E Ozdenoren and Oleg Rubanov
Profit Sharing and Incentives
Article

Ozdenoren, E and Rubanov, Oleg

(2022)

Profit Sharing and Incentives.

International Journal of Industrial Organization. p. 102857. ISSN 0167-7187

(In Press)

DOI: https://doi.org/10.1016/j.ijindorg.2022.102857

Elsevier
https://www-sciencedirect-com.lbs.idm.oclc.org/sci...

Journal Pre-proof

Profit Sharing and Incentives

Emre Ozdenoren, Oleg Rubanov

PII: S0167-7187(22)00033-9
DOI: https://doi.org/10.1016/j.ijindorg.2022.102857
Reference: INDOR 102857

To appear in: *International Journal of Industrial Organization*

Received date: 13 March 2021
Revised date: 5 May 2022
Accepted date: 10 May 2022

Please cite this article as: Emre Ozdenoren, Oleg Rubanov, Profit Sharing and Incentives, *International Journal of Industrial Organization* (2022), doi: https://doi.org/10.1016/j.ijindorg.2022.102857

Highlights

- We study how profits of a firm should be shared in the most efficient way among heterogeneously productive, risk averse workers, and an unproductive outside investor. We model the firm as a team production process subject to moral hazard. In addition to awarding shares, the firm can use incentive contracts based on noisy performance signals.

- We show that more productive agents with noisier performance signals are more likely to be motivated with shares as opposed to incentive contracts. Since the efforts of agents who have soft skills, such as managers, are more difficult to observe, this result provides a potential explanation for why managers in most firms are motivated by shares.

- Our results also suggest an explanation for why law or consulting firms, where agents' efforts are difficult to observe, are often organized as partnerships.

- We show that the unproductive outside investor holds shares only if all productive agents hold shares.

- We ask whether large firms are more or less likely to be owned by outside investors. We find that a firm that grows by opening branches is held almost entirely by outside investors when its output noise grows faster than the number of branches. Otherwise, insiders hold substantial amount of a large firm's shares.

Profit Sharing and Incentives

Emre Ozdenoren* London Business School

Oleg Rubanov† Nazarbayev University

May 25, 2022

Abstract

We model a firm as a team production process subject to moral hazard and derive the optimal profit sharing scheme between productive workers and outside investors together with incentive contracts based on noisy performance signals. More productive agents with noisier performance signals are more likely to receive shares which can explain why managers are motivated by shares, and law or consulting firms form partnerships. A firm that grows by opening branches is held almost entirely by outside investors when its output noise grows faster than the number of branches. Otherwise, insiders hold substantial amount of a large firm's shares.

1 Introduction

Firms are organized in a variety of ways. One common organizational form is a public corporation where external shareholders own shares in the firm. Often, a substantial amount of the shares in the firm are held by outside investors who do not participate in the productive activities of the firm. Another common organizational form is partnership. Unlike outside shareholders of a corporation, partners in a partnership are typically insiders who

*Contact: Department of Economics, London Business School, Regent's Park, NW1 4SA, London, United Kingdom. Email: eozdenoren@london.edu. Web: https://sites.google.com/site/ozdenoren/

†Contact: Department of Economics, Nazarbayev University, Qabanbay Batyr ave. 53, 010000, Nur-Sultan, Kazakhstan. Email: oleg.rubanov@nu.edu.kz. Web: https://sites.google.com/site/orubanov/

participate in productive activities of the firm. Hansmann (1996) reports that in the US employee ownership is uncommon in industrial sector of the American economy whereas it is quite common in the service professions like law, accounting, investment banking, management consulting, advertising, architecture, and medicine.[1] The pattern is similar in Europe and elsewhere. Although Hansmann's observation is broadly correct employee ownership is not an all or nothing decision. Most firms fall somewhere in between where some shares are held by productive insiders and the rest by outsiders. For example, in the services industry where partnership is common many workers are given performance based incentives either exclusively or in addition to ownership of the firm. Similarly, in industrial firms where ownership is not as common, often upper management receives ownership based incentives. This observation suggests that theories of employee ownership should also explain both who obtains shares within the firm as well as the extent to which outside investors receive shares.

This paper focuses on how profits of a firm should be shared in the most efficient way to incentivize their workers and share risks among them. In partnerships profits must be shared among the partners of the firm. In their path breaking work, Alchian and Demsetz (1972) and Holmström (1982), show that profit sharing leads to important incentive problems. This insight, in fact even more powerfully, applies to corporations where a substantial amount of the shares is held by outside shareholders. To solve the resulting incentive problems, in most firms, workers also receive contracts which condition pay on performance.[2] Therefore, the efficient allocation of ownership and the provision of performance based contracts must be jointly determined.

To derive the optimal allocation of ownership and performance contracts, in Section 2, we consider a firm that employs agents who are heterogeneous with respect to their productivities and levels of risk aversion. There is also an unproductive and potentially risk neutral outside investor. The firm's production depends stochastically on the efforts of the productive agents. Agents' efforts are not observable and subject to moral hazard. We assume that for each productive agent there is a contractible noisy performance signal of his effort level. The productive agents and the outside investor receive contracts based

[1]One can find partnerships in other industries. For example, John Lewis Partnership is a leading UK retail business where the partners are its permanent staff.

[2]Including both profit sharing and performance signals always improves surplus no matter how noisy output or the signals are as long as they are informative about performance. Generally, this follows from Holmström (1979) who shows that signals that are informative about performance should always be used in the contract no matter how noisy they might be.

on all the signals. For a productive agent, the contract he receives provides incentive to exert effort since it depends on his own effort. Profit of the firm is given by its production minus the contractual payments. Productive agents and the outside investor share the profit so that each agent receives a non-negative share of the firm and the shares add up to one.[3] For productive agents, receiving a share of the firm's profit provides additional incentive to work but also exposes the agent to the output risk. We abstract away from the bargaining process through which parties decide on share allocations and contracts, and assume that these are determined to maximize the total surplus.

In Section 3 we present our main model and characterize the optimal allocation of shares when the firm can write contracts based on performance signals. In the main model we restrict attention to linear contracts, normally distributed signals and CARA utility functions. In Appendix B we study a variation of the model with binary outcomes and signals, arbitrary concave utility functions and general contracts. Holmström (1982) shows that under budget balance first-best cannot be achieved but when there is a risk-neutral and unproductive outside investor who breaks the budget balance and acts as a residual claimant first-best is restored.[4] Our focus is slightly different than Holmström's. We preserve budget balance and introduce a risk neutral outside investor not as a residual claimant but to improve risk sharing. We derive an ownership parameter and rank the agents according to this parameter. An agent's ownership parameter increases in his productivity and the variance of his performance signal.[5] We show that only the agents with the k-highest ownership parameters own shares where k is determined endogenously. Hence more productive agents whose performances are poorly observable are more likely to be partially motivated with shares, rather than just a contract. Since the efforts of agents who have soft skills, such as managers, are more difficult to observe, this result provides a potential explanation for why managers in most firms are motivated by shares. The result also suggests an explanation for why law or consulting firms, where agents efforts are difficult to observe, are often organized as partnerships.

We show that the unproductive outside investor always has the lowest ownership

[3]Note that budget balance is automatically imposed since agents who receive profit shares effectively are liable for the contract payments in proportion to their shares.

[4]This solution has its issues. For example, when the total output is slightly less than the optimum, the principal's payoff is much higher. Hence, she has incentive to sabotage production. Also, see Eswaran and Kotwal (1984) who discuss why budget breaking schemes might be hard to implement. Legros and Matsushima (1991) discuss balanced transfers in partnerships.

[5]It also depends on the level of risk aversion of all agents but the ranking of the ownership parameter depends on the risk aversion parameters in a more subtle way.

parameter and holds shares only if *all* productive agents hold shares even though it is possible to write performance based contracts.[6] The reason is that due to budget balance, any improvement in risk sharing that is achieved by giving shares to the outside investor comes at the cost of reduced incentives for the productive agents, and when a productive agent does not have any shares, marginally increasing his shares has a first order effect on incentives and a second order effect on risk. To show that this result is robust to model specification and does not rely on the linear/normal structure, in Appendix B we present a variation of our main model where output and signals are binary and the contracts are unrestricted and show that the result still holds.

An interesting question is whether large firms are more or less likely to be owned by outside investors. As we emphasized above, in our model the outside investor is unproductive and the only purpose of giving her shares is to improve risk sharing which comes at a cost because it reduces the incentives of the insiders to exert effort. In Section 4, we ask whether, as the firm size grows, the insiders are able to avoid this cost by self-insuring or whether they need to give shares to an unproductive outside investor so that the large firm is owned partially or completely by outside investors. We model a large firm as a collection of identical branches. Each branch employs heterogenous productive worker types. We show that as the firm size (or the number of branches) grows, whether the insiders keep all the shares or they give some, and even almost all of the shares to the outside investor depends on the ratio of the variance of the noise in the total output to the number of branches. Roughly speaking, this ratio is determined by the correlation of output risk across branches. For example, when the output risk is perfectly and positively correlated across branches, the variance of the noise in the total output grows faster than the number of branches. In this case, the risk grows too fast for the risk averse insiders to shoulder and a large enough firm is held almost entirely by the outside investors. In contrast, when this ratio is constant (or vanishing), the outcome is determined by the tradeoff between risk and incentives and requires the agent types in each branch whose ownership parameters exceed a certain cutoff to be motivated by shares and shoulder the risk.[7] In particular, when the output risk vanishes relative to the size of the firm, shares are held by the agent type who has the highest ownership parameter. This result provides

[6] This result is about the extensive margin that determines whether an agent or outside investor holds shares. It is not about the intensive margin that determines how many shares they own. In particular, if the output is very noisy the outside investor might own most of the firm but the productive agents still own some shares.

[7] To be more precise, if the ratio is constant but large enough, some shares might be held by the outside investor.

one explanation for the observation that partnerships tend to form among individuals with similar characteristics such as lawyers in a law firm. The broader message is that larger firms are more likely to be partnerships only after controlling for output risk. If we do not control for risk, in industries where firms face common risks that increase rapidly as they expand, larger firms are more likely to be owned by outside investors.

In Section 5 we study how the optimal share allocation and contracts depend on the precision of the output noise, providing further intuition for the degree of outside ownership.

We discuss how the results of our paper compare to the literature in Section 6. Proofs missing from the main text can be found in the Appendix.

2 The Main Model: Linear Contracts

We model a firm as a team of agents who are engaged in the production of a good. There are n productive agents indexed by $i \in \{1, 2, \dots, n\}$ and an unproductive outside investor indexed by 0. The level of production of the firm depends on the efforts of the productive agents. We denote agent i's effort by $e_i \geq 0$.

The output of the firm is given by $y(e_1, \dots, e_n) = q(e_1, \dots, e_n) + \varepsilon_q$ where ε_q captures the uncertainty in the outcome of the production process. For tractability, we assume the production function to be the sum of individual efforts $q(e) = \sum_{i=1}^{n} e_i$.[8]

The efforts of the agents are not directly observable. Instead, for every agent i there is a signal $s_i = e_i + \varepsilon_{s_i}$ which is observable by everyone (including the court), and where ε_{s_i} are jointly independent.

Efforts are costly for the agents. The cost of effort is quadratic, and the cost functions are heterogenous. It costs agent i $C_i(e_i) = \mu_i \frac{e_i^2}{2}$ to exert effort e_i. The lower μ_i is, the less costly it is to exert effort for agent i. Sometimes we call agents with lower μ_i "more productive".

We assume that the agents have CARA utility functions. Agent i's von Neumann-Morgenstern utility function of consuming x and exerting effort e_i is given by $1 -$

[8]Here and below we often refer to the vector $(x_1, \dots, x_n)$ as x (for example, for λ, e and other variables).

$e^{-\gamma_i(x-C_i(e_i))}$, where $\gamma_i > 0$ is agent i's coefficient of absolute risk aversion. We assume that there is no limited liability, and the agents have deep pockets, so that they can be made to pay any amount of money if the contract requires them to do so.

All noises ε_q and ε_{s_i} are normally distributed where $\varepsilon_q \sim \mathcal{N}(0, \sigma_q^2)$, $\varepsilon_{s_i} \sim \mathcal{N}(0, \sigma_{s_i}^2)$, and signal noises ε_{s_i} and the output noise ε_q are jointly independent.

We consider two situations: with and without outside investment. To capture the possibility of outside investment, we consider a risk neutral agent 0 who represents all outside investors. We assume that exerting zero effort is costless but exerting strictly positive effort is extremely costly for the outside investor. Consequently, the outside investor does not directly participate in production and $e_0 = 0$. In the formulas below, when the outside investor is present we set $\gamma_0 = 0$ and $\mu_0 = \infty$. Although the external investor exerts zero effort and need not be incentivized, she might still own shares of the company and make transfers to the other agents based on their performances for risk sharing reasons. We denote the set of all agents by I which includes the outside investor agent 0 when there is one. Since the case with only one agent in the economy is trivial, we assume that there are at least two agents one of whom is potentially the outside investor, i.e. $|I| \geq 2$.

We assume that the agents can only benefit from the output of the firm by owning its shares, but cannot contract on the output otherwise. Hence, every agent owns share $\lambda_i \geq 0$ of the firm, and the total number of shares is equal to 1.[9] If agent i owns share λ_i of the company, and the profit of the company is π, then agent i receives $\lambda_i \pi$.[10]

Following Holmström and Milgrom (1987), we restrict attention to linear compensation contracts. Denote the compensation scheme of agent i as w_i. Since we assume agent i's compensation is linearly dependent on the available signals, denote $\tilde{\beta}_i^j$ the additional amount agent i receives for a unit increase in signal s_j. In addition, $\tilde{\beta}_i^0$ denotes the lump sum payment to agent i from the firm (or alternatively, the rest of the agents). Thus, for

[9] One reason for the non-contractibility of output is that contracts are short term, while some of the effects of efforts on output might be realized in the long term. Another reason is that the value of being a shareholder of a firm might provide non-tangible benefits and be much more important to the shareholders than just the production output of the company or its share price.

