(FS, \hat{c}_f(\delta))$.

Consider first $\delta = 0$.

$$\begin{aligned} & WF(FS, \hat{c}_f(0), 0) - WF(F, \hat{c}_f(0)) \\ &= \int_0^{pu(w)} p(y-w)f(c)dc + \int_0^{pu(w)} p\left(w - u^{-1}\left(\frac{c}{p}\right)\right) f(c)dc - K \\ &= \int_0^{pu(w)} p\left(w - u^{-1}\left(\frac{c}{p}\right)\right) f(c)dc + \left(pw - pu^{-1}(\frac{\hat{c}_f(0)}{p})\right) \\ &< 0 \end{aligned}$$

The last equality uses the definition of $\hat{c}_f(0)$, $K = \int_0^{pu(w)} p(y-w)f(c)dc + p(y-w) - p(y-u^{-1}(\frac{\hat{c}_f(0)}{p})$. The inequality follows from the fact that the first term is the joint surplus of hiring a stranger net of capital cost, which is always positive, and the second term is positive since the friend's incentive constraint is slack.

A similar argument holds for the case where $\delta = \infty$

(b) Suppose $\bar{K}_2 < K < \bar{K}_3$ as defined in Proposition C.3, and let $\hat{c}_f(\delta)$ be the level of c_f s.t. the entrepreneur is indifferent between F and SS (it exists by Proposition C.3). The proof is then similar to (a). If the organizational choice is F, welfare and entrepreneurial profits coincide. If the organizational choice is SS, $WF = \Pi(SS) + 2\int_0^{pu(w)} p\left(w - u^{-1}\left(\frac{c}{p}\right)\right) f(c)dc > \Pi(SS)$. Using the definition of $\hat{c}_f(\delta)$, $\Pi(SS) = \Pi(\hat{c}_f(\delta) - K(\hat{c}_f(\delta))$. It follows that there is a drop in welfare at $\hat{c}_f(\delta)$ from choosing F.

(c) Since $\hat{c}_f(\infty) > \hat{c}_f(0)$, $WF(\hat{c}_f(\infty), \infty) = WF(\hat{c}_f(\infty), 0)$. By the previous argument, $\lim^-_{c_f \to \hat{c}_f(\infty)} WF(c_f, \infty) < WF(\hat{c}_f(\infty), \infty)$, while $\lim^-_{c_f \to \hat{c}_f(\infty)} WF(c_f, 0) = WF(\hat{c}_f(\infty), 0) = WF(\hat{c}_f(\infty), \infty)$

This completes the proof. ■

Proof of Proposition 5.1. There are two possibilities for the organizational choice of entrepreneurs with c_f^L friends in the new equilibrium: $x(c_f^L; v^{*'}, \theta') = FS$ or $x(c_f^L; v^{*'}, \theta') = F$.

(1) In the baseline equilibrium, $\theta = \frac{(1-\chi)N + 2\chi N}{M + \chi N}$. Depending on whether $x(c_f^L; v^{*'}, \theta') = FS$ or $x(c_f^L; v^{*'}, \theta') = F$, $\theta' = \frac{(1-\chi)N}{M}$ or $\theta' = \frac{0}{M}$, both of which are lower than θ.

(2) In the baseline equilibrium

$$v^* = \frac{2\chi}{2\chi + (1-\chi)} \left[p\, u(s_1(SS, v^*, \theta) + (1-p)\, u(s_0(SS, v^*, \theta)\right] + \frac{(1-\chi)}{2\chi + (1-\chi)} \left[p\, u(s_1(FS, c_f^L, v^*, \theta) + (1-p)\, u(s_0(FS, c_f^L, v^*, \theta)\right]$$

If $x(c_f^L; v^{*'}, \theta') = F$, then $v^{*'} = 0$, which is trivially less than $v^{*'}$. If $x(c_f^L; v^{*'}, \theta') = FS$, then

$$v^* = \left[p\, u(s_1(FS, c_f^L, v^{*'}, \theta') + (1-p)\, u(s_0(FS, c_f^L, v^{*'}, \theta')\right]$$

By assumption, the low-cost friends are in the good friend region where the wage is lower than the pool wage, so $[p\, u(s_1(SS, v^*, \theta) + (1-p)\, u(s_0(SS, v^*, \theta)] >$ $\left[p\, u(s_1(FS, c_f^L, v^*, \theta) + (1-p)\, u(s_0(FS, c_f^L, v^*, \theta)\right]$. Further, the quality of the pool has increased with the removal of c_f^H workers which implies that $\left[p\, u(s_1(FS, c_f^L, v^{*'}, \theta') + (1-p)\, u(s_0(FS, c_f^L, v^{*'}, \theta')\right] <$ $\left[p\, u(s_1(FS, c_f^L, v^*, \theta) + (1-p)\, u(s_0(FS, c_f^L, v^*, \theta)\right]$.
Therefore, $v^{*'} < v^*$.

(3) Consider c_f^L such that the entrepreneur is close to indifferent between hiring her friend only and a friend and a stranger, $\pi(F, c_f^L, \theta, v^*) = \pi(FS, c_f^L, \theta, v^*) - K + \nu$ for some small ν. It is easy to verify that $\pi(F, c_f^L, \theta, v^*)$ is decreasing in both θ and v^*. Further, we show in the proof of Proposition C.4, that $\pi(F, c_f^L, \theta, v^*)$ is independent of c_f in this range of the parameter space. This extends easily to θ and v^*, since the proof relies on the fact that the friend constraint is slack. Since $v^{*'} < v^*$ and $\theta' < \theta$, it follows that $\pi(F, c_f^L, \theta', v^{*'}) > \pi(F, c_f^L, \theta, v^*)$ while the profits from hiring both a friend and a stranger are unchanged.

