Profit Sharing and Incentives

Emre Ozdenoren*
London Business School

Oleg Rubanov†
University of Bonn

February 21, 2018

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 participate in productive activities of the firm. Partnerships are commonly found

*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: Institute for Microeconomics, University of Bonn, Adenauerallee 24-42, 53113, Bonn, Germany. Email: orubanov@uni-bonn.de. Web: https://sites.google.com/site/orubanov/

in professional services such as investment banking, law, medicine, consulting. However, one can find partnerships in other industries. For example, John Lewis Partnership is a leading UK retail business where the partners are its 91,500 permanent staff. Most firms fall somewhere in between where some shares are held by productive insiders and the rest by outsiders.

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.[1] 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 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.[2] 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.

[1]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.

[2]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.

In Section 3 we characterize the optimal allocation of shares when the firm can write contracts based on performance signals. 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.[3] 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.[4] In particular, the unproductive outside investor always has the lowest ownership parameter. 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. The outsider holds shares only if *all* productive agents hold shares even though it is possible to write performance based contracts.[5] 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.

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,

[3]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.

[4]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.

[5]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.

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.[6] 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 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 other exogenous parameters such as the output noise, signal precision and productivity. Among others we show that all else constant agents who are more productive and whose performances are poorly observable (as in professional organizations, like law firms) are more likely to be partially motivated with shares, rather than solely by performance contracts.

Proofs missing from the main text can be found in the Appendix. The Online Appendix contains proofs of non-essential facts and descriptions of the firms illustrating Section 5.

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

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. 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.[7]

Rayo (2007) considers a similar problem of a group of agents producing a product together and incentivising each other for the effort they exert. However, Rayo studies relational contracts, i.e. the contracts which are enforced not by a court but rather by

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

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. Our work instead focuses on explicit contracts, where there is no future interaction between the parties (or they massively discount the future), and they solely rely on court enforceable contracts.[8]

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. Following Grossman and Hart (1986) a large literature looks at who should have ownership of the firm when ownership creates residual rights to a productive asset when it is prohibitively hard to specify all possible contingencies in the contract.

2 The Model

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 $i \in \{1, 2, \ldots, n\}$ 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 $y(e_1, \ldots, e_n) = q(e_1, \ldots, e_n) + \varepsilon_q$ 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 $q(e) = \sum_{i=1}^{n} e_i$.[9]

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-

[8]Ishiguro and Yasuda (2017) 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.

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

Morgenstern utility function of consuming x and exerting effort e_i is given by $1 - e^{-\gamma_i(x-C_i(e_i))}$, 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.[10] If agent i owns share λ_i of the company, and the profit of the company is π, then agent i receives $\lambda_i \pi$.[11]

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

[10] 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.

[11] 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.[12]

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.[13]

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 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

[12]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.

[13]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.

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

$$\max_{\lambda,w(s)} \sum_{i=1}^{n} \Big(\mathbb{E}\left(\lambda_i y(e) - C_i(e_i) + w_i(s)\right) - \frac{\gamma_i}{2}\mathbb{V}(\lambda_i y(e) + w_i(s))\Big),$$

$$\text{s.t. } \forall s \sum_{i=1}^{n} w_i(s) = 0, \forall i : \lambda_i \geq 0 \text{ and } \sum_{i=1}^{n} \lambda_i = 1.$$

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

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

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

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:

$$\mathcal{L} = \sum_{i=1}^{n} \left[\lambda_i \left(\sum_{j=1}^{n} \frac{\lambda_j + \beta_j^j}{\mu_j}\right) - \frac{\mu_i}{2}\left(\frac{\lambda_i + \beta_i^i}{\mu_i}\right)^2 - \frac{\gamma_i}{2}\left((\lambda_i)^2\sigma_q^2 + \sum_{j=1}^{n} \left(\beta_i^j\right)^2 \sigma_{s_j}^2\right)\right] +$$
$$\sum_{i=1}^{n} \lambda_i \theta_i + \chi\left(1 - \sum_{i=1}^{n} \lambda_i\right) - \sum_{j=0}^{n} \left(\eta_j \sum_{i=1}^{n} \beta_i^j\right). \quad (3)$$

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

Proposition 3.1. *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:*

$$\beta_j^j = \frac{\sum_{k\neq j} \lambda_k}{1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{i\neq j} \frac{1}{\gamma_i}}\right)}; \quad (4)$$

- *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:

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

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. In Appendix B.2 we show that using (4) and (5), the first-order condition for λ_i can be written as:

$$\frac{\partial\mathcal{L}}{\partial\lambda_i} = \left(\sum_{k=1}^{n}\lambda_k\left(\sum_{j\neq k,i}\frac{1}{\mu_j\left(1+\mu_j\sigma_{s_j}^2\left(\gamma_j+\frac{1}{\sum_{l\neq j}\frac{1}{\gamma_l}}\right)\right)}+\frac{1}{\mu_i}+\frac{1}{\mu_k}\mathbb{I}_{i\neq k}\right)\right) - \gamma_i\sigma_q^2\lambda_i+\theta_i-\chi = 0. \tag{6}$$

Equating χ in expressions (6) for i and m, we get:

$$\left(\sum_{k=1}^{n}\lambda_k\left(\sum_{j\neq k,i}B_j+A_i+A_k\mathbb{I}_{i\neq k}\right)\right)-\gamma_i\sigma_q^2\lambda_i+\theta_i=$$
$$\left(\sum_{k=1}^{n}\lambda_k\left(\sum_{j\neq k,m}B_j+A_m+A_k\mathbb{I}_{m\neq k}\right)\right)-\gamma_m\sigma_q^2\lambda_m+\theta_m. \quad (7)$$

where $A_i=\frac{1}{\mu_i}$ and $B_i=\frac{1}{\mu_i\left(1+\mu_i\sigma_{s_i}^2\left(\gamma_i+\frac{1}{\sum_{j\neq i}\frac{1}{\gamma_j}}\right)\right)}$.