[10] This includes the situations when π is negative. In such realizations of profit the owners split the obligations of the company, rather than the profit.

a given realization of signals $s = (s_1, \ldots, s_n)$ agent $i \in I$ receives:

$$\tilde{w}_i(s) = \tilde{\beta}_i^0 + \sum_{j=1}^{n} \tilde{\beta}_i^j s_j. \tag{1}$$

The profit of the company is equal to its output minus the cost of labour (we assume that the other costs are already incorporated in the production function):

$$\pi = y - \tilde{w}(s) = \sum_{i=1}^{n} e_i - \sum_{i \in I} \tilde{w}_i(s) + \varepsilon_q = \sum_{i=1}^{n} e_i - \sum_{i \in I} \sum_{j=1}^{n} \tilde{\beta}_i^j s_j - \sum_{i \in I} \tilde{\beta}_i^0 + \varepsilon_q. \tag{2}$$

There are four periods in this model. In the first period the shares are allocated among the agents. In the second period the agents sign contracts that determine transfers in every state of the world. In the third period, taking the allocation of shares and the contracts as given, agents exert effort. In the fourth period signals are realised, and the profit is shared according to the agents' shares in the company, and the transfers are made as specified in the contract given the state of the world.[11]

We assume that the agents choose the allocation of shares and the contracts to achieve the levels of effort that maximize the total surplus. This allows us to abstract away from the bargaining process through which parties decide on share allocations or contracts.[12]

3 Optimal ownership structure and contracts

The goal of the firm is to motivate the productive agents in the best possible way while allowing the agents to share risks optimally. The firm can use two tools to accomplish these objectives. These are ownership of shares (or profit sharing) and performance based incentive contracts. In this section we characterize the optimal mixture of ownership

[11] We do not assume that the transfers add up to zero in every state because, with profit sharing, budget balance is automatically satisfied. For example, if in a state the sum of contractual transfers is strictly positive then this amount is subtracted from output. The remainder, which is the profit of the firm that might be negative, is then shared among the agents.

[12] For example, a designated agent could make take-it-or-leave-it offers to all other agents. This agent would choose share allocations and contracts to maximize the total surplus, and make sure that all the other agents receive exactly their outside options so that they accept the offers. There are no outside options in our model, so any lump sum payments β_i^0 are optimal, because they do not affect the incentives of the agents.

and contracts. We then use this characterization to study how ownership and incentives depend on factors such as firm size, agents' productivities, and the riskiness of incentives and production.

As a preliminary observation consider an alternative formulation of the model in which, instead of sharing the profit, agents split the output of the firm and write contracts that are budget balanced, i.e. payments of the agents including the outside investor add up to zero for any realization of the signals.[13] In the Appendix A we study the relationship between the two formulations. Theorem A.2 in the Appendix shows that given profit shares $(\lambda_0, ..., \lambda_n)$ and a profile of contracts, identical output shares $(\lambda_0, ..., \lambda_n)$ and a *unique* profile of modified contracts that are budget balanced implement the same payoff for all agents. Conversely, given output shares $(\lambda_0, ..., \lambda_n)$ that sum to one and a profile of contracts that are budget balanced, the same profile of profit shares $(\lambda_0, ..., \lambda_n)$ and *multiple* profiles of modified contracts implement the same payoff for all agents. This implies that the optimal output shares and profit shares are identical but the contracts are unique only in the output setting. Therefore, in the remainder of the paper we study the optimal ownership structure using the output sharing formulation and the corresponding profit sharing arrangements can be derived in a straightforward manner.

As usual, we solve this problem backwards. The first step is to solve for optimal effort choices of the agents given a fixed allocation of output shares and contracts. In equilibrium given the efforts of other agents, agent i chooses e_i to maximize

$$\max_{e_i} \mathbb{E}\left(\lambda_i y(e) + w_i(s) - c_i(e_i)\right) - \frac{\gamma_i}{2}\mathbb{V}\left(\lambda_i y(e) + w_i(s) - c_i(e_i)\right),$$

where $e = (e_1, \ldots, e_n)$. Since the noise in output and signals are additive, the variance term is constant. Moreover, the other agents' efforts are taken as given, so the expression above is equivalent to:

$$\max_{e_i} \lambda_i e_i + \beta_i^i e_i - C_i(e_i).$$

Hence, optimal effort e_i satisfies $(\lambda_i + \beta_i^i) = C_i'(e_i)$ or $e_i = (\lambda_i + \beta_i^i)/\mu_i$.

In the Online Appendix we show that the optimal ownership and incentive structure maximizes the sum of certainty equivalents of the payoffs of all agents subject to the constraints that for any realization of signals payments among the agents are balanced,

[13]Notice that if the payments are budget balanced (add up to zero), then the profit is equal to output. So, output shares is equivalent to profit shares, when the contracts are budget balanced.

each agent receives a non-negative share of the firm, and the agents' shares in the firm add up to the whole firm:

$$\max_{\lambda, w(s)} \sum_{i \in I} \Big(\mathbb{E} \left(\lambda_i y(e) - C_i(e_i) + w_i(s) \right) - \frac{\gamma_i}{2} \mathbb{V}(\lambda_i y(e) + w_i(s)) \Big), \\ \text{s.t. } \forall s : \sum_{i \in I} w_i(s) = 0, \forall i : \lambda_i \geq 0 \text{ and } \sum_{i \in I} \lambda_i = 1. \tag{3}$$

Substituting the optimal effort levels and noticing that $\sum_{i=1}^{n} \mathbb{E} w_i(s) = 0$ we obtain:

$$\mathbb{E} \left(\lambda_i y(e) - C_i(e_i) \right) = \lambda_i \sum_{j=1}^{n} e_j - \frac{\mu_i e_i^2}{2} = \lambda_i \left(\sum_{j=1}^{n} \frac{\lambda_j + \beta_j^j}{\mu_j} \right) - \frac{\mu_i}{2} \left(\frac{\lambda_i + \beta_i^i}{\mu_i} \right)^2,$$

$$\mathbb{V}(\lambda_i y_i(e) + w_i(s)) = \lambda_i^2 \sigma_q^2 + \sum_{j=1}^{n} \left(\beta_i^j \right)^2 \sigma_{s_j}^2.$$

The next proposition characterizes the optimal contracts for a given ownership structure.

Proposition 3.1. *Without outside investment, given the share allocation λ, for every j,*

- *the increase in agent j's pay for a unit increase in his own performance signal s_j is given by:*
$$\beta_j^j = \frac{1 - \lambda_j}{1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{i \neq j} \frac{1}{\gamma_i}} \right)}; \tag{4}$$

- *for every $i \neq j$ the reduction in agent i's pay for a unit increase in agent j's signal s_j is given by:*

$$\beta_i^j = -\frac{\frac{1}{\gamma_i}}{\sum_{k \neq j} \frac{1}{\gamma_k}} \left(\frac{1 - \lambda_j}{1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{k \neq j} \frac{1}{\gamma_k}} \right)} \right). \tag{5}$$

With the outside investor, $\beta_j^j = \frac{1 - \lambda_j}{1 + \mu_j \sigma_{s_j}^2 \gamma_j}$, $\beta_0^j = -\beta_j^j$ and $\beta_i^j = 0$ for $j \neq 0, i$.

Proposition 3.1 allows us to make several observations about the incentive structure. First, agent j receives more powerful incentives if he holds fewer shares, or equivalently, the other agents own more of the firm. In addition, agent j receives more powerful incentives if he is more productive, has a less noisy performance signal and is less risk averse. Interestingly, agent j's incentives become more powerful if the other agents in the team become less risk averse. This is because agent j's payment comes from the other agents in the team and hence they are subject to the noise in agent j's payment. Finally, note that how much of agent j's payment comes from agent i does not depend on how much of the firm agent i owns. This is because the payments are shared to optimize the allocation of risk across agents (which is captured by the coefficient in front of the parenthesis in equation (5)). In fact, if there is a risk neutral outside agent 0 (for example, an insurance company) then only the outside investor (agent 0) makes payments to the agents and $\beta_i^j = 0$ for all $i \neq 0$ even if the outside investor does not own any shares in the firm.

Next we turn to the optimal allocation of shares. We define

$$D_i \equiv \frac{1}{\mu_i + \frac{1}{\sigma_{s_i}^2\left(\gamma_i + \frac{1}{\sum_{j\neq i}\frac{1}{\gamma_j}}\right)}}. \tag{6}$$

We refer to D_i as the *ownership parameter* of agent i.[14,15]

In the optimal share allocation problem, there is an extensive margin that determines whether an agent holds shares or is incentivized solely based on a performance contract. There is also an intensive margin that determines, conditional on holding shares, how many shares an agent holds. The ownership parameters D_i play a crucial role in determining both margins. To see this rank the agents according to their ownership parameters, so that $0 < D_1 \leq D_2 \leq \cdots \leq D_n$. If there is an outside investor, then by substituting $\mu_0 = \infty$ and $\gamma_0 = 0$ into the definition (6), we see that her ownership parameter $D_0 = 0$ which is strictly less than D_1. The following proposition shows that only the agents with highest ownership parameters hold shares in the firm. Agent i's ownership parameter increases, and the other agents' ownership parameters are constant, in his productivity and the variance in his performance signal. Hence, if an agent becomes more productive or his performance signal becomes more noisy, his rank will be higher and as we show

[14] The ownership parameters also feature in Rayo (2007) and play similar roles in both papers: they capture the sensitivity of the incentives based on the noisy signal of effort as the agent's level of ownership increases.

[15] In these derivations we do not assume that there is necessarily a risk neutral outside investor, and Equations (4), (5) and (6) all simplify when there is one. In particular with a risk neutral outside investor β_i^j's are all zero and (6) becomes $D_i = \frac{1}{\mu_i + \frac{1}{\sigma_{s_i}^2 \gamma_i}}$.

next he is more likely to hold shares in the firm.

Proposition 3.2. *If an agent with ownership parameter D_i has some shares of the firm, then all agents with at least as high ownership parameters $D_j \geq D_i$ also own shares of the firm (in both the case with and without the outside investor).*

For every two agents i and j with positive holdings of shares the following condition holds:

$$(D_i - D_m) - \lambda_i(D_i + \gamma_i\sigma_q^2) = -\lambda_m(D_m + \gamma_m\sigma_q^2). \tag{7}$$

The next proposition characterizes the optimal ownership structure in the firm.

Proposition 3.3. *Suppose m is the lowest k that satisfies*

$$\sum_{i=k+1}^{n} \frac{D_i - D_k}{D_i + \gamma_i\sigma_q^2} < 1. \tag{8}$$

Agent i holds shares if and only if $i \geq m$. That is, only the $n - m + 1$ agents with the highest ownership parameter D_i own shares of the firm.

If $m > 0$ then each of these $n - m + 1$ agents owns

$$\lambda_j = \frac{1 - \sum_{i=m}^{n} \frac{D_i - D_j}{D_i + \gamma_i\sigma_q^2}}{\sum_{i=m}^{n} \frac{D_j + \gamma_j\sigma_q^2}{D_i + \gamma_i\sigma_q^2}} \tag{9}$$

shares.

If condition (8) is satisfied for $k = 0$ (i.e. $m = 0$), then the outside investor's share is positive and is equal to:

$$\lambda_0 = 1 - \sum_{i=1}^{n} \frac{D_i}{D_i + \gamma_i\sigma_q^2}, \tag{10}$$

and agent j's share is

$$\lambda_j = \frac{D_j}{D_j + \gamma_j\sigma_q^2}. \tag{11}$$

Proposition 3.3 implies that if the outsider holds shares of the firm then every productive agent has shares of the firm. This result does not emerge as an artefact of the linear/normal model. In Appendix B we consider a variation of our model with arbitrary utility functions and compensation schemes where signals and output realizations are binary. We show that in that model the outsider also holds shares of the firm only if everyone else holds shares of the firm. This result is in line with many modern employment contracts, where employees are given equity as part of their compensation. However, there are many public firms in which most stakeholders are not shareholders, and yet most of the shares are held by outside investors. We provide one explanation for this observation in Section 4 where we show that if output noise grows faster than the size of the firm then in the limit almost all shares go to outsiders. Under this scenario, if there is a small fixed cost to introducing ownership based incentives into employment contracts than most agents should obtain no shares. A second explanation is that even agents who do not explicitly obtain shares in the firm may act *as if* they have some ownership (feeling company spirit and being part of the team) in which case optimal contract may provide them only performance based incentives.

To see why the outsider may be given some shares, imagine a situation, where all agents start with no shares, and then small amounts of shares are gradually allocated to the agents. Suppose at each step, the recipient of the next portion of shares is selected to maximize total welfare.[16] The agents getting their first portions of shares get a large boost to their incentive to work, but very little additional risk (the marginal cost of effort and the risk bearing are both almost zero relative to the marginal benefit). However, a risk averse agent bears a growing amount of risk as more shares are allocated to him. At some point, additional boost to performance is dominated by the additional risk the agent must bear. At this point, it is better to give all additional shares to the risk neutral outsider who gets a benefit proportional to output. Without the outsider, if the agents are sufficiently risk averse, the total amount of risk could exceed the benefit of production, in which case it would be beneficial for the agents to "burn shares". However, with the outsider, burning shares is no longer necessary. We elaborate further on the role of output risk on the allocation of shares between productive agents and the outsider in Section 5.

Inequality (8) always holds for $k = n - 1$ implying that the two agents with the highest productivity parameters always hold some shares. To see why this is the case consider the simpler problem of allocating shares of the entire firm to the two agents with the highest productivity parameter, i.e. agents n and $n-1$. For the moment assume that performance based incentives are not available. This problem can be written as maximizing $\sum_{i=n-1}^{n} \left(\lambda_i/\mu_i - C_i(e_i) - \lambda_i^2 \frac{\gamma_i}{2}\sigma_q^2\right)$ with $\lambda_{n-1} + \lambda_n = 1$. Each term in brackets

[16] This exercise is equivalent to solving the initial problem, but substituting constraint $\sum \lambda_i = 1$ with constraint $\sum \lambda_i = x$, where x changes between 0 and 1.

under the summation can be interpreted as the impact of agent i's ownership on the total surplus. It is easy to see that each term (on its own) is maximized at $0 < \lambda_i \leq 1$. That is, ignoring the adding up constraint, the planner would like each agent to hold some of the firm but when there is risk in the output not the entire firm.[17] Put differently, the term in brackets is increasing at 0 and weakly decreasing at 1. This means that starting from a situation where agent n owns the entire firm, reducing agent n's share slightly and increasing agent $n-1$'s share by exactly that amount always increases the total surplus.[18]

The above intuition also goes through when performance based incentives are available. We argued that, without performance based incentives, the impact of agent i's ownership on the total surplus (on its own) is maximized at $0 < \lambda_i \leq 1$. With performance based incentives, this number will weakly decrease since it is no longer necessary to motivate the agent solely by ownership but will still exceed zero since some ownership improves risk allocation. This means that, fixing the performance based incentives, starting from a situation where the most productive agent owns the entire firm, reducing the most productive agent's share slightly and increasing the second most productive agent's share by exactly that amount always increases the total surplus. Notice that while β_j^j is positive for all productive agents ($\sum_{k \neq j} \lambda_k > 0$ in equation (4), since we showed that agent j cannot be the sole shareholder of the firm), some of them might not get any shares. This is because although both shares and wages incentivize, the shares are in limited supply, so giving shares to one agent disincentivizes another agent. At the same time, paying one agent a higher wage has no effect on the other productive agents' incentives.