■

Proof of Proposition 5.2.

(1) In the baseline equilibrium, $\theta = \frac{2N}{M+N}$. Depending on whether $x(c_f^L; v^{*'}, \theta') = SS$ or $x(c_f^L; v^{*'}, \theta') = FS$, $\theta' = \frac{(1-\chi)N + 2\chi N}{M + \chi N}$ or $\theta' = \frac{N}{M}$, both of which are lower than θ.
(2) The wage offered to the general labor market is now independent of hiring choices and depends only on the quality of workers. By assumption, both c_f^H and c_f^L are of relatively low quality (entrepreneurs chose not to hire them in the baseline). When they are no longer in the pool of workers, the quality distribution improves and wages decline. It follows that $v^{*'} < v^*$
(3) Without fairness concerns, the choice between FS and SS is simply the choice between F and S. With a lower θ' and $v^{*'}$, the profits from hiring friends has increased, while the profits from hiring from the pool are unchanged.

■

Proof of Proposition 6.1. The proof proceeds in two steps.[28] Define $LR(z) := \frac{P(z|\theta=0)}{P(z|\theta=1)}$.
Step 1: Define a new martingale. Recall that the state of the world is $\theta = 1$. Let us consider a likelihood ratio against the truth. Define a prior likelihood ratio

$$\rho_0 := \frac{P(\theta = 0)}{P(\theta = 1)}.$$

Mechanically, Bayesian updating gives us at any state

$$\rho_{t+1} = \rho_t \cdot LR(z).$$

It can be calculated that[29]

$$\mathbb{E}[\rho_{t+1}] = \rho_t$$

and therefore this defines a martingale on $\mathbb{R}_{\geq 0}$. By hypothesis, we have an absorbing point $r*$ (a direct consequence of $q*$) and written this way for any $r > r*$ we have $x \in \mathcal{X}$, and so no signals are ever drawn in the future. This implies that ρ_t is a martingale for values above $r*$ as well and as calculated above, it is also a martingale in the interior $[0, r*)$ and therefore it is indeed a martingale throughout.
Step 2: Argue that the probability of absorption above $r*$ forever is less than 1. The first observation is $\{\rho_t\}$ is a martingale and as it satisfies the hypotheses by the martingale convergence theorem.

So it remains to look at where the martingale converges to. We will show that it does not almost surely converge past the boundary $r*$. Observe that if the martingale converges to some $r < r*$ then it must converge to 0. This is because if there is uncertainty that the agent faces, any experiment will move the beliefs. By definition then the resting point can only be on the truth.

As a consequence, the only resting points are 0 and some collection of $r \geq r*$. But because we began with the fact that $\{\rho_t\}$ is a martingale and ρ_0 was less than $r*$, which was a necessary condition to even begin to draw signals, then to construct a limit expectation in the interior there must be mass on 0. ■

Proof of Proposition 6.3.

(1) In the case of a one-period lived entrepreneur, q^* is defined as $\Pi(FS, q^*, c_f, \delta) = \Pi(F, c_f, \delta)$. First note that $\Pi(FS, q^*, c_f, \delta)$ is non-decreasing in q everywhere. Consider $q' > q$. Since F_0 first-order stochastically dominates F_1,
$\max_s \{[p(y - s)][1 + q'F_1(pu(s)) + (1 - q')F_0(pu(s))] - 2K\} >$
$\max_s \{[p(y - s)][1 + qF_1(pu(s)) + (1 - q)F_0(pu(s))] - 2K\}$. $\Pi(F, c_f, \delta)$ on the other hand is independent of q. The existence of q^* then follows from the assumed boundary conditions and continuity of the profit function. In the case of a forward-looking entrepreneur,

[28]We are extremely grateful to Ben Golub for helping us with this argument.
[29]That is, we are taking expectations as an external observer who knows the state of the world.

note that the value of hiring the friend only is simply equal to $1/(1-\beta)\Pi(F, c_f, \delta)$, which is independent of q. This follows from the fact that if $x_t = F$, then $q_{t+1} = q_t$ and all subsequent choices must be hiring the friend only as well. The value of hiring a stranger in addition to the friend on the other hand is increasing in q^*, since $\Pi(FS, q, c_f, \delta)$ is increasing in q everywhere.

(2) Totally differentiating the definition of q^*, $\frac{\partial\Pi(FS,c_f,q^*)}{\partial c_f}dc_f + \frac{\partial\Pi(FS,c_f,q^*)}{\partial q^*}dq^* = \frac{\partial\Pi(F,c_f)}{\partial c_f}dc_f$ or $\frac{dq^*}{dc_f} = \frac{\frac{\partial\Pi(F,c_f)}{\partial c_f} - \frac{\partial\Pi(FS,c_f,q^*)}{\partial c_f}}{\frac{\partial\Pi(FS,c_f,q^*)}{\partial q^*}}$. By (1), the denominator is > 0. As shown in C.4, the numerator is $< 0 \; \forall c_f < \hat{c}$.

(3) We show existence by example. The following figure illustrates.