Next we define

$$D_i\equiv A_i-B_i=\frac{1}{\mu_i+\frac{1}{\sigma_{s_i}^2\left(\gamma_i+\frac{1}{\sum_{j\neq i}\frac{1}{\gamma_j}}\right)}}. \quad (8)$$

We refer to D_i as the *ownership parameter* of agent i.[14] Since $\sum_{j=1}^{n}\lambda_j=1$ we can further simplify (7):

$$(D_i-D_m)-\lambda_i(D_i+\gamma_i\sigma_q^2)=\theta_m-\theta_i-\lambda_m(D_m+\gamma_m\sigma_q^2). \quad (9)$$

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 $0<D_1\leq D_2\leq\cdots\leq D_n$.[15] If there is an outside investor, then by substituting $\mu_0=\infty$ and $\gamma_0=0$ into the definition (8), 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 next he is more likely to hold shares in the firm.

[14]In these derivations we do not assume that there is necessarily a risk neutral outside investor, and Equations (4), (5) and (8) all simplify when there is one. In particular with a risk neutral outside investor β_i^j's are all zero and (8) becomes $D_i=\frac{1}{\mu_i+\frac{1}{\sigma_{s_i}^2\gamma_i}}$.

[15]It is easy to see that substituting positive numbers μ_i, $\sigma_{s_i}^2$, and γ_j yields a positive number in definition (8).

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.*

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

$$(D_i - D_m) - \lambda_i(D_i + \gamma_i \sigma_q^2) = -\lambda_m(D_m + \gamma_m \sigma_q^2). \tag{10}$$

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

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

$$\sum_{i=k+1}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} < 1. \tag{11}$$

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

$$\lambda_j = \frac{1 - \sum_{i=k}^{n} \frac{D_i - D_j}{D_i + \gamma_i \sigma_q^2}}{\sum_{i=k}^{n} \frac{D_j + \gamma_j \sigma_q^2}{D_i + \gamma_i \sigma_q^2}} \tag{12}$$

shares.

If condition (11) is satisfied for $k = 0$ (i.e. $m = 0$), then the outside investor's share is positive and is equal to:

$$\lambda_0 = 1 - \sum_{i=1}^{n} \frac{D_i}{D_i + \gamma_i \sigma_q^2}, \tag{13}$$

and agent j's share is

$$\lambda_j = \frac{D_j}{D_j + \gamma_j \sigma_q^2}. \tag{14}$$

Inequality (11) 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 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.[16] 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.[17]

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.

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'

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

[17] 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.

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.[18] 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:

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

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. In other words, we need to specify the correlations between different ε_{q_k}.

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

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 Comparative Statics

In this section we study how the optimal ownership and incentive structures change with the exogenous variables of the model which include the productivity and risk aversion parameters of the agents (μ_i and γ_i), noisiness of the performance signals and the output of the firm ($\sigma_{s_i}^2$ and σ_q^2). These comparative statics provide interesting testable implications. These results also help us understand in which industries we should expect to find concentrated vs. diffuse ownership of shares.

As can be seen from definition (8), D_i is a continuous function of all the variables. Also, the extensive margin (condition (11) determining, whether agent i has shares of the firm) and the intensive margin (equations (12), (13) and (14), determining how many shares the agents hold) continuously depend on all the parameters. In this section we show how the extensive and the intensive margins depend on the parameters of the model.

5.1 Productivity

First, we study how the optimal share allocation changes with productivity. In this and following sub-sections we supplement the theoretical results with a numerical example to illustrate how the ownership structure varies with respect to different characteristics of the firm.

Propositions 5.1 and 5.2 describe how the profit shares changes when one of the agents become less productive.

Proposition 5.1 (Extensive margin). *Fix the parameters of all agents except the productivity parameter of agent i. Suppose $\mu_i < \mu_i'$.*
A) If agent i does not own shares when his productivity is μ_i, he does not own shares when his productivity is μ_i'. If agent i owns shares when his productivity is μ_i', he owns shares when his productivity is μ_i.
B) If agent j owns shares when agent i's productivity is μ_i, he owns shares when agent i productivity is μ_i'. If agent j does not own shares when agent i's productivity is μ_i', he does not own shares when agent i productivity is μ_i.

Part A of Proposition 5.1 says that if an agent does not own shares when he is more productive, he should not own shares if he becomes less productive. Similarly, if an agent owns shares when he is less productive, he should also own shares when he becomes more productive. Part B of the proposition considers the impact of agent i's productivity change on the composition of ownership by other agents. If agent i becomes less productive, and agent who owns shares should still own shares. Similarly, if agent i becomes more productive, any agent who does not own shares should still not own shares.[19]

Proposition 5.2 (Intensive margin). *Fix the parameters of all agents except the productivity parameter of agent i. Suppose $\mu_i < \mu_i'$ and the set of agents for whom it is optimal to hold shares is the same for μ_i' and μ_i.*
A) If the outside investor has shares of the firm, then the shares of all agents j, s.t. $j \notin \{0, i\}$, are the same, regardless whether agent i's productivity is μ_i or μ_i'. The outside investor's share is higher for μ_i' than for μ_i and agent i's share is smaller for μ_i' by the same amount.