As is typical in moral hazard problems, all agents' efforts are lower than the first best.[19] Indeed, the first best level of effort is $e_i^* = \frac{1}{\mu_i}$, but since effort is unobservable, the optimal effort is $e_i = \frac{\lambda_i + \beta_i^i}{\mu_i}$, which is lower than e_i^*, as it follows from equation (4) that $\beta_i^i < 1 - \lambda_i$. It is natural to assume that every agent's effort without an outside investor is smaller than with an outside investor, and it is true in most cases, but we provide a

[17]When there is no risk in the output, ignoring the adding up constraint, the planner would like each agent to hold the entire firm.

[18]This argument does not imply that all agents must hold positive shares. Consider a situation with $k \geq 2$ agents such that these agents optimally get positive shares in the firm. Now suppose a new agent (with a lower ownership parameter) joins the team. The new agent may optimally get zero shares in the firm: the impact of the new agent's ownership on the total surplus will be strictly increasing in his ownership (since he starts at zero). But at the previous optimum, the shares are optimally allocated so that the impact of each existing agents' ownership on the total surplus is also increasing that agent's ownership. Thus allocating shares to the new agent from the existing ones might sometimes reduce the overall surplus.

[19]The opposite can happen in some models as well, see Rauh (2014).

counter-example in an online appendix.

Proposition 3.3 characterizes both the extensive margin (who holds shares) and the intensive margin (how much of the firm a shareholder holds). Once these are determined, Proposition 3.1 tells us the optimal performance based incentives each agent receives. In the next section, we use these two propositions to provide comparative statics results on the ownership and incentive structures in the firm.

4 Profit Sharing and Firm size

We have seen that sharing profits with an outside investor improves risk sharing, but reduces the incentives of the insiders to exert effort. As the firm size grows the insiders' ability to share risks among themselves increases but the risk that they face may also grow. Is the outside investor more likely to own profit shares in a large firm? Could the outside investor own a large firm entirely? To answer these questions we look at the optimal profit sharing in a large firm. We model a large firm as a collection of N identical branches. Each branch employs b heterogenous productive worker types indexed by $\{1, \ldots, b\}$. In addition, we assume that there is an unproductive and risk neutral agent 0.[20] Agent j in each branch is characterised by a triplet of parameters $(\mu_j^N, (\sigma_{s_j}^N)^2, \gamma_j^N)$ corresponding to the cost of effort coefficient, variance of the performance signal and the coefficient of risk aversion. Note that we allow for the parameters to depend on the number of branches N, allowing larger firms to employ workers of different characteristics. For example, larger firms might have access to better monitoring technologies that would lower the variance of the performance signal as N gets larger. The parameters for the agents within a branch can be different, so our specification allows for heterogeneity within a branch but requires the composition of the branches to be the same. We also assume that, as N grows, for all type j, productivity parameter, variance of the performance signal and coefficient of risk aversion, $(\mu_j^N, (\sigma_{s_j}^N)^2, \gamma_j^N)$, converge to finite positive limits given by $(\mu_j, \sigma_{s_j}^2, \gamma_j) > 0$.

As in our main model we specify the production of branch k as:

$$y_k = \sum_{i=1}^{b} e_i^k + \varepsilon_{q_k}$$

[20]Hence, there are $n = Nb$ productive agents altogether.

where e_i^k is the effort of agent i who works for branch k and ε_{q_k} is the noise in the production of branch k. The production of the firm is equal to the sum of productions of individual branches:

$$Y = \sum_{k=1}^{N} y_k = \sum_{k=1}^{N} \left(\sum_{i=1}^{b} e_i^k \right) + \varepsilon_q^2$$

where ε_q is the sum of the output noise in the branches. We assume that ε_{q_k}'s are identically and jointly normally distributed (but not necessarily independent), so their sum is also Gaussian:

$$\varepsilon_q = \sum_{j=1}^{N} \varepsilon_{q_k} \sim \mathcal{N}(0, \sigma_q^2).$$

To complete our description of the firm we need to specify how production is correlated across branches. Depending on the correlation structure, as the firm grows, the variance in the firm's output can be vanishing relative to N (for example, if $\varepsilon_{q_k} = -\varepsilon_{q_{k+1}}$ for every k), or growing relative to N (for example, if $\varepsilon_{q_k} = \varepsilon_{q_{\hat{k}}}$ for every k and $\hat{k}$, then $\sigma_q^2 = N^2 \sigma_k^2$). For the special case of no correlation between ε_k (independent noises), total variance is equal to $\sigma_q^2 = N\sigma_{q_k}^2$ and the variance in the firm's output grows at the same rate as N.

We define the ownership parameter for type j agent as:

$$D_j^N \equiv \frac{1}{\mu_j^N + \frac{1}{(\sigma_{s_j}^N)^2 \gamma_j^N}}.$$

Since all the parameters converge to finite positive limits, D_j^N converges to a finite positive value given by

$$D_j = \frac{1}{\mu_j + \frac{1}{\sigma_{s_j}^2 \gamma_j}}.$$

We assume that the limiting values are distinct and we order the types by their ownership parameter in the limit, i.e., $D_b > D_{b-1} > \cdots > D_1 > 0$. As before the ownership parameter of the unproductive outside investor, D_0, is zero.

The following proposition characterizes the distribution of shares, as the firm grows, depending on how fast σ_q^2 grows.

Proposition 4.1. *Let* $\alpha \equiv \lim_{N\to\infty} \frac{\sigma_q^2}{N} \in [0, \infty]$. *There exist* $(R_0, \ldots, R_{b-1})$, $\infty > R_0 = \sum_{j=1}^{b} \frac{D_j}{\gamma_j} > R_1 \geq \cdots \geq R_{b-1} > 0$, *such that*
(i) if $0 \leq \alpha \leq R_{b-1}$ *then for* N *large enough only the agents with the highest ownership parameter* D_b *own shares of the firm.*

(ii) if for some $j \in \{1, \ldots, b-1\}$, $R_j < \alpha \leq R_{j-1}$, *then for* N *large enough types* $j, j+1, \ldots, b$ *have positive shares of the firm, and the other types (including the outside investor) do not have any shares of the firm.*
(iii) if $\alpha \in (R_0, \infty)$, *then for* N *large enough agents of all types (including the outside investor) have strictly positive shares of the firm.*
(iv) if $\alpha = \infty$ *then in the limit the firm is owned entirely by the outside investor (although for any fixed* N *insiders always own a positive share).*

To understand the intuition for Proposition 4.1 let's start with the case where α goes to zero as N grows, or equivalently σ_q^2 is $o(N)$. Part (i) of the proposition implies that in this case a large enough firm is owned by only one type of insider. This is because as the firm grows output risk per agent vanishes and the agents with the highest ownership parameter are not only the best to own shares but are also able to shoulder the vanishing output risk. In fact, part (i) of Proposition 4.1 says that agents with the highest ownership parameter should hold the shares as long as α is below R_{b-1} in the limit. This sheds light on why partnerships tend to form among individuals with similar characteristics (e.g. lawyers in a law firm.) If we assume that insiders are similar in terms of their risk aversion parameters then agents with the largest ownership parameters have largest productivity parameters and their performances are difficult to observe. Indeed, lawyers in a law firm are more likely to have these characteristics than other staff.

When α goes to a strictly positive number above R_{b-1}, or equivalently σ_q^2 is $O(N)$, the output risk does not vanish and needs to be shared among the agents. In these cases, the outcome is determined by the tradeoff between risk and incentives and requires the agents with the highest ownership parameters to be motivated by shares and shoulder the risk although if the risk is large enough as in case (iii) some shares might be held by the outside investor. In case (iv) σ_q^2 grows faster than N and the risk grows too fast for the risk averse insiders to shoulder. In this case, a large enough firm is held almost entirely by the outside investor.

5 Output noise

In this section we study how the optimal share allocation changes with the precision of the output noise.[21] This comparative static crucially depends on whether the outside investor owns any shares of the firm. Everything else constant, the outside investor owns shares only if there is enough output risk. We state this formally in the next proposition.

Proposition 5.1. *There exists a cutoff $\bar{\sigma}_q^2 > 0$ such that the outside investor does not own any shares if $\sigma_q^2 \leq \bar{\sigma}_q^2$ and owns a positive share of the firm $\sigma_q^2 > \bar{\sigma}_q^2$. If the outside investor owns positive share of the firm then agent i owns,*

$$\lambda_i = \frac{D_i}{D_i + \gamma_i \sigma_q^2}. \tag{12}$$

From, Proposition 5.1 we see that if the outside investor owns positive shares of the firm, then her share increases and all the insiders' shares decrease as the variance of the output noise σ_q^2 increases. In the limit, the outside investor holds almost all the shares.

The behavior of the optimal ownership structure for $\sigma_q^2 \in [0, \bar{\sigma}_q^2]$ is more nuanced. For $\sigma_q^2 = 0$ not all agents might hold some shares of the firm.[22] Agents $\{m, \dots, n\}$ (in total, $n - m + 1$ of them) participate if condition

$$\frac{1}{D_k} < \frac{1}{n-k-1} \sum_{i=k+1}^{n} \frac{1}{D_i}$$

is satisfied for $k = m$ and is not satisfied for $k = m - 1$.[23]

As the output noise σ_q^2 increases, the left hand side of condition (8) decreases.[24] This implies that if agent i holds a positive amount of shares for a lower σ_q^2, she will also hold a positive amount of shares for a higher σ_q^2. Consider two agents i and m who have positive

[21] In an online appendix, we study comparative statics with respect to the productivity and risk aversion parameters of the agents (μ_i and γ_i) and noisiness of the performance signals ($\sigma_{s_i}^2$).

[22] Too see why both all and not all agents might hold shares of the firm, consider two examples. In the first example we have n identical agents. Then they all must hold $\frac{1}{n}$ shares of the firm. In the second example consider a firm with some of the agents being similar in their parameters to the outside investor (unproductive, large μ_i). Then it is easy to see that they will not hold any shares of the firm.

[23] Such m exists, because for $k = n - 1$ it is satisfied (so for no noise in the output there will be at least two owners of the firm), and for $k = 0$ it is not satisfied.

[24] $D_i \geq D_k$ for every $i > k$ and D_i does not depend on σ_q^2. As σ_q^2 appears only in the denominator, and all terms in the sum a non-negative, the whole expression in the left-hand side of condition (8) decreases.

amounts of shares, and differentiate (7) with respect to σ_q^2:

$$\frac{d\lambda_i}{d\sigma_q^2}(D_i + \gamma_i \sigma_q^2) + \gamma_i \sigma_q^2 = \frac{d\lambda_m}{d\sigma_q^2}(D_m + \gamma_m \sigma_q^2) + \gamma_m \sigma_q^2.$$

Since it cannot be the case that everyone's share holdings increase, for some agents $\frac{d\lambda_j}{d\sigma_q^2}$ is non-negative, and for some it is non-positive. If it is negative for a less risk averse agent (share holding decreases with σ_q^2), then share holding of a more risk averse agent must also decrease. So, as σ_q^2 increases, share holdings of agents who are more risk averse then a threshold decreases and less risk averse than a threshold increases.

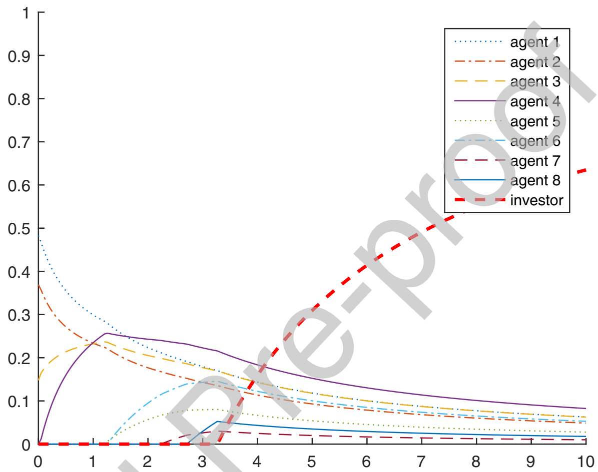

Figure 1: Changing output noisiness σ_q^2.

Figure 1 illustrates the situation with an outside investor. Notice that for low values of output noise the outside investor does not own shares of the firm, and the optimal allocation with or without the external investor is the same. However, as the output noise becomes greater, eventually the outside investor owns positive shares. After this point, if output noise increases further outside investor owns more and all other agents own fewer shares.

6 Discussion and Related Literature

In this section we discuss how our results on the optimal ownership structure compare to those in the literature.

In Proposition 3.2 we showed that agents with higher ownership parameter D_i (other things equal, these are productive risk averse agents whose performance signals are noisy) own shares of the company. We also demonstrated that a variety of organizational forms can be possible (see, for example, Proposition 4.1 or Figure 1), depending on the production structure of the company: partnerships with only few partners, cooperatives where all or a big share of employees own all shares, companies partially or almost completely owned by outside investors. These results are consistent with the findings in the literature.

Hansmann (1996) discusses various explanations of relative prevalence of employee ownership in the services industry. One critical factor is the difficulty of monitoring performance (which corresponds to the noisiness of performance signals in our paper). Hansmann specifically discusses law firms and is skeptical whether the ownership pattern in law firms is due to the difficulty of monitoring lawyers' performances. Hansmann argues that lawyers tend to provide detailed reports of their time use. On the other hand, Galanter and Palay (1990) argue that monitoring output in the provision of legal services to clients is difficult and costly. They point out that while law firms can measure the number of hours a lawyer puts in, it is more difficult to assess how many "quality" hours a lawyer has worked. Once again our theory is consistent with the pattern that is typically observed in law firms: senior partners whose contributions are more difficult to measure obtain shares, more junior lawyers whose contributions are easier to quantify obtain more performance based incentives, and outside investors who do not provide productive effort rarely obtain shares. Coram and Robinson (2017) interview nine participants, all partners from Big 4 firms and larger mid-tier firms in 2012 in Australia. Their study points out that "given the nature and scope of partner responsibilities in accounting firms, it is nearly impossible to accurately measure total effort or contribution on an individual basis." They find that these firms use a combination of profit-sharing schemes and performance incentives. Although they do not study the relative weight on profit sharing versus performance incentives as a function of difficulty of measurement, their study is consistent with the results in Proposition 3.2 – the agents with high ownership parameters D_i get both shares of the company and incentives, and the agents with high ownership parameters are those who perform difficult to measure tasks.