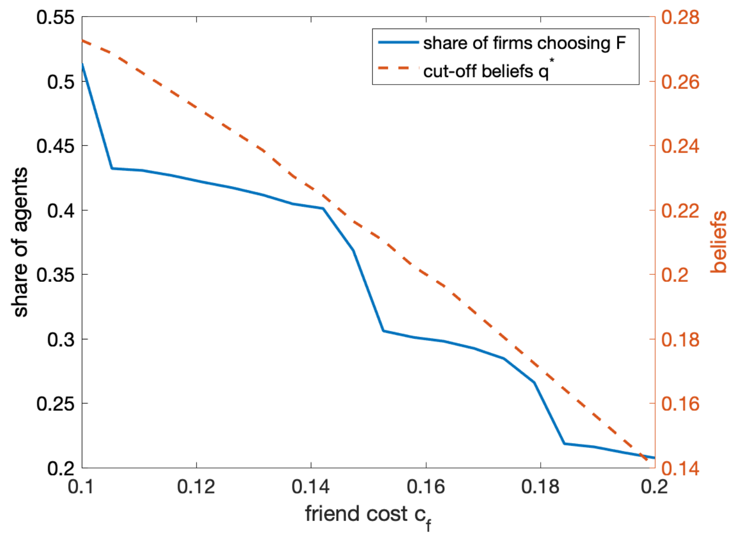


FIGURE A.1. This figures plots the share of entrepreneurs who organize as a small and networked firm in the long run (left axis), as well as the threshold level of beliefs such that entrepreneurs stop experimenting (right axis). The x-axis measures the cost of their network friend. The two lines are based on the following parametrization: $u(c) = c^{1/2}$, $F_c(1) \sim U(0,1), F_c(0) \sim U(0,1)$, $K = .1$, $\delta = 0$, $\beta = .9$ c_f ranges from .1 to .2. For the share of firms who choose F, 2000 entrepreneurs were simulated over 40 periods each.

(4) We show this in two steps. First, $\forall\, c_f < \tilde{c}$, the wage is independent of c_f. This follows from the definition of $\tilde{c}$ as the lowest c_f for which the friend's constraint is binding. Using Proposition C.2, whenever $\lambda = \mu = 0$, the optimal solution is independent of c_f. Since wages are independent of c_f, so is the distribution of the binary signal z_t and the updating rule. Therefore, conditional on $q > q^*$, the path of beliefs are independent of c_f. Second, as shown above, q^* is decreasing in c_f. Therefore, as long as $c_f < \tilde{c}$, c_f' creates the same process for beliefs as $c_f > c_f'$ but with a higher absorbing threshold.

■

Proof of Proposition 6.4.

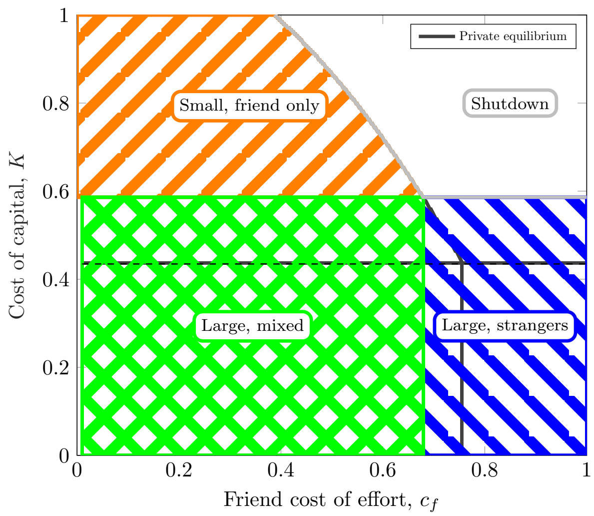
(1) Consider a case where, even though the entrepreneur optimally chooses F if the state of the world were truly bad, the social planner still prefers to hire the stranger as well. Since welfare is strictly larger than profits whenever a stranger is hired, such a case exists. (a) and (b) then follow immediately, since $q^*_{SP} = 0$ and all entrepreneurs learn the state of the world under social planner experimentation.

(2) (a) This follows directly from the definition of q^* as well as the fact that $\Pi(F, c_f, \delta) = WF(F, c_f, \delta)$, but $\Pi(FS, q, c_f, \delta) < WF(FS, q, c_f, \delta)$.

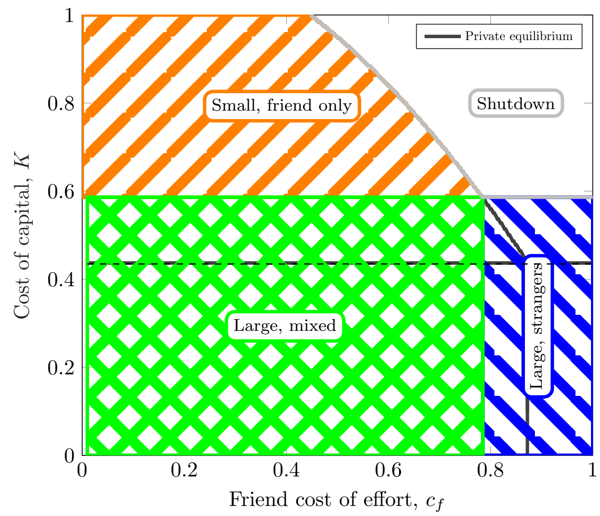
 (b) This follows from (a) and the observation that since wages are still chosen by the entrepreneur, the resulting learning process is identical to the private equilibrium.

■

Appendix B. Figures

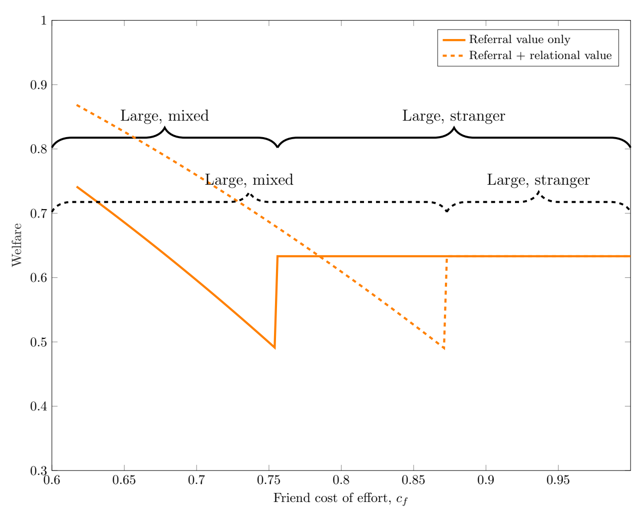


(A) Referral value only: Independent wages



(B) Referral and relational value: Independent wages

Figure B.1. This figure shows the social planner's organization choice, taking the wage contract as given, in an environment with fully independent wages. Panel (A) shows the case where the network has referral value only, panels (B) shows the case where the network has referral and relational value. In each panel, the private equilibrium is shown as a black line. The dashed horizontal line shows the maximum cost of capital before the stranger-only firm would shut down in the private equilibrium.