[19]Note that the first statement in part B does not rule out that an agent who did not own shares initially might own shares when agent i becomes less productive. Similarly, the second statement in part B does not rule out that an agent who initially own shares may might not own shares when agent i becomes more productive.

B) If the outsider has no shares of the firm and agent i has shares of the firm, then all agents $j \neq i$ who have shares of the firm have strictly more shares when the productivity of agent i is μ_i' than when his productivity is μ_i, and agent i has more shares when his productivity is μ_i.
C) If agent i has no shares of the firm, then every agent's shares are the same, regardless whether agent i's productivity is μ_i or μ_i'.

According to Proposition 3.2, if the outside investor is an owner of the firm, so are all other agents, including agent i, so in Part C of Proposition 5.2 the outside investor has no shares, and in Part A all agents have shares of the firm.

Together the two propositions explain the change of the ownership structure of the firm.

In accordance with Proposition 5.2, there are three regimes of ownership:
Regime A, when all agents including agent i and the outside investor have shares of the firm;
Regime B, when the outside investor has no shares, but agent i has shares of the firm;
Regime C, when neither agent i, nor the outside investor has shares of the firm.

Lets analyze how the ownership structure changes when we start from a situation where agent i is extremely productive (i.e. μ_i is close to zero) and gradually becomes less productive holding rest of the parameters fixed. First, note that even when agent i is extremely productive depending on the other parameters of the model, we could still be in any of the three regimes.[20]

Given that, suppose initially we are in Regime A. Then the less productive agent i is (the higher μ_i is) the smaller is his share λ_i and the higher is the outside investor's share λ_0. For all greater μ_i we will remain in Regime A. Also, each of the other agents holds the same share, no matter how unproductive agent i is. Notice that even though agent i's share gets smaller, it is positive.

[20] When $\mu_i \to 0$, agent i's ownership parameter $D_i = \frac{1}{\mu_i + \frac{1}{\gamma_i \sigma_{s_i}^2}}$ converges to $\gamma_i \sigma_{s_i}^2$. According to Proposition 3.3, we are in Regime A if $\frac{\sigma_{s_i}^2}{\sigma_{s_i}^2 + \sigma_q^2} + \sum_{j=1, j \neq i}^n \frac{D_j}{D_j + \gamma_j \sigma_q^2} < 1$, or, equivalently, $\sum_{j=1, j \neq i}^n \frac{D_j}{D_j + \gamma_j \sigma_q^2} < \frac{\sigma_q^2}{\sigma_{s_i}^2 + \sigma_q^2}$. We are in Regime C if $\sum_{j: D_j > \gamma_i \sigma_{s_i}^2} \frac{D_j - \gamma_i \sigma_{s_i}^2}{D_j + \gamma_j \sigma_{s_j}^2} \geq 1$. If neither of the two conditions is satisfied, we are in Regime B.

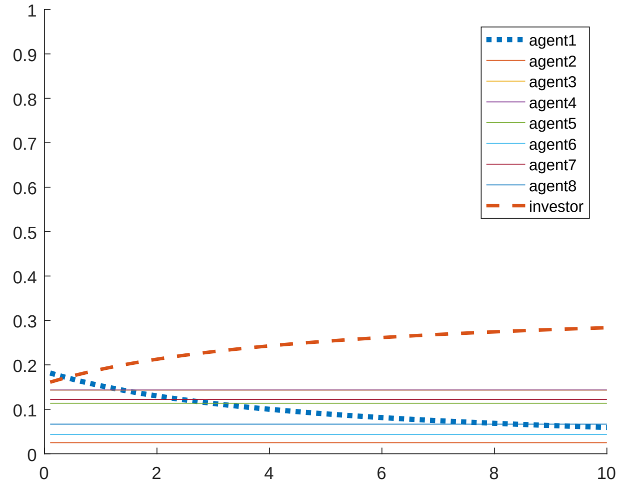

Figure 1: Changing agent 1's productivity μ_1. Regime A.

If initially we are in Regime C for μ_i close to 0, then when agent i is less productive (higher μ_i) we remain in Regime C, and no one's shares change, so they are independent of μ_i.

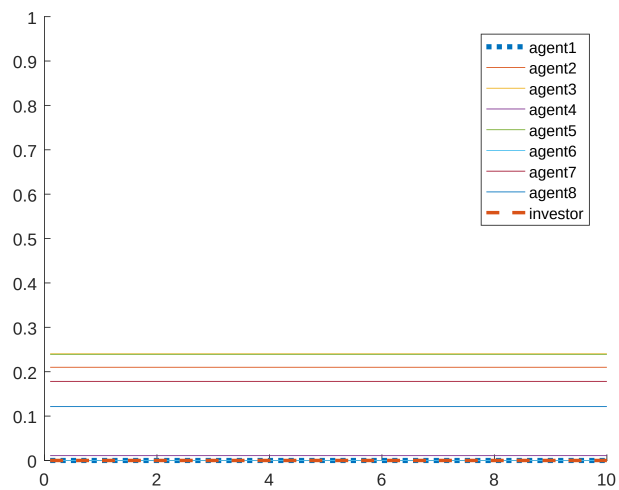

Figure 2: Changing agent 1's productivity μ_1. Regime C.