Testing theories of ownership directly is interesting but also quite difficult due to lack of data, number of factors affecting incentives and performance and related endogeneity problems. Nevertheless, empirical findings from two widely cited studies are consistent with our theory. Anderson and Schmittlein (1984) use survey data from the electronic components industry to look at what factors influence companies' decisions to use sales reps (who are given incentive contracts) as opposed to direct sales force (who are employed by the company). They ask sales managers whether measuring the results of individual sales people are difficult. They find that companies that report that performance measurement is difficult are more likely to use direct sales force suggesting that measurement plays a role in incentive design.[25] Lafontaine (1992) studies franchisors decisions' to use company owned versus franchised outlets. If we view managers of a franchise as productive agents and the franchisor as an outside investor, then in our theory independent franchises are motivated through shares and company owned outlets are motivated through incentive contracts. Lafontaine uses geographical dispersion as a measure of difficulty of performance measurement and, consistent with our theory, finds that this variable is correlated with more independent outlets. Literature also acknowledges that workers in managerial roles whose performances are harder to measure tend to receive more ownership based incentives. Although the aforementined studies are suggestive, we are not aware of a direct test of how measurability of performance in different tasks affect the composition of incentives in organizations. Hence, this seems to be an open question for future empirical research.

The paper that is most closely related to ours is Rayo (2007) who considers incentive provision in a team production setting. Rayo studies relational contracts, i.e. the contracts which are enforced not by a court but rather by mutual trust between the parties. The parties do not deviate from the specified payments because they can be excluded from the joint production in the future.[26] In Rayo (2007), with risk-neutral and infinitely patient agents, any relational contract which gives expected value greater than the outside option can be maintained with a grim trigger strategy (following the argument by Levin (2003), which is a special case of Rayo (2007) with one productive agent). In particular, the first best levels of effort can be implemented. In our model, agents are risk averse, but in

[25]Although suggestive, the study does not investigate the incentive contracts of direct sales force.

[26]Ishiguro and Yasuda (2021) study a static model with a principle and multiple agents and without explicit contracts. They show that when there are at least two risk neutral agents who can be interpreted as shareholders, second best outcome can be implemented. Their focus is very different because in their model there is no team production, and profit sharing does not motivate the agents.

the special case with risk neutrality, first best effort can also be implemented using solely noisy performance signals. In this sense, the outcome of our model with risk neutral agents coincides with Rayo's model when agents are infinitely patient. Differently from Rayo (2007), our focus is on the case with risk-averse agents where optimal contracts balance the amount of risk coming from two sources: shares of output and payment based on individual performance. We show that if the agents are collectively relatively risk averse, they cannot bear the risk resulting from shares and have to give some shares to the outsider.

The results on the concentration of shares in our paper and Rayo's work also have some similarities. Rayo (2007) assumes that the vector of signals is a sufficient statistic for output, conditional on the efforts of the agents. In our paper, signals and output are conditionally independent, so using shares has additional informational value. In Rayo (2007), when there is enough noise in the signals, only one person gets all shares of the company. In contrast, we show that all agents are shareholders in this situation. The other extreme is similar: if the signals are not noisy (observable effort), then output becomes a bad measure of performance (compared to the signals) and it is best to provide all agents with some shares, but give most shares to the outsider (unless we are at the exact limit with signal noise equal to zero, where the outsider receives all shares of the company). In Rayo (2007) there is more than one owner in this situation.

In the literature there are alternative theories about profit sharing based on various forms of adverse selection. Most of these theories focus on professional services such as investment banking or law firms. Levin and Tadelis (2005) study why profit-sharing partnerships are common in professional services. They find that it is optimal to use a partnership when clients are at a disadvantage in determining the average ability of the workers in the firm. They argue that this informational asymmetry is especially important in professional services relative to firms in other industries. Focusing on professional services Morrison and Wilhelm (2004) and Morrison and Wilhelm (2008) argue that partnerships foster the formation of human capital through mentoring and on-the-job training. Kandel and Lazear (1992) argue that profit sharing might increase motivation through peer pressure. Poblete (2015), in a career concerns framework, study agents' choice between working for firms with profit sharing and firms in which pay is based on individual productivity. Profit sharing makes it easier for agents to signal their productivity, but suffers from free riding. In equilibrium skilled agents are more likely to belong to profit sharing organizations. Garicano and Santos (2004) suggest profit sharing provides incentives to allocate client work efficiently within a diverse group of partners. Relative to these papers, our paper provides a complementary rationale for observing partnerships in professional services. We argue that it is the relative importance of output risk and ease of monitoring of the agents' effort levels that determines the organizational form. We predict that partnerships emerge in industries where the agents' performance metrics are noisy relative to the output.

Like us Heywood and Jirjahn (2009) consider the relationship between profit sharing and firm size. They cite many studies showing no significant relationship between firm size and profit sharing. They find this surprising because, with team production, larger firms would avoid profit sharing since they are subject to more free riding. These empirical results are consistent with our findings, since in our model larger firms are less likely to use profit sharing only if output risk grows faster than the number of agents (see Proposition 4.1). Otherwise, larger firms are able to self insure. This means that one needs to control for output risk when testing for the relationship between firm size and profit sharing.[27]

Weitzman (1980) studies the optimal cost splitting between a buyer and a contractor. In his model, the ambiguity is in the contractor's private benefit resulting from the project, and it is unknown to both parties at the time when the sign the contract. Weitzman shows that the buyer covers a greater part of the cost when the contractor is relatively more risk averse and when the contractor's ability to change the cost is more limited.

We focus on profit sharing and its impact on incentives and risk sharing. Often profit sharing involves ownership which also has implications on control rights and decision making in the firm. Starting with Williamson (1985) a large literature looks at whether a firm should vertically integrate certain functions or provide market based incentives. Williamson (1985) and Grossman and Hart (1986) argue that vertical integration is superior to market based incentives when there are relationship specific assets and ownership creates residual rights to the asset when it is prohibitively hard to specify all possible contingencies in the contract. Our theory provides a complementary view: without asset specificity, relative difficulty in performance measurement drives the distribution of ownership of shares within a firm. Hence, our theory is not only about inside ownership but also about how shares will be distributed within the firm.

[27] Li (2016) also studies profit sharing in a firm with many agents, but his focus is on information acquisition.

7 Conclusion

Performance contracts and profit sharing are often used jointly to incentivize productive agents and share risks both within the firm and with outside investors. Our paper provides a simple framework to study how profits should be shared among insiders and outsiders. As usual there is a tradeoff between risk sharing and incentives. When output is risky insiders would like to share risks with outside investors but this reduces the incentives of the insiders to exert costly effort. The firm can counter this by writing more powerful incentive contracts. In spite of this, we show that outsiders hold shares only if all insiders hold shares in the company although insiders' shares might be very small if output is very risky. Our paper provides several testable hypothesis. For example, we show that insiders in larger firms are more likely to share profit if the output risk is unchanged, but if the output risk grows too fast, larger firms are more likely to share their profits with the outside investors.

Appendix

A Equivalence of the profit and output-sharing problems

Consider a setting identical to the problem described in Section 2, but where we allocate shares of *output* rather than *profit*. Agent i holds a claim to share λ_i of output and has a contract paying $w_i(s)$ given by:

$$w_i(s) = \beta_i^0 + \sum_{j=1}^{n} s_j \beta_i^j.$$

Therefore, the agent's total payment equals

$$\lambda_i y + w_i(s). \tag{13}$$

Since the sum of payments to all agents must equal the total output, we obtain:

$$\sum_i w_i(s) = \sum_i \left(\beta_i^0 + \sum_{j=1}^{n} \beta_i^j s_j \right) = 0$$

for every realisation of signals s. This is satisfied if and only if $\sum_i \beta_i^j = 0$ for every $j = 0, \dots n$.[28]

Proposition A.1. *Consider two compensation schemes:*

- profit sharing*, where each agent i gets a share of profit λ_i and is additionally paid according to a contract $\tilde{w}_i(s) = \tilde{\beta}_i^j s_j + \tilde{\beta}_i^0$;*
- *and* output sharing*, where each agent i gets a share of output λ_i and is additionally paid according to a contract $w_i(s) = \beta_i^j s_j + \beta_i^0$.*

If $\tilde{\beta}$ and β are such that for every $i \in \{0, \dots, n\}$ and $j \in \{0, \dots, n\}$,

$$\beta_i^j \equiv \tilde{\beta}_i^j - \lambda_i \sum_{k \in I} \tilde{\beta}_k^j. \tag{14}$$

then each agent receives the same payment under the two payment schemes for any realization of output y and signals s.

Proof. To prove the proposition, we consider a realisation of signals $\{s_j\}_{j=1}^n$ and calculate how much agent i gets under the two compensation schemes:

$$\lambda_i y + w_i(s) = \lambda_i y + \beta_i^0 + \sum_{j=1}^{n} \beta_i^j s_j = \lambda_i y + \tilde{\beta}_i^0 - \lambda_i \sum_{k=1}^{n} \tilde{\beta}_k^0 + \sum_{j=1}^{n} \left(\tilde{\beta}_i^j s_j - \lambda_i \sum_{k=1}^{n} \tilde{\beta}_k^j \right) =$$
$$\lambda_i \left(y - \sum_{k=1}^{n} \left(\tilde{\beta}_k^0 + \sum_{j=1}^{n} \tilde{\beta}_k^j s_j \right) \right) + \tilde{\beta}_i^0 + \sum_{j=1}^{n} \tilde{\beta}_i^j s_j =$$
$$\lambda_i \left(y - \sum_{k=1}^{n} \tilde{w}_k(s) \right) + \tilde{w}_i(s) = \lambda_i \pi + \tilde{w}_i(s). \tag{15}$$

□

[28]To see that $\sum_i \beta_i^0 = 0$, consider the realisation of signals $s = (0, \dots, 0)$. Since $\sum_i \beta_i^0 = 0$, consider a realisation of signals where for some j, $s_j = 1$, and for all $k \neq j$, $s_k = 0$. For such combination of signals s, $\sum_i w_i(s) = \sum_i \left(\beta_i^0 + \sum_{j=1}^n \beta_i^j s_j \right) = \sum_i \beta_i^0 + \sum_i \beta_i^j = 0$, so $\sum_i \beta_i^j = 0$.

This equivalence shows that the *profit* and the *output* formulations are *payoff equivalent*. Since the payoffs are exactly the same, the incentives of the agents in the two settings are exactly the same.

Each vector of coefficients $(\tilde{\beta}_j^0, \dots, \tilde{\beta}_j^n)$ uniquely defines $(\beta_j^0, \dots, \beta_j^n)$:

$$\begin{pmatrix} \beta_0^j \\ \beta_1^j \\ \dots \\ \beta_n^j \end{pmatrix} = \begin{pmatrix} 1-\lambda_0 & -\lambda_0 & \dots & -\lambda_0 \\ -\lambda_1 & 1-\lambda_1 & \dots & -\lambda_1 \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_n & -\lambda_n & \dots & 1-\lambda_n \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} \equiv \Lambda \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix}. \tag{16}$$

Notice that

$$\sum_{i=0}^{n} \beta_i^j = 0. \tag{17}$$

Lemma A.1. *If $\sum_{j=0}^n \lambda_i = 1$, then the rank of the $(n+1) \times (n+1)$ matrix Λ is n. If $\sum_{j=0}^n \lambda_i \neq 1$, then the rank of matrix Λ is $n+1$.*

Proof. First, let us show that if $\sum_{i=0}^n \lambda_i = 1$ the rank of matrix Λ is not full (not $n+1$). Indeed, the sum of its rows is 0. Indeed, the sum of elements in each column equals

$$1 - \sum_{i=0}^{n} \lambda_i = 0.$$

Now, let us show that if we remove one row, then the remaining n rows are linearly independent. Removing the first row from matrix Λ leaves

$$\begin{pmatrix} -\lambda_1 & 1-\lambda_1 & -\lambda_1 & \dots & -\lambda_1 \\ -\lambda_2 & -\lambda_2 & 1-\lambda_2 & \dots & -\lambda_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\lambda_n & -\lambda_n & -\lambda_n & \dots & 1-\lambda_n \end{pmatrix}.$$

Denote r_i row i of this equation. If the rank of this matrix is n, then all rows are linearly independent. It means, that there does not exist a non-zero vector $(a_1, \dots, a_n)$, such that $\sum_{i=1}^n a_i r_i = 0$. Indeed, if such vector exists, then the sum of elements in the first column with weights a_i equals

$$\sum_{i=1}^{n} a_i(-\lambda_i) = 0.$$

The sum of elements in column j equals

$$a_j - \sum_{i=1}^{n} a_i(-\lambda_i) = 0.$$

But the two equations together imply that $a_j = 0$ for every j. So, rows $r_1, \dots r_n$ are linearly independent, and the rank of matrix Λ is n (one less than the full rank).

Similarly, if $\sum_{i=0}^{n} \lambda_i \neq 1$, let us show that any linear combination of its columns is non-zero. Denote the columns of matrix Λ as $c_0, c_1, \dots, c_n$. The columns are linearly dependent if and only if there exist real numbers $a_0, a_1, \dots, a_n$, such that $\sum_{i=0}^{n} a_i c_i = 0$. For element i of the columns it means that

$$a_i - \lambda_i \sum_{k=0}^{n} a_k = 0. \tag{18}$$

Let us sum equations (18) for i from 0 to n:

$$\sum_{i=0}^{n} a_i - \sum_{i=0}^{n} \lambda_i \sum_{k=0}^{n} a_k = \left(1 - \sum_{i=0}^{n} \lambda_i\right) \sum_{k=0}^{n} a_k = 0. \tag{19}$$

Since $\sum_{i=0}^{n} \lambda_i \neq 1$, equation (19) implies that $\sum_{k=0}^{n} a_k = 0$. Then, from equation (18) it follows that $a_i = 0$ for every i. Therefore, the columns of matrix Λ are independent, and it has full rank. $\square$

Even though $(\tilde{\beta}_j^0, \dots, \tilde{\beta}_j^n)$ uniquely defines $(\beta_j^0, \dots, \beta_j^n)$, $(\beta_j^0, \dots, \beta_j^n)$ does not uniquely define $(\tilde{\beta}_j^0, \dots, \tilde{\beta}_j^n)$, because the square matrix in equation (16) has rank n. On the other hand, the payments β_i^j must satisfy condition (17). Theorem A.2 shows how different profit-sharing contracts $\tilde{\beta}$ corresponding to the same output-sharing contract β relate to each other.