(A) Independent wages

Figure B.2. This figure shows the welfare achieved by the private organizational structure in the case of independent wages. The range of c_f is chosen to show the jump where the market allocation switches from running a large networked firm to running a large stranger-only firm.

Appendix C. Auxiliary Results: Optimal Firm Size and Workforce Composition

C.1. **Optimal Contracts.**

Proposition **C.1.** *Under the above assumptions, the following hold. Workers always accept the contract and exert effort iff* $c \leq \bar{c} = p\left(u(s_1) - u(s_0)\right)$.

	s_0 *solves*	s_1 *solves*
(1) small, stranger	0	$F(\bar{c}) = f(\bar{c})u'(s_1)p(y - s_1)$
(2) small, friend	$u(s_0) = \min\{\delta, c_f\}$	$u(s_1) = \max\{\frac{c_f-(1-p)\delta}{p}, c_f\}$
(3) large, 2 strangers	*see (1)*	*see (1)*
(4) large, mixed	$(1-p)\left[1 + F(\bar{c})\right] +$ $\{[p(y - s_1) - (1-p)s_0]f(\bar{c})pu'(s_0) - (1 - F(\bar{c}))\}$ $+u'(s_0)\left(\lambda(1-p) - \mu\right) = 0$	$p\left[1 + F(\bar{c})\right] = s_0 f(\bar{c})pu'(s_1)$ $+[p(y - s_1) - (1-p)s_0]pu'(s_1)\left[f(\bar{c})\right]$ $+pu'(s_1)(\lambda + \mu)$

λ is the Lagrange multiplier on the friend's IC constraint $pu(s_1) + (1-p)u(s_0) - c_f \geq u(s_0) - \delta_f$. It binds whenever δ is sufficiently low. μ is Lagrange multipliers on the friend's IR constraint $pu(s_1) + (1-p)u(s_0) - c_f \geq 0$. It binds whenever c_f is sufficiently high.

	$\mathbb{E}\left[\Pi\right]$
(1) small, stranger	$[p(y - s_1)]\left[F(\bar{c})\right] - K$
(2) small, friend	$[p(y - s_1) - (1-p)s_0] - K$
(3) large, 2 strangers	$2\left\{[p(y - s_1)]\left[F(\bar{c})\right] - K\right\}$
(4) large, mixed	$[p(y - s_1) - (1-p)s_0]\left[1 + F(\bar{c})\right] - s_0[1 - F(\bar{c})] - 2K$

C.2. **Optimal Contracts without Moral Hazard.**

For completeness, we study the case with no moral hazard, that is both friend and stranger have $\delta = \infty$. The equilibrium wage offered, given the parameters, in such a case is of course state-independent. The nature of the contract offered depends on both the selected size of the firm as well as the quality of the friend, if one is available. We also consider the case where the entrepreneur has no friend, which can be thought of as having a friend of infinite (or sufficiently high) cost of applying effort.

Proposition **C.2.** *Assume $\delta_f = \delta = \infty$. Under the above assumptions, the following hold. Strangers accept the contract as long as their individual rationality constraint is satisfied. The cutoff stranger who is indifferent between accepting and rejecting is given by $\bar{c} = u(s)$.*

	$s_0 = s_1 \equiv s$ *solves*
(1) small, stranger	$F(\bar{c}) = (py - s)\, u'(s)\, f(\bar{c})$
(2) small, friend	$u(s) = c_f$
(3) large, 2 strangers	$F(\bar{c}) = (py - s)\, u'(s)\, f(\bar{c})$
(4) large, mixed	$1 + F(\bar{c}) = (py - s)\, u'(s)\, f(\bar{c}) + \lambda u'(s)$

λ is the Lagrange multipliers on the friend's IR constraint $u(s) - c_f \geq 0$.
It binds whenever c_f is sufficiently high.

	$\mathbb{E}[\Pi]$
(1) small, stranger	$(py - s)F(\bar{c}) - K$
(2) small, friend	$(py - u^{-1}(c_f)) - K$
(3) large, 2 strangers	$2\left((py - s)F(\bar{c}) - K\right)$
(4) large, mixed	$(py - s)(1 + F(\bar{c})) - 2K$

Proof of Proposition C.2. Standard calculations left to the reader. ■

An entrepreneur operating a firm with only her friend simply needs to satisfy the friend's participation constraint which depends on his c_f. When trying to hire a stranger, the optimal wage offered trades off the benefits of a marginal increase in the probability of the worker accepting (the right-hand side of equation (i)) against extra the payments that have to be made to all workers who accept the contract (left-hand side). Because the model has linear payoffs, a entrepreneur who hires two strangers offers the same wage and payoffs simply scale linearly. Since all workers in a firm need to be paid the same wage, there are interesting interactions between friends and strangers in a large, mixed firm. First, if the friend is of low quality, so c_f is sufficiently high, then the wage is set to satisfy the friend's participation constraint. Second, if the friend is of high enough quality, there is an additional cost of offering a higher wage, as this higher wage also has to be paid to friend (see equation (iv)). The optimal wage is lower than in a firm with only strangers. In other words, the presence of a high-quality friend — who must be paid the same wage offered to the pool — makes the entrepreneur more selective in terms of hiring from outside.