If initially we are in Regime B for μ_i close to 0, then there are three possibilities: 1) we remain in Regime B as agent i becomes less productive (higher μ_i);

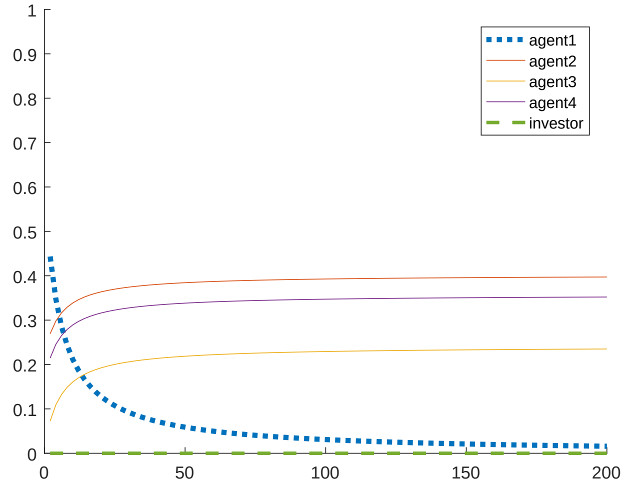

Figure 3: Changing agent 1's productivity μ_1. Regime B.

2) there is a threshold value $\bar{\mu}_i^A$, such that we are in Regime B when agent i is relatively productive ($\mu_i \leq \bar{\mu}_i^A$) and in Regime A when agent i is relatively unproductive ($\mu_i > \bar{\mu}_i^A$);

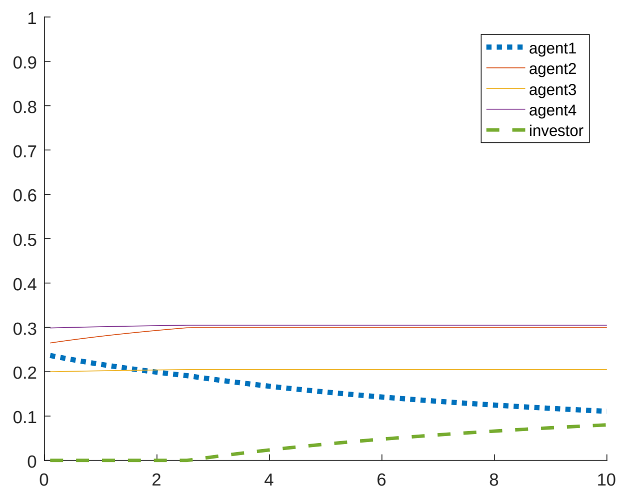

Figure 4: Changing agent 1's productivity μ_1. Regime B and Regime A.

3) there is a threshold value $\bar{\mu}_i^C$, such that we are in Regime B when agent i is relatively productive ($\mu_i < \bar{\mu}_i^C$) and in Regime C when agent i is relatively unproductive ($\mu_i \geq \bar{\mu}_i^C$).

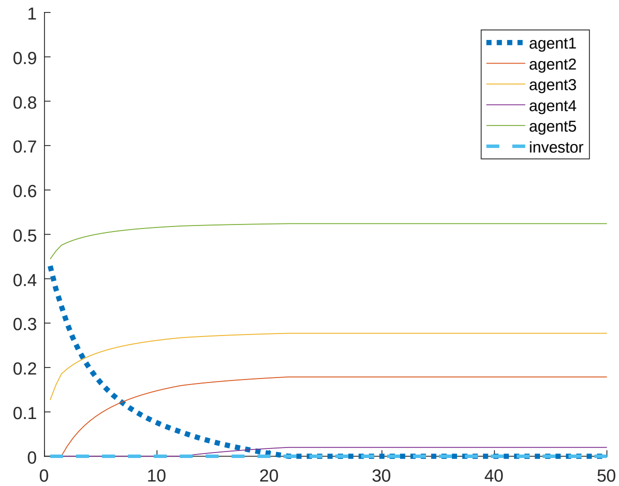

Figure 5: Changing agent 1's productivity μ_1. Regime B and Regime C.

For all the values of μ_i while we are in Regime B, the more productive agent i is the fewer shares he owns, and all of the other agents who own shares when agent i is more productive, own even more shares when agent i is less productive. Additionally, some of the other agents who do not have shares might get shares when agent i becomes less productive. As agent i becomes extremely unproductive ($\mu_i \to \infty$), he either stops being an owner at some point $\mu_i = \bar{\mu}_i^A$ (then we are in case 3), or remains positive for all μ_i (then we are in cases 1) or 2)). If at some point $\mu_i = \bar{\mu}_i^C$ the outside investor becomes one of the owners, then we are in case 2, otherwise we are in case 1. When the agent is relatively unproductive (his productivity parameter μ_i is above the thresholds $\bar{\mu}_i^A$ and $\bar{\mu}_i^C$), the behavior of the agents' shares is as explained above for Regimes A and C.

Table 1 summarizes the change in the agents' ownership, as agent i becomes less productive.[21]

[21]The thresholds are derived by substituting the limit values of D_i ($D_i \to 0$ when $\mu_i \to \infty$ and $D_i \to \gamma_i \sigma_{s_i}^2$ when $\mu_i \to \infty$) into Proposition 3.3.