Theorem A.2. *Given a vector $(\beta_0^j, \dots, \beta_n^j)$ satisfying condition (17), the system of equations (16) has a solution and any pair of its solutions $(\tilde{\beta}_0^j, \dots, \tilde{\beta}_n^j)$ and $(\tilde{\beta}_0^{j*}, \dots, \tilde{\beta}_n^{j*})$ satisfy:*

$$\begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \dots \\ \tilde{\beta}_n^{j*} \end{pmatrix} + \begin{pmatrix} x\lambda_0 \\ x\lambda_1 \\ \dots \\ x\lambda_n \end{pmatrix} \tag{20}$$

for some real number x. Also, if $(\tilde{\beta}_0^{j*}, \ldots, \tilde{\beta}_n^{j*})$ is a solution to problem (16)*, then any $(\tilde{\beta}_0^{j}, \ldots, \tilde{\beta}_n^{j})$ is also a solution to problem* (16) *for any real x.*

Proof. **Existence**

Since $\sum_{i=0}^{n} \lambda_i = 1$, there is an agent with a strictly positive share λ_i. Without loss of generality, let it be agent n. We will show that there is a solution, where $\tilde{\beta}_n^j = 0$. Since $\tilde{\beta}_n^j = 0$, we can remove the last column of Λ and the equation (16) remains correct

$$\begin{pmatrix} 1-\lambda_0 & -\lambda_0 & \ldots & -\lambda_0 \\ -\lambda_1 & 1-\lambda_1 & \ldots & -\lambda_1 \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_{n-1} & -\lambda_{n-1} & \ldots & 1-\lambda_{n-1} \\ -\lambda_n & -\lambda_n & \ldots & -\lambda_n \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \ldots \\ \tilde{\beta}_{n-1}^j \end{pmatrix} = \begin{pmatrix} \beta_0^j \\ \beta_1^j \\ \ldots \\ \beta_{n-1}^j \\ \beta_n^j \end{pmatrix}.$$

Now, remove the last equation from this system of equation:

$$\begin{pmatrix} 1-\lambda_0 & -\lambda_0 & \ldots & -\lambda_0 \\ -\lambda_1 & 1-\lambda_1 & \ldots & -\lambda_1 \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_{n-1} & -\lambda_{n-1} & \ldots & 1-\lambda_{n-1} \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \ldots \\ \tilde{\beta}_{n-1}^j \end{pmatrix} = \begin{pmatrix} \beta_0^j \\ \beta_1^j \\ \ldots \\ \beta_{n-1}^j \end{pmatrix}. \tag{21}$$

By Lemma A.1 the $n \times n$ matrix in this equation has full rank, and there exists a unique solution

$$\begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \ldots \\ \tilde{\beta}_{n-1}^{j*} \end{pmatrix}$$

of system (21). Let us show that then

$$\begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \ldots \\ \tilde{\beta}_{n-1}^j \\ \tilde{\beta}_n^j \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \ldots \\ \tilde{\beta}_{n-1}^{j*} \\ 0 \end{pmatrix}$$

is a solution of equation (16). We already know that it is true for the first n equations of this system. Now, let us check if

$$\beta_n^j = \begin{pmatrix} -\lambda_n & -\lambda_n & \dots & 1-\lambda_n \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \dots \\ \tilde{\beta}_{n-1}^{j*} \\ 0 \end{pmatrix}. \tag{22}$$

Indeed, adding all the rows in equation (21), we obtain

$$\left(1 - \sum_{i=0}^{n-1} \lambda_i\right) \sum_{k=0}^{n-1} \tilde{\beta}_i^{j*} = \sum_{k=0}^{n-1} \beta_i^j. \tag{23}$$

Since $\sum_{i=0}^{n} \beta_i^j = 0$, $\beta_n^j = -\sum_{i=0}^{n-1} \beta_i^j$, and equation (23) can be rewritten as:

$$\lambda_n \sum_{k=0}^{n-1} \tilde{\beta}_i^{j*} = \sum_{k=0}^{n-1} \beta_i^j = -\beta_n^j.$$

The last equation is equivalent to equation (22).

Relation between solutions

Take any two solutions $\tilde{\beta}^j$ and $\tilde{\beta}^{j*}$, which are solutions of equation (16), so that

$$\beta^j = \Lambda \tilde{\beta}^j = \Lambda \tilde{\beta}^{j*}.$$

Then

$$\Lambda(\tilde{\beta}^j - \tilde{\beta}^{j*}) = 0.$$

The statement of the theorem is therefore equivalent to saying that the only solutions of the system

$$\Lambda z = 0 \tag{24}$$

are vectors

$$z = \begin{pmatrix} z_0 \\ z_1 \\ \dots \\ z_n \end{pmatrix} = \begin{pmatrix} x\lambda_0 \\ x\lambda_1 \\ \dots \\ x\lambda_n \end{pmatrix}. \tag{25}$$

Equation i in system $\Lambda z = 0$ states that

$$(1-\lambda_i)z_i - \lambda_i \sum_{k\neq i} z_k = z_i - \lambda_i \sum_{k=0}^{n} z_k = 0,$$

so, if $\sum_{k=0}^{n} z_k = 0$, then $z_i = 0$ for every i. Also, if $\lambda_i = 0$, then $z_i = 0$. Otherwise, for all i and k, such that $\lambda_i, \lambda_k \neq 0$:

$$\frac{z_i}{\lambda_i} = \frac{z_k}{\lambda_k} \equiv x. \tag{26}$$

Then all solutions z of system (24) satisfy condition (25).

It is straightforward to check that any such solution satisfies condition (25) also solves system (24), so if $\tilde{\beta}^{j*}$ is a solution of system (16) then the set of solutions of system (16) is given by

$$\left\{ \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} \middle| \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \dots \\ \tilde{\beta}_n^{j*} \end{pmatrix} + \begin{pmatrix} x\lambda_0 \\ x\lambda_1 \\ \dots \\ x\lambda_n \end{pmatrix}, x \in \mathbb{R} \right\}. \tag{27}$$

□

Indeed, adding a payment proportional to the allocation of shares to all agents leads to the exact same payment to each agent in every state of the world, so it does not change the incentives or utilities of any of the agents.

It means that the firm where *profit* is split according to shares λ_i (agent i gets share of the profit equal to $\lambda_i \pi$), and the contracts are given by a combination of parameters $\{\tilde{\beta}_i^j\}_{i,j\in\{0,\dots,n\}}$ is equivalent to a firm where *output* is split according to shares λ_i (agent i gets share of the profit equal to $\lambda_i y$) and the contracts are given by a combination of parameters $\{\beta_i^j\}_{i,j\in\{0,\dots,n\}}$ (satisfying the budget constraint (17)) defined by a system of equations (16) for each $j \in \{0,\dots,n\}$.

In the remaining sections we speak only in terms of the *output-sharing setting*. The reason is that the solution in this setting is unique, while there is a number of payoff-equivalent solutions in the *profit-sharing setting*. However, Proposition A.1 proves that all these solutions are also payoff equivalent to the unique solution in the *output-sharing setting*, and the optimal shares are the same in the unique optimal allocation in the *output-sharing setting* and all optimal allocations in the *profit-sharing setting*.

B Binary model

In this section we develop a model which is slightly different from the model in the main part of the paper. We show that one of the main results from Proposition 3.3 that if the outsider holds shares of the firm, then every insider holds shares of the firm is still valid in a different setting.

The model is more general in that it allows for any concave utility functions, however here we assume that output and all signals have only binary realisations.

B.1 Setting

Consider a universe with $n+1$ agents indexed $0, 1, \ldots, n$.

Agent 0 represents all non-productive people in the society. Agent 0 cannot exert effort, but can own shares of the firm. We assume that she is risk neutral, so that her utility from getting amount of money x is normalized to $u_0(x) = x$.

Agent i (for $i \geq 1$) exerts effort e_i at cost $c_i(e_i)$ ($c_i'(e_i) > 0$ and $c_i''(e_i) < 0$ for every e_i). The agent's effort increases the probability of high output, that is we assume that the output can have two realisations: high=1 and low=0. The probability of high output is:

$$\mathbb{P}(y = 1|e) = \min\left\{\sum_j e_j, 1\right\}.$$

We assume that the functions and other parameters are such that it is never optimal to exert level of efforts e_i with $\sum_{i=1}^n e_i \geq 1$,[29] so that the probability above can be written as just e.

In addition, there is a signal of effort of agent 1:[30]

$$s_i = \begin{cases} 0, & \text{with probability } 1 - t_i e_i; \\ 1, & \text{with probability } t_i e_i. \end{cases}$$

λ_i denotes agent i's share of the company (claim of output), and the rest of the

[29] For example, if $c_i(\cdot)$ is such that $c_i(e_i) \to \infty$ when $e_i \to \frac{1}{n}$.

[30] Again, we choose parameters such that $t_i e_i$ is always in $[0, 1]$, similar to the comment in the previous footnote.

company (share $\lambda_0 = 1 - \sum_{j\neq i} \lambda_j$) belongs to an unproductive risk neutral agent 0. All agents signs a contract based on signal s_i with agent 0 (or he signs it with the firm, but the firm itself buys an insurance contract from agent 0), who pays agent i wage[31]

$$w_i = \begin{cases} \bar{w}_i, & \text{if } s_i = 1; \\ \underline{w}_i, & \text{if } s_i = 0. \end{cases}$$

We assume that no agent can get a negative share of output, so $\lambda_i \geq 0$ for $i \in \{0, \ldots, n\}$.

Let the utility function of agent i be $u_i(x_i) - c_i(e_i)$, where x_i is the amount of money he gets and u_i with $u_i' > 0$ and $u_i'' < 0$ (for example, $1 - e^{-x_i}$).

B.2 Results

Proposition B.1. *The optimal distribution of shares is such that the outsider (agent 0) owns shares only if all productive agents also own some shares of the firm.*

Proof. Every productive agent i maximizes (assuming that equilibrium efforts are such that $\sum_{i=1}^n e_i < 1$):

$$\max_{e_i} t_i e_i \left(\sum_{j=1}^n e_j\right) u_i(\lambda_i + \bar{w}_i) + (1 - t_i e_i) \left(\sum_{j=1}^n e_j\right) u_i(\lambda_i + \underline{w}_i) + \\ t_i e_i \left(1 - \sum_{j=1}^n e_j\right) u_i(\bar{w}_i) + (1 - t_i e_i) \left(1 - \sum_{j=1}^n e_j\right) u_i(\underline{w}_i) - c_i(e_i) \quad (28)$$

[31] We assume that the wages are paid by agent 0 for brevity here. In principle, we can think of this as the firm paying wages to workers and then buying an insurance contract from agent 0. It is better for agent 0 to carry the risk of the wage associated with signal s_i, because other agents (not 0 or i) cannot affect signal s_i, so it does not have a positive effect on their incentives, but they prefer a certain outcome over a lottery based on random variable s_i, so it is better if agent 0 bears this risk.

FOC:

$$V_i \equiv \frac{dU_i}{de_i} = t_i\left(e_i + \sum_{j=1}^{n} e_j\right)u_i(\lambda_i + \bar{w}_i) + \left(1 - t_i e_i - t_i\sum_{j=1}^{n} e_j\right)u_i(\lambda_i + \underline{w}_i) +$$
$$t_i\left(1 - e_i - \sum_{j=1}^{n} e_j\right)u_i(\bar{w}_i) - \left(t_i\left(1 - \sum_{j=1}^{n} e_j\right) + 1 - t_i e_i\right)u_i(\underline{w}_i) - c_i'(e_i) = 0. \quad (29)$$

The second derivative is:

$$2t_i\left(u_i(\lambda_i + \bar{w}_i) - u_i(\lambda_i + \underline{w}_i) - u_i(\bar{w}_i) + u_i(\underline{w}_i)\right) - c_i''(e_i) < 0,$$

so there is a unique maximum. It is also greater than 0 (if $\sum_{j\neq i} e_j < 1$), because the LHS of equation (29) is positive when $e_i = 0$:

$$t_i\sum_{j\neq i} e_j u_i(\lambda_i + \bar{w}_i) + \left(1 - t_i\sum_{j\neq i} e_j\right)u_i(\lambda_i + \underline{w}_i) +$$
$$t_i\left(1 - \sum_{j\neq 0,i} e_j\right)u_i(\bar{w}_i) - \left(t_i\left(1 - \sum_{j\neq 0,i} e_j\right) + 1\right)u_i(\underline{w}_i) - c_i'(e_i) =$$
$$t_i\left(1 - \sum_{j\neq i} e_j\right)\left(-u_i(\lambda_i + \bar{w}_i) + u_i(\lambda_i + \underline{w}_i) + u_i(\bar{w}_i) - u_i(\underline{w}_i)\right) +$$
$$t_i(u_i(\lambda_i + \bar{w}_i) - u_i(\lambda_i + \underline{w}_i)) + (u_i(\lambda_i + \underline{w}_i) - u_i(\underline{w}_i)). \quad (30)$$

The first term above is non-negative because function u_i is concave and $\sum_{j\neq i} e_j < 1$, the second term is positive if $\bar{w}_i > \underline{w}_i$ and the third term is positive if $\lambda_i > 0$, because u_i is an increasing function, and we assumed that $c_i'(e_i) = 0$. So, the expression above is non-negative and only equals zero if $\lambda_i = 0$ and $\bar{w}_i = \underline{w}_i$. So, as long as agent i is motivated with either shares ($\lambda_i > 0$) or a wage bonus ($\bar{w}_i > \underline{w}_i$) agent to exert positive effort. If he is not motivated with either of them, he puts in effort 0.

All Pareto optimal compensation schemes solve the following maximization problem:

$$\max_{\lambda_0,\{\underline{w}_i,\bar{w}_i,\lambda_i\}_{i=1}^{n}} \sum_{i=0}^{n} \nu_i U_i, \text{s.t. } \forall i: \lambda_i \geq 0, \sum_{i=0}^{n} \lambda_i \leq 1 \quad (31)$$

Here the utility of all productive agents i, $i \geq 1$ is given by the expression in problem (28) and utility of agent 0 is:

$$U_0 \equiv \lambda_0 \sum_{j=1}^{n} e_j - \sum_{j=1}^{n} \left((1 - t_j e_j) \underline{w}_j + t_j e_j \bar{w}_j \right).$$

Let us show that it cannot be optimal for the outsider to have shares of the company if at least one of the insiders has no shares of the company. In other words $\lambda_0 > 0$ and $\lambda_i = 0$ for some i cannot be the solution of the optimization problem (31).