C.3. **Optimal Firm Organization without Moral Hazard.** When deciding on the composition of her workforce, a entrepreneur in principle has five choices: she could run a small firm in which she hires only her friend (F) or only a stranger (S) or she could run a large firm consisting either of two stranger (SS) or her friend and a stranger (FS). In addition, she could also decide to shut down and not pay the fixed cost K.

C.3.1. *Choice of composition.* We start by analyzing, conditional on size, the optimal choice of workforce composition as a function of c_f, the quality of the entrepreneurs network connection. For small firms the choice is between hiring a friend or a stranger: $x_s : \mathbb{R}^0_+ \to \{F, S\}$. For large firms the choice is between hiring a friend and a stranger or two stranger: $x_l : \mathbb{R}^0_+ \to \{FS, SS\}$. We use $\pi(i, c_f),\ i \in \{S, F, SS, FS\}$ to denote the optimal expected profits *gross* of the capital cost of choosing to run each possible type of firm.

First consider the case where the entrepreneur has a network connection of high enough quality such that the wage she must pay him is lower than what she would optimally offer to the pool. Clearly, hiring her friend is dominant in this region of c_f, which we refer to as the *cost effect*. Beyond that cost of effort, there is a region where the entrepreneur chooses to hire her friend despite the fact that the wage is higher than she would offer to the pool. We refer to this as the *certainty effect*. When the friend's cost of effort is too high relative to the certainty effect, the entrepreneur chooses to hire from the pool instead. For large firms, the range of c_f such that the certainty effect outweighs the higher cost is smaller. This is due to the fact that having a high-cost friend also distorts the contract that has to be offered to the other worker. Proposition C.3 formalizes the relationship between c_f and firm composition.

Proposition **C.3.** *Let c_f^* be the quality level of friend s.t. the entrepreneur is indifferent between hiring a friend or a stranger, $\pi(F, c_f^*) = \pi(S, c_f^*)$. Let c_f^{**} be the quality level of friend s.t. the entrepreneur is indifferent between hiring a friend and a stranger, or 2 strangers $\pi(FS, c_f^{**}) = \pi(SS, c_f^{**})$.*

(1) *c_f^* and c_f^{**} are unique;* $x_s(c_f) = \begin{cases} F & \forall c \leq c_f^* \\ S & otherwise \end{cases}$; $x_l(c_f) = \begin{cases} FS & \forall c \leq c_f^{**} \\ SS & otherwise \end{cases}$
(2) $w(S) < u^{-1}(c_f^*)$, $w(SS) < u^{-1}(c_f^{**})$.
(3) $u(w(S)) < c_f^{**} < c_f^*$.

Proof of Proposition C.3.

(1) Easy to show that $\pi(F, c_f)$ is decreasing in c_f while $\pi(S, c_f)$ is independent of c_f. Same applies to $\pi(FS, c_f)$ and $\pi(SS, c_f)$.
(2) Assume the contrary. Then there exists a $c_f > c^*$ s.t. $u(w(S)) > c_f$. But offering that $w(S)$ to a friend must give higher expected profits, since they accept with certainty. This contradicts $c_f >^*$. The same argument holds for c^{**}.
(3) $u(w(S)) < c^{**}$: Assume the contrary. Then there exists a value of c_f such that $u(w(SS)) > c_f$ and $c_f > c^{**}$, that is $x_l(c_f) = SS$. But offering $w(SS)$ to the friend must give higher profits, since they accept with certainty and so $\pi(FS, c_f) > \pi(SS, c_f)$.
$c^{**} < c^*$: First notice that $u(w(FS, c^{**})) = c^{**}$, that is the friend's participation constraint is binding at c^{**}. This is because, as long as the constraint is not binding, $\pi(FS)$ is independent of c_f and higher than $\pi(SS)$. Therefore the constraint must bind at c^{**} where the two cross.
This implies that $\pi(FS, c^{**}) = \pi(F, c^{**}) + \pi(S)|_{w=u^{-1}(c^{**})}$. Also, $\pi(FS, c^{**}) = \pi(SS) = \pi(S) + \pi(S)$ by the definition of c^{**}. $\pi(S)|_{w=u^{-1}(c_f^{**})} < \pi(S)$ since $\pi(S)$ is the unrestricted optimum, and therefore it must be that $\pi(F, c^{**}) < \pi(S)$. By monotonicity of $\pi(F, c_f)$, $c^{**} < c^*$ follows.

■

C.3.2. *Choice of size and composition.* Next we analyze the choice of size and organizational form of entrepreneurs, as a function of c_f and the fixed cost K. $x : \mathbb{R}_+^0 \times \mathbb{R}_+^0 \to \{0, F, S, FS, SS\}$. $x(c_f, K) = 0$ corresponds to the entrepreneur choosing not to enter the market. Figure C.1 plots the optimal organizational choice for an entrepreneur facing a general market distribution with relatively low variance (left panel) and relatively high variance (right panel).