Value of $\sum_{j=1,j\neq i}^{n}\frac{D_j}{D_j+\gamma_j\sigma_q^2}$	Regimes	Shares		
		λ_i	λ_j, $j \neq i$	λ_0
$< \frac{\sigma_q^2}{\sigma_q^2+\sigma_{s_i}^2}$	Regime A	decreases	constant	increases
$\in \left[\frac{\sigma_q^2}{\sigma_q^2+\sigma_{s_i}^2};1\right)$	$\mu_i \leq \bar{\mu}_i^A$: Regime B	decreases	zero and stays zero OR zero and then increases OR increases	zero
	$\mu_i > \bar{\mu}_i^A$: Regime A	decreases	constant	increases
$=1$	Regime B	decreases	zero and stays zero OR zero and then increases OR increases	zero
>1 and $< 1+\sum_{j\neq 0,i}\frac{\gamma_i\sigma_{s_i}^2}{D_j+\gamma_j\sigma_q^2}$	$\mu_i < \bar{\mu}_i^C$: Regime B	decreases	zero and stays zero OR zero and then increases OR increases	zero
	$\mu_i \geq \bar{\mu}_i^C$: Regime C	zero	constant	zero
$\geq 1+\sum_{j\neq 0,i}\frac{\gamma_i\sigma_{s_i}^2}{D_j+\gamma_j\sigma_q^2}$	Regime C	zero	constant	zero

Table 1: Changes in shares when agent i becomes less productive (increase in μ_i).

This analysis has an important testable implication. Suppose the productivity of agents in a firm changes (e.g., due to a technological change). The increasing use of computer technology and robots have an asymmetric impact on the productivity of agents who do different jobs. While some agents become less productive (such as factory workers whose jobs are largely performed by robots), others become more productive (such as surgeons who can use robots to assist them in surgery). This section provides a clear prediction about how the ownership of these different types of agents respond to the changes in their productivity.

5.2 Signal precision

In this section we study how the optimal share allocation changes with the precision of performance signals. The next proposition describes how the ownership of the firm changes when the performance signal of one of the agents become less precise.

Proposition 5.3 (Extensive margin)**.** *Fix the parameters of all agents except the variance of agent i's signal. Suppose* $\sigma_{s_i}^2 < (\sigma'_{s_i})^2$.
A) If agent i does not own shares when the variance of his performance signal is $(\sigma'_{s_i})^2$*, he also does not own shares when when the variance of his performance signal is* $\sigma_{s_i}^2$*. If agent i owns shares when the variance of his performance signal is* $(\sigma_{s_i})^2$*, he also owns shares when the variance of his performance signal is* $(\sigma'_{s_i})^2$*.*
B) If agent j owns shares when the variance of agent i's performance signal is $(\sigma'_{s_i})^2$*, he also owns shares when the variance of agent i's performance signal is* $(\sigma_{s_i})^2$*. If agent j does not own shares when the variance of agent i's performance signal is* $(\sigma_{s_i})^2$*, he does not own shares when the variance of agent i's performance signal is* $(\sigma'_{s_i})^2$*.*

Part A of Proposition 5.3 says that if an agent does not own shares when his effort is difficult to observe, he should not own shares when more precise information about his performance becomes available (it becomes more efficient to motivate such agent with a regular contract). Similarly, if an agent owns shares when his effort is easy to observe, he should not own shares when information about his performance is more obscure. Part B of the proposition considers the impact of change in agent i's signal noisiness on the composition of ownership by other agents. If agent i's signal becomes noisier, then any agent who owns shares should still own shares. Similarly, if agent i's signal becomes less noisy, any agent who does not own shares should still not own shares.

Proposition 5.4 (Intensive margin)**.** *Fix the parameters of all agents except the variance of agent i's signal. Suppose* $\sigma_{s_i}^2 < (\sigma'_{s_i})^2$ *and the set of agents for whom it is optimal to hold shares is the same for* $\sigma_{s_i}^2$ *and* $(\sigma'_{s_i})^2$*.*
A) If the outside investor has shares of the firm, then the shares of all agents j, s.t. $j \notin \{0, i\}$*, are the same, regardless how noisy agent i's signal is (*$\sigma_{s_i}^2$ *or* $(\sigma'_{s_i})^2$*). The outside investor's share is higher for* $\sigma_{s_i}^2$ *than for* $(\sigma'_{s_i})^2$ *and agent i's share is smaller for* $\sigma_{s_i}^2$ *by the same amount.*
B) If the outsider is not an owner of the firm and agent i has shares of the firm, then all agents $j \neq i$ *who have shares of the firm have strictly smaller shares when agent i's effort*

is less observable (noisiness $(\sigma'_{s_i})^2$, rather than $\sigma^2_{s_i}$), and agent i has a larger share when his effort is harder to observe.
C) If agent i has no shares of the firm, then every agent's shares are the same, regardless whether the noisiness of agent i's signal is $\sigma^2_{s_i}$ or $(\sigma'_{s_i})^2$.

The change in ownership in response to a change in the noisiness of agent i's signal is analogous to the effect of a change in his productivity, but with an opposite sign.

The definitions of Regimes A, B and C are as in Section 5.1.

Table 2 summarizes when the firm is in each of the regimes and how every agent's shares change in response to a change in the noisiness of agent i's signal.