The Lagrangian of this problem is:

$$\mathcal{L} \equiv \sum_{i=0}^{n} \nu_i U_i + \sum_{i=0}^{n} \theta_i \lambda_i + \chi \left(1 - \sum_{i=0}^{n} \lambda_i \right).$$

We normalize $\nu_0 = 1$.

In order to proceed with first order conditions, we introduce additional notation to make the following derivations more concise.

Denote ϵ_i an n-dimensional column-vector, such that component i is equal to one, and all other components are zero.

Notice that conditions $V_i = 0$ from equation (29) determine effort levels e_j. Thinking of the vector of V_is as a vector-valued function, we can rewrite the conditions, determining the link between different derivatives

$$\frac{dV_i}{dx} = \frac{\partial V_i}{\partial x} + \sum_{j=1}^{n} \frac{\partial V_i}{\partial e_j} \frac{de_j}{dx} = 0, \tag{32}$$

where x is one of the parameters (for example, λ_i or $\bar{w}_k$), using the Jacobian matrix

$$J \equiv \begin{pmatrix} \frac{\partial V_1}{\partial e_1} & \frac{\partial V_1}{\partial e_2} & \cdots & \frac{\partial V_1}{\partial e_n} \\ \frac{\partial V_2}{\partial e_1} & \frac{\partial V_2}{\partial e_2} & \cdots & \frac{\partial V_2}{\partial e_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial e_1} & \frac{\partial V_n}{\partial e_2} & \cdots & \frac{\partial V_n}{\partial e_n} \end{pmatrix}; J \cdot \begin{pmatrix} \frac{de_1}{dx} \\ \frac{de_2}{dx} \\ \vdots \\ \frac{de_n}{dx} \end{pmatrix} = \begin{pmatrix} -\frac{\partial V_1}{\partial x} \\ -\frac{\partial V_2}{\partial x} \\ \vdots \\ -\frac{\partial V_n}{\partial x} \end{pmatrix} \equiv -\frac{\partial V}{\partial x}.$$

In lemma B.2 we show that matrix J is invertible.

Since matrix J is invertible, $\frac{de_i}{dx} = -\epsilon_i^T J^{-1} \frac{\partial V}{\partial x}$. If variable x has index j (like $x = \bar{w}_j$), then first order conditions are such that V_k does not depend on x for $k \neq j$, so $\frac{\partial V_k}{\partial x} = 0$. Thus, $\frac{de_i}{dx} = -\epsilon_i^T J^{-1} \epsilon_j \frac{\partial V_j}{\partial x}$.

Finally, we need to calculate derivatives (for $j \neq k$):

$$\frac{\partial U_j}{\partial e_k} = t_j e_j (u_j(\lambda_j + \bar{w}_j) - u_j(\bar{w}_j)) + (1 - t_j e_j)(u_j(\lambda_j + \underline{w}_j) - u_j(\underline{w}_j)) \equiv S_j.$$

Remember that $\frac{\partial U_j}{\partial e_j} = 0$, because e_j is defined as the solution of optimization problem (28). Also, $S_j > 0$ if $\lambda_j > 0$ and $S_j = 0$ if $\lambda_j = 0$.

First order conditions:

$$\begin{aligned}\frac{d\mathcal{L}}{d\bar{w}_i} = \sum_{j=0}^{n} \nu_j \left(\frac{\partial U_j}{\partial \bar{w}_i} + \sum_{k \neq 0,j} \frac{\partial U_j}{\partial e_k} \frac{de_k}{d\bar{w}_i} \right) = \nu_i \frac{\partial U_i}{\partial \bar{w}_i} + \frac{\partial U_0}{\partial \bar{w}_i} + \sum_{j=0}^{n} \nu_j \sum_{k \neq 0,j} \frac{\partial U_j}{\partial e_k} \frac{de_k}{d\bar{w}_i} = \\ \nu_i t_i e_i \left(\sum_{j=1}^{n} e_j u_i'(\lambda_i + \bar{w}_i) + \left(1 - \sum_{j=1}^{n} e_j\right) u_i'(\bar{w}_i) \right) - t_i e_i + \\ + \sum_{k=1}^{n} \frac{de_k}{d\bar{w}_i} \left((\lambda_0 - t_k(\bar{w}_k - \underline{w}_k)) + \sum_{j \neq 0,k} \nu_j S_j \right) = 0. \quad (33)\end{aligned}$$

$$\begin{aligned}\frac{d\mathcal{L}}{d\underline{w}_i} = \sum_{j=0}^{n} \nu_j \left(\frac{\partial U_j}{\partial \underline{w}_i} + \sum_{k \neq 0,j} \frac{\partial U_j}{\partial e_k} \frac{de_k}{d\underline{w}_i} \right) = \nu_i \frac{\partial U_i}{\partial \underline{w}_i} + \frac{\partial U_0}{\partial \underline{w}_i} + \sum_{j=0}^{n} \nu_j \sum_{k \neq 0,j} \frac{\partial U_j}{\partial e_k} \frac{de_k}{d\underline{w}_i} = \\ \nu_i (1 - t_i e_i) \left(\sum_{j=1}^{n} e_j u_i'(\lambda_i + \underline{w}_i) + \left(1 - \sum_{j=1}^{n} e_j\right) u'(\underline{w}_i) \right) - (1 - t_i e_i) + \\ + \sum_{k=1}^{n} \frac{de_k}{d\underline{w}_i} \left((\lambda_0 - t_k(\bar{w}_k - \underline{w}_k)) + \sum_{j \neq 0,k} \nu_j S_j \right) = 0. \quad (34)\end{aligned}$$

$$\left.\frac{d\mathcal{L}}{d\lambda_i}\right|_{i\neq 0} = \theta_i - \chi + \sum_{j=0}^{n} \nu_j \left(\frac{\partial U_j}{\partial \lambda_i} + \sum_{k\neq 0,j} \frac{\partial U_j}{\partial e_k} \frac{de_k}{d\lambda_i} \right) =$$

$$\theta_i - \chi + \nu_i \frac{\partial U_i}{\partial \lambda_i} + \frac{\partial U_0}{\partial \lambda_i} + \sum_{j=0}^{n} \nu_j \sum_{k\neq 0,j} \frac{\partial U_j}{\partial e_k} \frac{de_k}{d\lambda_i} =$$

$$\theta_i - \chi + \nu_i \left(\sum_{j=1}^{n} e_j \right) (t_i e_i u_i'(\lambda_i + \bar{w}_i) + (1 - t_i e_i) u'(\lambda_i + \underline{w}_i)) +$$

$$+ \sum_{k=1}^{n} \frac{de_k}{d\lambda_i} \left((\lambda_0 - t_k(\bar{w}_k - \underline{w}_k)) + \sum_{j\neq 0,k} \nu_j S_j \right) = 0. \quad (35)$$

$$\frac{d\mathcal{L}}{d\lambda_0} = \sum_{j=1}^{n} e_j + \theta_0 - \chi = 0. \quad (36)$$

Denote

$$R_k \equiv (\lambda_0 - t_k(\bar{w}_k - \underline{w}_k)) + \sum_{j\neq 0,k} \nu_j S_j. \quad (37)$$

Lemma B.1. *If there is an agent i, $i \geq 1$ with no shares ($\lambda_i = 0$), then the outsider has no shares as well ($\lambda_0 = 0$).*

Proof. If $\lambda_i = 0$, then condition (33) can be simplified into:

$$\nu_i t_i e_i u_i'(\bar{w}_i) - t_i e_i + \sum_{k=1}^{n} \frac{de_k}{d\bar{w}_i} R_k = 0. \quad (38)$$

In turn, since $\frac{de_j}{d\bar{w}_i} = -\epsilon_j^T J^{-1} \epsilon_i \frac{\partial V_i}{\partial \bar{w}_i}$, this condition can be written as:

$$\nu_i t_i e_i u_i'(\bar{w}_i) - t_i e_i - \frac{\partial V_i}{\partial \bar{w}_i} \sum_{k=1}^{n} \epsilon_k^T J^{-1} \epsilon_i R_k = 0. \quad (39)$$

Since $\lambda_i = 0$:

$$\frac{\partial V_i}{\partial \bar{w}_i} = t_i \left(e_i + \sum_{j=1}^{n} e_j \right) u_i'(\lambda_i + \bar{w}_i) + t_i \left(1 - e_i - \sum_{j=1}^{n} e_j \right) u_i'(\bar{w}_i) \underset{\lambda=0}{=} t_i u_i'(\bar{w}_i).$$

Analogously,

$$\frac{\partial V_i}{\partial \underline{w}_i} \underset{\lambda=0}{=} -t_i u_i'(\underline{w}_i),$$

$$\frac{\partial V_i}{\partial \lambda_i} = t_i \left(e_i + \sum_{j=1}^{n} e_j \right) u_i'(\lambda_i + \bar{w}_i) + \left(1 - t_i e_i - t_i \sum_{j=1}^{n} e_j \right) u_i'(\lambda_i + \underline{w}_i) \underset{\lambda=0}{=}$$
$$t_i \left(e_i + \sum_{j=1} e_j \right) u_i'(\bar{w}_i) + \left(1 - t_i e_i - t_i \sum_{j=1}^{n} e_j \right) u_i'(\underline{w}_i).$$

First order conditions (33), (34) and (35) can be written as:

$$\nu_i t_i e_i u_i'(\bar{w}_i) - t_i e_i - t_i u'(\bar{w}_i) \sum_{k=1}^{n} \epsilon_k^T J^{-1} \epsilon_i R_k = 0, \tag{40}$$

$$\nu_i (1 - t_i e_i) u_i'(\underline{w}_i) - (1 - t_i e_i) + t_i u'(\underline{w}_i) \sum_{k=1}^{n} \epsilon_k^T J^{-1} \epsilon_i R_k = 0, \tag{41}$$

and

$$\theta_i - \theta_0 - \sum_{j=1}^{n} e_j + \nu_i \left(\sum_{j=1}^{n} e_j \right) (t_i e_i u_i'(\bar{w}_i) + (1 - t_i e_i) u'(\underline{w}_i)) -$$
$$\left(t_i \left(e_i + \sum_{j} e_{j=1}^{n} \right) u_i'(\bar{w}_i) + \left(1 - t_i e_i - t_i \sum_{j=1}^{n} e_j \right) u_i'(\underline{w}_i) \right) \sum_{k=1}^{n} \epsilon_k^T J^{-1} \epsilon_i R_k = 0. \tag{42}$$

Subtract equation (40)$\times \left(\sum_{j=1}^{n} e_j \right)$ and equation (41)$\times \left(\sum_{j=1}^{n} e_j \right)$ from equation (42) to get:

$$\theta_i - \theta_0 - (t_i e_i u'(\bar{w}_i) + (1 - t_i e_i) u_i'(\underline{w}_i)) \sum_{k} \epsilon_k^T J^{-1} \epsilon_i R_k = 0. \tag{43}$$

In lemma B.3 we show that

$$\sum_{k} \epsilon_k^T J^{-1} \epsilon_i R_k < 0,$$

so equation (43) implies that $\theta_i < \theta_0$. If $\lambda_0 > 0$, then $\theta_0 = 0$. But this is impossible, because $\theta_i \geq 0$ for all i. □

So, agent 0 has shares only if all other agents have shares. □

Lemma B.2. *Matrix J is invertible.*

Proof. First, we calculate the elements of matrix J:

$$\frac{\partial V_i}{\partial e_i} = 2t_i\left(u_i(\lambda_i + \bar{w}_i) - u_i(\lambda_i + \underline{w}_i) - u_i(\bar{w}_i) + u_i(\underline{w}_i)\right) - c_i''(e_i).$$

$$\left.\frac{\partial V_i}{\partial e_j}\right|_{i \neq j} = t_i\left(u_i(\lambda_i + \bar{w}_i) - u_i(\lambda_i + \underline{w}_i) - u_i(\bar{w}_i) + u_i(\underline{w}_i)\right).$$

Denote the expression from above as T_i. Notice that $T_i \leq 0$ and $T_i = 0$ if and only if either $\lambda_i = 0$ or $\bar{w}_i = \underline{w}_i$.

$$J = \begin{pmatrix} 2T_1 - c_1''(e_1) & T_1 & \dots & T_1 \\ T_2 & 2T_2 - c_2''(e_2) & \dots & T_2 \\ \vdots & \vdots & \ddots & \vdots \\ T_n & T_n & \dots & 2T_n - c_n''(e_n) \end{pmatrix}$$

J has full rank if and only if its columns (denote them $J_1, \dots, J_n$) are linearly independent:

$$\sum_{i=1}^{n} b_i \cdot J_i = 0 \Leftrightarrow \forall i : b_i = 0.$$

Condition $\sum_{i=1}^{n} b_i \cdot J_i = 0$ implies

$$T_i \sum_{j=1}^{n} b_j = b_i(c_i''(e_i) - T_i).$$

If $T_i = 0$, then this condition cannot be satisfied, unless $b_i = 0$, because $c_i''(e_i) > 0$. If $T_i < 0$ (notice that T_i cannot be positive), then either $b_i = \sum_{j=1}^{n} b_j = 0$ or b_i and $\sum_{j=1}^{n} b_j$ have different signs. But all b_i and $\sum_j b_j$ cannot have different signs (if the sum is positive, all b_i cannot be negative, and if the sum is negative, all b_i cannot be positive), so $b_i = 0$ for all i. Thus, matrix J has the full rank. $\square$

Lemma B.3.

$$\sum_k \epsilon_k^T J^{-1} \epsilon_i R_k < 0,$$

Proof. Notice that $\sum_{k=1}^{n} \epsilon_k^T J^{-1} \epsilon_i R_k \leq 0$. Indeed, subtract equation (41)$\times t_i e_i$ from equation (40)$\times(1 - t_i e_i)$:

$$\nu_i t_i e_i (1 - t_i e_i)\left(u_i'(\bar{w}_i) - u_i'(\underline{w}_i)\right) = t_i\left((1 - t_i e_i) u_i'(\bar{w}_i) + t_i e_i u_i'(\underline{w}_i)\right) \sum_k \epsilon_k^T J^{-1} \epsilon_i R_k. \quad (44)$$

If $\bar{w}_i > \underline{w}_i$, then the left hand side of this equation is negative ($t_i e_i \in (0,1)$ and $u'(\bar{w}_i) < u'(\underline{w}_i)$, and $t_i\left((1-t_ie_i)u'(\bar{w}_i) + t_ie_iu'(\underline{w}_i)\right) \geq 0$, so

$$\sum_k \epsilon_k^T J^{-1}\epsilon_i R_k < 0.$$

Let us show that it cannot be the case that $\bar{w}_i = \underline{w}_i$ (so that the agent is not motivated not only with shares but also with a wage bonus).