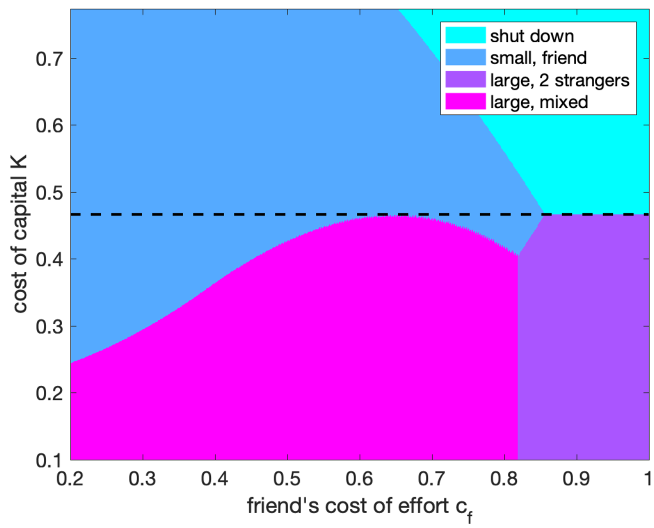


FIGURE C.1. Optimal Firm Organization without Moral Hazard

Given the linearity of payoffs, S is never chosen as it is dominated either by $\Pi(SS) = 2\Pi(S)$ or 0. When the cost of capital is prohibitively high (above the dashed line in Figure C.1), entrepreneurs never hire from the general market and enter only when their network connection has exceptionally low cost. In that region, entrepreneurs with better (or more) connections run larger businesses. On the other hand, when the complementary capital investment is very low, entrepreneurs always run a big firm. The composition of that firm is solely a function of the quality of their network connection.

For intermediate levels of capital, firm size is a non-monotonic function of the quality of the network connection. There is always a single threshold level of quality such that for any worse quality (higher cost), the entrepreneur chooses to hire from the pool only. For lower values of the cost, the entrepreneur either chooses to hire only their friend or to run a mixed firm consisting of one friend and one relative. The value of hiring a stranger is highest when the friend is most similar to the stranger, in the sense that the optimal wage offered to the friend coincides with the optimal wage paid to a stranger. This corresponds to the peak of the FS region in Figure C.1). For network connection with lower cost than that, hiring a stranger is less valuable since the wage the entrepreneur would like to offer them is higher than what they pay their friend. This is closely related to the *cost effect* introduced above.

For network connection with higher cost than at the peak, entrepreneurs would need to offer a higher-than-optimal wage to the general market, in order to also ensure that their friend accepts. This again reduces the value of hiring a stranger, who might or might not accept even the higher wage. This is closely related to the *certainty effect* introduced above.

In summary, for intermediate values of the capital cost, a better network connection is in general associated with choosing a smaller firm consisting of fewer strangers. If the friend's effort cost is also intermediate, there are situations in which entrepreneurs would choose to run a large firm both with a higher- and a lower- quality friend. Proposition C.4 shows the result formally.

Proposition **C.4.** *[i]*

(1) $\{c_f, K : x(c_f, K) = S\} = \varnothing$

(2) *Let* $\bar{K}_1 = \pi(FS, 0) - \pi(F, 0)$*,* $\bar{K}_2 = \pi(FS, c_F^{**}) - \pi(F, c_f^{**})$*, and* $\bar{K}_3 = \pi(S))$*. Suppose that* $\bar{K}_1 < \bar{K}_2$*.* [30] *Then* $\bar{K}_2 < \bar{K}_3$ *and the following holds:*

(a) $\forall\, K < \bar{K}_1,\ x(c_f, K) = \begin{cases} FS & \forall\, c \leq c^{**} \\ SS & otherwise \end{cases}$

(b) $\forall\, \bar{K}_1 < K < \bar{K}_2,\ x(c_f, K) = \begin{cases} F & \forall\, c_f \leq \bar{c_1}(K) \\ FS & \forall\, \bar{c_1}(K) \leq c_f \leq c^{**} \\ SS & otherwise \end{cases}$

(c) $\forall\, \bar{K}_2 < K < \bar{K}_3,\ x(c_f, K) = \begin{cases} F & \forall\, c_f \leq \bar{c_1}(K) \\ FS & \forall\, \bar{c_1}(K) \leq c_f \leq \bar{c_2}(K) \\ F & \forall\, \bar{c_2}(K) \leq c_f \leq \bar{c_3}(K) \\ SS & otherwise \end{cases}$

(d) $\forall\, K > \bar{K}_3,\ x(c_f, K) = \begin{cases} F & \forall\, c \leq \bar{c_4}(K) \\ 0 & otherwise \end{cases}$

Both $\bar{c_1}(K)$ *and* $\bar{c_2}(K)$ *are implicitly defined by* $\pi(FS, \bar{c_i}(K)) - \pi(F, \bar{c_i}(K)) = K$*.* $\bar{c_3}(K) = u\left[py - (\pi(SS) - K)\right]$ *and* $\bar{c_4}(K) = u(py - K)$

Proof of Proposition C.4.

(1) Follows immediately from linearity of payoffs: $\Pi(SS) = 2\Pi(S)$ so either $\Pi(SS) > \Pi(S)$ or $\Pi(S) < 0$.

(2) Let $\widetilde{c_f}$ denote the minimum level of c_f such that the friend's participation constraint is binding, so $w(FS, \widetilde{c_f}) = u^{-1}(\widetilde{c_f})$. Let $\widehat{c_f}$ denote the minimum level of c_f such that the wage offered in FS is the same as in SS, so $w(FS, \widehat{c_f}) = w(SS)$. Note that $\widetilde{c_f} < \widehat{c_f} < c^{**}$, since the presence of a low-cost friend reduces the optimal wage in the FS firm. It will be helpful to show the following:

[30] If $\bar{K}_1 > \bar{K}_2$, then $\forall\, \bar{K}_1 > K > \bar{K}_2$, $x(c_f, K) = \begin{cases} FS & \forall\, c_f \leq \bar{c_1}(K) \\ F & \forall\, \bar{c_1}(K) \leq c_f \leq c^{*} \\ SS & \text{otherwise} \end{cases}$

Proposition, proof and intuition are otherwise similar and therefore omitted.