Value of $\sum_{j=1,j\neq i}^{n} \frac{D_j}{D_j+\gamma_j\sigma_q^2}$	Regimes	Shares		
		λ_i	$\lambda_j,\ j\neq i$	λ_0
$< 1 - \frac{1}{1+\mu_i\gamma_i\sigma_q^2}$	Regime A	increases	constant	decreases
$\in \left[1 - \frac{1}{1+\mu_i\gamma_i\sigma_q^2}; 1\right)$	$\sigma^2_{s_i} < (\bar{\sigma}^A_{s_i})^2$: Regime A	increases	constant	decreases
	$\sigma^2_{s_i} \geq (\bar{\sigma}^A_{s_i})^2$: Regime B	increases	always zero OR decreases and becomes zero OR decreases	zero
$= 1$	Regime B	increases	always zero OR decreases and becomes zero OR decreases	zero
> 1 and $< 1+\frac{1}{\mu_i}\sum_{j\neq 0,i}\frac{1}{D_j+\gamma_j\sigma_q^2}$	$\sigma^2_{s_i} \leq (\bar{\sigma}^C_{s_i})^2$: Regime C	zero	constant	zero
	$\sigma^2_{s_i} > (\bar{\sigma}^C_{s_i})^2$: Regime B	increases	always zero OR decreases and becomes zero OR decreases	zero
$\geq 1+\frac{1}{\mu_i}\sum_{j\neq 0,i}\frac{1}{D_j+\gamma_j\sigma_q^2}$	Regime C	zero	constant	zero

Table 2: Changes in shares when agent i becomes less productive (increase in μ_i).

Figure 6 illustrates the effect of agent 1's signal precision change (in this example, it is Regime A for small values of $\sigma^2_{s_i}$ and Regime B for greater values of $\sigma^2_{s_i}$).

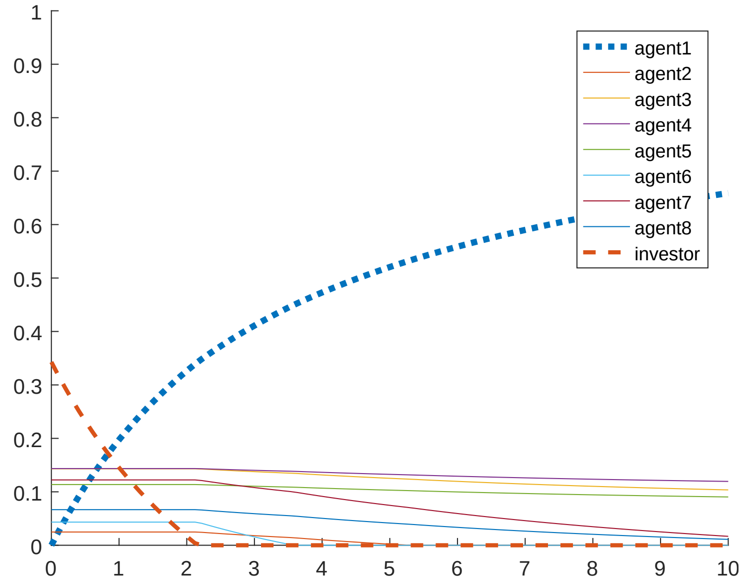

Figure 6: Changing agent 1's signal precision $\sigma^2_{s_1}$.

5.3 Output noise

In this section we study how the optimal share allocation changes with the precision of the output noise. 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 5.5. *There exists a cutoff $\bar{\sigma}^2_q > 0$ such that the outside investor does not own any shares if $\sigma^2_q \leq \bar{\sigma}^2_q$ and owns a positive share of the firm $\sigma^2_q > \bar{\sigma}^2_q$. If the outside investor owns positive share of the firm then agent i owns,*

$$\lambda_i = \frac{D_i}{D_i + \gamma_i \sigma^2_q}. \tag{15}$$

From, Proposition 5.5 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^2_q increases. In the limit, the outside investor holds almost all the shares.

The behavior of the optimal ownership structure for $\sigma^2_q \in [0, \bar{\sigma}^2_q]$ is more nuanced. For

$\sigma_q^2 = 0$ not all agents might hold some shares of the firm.[22] Agents $\{m, \ldots, n\}$ (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 (11) 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 amounts of shares, and differentiate (9) with respect to σ_q^2 (since $\theta_i = \theta_m = 0$):

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

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.

[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 (11) decreases.

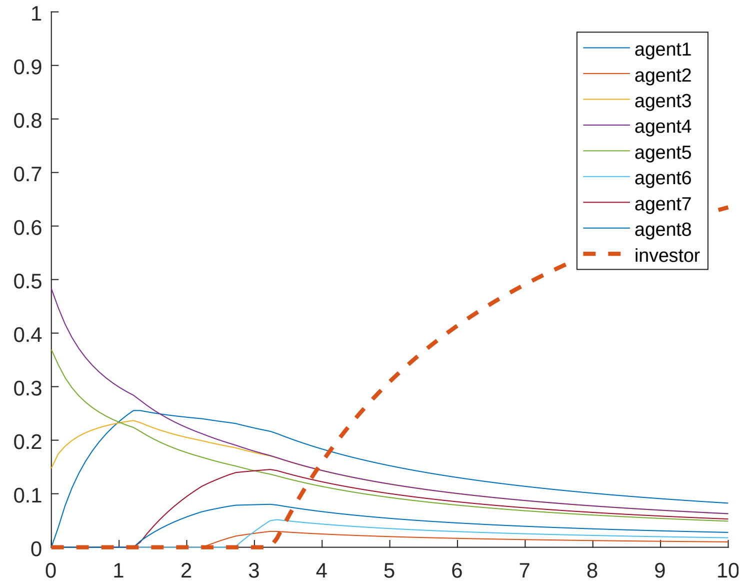

Figure 7: Changing output noisiness σ_q^2.