If $\bar{w}_i = \underline{w}_i$ then equation (44) implies that

$$\sum_k \epsilon_k^T J^{-1}\epsilon_i R_k = 0.$$

Denote

$$\begin{pmatrix} X_1 & X_2 & \dots & X_N \end{pmatrix} \equiv \begin{pmatrix} R_1 & R_2 & \dots & R_N \end{pmatrix} \cdot J^{-1}.$$

Then $X_i = \sum_k \epsilon_k^T J^{-1}\epsilon_i R_k$ and

$$\begin{pmatrix} X_1 & X_2 & \dots & X_N \end{pmatrix} \cdot J = \begin{pmatrix} R_1 & R_2 & \dots & R_N \end{pmatrix}$$

Hence,

$$R_j - R_i = \begin{pmatrix} X_1 & X_2 & \dots & X_N \end{pmatrix} \cdot \begin{pmatrix} T_1 - T_1 \\ \vdots \\ T_i - 2T_i - c_i''(e_i) \\ \vdots \\ 2T_j - c_j''(e_j) - T_j \\ \vdots \\ T_n - T_n \end{pmatrix} = \\ - X_i(T_i - c_i''(e_i)) + X_j(T_j - c_j''(e_j)). \quad (45)$$

Using the definition of R in equation (37):

$$R_j - R_i = -t_j(\bar{w}_j - \underline{w}_j) + t_i(\bar{w}_i - \underline{w}_i) + \nu_i S_i - \nu_j S_j.$$

If $X_i = 0$, then $\bar{w} = \underline{w}_i$ and therefore $e_i = 0$ should hold, and also $S_i = 0$ since $\lambda_i = 0$.

Then the equation above is:

$$R_j - R_i = -t_j(\bar{w}_j - \underline{w}_j) - \nu_j S_j.$$

There is an agent with positive shares $\lambda_j > 0$ or $\bar{w}_j - \bar{w}_i > 0$, then $R_j - R_i < 0$. If $X_i = 0$, then equation (45) becomes:

$$0 > R_j - R_i = X_j(T_j - c_j''(e_j)).$$

As $T_j - c_j''(e_j) < 0$, $X_j > 0$. So, for all agents who have either shares or bonuses $X_j > 0$ and for the ones who have neither shares, nor bonuses $X_i = 0$ (implied by equation (44)). But

$$R_i = \begin{pmatrix} X_1 & X_2 & \dots & X_N \end{pmatrix} \cdot \begin{pmatrix} T_1 \\ \vdots \\ 2T_i - c_i''(e_i) \\ \vdots \\ T_n \end{pmatrix} =$$

$$X_i(T_i - c_i''(e_i)) + \sum_k X_j T_k = \sum_k X_k T_k < 0. \quad (46)$$

The last inequality is true because $X_i = 0$, $X_j > 0$ and $T_j < 0$, and for all k: $X_k \geq 0$ and $T_k \leq 0$. But

$$R_i = \lambda_0 - t_i(\bar{w}_i - \underline{w}_i) + \sum_{j \neq i} \nu_j S_j = \lambda_0 + \sum_{j \neq i} \nu_j S_j \geq 0. \quad (47)$$

Equations (46) and (47) contradict each other, therefore, the assumption that $X_i = 0$ is wrong, so $X_i < 0$.

□

C Proofs

C.1 Proof of Proposition 3.1

Proof. Consider the optimization problem in (3). Denote the Lagrange multipliers for $\lambda_i \geq 0$ by θ_i, for $\sum \lambda_i = 1$ by χ, and for $\sum_i \beta_i^j = 0$ by η_j. Then the Lagrangian is:[32]

$$\mathcal{L} = \sum_{i \in I} \left[\lambda_i \left(\sum_{j=1}^{n} \frac{\lambda_j + \beta_j^j}{\mu_j} \right) - \frac{\mu_i}{2} \left(\frac{\lambda_i + \beta_i^i}{\mu_i} \right)^2 - \frac{\gamma_i}{2} \left((\lambda_i)^2 \sigma_q^2 + \sum_{j \in I} \left(\beta_i^j\right)^2 \sigma_{s_j}^2 \right) \right] + \sum_{i \in I} \lambda_i \theta_i + \chi \left(1 - \sum_{i \in I} \lambda_i \right) - \sum_{j=1}^{n} \left(\eta_j \sum_{i \in I} \beta_i^j \right). \quad (48)$$

Notice that the objective function is concave in the variables β_i^j. Indeed, all cross derivatives $\frac{\partial^2 \mathcal{L}}{\partial \beta_i^j \partial \beta_k^l}$ are equal to zero and all second derivatives $\frac{\partial^2 \mathcal{L}}{\partial \left(\beta_i^j\right)^2}$ are negative. So first order conditions define the unique maximum of this problem. Differentiating the Lagrangian from expression (48) with respect to choice variables β_i^j yields first order conditions (equation (49) is for $j \neq i$):

$$\frac{\partial \mathcal{L}}{\partial \beta_i^j} = -\beta_i^j \gamma_i \sigma_{s_j}^2 - \eta_j = 0. \quad (49)$$

$$\frac{\partial \mathcal{L}}{\partial \beta_j^j} = \sum_{i=1}^{n} \frac{\lambda_i}{\mu_j} - \frac{\lambda_j + \beta_j^j}{\mu_j} - \beta_j^j \gamma_j \sigma_{s_j}^2 - \eta_j = \sum_{i \neq j} \frac{\lambda_i}{\mu_j} - \frac{\beta_j^j}{\mu_j} - \beta_j^j \gamma_j \sigma_{s_j}^2 - \eta_j = 0. \quad (50)$$

Comparing equation (49) for pairs (i, j) and (k, j), we obtain:

$$\frac{\beta_i^j}{\beta_k^j} = \frac{\gamma_k}{\gamma_i}.$$

Equating η_j from equations (49) and (50) yields:

$$\frac{\beta_j^j}{\mu_j} - \beta_i^j \gamma_i \sigma_{s_j}^2 = \frac{1}{\mu_j} \sum_{k \neq j}^{n} \lambda_k - \beta_j^j \gamma_j \sigma_{s_j}^2.$$

[32] In the case with the outside investor the first term under the sum $\sum_{i \in I}$ for $i = 0$ is simply $\lambda_0 \left(\sum_{j=1}^{n} \frac{\lambda_j + \beta_j^j}{\mu_j} \right)$.

Thus,

$$\beta_i^j = \frac{1}{\gamma_i}\left(\beta_j^j\left(\gamma_j + \frac{1}{\mu_j \sigma_{s_j}^2}\right) - \frac{\sum_{k\neq j}\lambda_k}{\mu_j \sigma_{s_j}^2}\right). \tag{51}$$

Sum expression (51) for all $i \neq j$:

$$-\beta_j^j = \sum_{i\neq j}\beta_i^j = \beta_j^j\left(\gamma_j + \frac{1}{\mu_j\sigma_{s_j}^2}\right)\sum_{i\neq j}\frac{1}{\gamma_i} - \frac{\sum_{i\neq j}\frac{1}{\gamma_i}\sum_{k\neq j}\lambda_k}{\mu_j\sigma_{s_j}^2}.$$

Therefore, the coefficient β_j^j which determines how agent i's payment depends on her effort is given by:

$$\beta_j^j = \frac{\sum_{i\neq j}\frac{1}{\gamma_i}\sum_{k\neq j}\lambda_k}{\mu_j\sigma_{s_j}^2\left(\left(\gamma_j + \frac{1}{\mu_j\sigma_{s_j}^2}\right)\sum_{i\neq j}\frac{1}{\gamma_i} + 1\right)} = \frac{\sum_{k\neq j}\lambda_k}{\mu_j\sigma_{s_j}^2\left(\gamma_j + \frac{1}{\mu_j\sigma_{s_j}^2} + \frac{1}{\sum_{i\neq j}\frac{1}{\gamma_i}}\right)} =$$
$$\frac{\sum_{k\neq j}\lambda_k}{1+\mu_j\sigma_{s_j}^2\left(\gamma_j + \frac{1}{\sum_{i\neq j}\frac{1}{\gamma_i}}\right)} = \frac{1-\lambda_j}{1+\mu_j\sigma_{s_j}^2\left(\gamma_j + \frac{1}{\sum_{i\neq j}\frac{1}{\gamma_i}}\right)}. \tag{52}$$

Now, substitute β_j^j into expression (51) to obtain β_i^j:

$$\beta_i^j = \frac{1}{\gamma_i}\left(\frac{\sum_{k\neq j}\lambda_k}{1+\mu_j\sigma_{s_j}^2\left(\gamma_j + \frac{1}{\sum_{k\neq j}\frac{1}{\gamma_k}}\right)}\left(\gamma_j + \frac{1}{\mu_j\sigma_{s_j}^2}\right) - \frac{\sum_{k\neq j}\lambda_k}{\mu_j\sigma_{s_j}^2}\right) =$$
$$\frac{1}{\gamma_i}\left(\frac{\sum_{k\neq j}\lambda_k}{1+\mu_j\sigma_{s_j}^2\left(\gamma_j + \frac{1}{\sum_{k\neq j}\frac{1}{\gamma_k}}\right)}\left(\gamma_j + \frac{1}{\mu_j\sigma_{s_j}^2} - \gamma_j - \frac{1}{\sum_{k\neq j}\frac{1}{\gamma_k}} - \frac{1}{\mu_j\sigma_{s_j}^2}\right)\right) =$$
$$-\frac{\frac{1}{\gamma_i}}{\sum_{k\neq j}\frac{1}{\gamma_k}}\left(\frac{1-\lambda_k}{1+\mu_j\sigma_{s_j}^2\left(\gamma_j + \frac{1}{\sum_{k\neq j}\frac{1}{\gamma_k}}\right)}\right). \tag{53}$$

With the outside investor, the first order condition (49) for $i = 0$ means that $\eta_j = 0$. So, $\beta_i^j = 0$ for all $i \neq 0, j$ and $\beta_0^j = -\beta_j^j$ (because $\sum_i \beta_i^j = 0$). Since $\eta_j = 0$, equation (50) simplifies into $\beta_j^j = \frac{\sum_{k\neq j}\lambda_k}{1+\mu_j\sigma_{s_j}^2\gamma_j} = \frac{1-\lambda_j}{1+\mu_j\sigma_{s_j}^2\gamma_j}$. □

C.2 Proof of Proposition 3.2

Proof. In order to find the optimal allocation of shares, we derive the first order condition on the Lagrangian:

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = \sum_{j=1}^{n} \frac{\lambda_j + \beta_j^j}{\mu_j} + \left(\sum_{j\in I} \lambda_j\right) \frac{1}{\mu_i} - \mu_i \frac{1}{\mu_i} \left(\frac{\lambda_i + \beta_i^i}{\mu_i}\right) - \gamma_i \sigma_q^2 \lambda_i + \theta_i - \chi = 0. \tag{54}$$

Notice, that the equation above is valid in both the case with and without the outside investor. Indeed, $\frac{\partial \mathcal{L}}{\partial \lambda_0} = \sum_{j\in I} \frac{\lambda_j + \beta_j^j}{\mu_j} + \theta_0 - \chi$, which is the same as equation (54) when $\mu_0 \to 0$ and $\gamma_0 = 0$. After some algebra we rearrange the first order condition as,

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = \left(\sum_{k\in I} \lambda_k \left(\sum_{j\neq k,i} \frac{1}{\mu_j \left(1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{l\neq j} \frac{1}{\gamma_l}}\right)\right)} + \frac{1}{\mu_i} + \frac{1}{\mu_k} \mathbb{I}_{i\neq k}\right)\right) - \gamma_i \sigma_q^2 \lambda_i + \theta_i - \chi = 0. \tag{55}$$

In the Online Appendix we show that the problem above is concave on the hyperplane $\sum_{i=1}^{n} \lambda_i = 1$ (feasible allocations of shares).

Denote $A_i = \frac{1}{\mu_i}$ and $B_i = \frac{1}{\mu_i \left(1+\mu_i \sigma_{s_i}^2 \left(\gamma_i + \frac{1}{\sum_{j\neq i} \frac{1}{\gamma_j}}\right)\right)}$. In the case with the outside investor, $B_i = \frac{1}{\mu_i \left(1+\mu_i \sigma_{s_i}^2 \gamma_i\right)}$ and $A_0 = B_0 = 0$.

Equating χ in expressions (55) for i and m, we get:

$$\left(\sum_{k\in I} \lambda_k \left(\sum_{j\neq k,i} B_j + A_i + A_k \mathbb{I}_{i\neq k}\right)\right) - \gamma_i \sigma_q^2 \lambda_i + \theta_i = \left(\sum_{k\in I} \lambda_k \left(\sum_{j\neq k,m} B_j + A_m + A_k \mathbb{I}_{i\neq k}\right)\right) - \gamma_m \sigma_q^2 \lambda_m + \theta_m. \tag{56}$$

Rearrange this expression:

$$(A_i - A_m - B_i + B_m) \sum_{k\neq i,m} \lambda_k + \lambda_m (A_i - B_i) - \theta_m - \gamma_i \sigma_q^2 \lambda_i = \lambda_i (A_m - B_m) - \theta_i - \gamma_m \sigma_q^2 \lambda_m.$$

Denote $D_i \equiv A_i - B_i$. Since $\sum_{j\in I} \lambda_j = 1$,

$$(D_i - D_m) - \lambda_i(D_i + \gamma_i \sigma_q^2) = \theta_m - \theta_i - \lambda_m(D_m + \gamma_m \sigma_q^2). \tag{57}$$

If both agents i and m hold positive shares of the company, then $\theta_i = \theta_m = 0$, and condition (7) follows. □

C.3 Proof of Proposition 3.3

Proof. First, let us assume that $D_i \geq D_k$. Let us show that if $\lambda_k > 0$, then $\lambda_i > 0$ as well.