- $\frac{\partial \Pi(FS,c_f)-\Pi(F,c_f)}{\partial c_f} = \begin{cases} >0 & \text{if } 0 \leq c_f < \widehat{c_f} \\ =0 & \text{if } c_f = \widehat{c_f} \\ <0 & \text{if } c_f > \widehat{c_f} \end{cases}$
  For all $c_f < \widetilde{c_F}$, $\frac{\partial \Pi(FS,c_f)}{\partial c_f} = 0$ whereas $\frac{\partial \Pi(F,c_f)}{\partial c_f} < 0 \forall c_f$. When the effort cost is low enough such that the friend's constraint is slack, the cost has no effect on profits in an FS firm. For all $c_f > \widetilde{c_f}$, $\pi(FS,c_f) = \pi(F,c_f) + \pi(S)|_{w=w(FS,c_f}$. Therefore $\frac{\partial \Pi(FS,c_f)}{\partial c_f} = \frac{\partial \Pi(F,c_f)}{\partial c_f} + \frac{\partial \Pi(S)}{\partial c_f}|_{w=w(FS,c_f)}$. Since $w(FS,c_f) < w(SS) \forall c_f < \widetilde{c_f}$, $\frac{\partial \Pi(S)}{\partial c_f}|_{w=w(FS,c_f} > 0$ as the wage is getting closer to the optimal wage offered to the pool. This proves the first interval. By a similar argument, $\frac{\partial \Pi(S)}{\partial c_f}|_{w=w(FS,\widehat{c_f})} = 0$ and the derivative becomes negative as soon as $c_f > \widehat{c_f}$.
- $\bar{K}_1 < \bar{K}_2 \iff \pi(FS,0) - \pi(F,0) < \pi(FS,c^{**}) - \pi(F,c^{**})$ and $\bar{K}_1 \geq 0$ since the entrepreneur can always offer $w(F,0) = 0$ to the pool at no extra cost.
- $\bar{K}_2 < \bar{K}_3$ follows from $\bar{K}_2 = \pi(FS,c^{**}) - \pi(F,c^{**}) < \pi(F,c^{**}) + \pi(S) - \pi(F,c^{**}) = \bar{K}_3$

(a) Since $K < \bar{K}_1 = \pi(FS,0) - \pi(F,0)$, $x(0,K) = FS$. The choice between FS and SS follows from Proposition C.3. Since $\Pi(FS,0) - \Pi(F,0) > \Pi(FS,c^{**}) - \Pi(F,c^{**})$, $K < \bar{K}_1$ guarantees that F will never be optimal at high levels of c_f either.

(b) Since $K > \bar{K}_1$, $x(0,K) = F$. Further, as long as $K < \bar{K}_3$, $x(\widehat{c_f},K) = FS$. $\widehat{c_f} < c^{**}$, so $\Pi(FS,\widehat{c_f}) > \Pi(SS)$. Also, by definition of $\widehat{c_f}$, $\pi(FS,\widehat{c_f}) = \pi(S) + \pi(F,\widehat{c_f}) > \pi(F,\widehat{c_f})$ as long as $K < \bar{K}_3 = \pi(S)$. $\pi(FS,c_f) - \pi(F,c_f)$ is increasing, so there exists a value $\bar{c_f} \in [0,\widehat{c_f}]$ s.t. $\pi(FS,\bar{c_f}) = \pi(F,\bar{c_f})$. From Proposition C.3, we know that FS dominates SS for all $c_f > c^{**}$. It remains to show that F cannot be optimal for any $c_f \in [\widehat{c_f}, c^{**}]$. Since $\frac{\partial \Pi(FS,c_f)-\Pi(F,c_f)}{\partial c_f} < 0$ in that region, it suffices to show that F is not optimal at c^{**}. This is precisely the definition of $\bar{K}_2$.

(c) In contrast to (c), we now have $K > \pi(FS,c^{**}) - \pi(F,c^{**})$. Because $\pi(FS,c^{**}) - \pi(F,c^{**})$ is decreasing in this region, there exists a value $c_f < c^{**}$ s.t. $\Pi(FS,c_f) = \pi(FS,c_f) - 2K < \Pi(F,c_f) = \pi(F,c_f) - K$. Let $\bar{c_2}(K)$ denote the smallest such value for each K. As c_f increases beyond $u(py - (\pi(SS) - K))$, $x(c_f,K) = SS$.

(d) Since $K > \bar{K}_3 = \pi(S)$, hiring two strangers gives negatives profits. Because $\pi(FS) \leq \pi(F) + \pi(S)$, FS is never optimal either. In this region, entrepreneurs choose to hire their friend iff $c_f < u(py - K)$.

■

C.4. **Optimal Firm Organization with Moral Hazard.** This section analyzes organizational choices when δ of the stranger in ∞, for both cases of the friend δ and the entire range of friend quality c_f.

C.4.1. *Choice of Composition.*

Proposition C.5. *Let $c_f^*(\delta)$, $\delta \in \{0, \infty\}$ be the quality level of friend s.t. the entrepreneur is indifferent between hiring a friend or a stranger, $\pi(F, c_f^*(\delta), \delta)) = \pi(S, c_f^*(\delta), \delta))$. Let $c_f^{**}(\delta), \delta \in \{0, \infty\}$ be the quality level of friend s.t. the entrepreneur is indifferent between hiring a friend and a stranger, and 2 strangers $\pi(FS, c_f^{**}(\delta), \delta)) = \pi(SS, c_f^{**}(\delta), \delta))$.*

(1) $\forall \delta$, $c_f^*(\delta)$ *and* $c^{**}(\delta)$ *are unique;*

$$x_s(c_f, \delta) = \begin{cases} F & \forall c \leq c_f^*(\delta) \\ S & otherwise \end{cases}; \quad x_l(c_f, \delta) = \begin{cases} FS & \forall c \leq c^{**}(\delta) \\ SS & otherwise \end{cases}$$

(2) $\forall \delta$, $w(S) < u^{-1}(c_f^*(\delta))$, $w(SS) < u^{-1}(c_f^{**}(\delta)$.