Figure 7 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 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

$$\lambda_i y + w_i(s). \tag{16}$$

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 $j = 0, \ldots n$.[25]

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 $\tilde{w}_i(s) = \tilde{\beta}_i^j s_j + \tilde{\beta}_i^0$;*

[25] To see that $\sum_i \beta_i^0 = 0$, consider the realisation of signals $s = (0, \ldots, 0)$. 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 $\sum_i \beta_i^j = 0$.

- *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 $i \in \{0, \ldots, n\}$ and $j \in \{0, \ldots, n\}$,

$$\beta_i^j \equiv \tilde{\beta}_i^j - \lambda_i \sum_{k \in I} \tilde{\beta}_k^j. \tag{17}$$

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:

$$\lambda_i y + w_i(s) = \lambda_i y + \beta_i^0 + \sum_{j=1}^n \beta_i^j s_j = \lambda_i y + \tilde{\beta}_i^0 - \lambda_i \sum_{k=1}^n \tilde{\beta}_k^0 + \sum_{j=1}^n \left(\tilde{\beta}_i^j s_j - \lambda_i \sum_{k=1}^n \tilde{\beta}_k^j \right) =$$
$$\lambda_i \left(y - \sum_{k=1}^n \left(\tilde{\beta}_k^0 + \sum_{j=1}^n \tilde{\beta}_k^j s_j \right) \right) + \tilde{\beta}_i^0 + \sum_{j=1}^n \tilde{\beta}_i^j s_j =$$
$$\lambda_i \left(y - \sum_{k=1}^n \tilde{w}_k(s) \right) + \tilde{w}_i(s) = \lambda_i \pi + \tilde{w}_i(s). \tag{18}$$

□

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 $(\tilde{\beta}_j^0, \ldots, \tilde{\beta}_j^n)$ uniquely defines $(\beta_j^0, \ldots, \beta_j^n)$:

$$\begin{pmatrix} \beta_0^j \\ \beta_1^j \\ \ldots \\ \beta_n^j \end{pmatrix} = \begin{pmatrix} 1-\lambda_0 & -\lambda_0 & \ldots & -\lambda_0 \\ -\lambda_1 & 1-\lambda_1 & \ldots & -\lambda_1 \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_n & -\lambda_n & \ldots & 1-\lambda_n \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \ldots \\ \tilde{\beta}_n^j \end{pmatrix} \equiv \Lambda \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \ldots \\ \tilde{\beta}_n^j \end{pmatrix}. \tag{19}$$

Notice that

$$\sum_{i=0}^n \beta_i^j = 0. \tag{20}$$

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

$$a_i - \lambda_i \sum_{k=0}^{n} a_k = 0. \tag{21}$$

Let us sum equations (21) for i from 0 to n:

$$\sum_{i=0}^{n} a_i - \sum_{i=0}^{n} \lambda_i \sum_{k=0}^{n} a_k = \left(1 - \sum_{i=0}^{n} \lambda_i\right) \sum_{k=0}^{n} a_k = 0. \tag{22}$$

Since $\sum_{i=0}^{n} \lambda_i \neq 1$, equation (22) implies that $\sum_{k=0}^{n} a_k = 0$. Then, from equation (21) it follows that $a_i = 0$ for every i. Therefore, the columns of matrix Λ are independent, and it has full rank. □

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 (19) has rank n. On the other hand, the payments β_i^j must satisfy condition (20). 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 (20), the system of equations (19) 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:*

$$\begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \dots \\ \tilde{\beta}_n^{j*} \end{pmatrix} + \begin{pmatrix} x\lambda_0 \\ x\lambda_1 \\ \dots \\ x\lambda_n \end{pmatrix} \tag{23}$$

for some real number x. Also, if $(\tilde{\beta}_0^{j*}, \dots, \tilde{\beta}_n^{j*})$ is a solution to problem (19), then any $(\tilde{\beta}_0^j, \dots, \tilde{\beta}_n^j)$ is also a solution to problem (19) 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 (19) remains correct

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

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

$$\begin{pmatrix} 1-\lambda_0 & -\lambda_0 & \dots & -\lambda_0 \\ -\lambda_1 & 1-\lambda_1 & \dots & -\lambda_1 \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_{n-1} & -\lambda_{n-1} & \dots & 1-\lambda_{n-1} \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_{n-1}^j \end{pmatrix} = \begin{pmatrix} \beta_0^j \\ \beta_1^j \\ \dots \\ \beta_{n-1}^j \end{pmatrix}. \tag{24}$$

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

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

of system (24). Let us show that then

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

is a solution of equation (19). We already know that it is true for the first n equations of this system. Now, let us check if

$$\beta_n^j = \begin{pmatrix} -\lambda_n & -\lambda_n & \dots & 1-\lambda_n \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \dots \\ \tilde{\beta}_{n-1}^{j*} \\ 0 \end{pmatrix}. \tag{25}$$

Indeed, adding all the rows in equation (24), we obtain

$$\left(1 - \sum_{i=0}^{n-1} \lambda_i\right) \sum_{k=0}^{n-1} \tilde{\beta}_i^{j*} = \sum_{k=0}^{n-1} \beta_i^j. \tag{26}$$

Since $\sum_{i=0}^{n} \beta_i^j = 0$, $\beta_n^j = -\sum_{i=0}^{n-1} \beta_i^j$, and equation (26) 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 (25).