Indeed, then θ_k is 0, and the right-hand side of condition (if we substitute m with k) (57) is non-positive.[33] If λ_i was zero, then the left-hand side of condition (57) is either positive (if $D_i > D_k$), or zero (if $D_i = D_k$). If it is positive, then we get a contradiction. If it is zero, then the right-hand side has to also be zero, but it can be zero only if $D_i = 0$, which is not the case for any productive agent i. Therefore, λ_i has to be positive.

Hence, there is an ordering of the agents, and if agents with a lower D_i have a share in the firm, then agents with a higher share D_m have a share as well. In particular, if there is an agent with a noiseless signal, while there are other agents with noisy signals about their performances, then she should not hold any shares of the company. There will be a threshold type, such that all agents with a higher D_i own stock in the company, and all agents below do not.

Consider a situation with an outside investor. Let us check when it optimal for everyone, including the outside investor to hold shares of the company. Since the outside investor is risk neutral ($\gamma_0 = 0$) and $D_0 = 0$, all other agents have higher D and also have shares of the company. It means that $\theta_i = 0$ for every i. Substitute $m = 0$ into equation (57), then we immediately get:

$$\lambda_i = \frac{D_i}{D_i + \gamma_i \sigma_q^2}$$

for every $i \in \{1, \ldots, n\}$. As the sum of all shares should equal to 0, the outsider's share is given by:

$$\lambda_0 = 1 - \sum_{i=1}^{n} \lambda_i = 1 - \sum_{i=1}^{n} \frac{D_i}{D_i + \gamma_i \sigma_q^2}. \tag{58}$$

[33] In fact, it is strictly negative, unless agent k is the outside investor with $D_k = 0$ and $\gamma_k = 0$. For all other agents $D_k + \gamma_k \sigma_q^2 > 0$.

Thus, the outsider owns shares of the company if and only if λ_0 from expression (58) is positive.

If it is negative, it is not optimal for the outsider to own shares of the company.

Now we will check when it is optimal for agents $k, k+1, \dots, n$ to own some stock of the company. If in the optimal allocation they all have positive shares, then all $\theta_k = \theta_{k+1} = \dots = \theta_n = 0$. Solving the system of first order conditions (57) and the feasibility condition $\sum \lambda_i = 0$ we will find λ_is.

$$\lambda_i = \frac{D_i - D_j + \lambda_j(D_j + \gamma_j \sigma_q^2)}{D_i + \gamma_i \sigma_q^2}.$$

Assume that agent k is the one with the lowest index who has shares of the company. It follows from above that all agents i with $D_i \geq D_k$ have shares of the company, so all agents $k, k+1, \dots, n$ also have positive shares of the company. The sum of their shares is equal to 1:

$$1 = \sum_{i=k}^{n} \lambda_i = \sum_{i=k}^{n} \frac{D_i - D_j}{D_i + \gamma_i \sigma_q^2} + \lambda_j \sum_{i=k}^{n} \frac{D_j + \gamma_j \sigma_q^2}{D_i + \gamma_i \sigma_q^2}.$$

Then for all $j \in \{k, k+1, \dots, n\}$:

$$\lambda_j = \frac{1 - \sum_{i=k}^{n} \frac{D_i - D_j}{D_i + \gamma_i \sigma_q^2}}{\sum_{i=k}^{n} \frac{D_j + \gamma_j \sigma_q^2}{D_i + \gamma_i \sigma_q^2}}. \tag{59}$$

If all λ_j are non-negative, then due of concavity the allocation of λs is optimal under the assumption that agents with indices lower than m do not get shares. If some of the λ_js are negative, then it is not optimal for all agents $k, k+1, \dots, n$ to have positive shares, and fewer agents should be owners of the firm. In order to check that all λ_js are positive, it is sufficient to check that $\lambda_k > 0$. Indeed, notice that the numerator in this formula positively depends on D_j, so if λ_k is positive, so are λ_j for $j > k$. $\lambda_k > 0$ if and only if

$$\sum_{i=k}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} < 1.$$

We showed that there is a threshold type, so we need to find this threshold type. We start by checking if all agents have positive shares of the company. If this is optimal (condition (8) is satisfied for the $k = 1$ or $k = 0$, if there is an outsider), then we found the best solution (the unconstrained maximum satisfies all constraints $\lambda_i \geq 0$ are satisfied). If this is not optimal, check if it can be optimal for $n-1$ agents to have shares of the company, so, check condition (8) for $k = 2$ (or n agents to be owners of the company and check condition (8) for $k = 1$ if there is an outsider). Continue this process till we find $k = m$ where all agents with indices at least k have positive λs.[34] This is the optimal allocation, because we showed that for a greater number of agents the allocation is not optimal, and this is the optimal allocation under the assumption that agents $j = 0, \dots, k-1$ have no shares. □

C.4 Proof of Proposition 4.1

Proof. To distinguish between an agent of type j for a given N, an agent of type j in the limit, and an agent number i for a given number of branches N, we denote the variables related to an agent of type j for a given N by $(\mu_j^N, \gamma_j^N, (\sigma_{s_j}^N)^2)$, the variables related to the limit type j by $(\mu_j, \gamma_j, \sigma_{s_j}^2)$, and $(\hat{\mu}_i, \hat{\gamma}_i, \hat{\sigma}_{s_i}^2)$ denote the variables related to agent i for a finite N. Then, for example, $\hat{\mu}_i = \mu^N_{\lceil \frac{i}{N} \rceil}$ and $\mu_j^N \to \mu_j$.

Let us check which of the workers will have shares of the firm, as it consists of more branches. According to Proposition 3.3, agent k has shares if and only if[35]

$$\sum_{i=k+1}^{n} \frac{\hat{D}_i - \hat{D}_k}{\hat{D}_i + \hat{\gamma}_i \sigma_q^2} = \sum_{i=N\cdot\lceil \frac{k}{N} \rceil + 1}^{n} \frac{\hat{D}_i - \hat{D}_k}{\hat{D}_i + \hat{\gamma}_i \sigma_q^2} < 1. \tag{60}$$

We can order limit types by

$$D_l = \frac{1}{\mu_l + \frac{1}{\sigma_{s_l}^2 \left(\gamma_l + \frac{1}{\sum_{j \neq l \cdot N} \frac{1}{\gamma_j}} \right)}}.$$

If there is a risk neutral outsider, then the sum of inverse coefficients of risk aversion in the denominator is equal to zero, so the expression above can be rewritten as:

$$D_l = \frac{1}{\mu_l + \frac{1}{\sigma_{s_l}^2 \gamma_l}}.$$

[34] The process will end, because condition (8) is satisfied for $k = n - 1$.

[35] Notice that all the terms between $i = k$ and $i = N \cdot \lceil \frac{k}{N} \rceil$ are equal to 0, since then $\hat{D}_i = \hat{D}_k$.

For a big enough N the types are sorted the same way as in the limit.[36] Since there are N identical agents of each type l, the left hand side of inequality (60) equals:[37]

$$N \sum_{j=l+1}^{b} \frac{D_j^N - D_l^N}{D_j^N + \gamma_j^N \sigma_q^2}$$

If $(\mu_l^N, (\sigma_{s_l}^N)^2, \gamma_l^N) \to (\mu_l, \sigma_{s_l}^2, \gamma_l) > 0$ and $\sigma_q^2 \to \infty$, as $N \to \infty$, then the expression behaves asymptotically as

$$\frac{N}{\sigma_q^2} \sum_{i=l+1}^{b} \frac{D_i^N - D_l^N}{\gamma_i}. \tag{61}$$

Hence, if σ_q^2 grows faster than N ($N = o(\sigma_q^2)$), then this expression converges to 0, and every type l will participate (all workers have shares of the firm), and the outsider participates as well. In fact, the outsider's share in this case converges to 1:

$$\lambda_0 = 1 - N \sum_{i=1}^{b} \frac{D_i^N}{D_i^N + \gamma_i^N \sigma_q^2} \to 1. \tag{62}$$

If σ_q^2 grows slower than N ($\sigma_q^2 = o(N)$), then expression (61) converges to infinity for any $l < \bar{n}$. It means that in the limit on the type with the highest $\hat{D}$ in each branch will have shares of the firm (e.g. the top management), but not the other types of workers.

In the threshold situation, when $\sigma_q^2 = O(N)$ (for example, in case of independent noises ε_{q^i}), it depends on how big the limit of the ratio $\frac{\sigma_q^2}{N}$ is.

For every $l \in \{0, 1, \ldots, b-1\}$ denote

$$R_l \equiv \sum_{i=l+1}^{b} \frac{D_i - D_l}{\gamma_i}.$$

Thus, if $\alpha > R_l$, then condition (60) is satisfied:

$$\frac{N}{\sigma_q^2} \sum_{i=l+1}^{b} \frac{D_i^N - D_l^N}{\gamma_i^N} \to \frac{\sum_{i=l+1}^{b} \frac{D_i - D_l}{\gamma_i}}{\alpha} = \frac{R_l}{\alpha} < 1,$$

so, type l agents have positive shares of the firm. Notice that if $D_i > (\geq) D_j$, then

[36]The only issue can happen if the limit some types have the same parameter D_j. Then one of the types' D_j^N might be higher than the other for any N, but they converge to the same D_j. The behavior in this case is not very different, but we assume this case away to simplify the proof.

[37]Here l is the type corresponding to agent i in inequality (60), that is $l = \lceil \frac{k}{N} \rceil$.

$R_i < (\leq) R_j$ (each term in the sum is smaller, and there are fewer of them). Effectively, R_is are thresholds determining which types of agents hold shares of the firm in the limit. Notice that if $R_0 \geq \alpha > R_l$ that in the limit all agents of type l together own a strictly positive share of the firm in the limit. Indeed, the sum of their shares (from equation (9)) is equal to:

$$N\lambda_l = N \frac{1 - N\sum_{i=k}^{b} \frac{D_i^N - D_l^N}{D_i^N + \gamma_i^N \sigma_q^2}}{N\sum_{i=k}^{b} \frac{D_l^N + \gamma_l^N \sigma_q^2}{D_i^N + \gamma_i^N \sigma_q^2}} \to \frac{1 - \frac{R_l}{\alpha}}{\gamma_l \sum_{i=k}^{b} \frac{1}{\gamma_i}} > 0.$$

If $\infty > \alpha > R_0$, then

$$N\lambda_j = N \frac{D_j^N}{D_j^N + \gamma_j^N \sigma_q^2} \to \frac{D_j}{\gamma_j \alpha} > 0.$$

□

C.5 Proof of Proposition 5.1

Proof. By setting $k = 0$ in (8) we see that the outside investor owns a positive number of shares if and only if:

$$\sum_{i=1}^{n} \frac{D_i}{D_i + \gamma_i \sigma_q^2} < 1. \tag{63}$$

The left hand side of condition (63) strictly decreases with σ_q^2, and when $\sigma_q^2 = 0$, it is equal to n, so the condition is not satisfied. When $\sigma_q^2 \to \infty$, the left-hand side of condition (63) monotonically converges to zero, therefore there is a unique value of $\sigma_q^2 = \bar{\sigma}_q^2$, such that if $\sigma_q^2 \leq \bar{\sigma}_q^2$ the outsider does not own any shares of the company, and if $\sigma_q^2 > \bar{\sigma}_q^2$, then the outsider owns a positive share. If the outside investor owns a positive share of the firm, then substituting $m = 0$ in equation (7) yields (12). □

References

Alchian, A. A. and Demsetz, H. (1972). Production, information costs, and economic organization. *The American Economic Review*, **62** (5), 777–795.

Anderson, E. and Schmittlein, D. C. (1984). Integration of the sales force: An empirical examination. *The Rand Journal of Economics*, pp. 385–395.

Coram, P. J. and Robinson, M. J. (2017). Professionalism and performance incentives in accounting firms. *Accounting Horizons*, **31** (1), 103–123.

Eswaran, M. and Kotwal, A. (1984). The moral hazard of budget-breaking. *The Rand Journal of Economics*, pp. 578–581.

Galanter, M. and Palay, T. M. (1990). Why the big get bigger: The promotion-to-partner tournament and the growth of large law firms. *Virginia Law Review*, pp. 747–811.

Garicano, L. and Santos, T. (2004). Referrals. *The American Economic Review*, pp. 499–525.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *The Journal of Political Economy*, pp. 691–719.

Hansmann, H. (1996). *The ownership of enterprise*. Harvard University Press.

Heywood, J. S. and Jirjahn, U. (2009). Profit sharing and firm size: The role of team production. *Journal of Economic Behavior and Organization*, **71**, 246–258.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, pp. 74–91.

Holmström, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, pp. 324–340.

Holmström, B. and Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, **55** (2), 303–328.

Ishiguro, S. and Yasuda, Y. (2021). Moral hazard and subjective evaluation. *Working Paper*.

KANDEL, E. and LAZEAR, E. P. (1992). Peer pressure and partnerships. *Journal of Political Economy*, **100** (4), 801–817.

LAFONTAINE, F. (1992). Agency theory and franchising: some empirical results. *The rand journal of economics*, pp. 263–283.

LEGROS, P. and MATSUSHIMA, H. (1991). Efficiency in partnerships. *Journal of Economic Theory*, **55** (2), 296–322.

LEVIN, J. (2003). Relational incentive contracts. *The American Economic Review*, **93** (3), 835–857.

— and TADELIS, S. (2005). Profit sharing and the role of professional partnerships. *The Quarterly Journal of Economics*, pp. 131–171.

LI, J. (2016). Profit-sharing, wisdom of the crowd, and theory of the firm. *Working Paper*.

MORRISON, A. D. and WILHELM, W. J. (2004). Partnership firms, reputation, and human capital. *The American Economic Review*, **94** (5), 1682–1692.

— and — (2008). The demise of investment banking partnerships: Theory and evidence. *The Journal of Finance*, **63** (1), 311–350.

POBLETE, J. (2015). Profit sharing and market structure. *International Journal of Industrial Organization*, **39**, 10–18.

RAUH, M. T. (2014). Incentives, wages, employment, and the division of labor in teams. *The RAND Journal of Economics*, **45** (3), 533–552.

RAYO, L. (2007). Relational incentives and moral hazard in teams. *The Review of Economic Studies*, **74** (3), 937–963.

WEITZMAN, M. L. (1980). Efficient incentive contracts. *The Quarterly Journal of Economics*, **94** (4), 719–730.

WILLIAMSON, O. E. (1985). *The economic institutions of capitalism*. The Free Press, New York.

Credit Author Statment

Both authors have been equally involved in all aspects of writing this article.