(3) $\forall \delta$, $u(w(S)) < c_f^{**}(\delta) < c_f^*(\delta)$.

(4) $c_f^*(0) < c_f^*(\infty)$ *and* $c_f^{**}(0) < c_f^{**}(\infty)$

C.4.2. *Choice of size and composition.*

Proposition C.6. *[i]*

(1) $\forall \delta$, $\{c_f, K : x(c_f, K, \delta) = S\} = \varnothing$

(2) *Let*

- $\overline{K}_1(\delta) = \pi(FS, 0, \delta) - \pi(F, 0, \delta)$
- $\overline{K}_2(\delta) = \pi(FS, c^{**}(\delta), \delta) - \pi(F, c^{**}(\delta), \delta)$
- $\overline{K}_3(\delta) = \pi(FS, \hat{c}, \delta) - \pi(F, \hat{c}, \delta)$
- $\overline{K}_4(\delta) = \pi(S)$.

where $c^{**}(\delta)$ is as defined in Proposition C.5 and $\hat{c} = argmax_{c_f} \pi(FS, c_f, \delta) - \pi^{MH}(F, c_f, \delta)$.
Suppose that $\overline{K}_1(\delta) < \overline{K}_2(\delta)$.
Then $\overline{K}_2(\delta) < \overline{K}_3(\delta) \leq \overline{K}_4(\delta)$ and the following holds:

(a) $\forall\, K < \overline{K}_1(\delta)$, $x(c_f, K, \delta)) = \begin{cases} FS & \forall\, c \leq c^{**}(\delta) \\ SS & otherwise \end{cases}$

(b) $\forall\, \overline{K}_1(\delta) < K < \overline{K}_2(\delta)$, $x(c_f, K, \delta) = \begin{cases} F & \forall\, c_f \leq \bar{c_f}(K, (\infty)) \\ FS & \forall\, \bar{c_1}(K, \delta) \leq c_f \leq c^{**}(\delta) \\ SS & otherwise \end{cases}$

(c) $\forall\, \overline{K}_2(\delta) < K < \overline{K}_3(\delta)$, $x(c_f, K, \delta) = \begin{cases} F & \forall\, c_f \leq \bar{c_1}(K, \delta) \\ FS & \forall\, \bar{c_1}(K, \delta) \leq c_f \leq \bar{c_2}(K, \delta) \\ F & \forall\, \bar{c_2}(K, \delta) \leq c_f \leq \bar{c_3}(K, \delta) \\ SS & otherwise \end{cases}$

(d) $\forall\, \overline{K}_3(\delta) < K < \overline{K}_4(\delta)$, $x(c_f, K, \delta) = \begin{cases} F & \forall\, c \leq \bar{c_3}(K, \delta) \\ SS & otherwise \end{cases}$

(e) $\forall\, K > \overline{K}_4(\delta)$, $x(c_f, K, \delta) = \begin{cases} F & \forall\, c \leq \bar{c_4}(K, \delta) \\ 0 & otherwise \end{cases}$

Both $\bar{c_1}(K,\delta)$ *and* $\bar{c_2}(K,\delta)$ *are implicitly defined by* $\pi(FS,\bar{c_i}(K),\delta)-\pi(F,\bar{c_i}(K),\delta)=K$. $\bar{c_3}(K,0)=pu\left[y-\frac{(\pi(SS)-K)}{p}\right]$, $\bar{c_4}(K,0)=pu\left(y-\frac{K}{p}\right)$, $\bar{c_4}(K,\infty)=u(py-(\pi(SS)-K))$, *and* $\bar{c_4}(K,\infty)=u(py-K)$

(3) *The following comparisons hold*
  (a) $\overline{K}_3(0)=\overline{K}_4(0)$, $\overline{K}_3(\infty)<\overline{K}_4(\infty)$ *and* $\overline{K}_4(0)=\overline{K}_4(\infty)$
  (b) $\overline{K}_3(0)=\overline{K}_3$
  (c) $\bar{c_4}(K,0)<\bar{c_4}(K)$
  (d) $\bar{c_4}(K,0)<\bar{c_4}(K,\infty)$
  (e) $\bar{c_3}(K,0)<\bar{c_3}(K,\infty)$
  (f) $\bar{c_3}(K,\infty)>\bar{c_3}(K)$

Proof of Proposition C.6.

(1) As before, profits are linear and hiring two strangers always dominates hiring one.
(2) For $\delta=0$, the proof of Proposition C.4 applies. For $\delta=\infty$, note that $\max_{c_f}\pi(FS)-\pi(F)<\pi(S)$. Since now also the slope of the optimal wage schedule is different for friends and strangers, the two wages can never coincide and running a mixed firm is always less profitable than hiring a friend and a stranger separately. It follows that $K_4>K_3$.
(3) Straightforward calculation.

■

Appendix D. Quotes

"I pay a piece rate for each cloth worked on by my workers (INR 3 per piece). I will pay both a stranger and a friend the same rate. . . If I differentiate now, God will give me less later."

–Dry cleaning shop owner, Bangalore

"You have to pay them the same for the same work. If I pay one person INR 15,000 and another person INR 10,000 they will talk to each other and get to know one person is being paid less. This person will then come fight with me. People will not stay if you pay less."

–Paint shop owner, Bangalore

"To known and unknown workers we don't pay different salary because they will then not agree to work."

–Shoe shop owner, Bangalore

"One cannot differentiate between two eyes, they see the same. If we pay more for one person, another will not work properly"

–Clothes shop owner, Bangalore

"Issues will happen (if we pay them differently), co-workers will discuss with each other and may get jealous of each other. This may lead to misunderstandings/fights, my business will not go good."

– Watch shop owner, Bangalore