Relation between solutions

Take any two solutions $\tilde{\beta}^j$ and $\tilde{\beta}^{j*}$, which are solutions of equation (19), 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

$$\Lambda z = 0 \tag{27}$$

are vectors

$$z = \begin{pmatrix} z_0 \\ z_1 \\ \dots \\ z_n \end{pmatrix} = \begin{pmatrix} x\lambda_0 \\ x\lambda_1 \\ \dots \\ x\lambda_n \end{pmatrix}. \tag{28}$$

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 $\lambda_i, \lambda_k \neq 0$:

$$\frac{z_i}{\lambda_i} = \frac{z_k}{\lambda_k} \equiv x. \tag{29}$$

Then all solutions z of system (27) satisfy condition (28).

It is straightforward to check that any such solution satisfies condition (28) also solves system (27), so if $\tilde{\beta}^{j*}$ is a solution of system (19) then the set of solutions of system (19) is given by

$$\left\{ \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} \middle| \begin{pmatrix} \tilde{\beta}_0^j \\ \tilde{\beta}_1^j \\ \dots \\ \tilde{\beta}_n^j \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0^{j*} \\ \tilde{\beta}_1^{j*} \\ \dots \\ \tilde{\beta}_n^{j*} \end{pmatrix} + \begin{pmatrix} x\lambda_0 \\ x\lambda_1 \\ \dots \\ x\lambda_n \end{pmatrix}, x \in \mathbb{R} \right\}. \tag{30}$$

□

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 (20)) defined by a system of equations (19) 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 Proofs

B.1 Proof of Proposition 3.1

Proof. 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 (3) with respect to choice variables β_i^j yields first order conditions (equation (31) is for $j \neq i$):

$$\frac{\partial \mathcal{L}}{\partial \beta_i^j} = -\beta_i^j \gamma_i \sigma_{s_j}^2 - \eta_j = 0. \tag{31}$$

$$\frac{\partial \mathcal{L}}{\partial \beta_j^j} = \sum_{i=1}^{n} \frac{\lambda_i}{\mu_j} - \frac{\lambda_j + \beta_j^j}{\mu_j} - \beta_j^j \gamma_j \sigma_{s_j}^2 - \eta_j = \sum_{i \neq j} \frac{\lambda_i}{\mu_j} - \frac{\beta_j^j}{\mu_j} - \beta_j^j \gamma_j \sigma_{s_j}^2 - \eta_j = 0. \tag{32}$$

Comparing equation (31) 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 (31) and (32) yields:

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

Thus,

$$\beta_i^j = \frac{1}{\gamma_i} \left(\beta_j^j \left(\gamma_j + \frac{1}{\mu_j \sigma_{s_j}^2} \right) - \frac{\sum_{k \neq j} \lambda_k}{\mu_j \sigma_{s_j}^2} \right). \tag{33}$$

Sum expression (33) 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:

$$\beta_j^j = \frac{\sum_{i \neq j} \frac{1}{\gamma_i} \sum_{k \neq j} \lambda_k}{\mu_j \sigma_{s_j}^2 \left(\left(\gamma_j + \frac{1}{\mu_j \sigma_{s_j}^2} \right) \sum_{i \neq j} \frac{1}{\gamma_i} + 1 \right)} = \frac{\sum_{k \neq j} \lambda_k}{\mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\mu_j \sigma_{s_j}^2} + \frac{1}{\sum_{i \neq j} \frac{1}{\gamma_i}} \right)} = \frac{\sum_{k \neq j} \lambda_k}{1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{i \neq j} \frac{1}{\gamma_i}} \right)}. \tag{34}$$

Now, substitute β_j^j into expression (33) to obtain β_i^j:

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

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

$$-\frac{\frac{1}{\gamma_i}}{\sum_{k\neq j}\frac{1}{\gamma_k}}\left(\frac{\sum_{k\neq j}\lambda_k}{1+\mu_j\sigma_{s_j}^2\left(\gamma_j+\frac{1}{\sum_{k\neq j}\frac{1}{\gamma_k}}\right)}\right). \quad (35)$$

□

B.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:

$$\frac{\partial\mathcal{L}}{\partial\lambda_i}=\sum_{j=1}^{n}\frac{\lambda_j+\beta_j^j}{\mu_j}+\left(\sum_{j=1}^{n}\lambda_j\right)\frac{1}{\mu_i}-\mu_i\frac{1}{\mu_i}\left(\frac{\lambda_i+\beta_i^i}{\mu_i}\right)-\gamma_i\sigma_q^2\lambda_i+\theta_i-\chi=0. \quad (36)$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_i}=\left(\sum_{j=1}^{n}\frac{\lambda_j+\beta_j^j}{\mu_j}+\frac{1}{\mu_i}\sum_{j=1}^{n}(\lambda_j+\beta_j^j)-\frac{1}{\mu_i}\lambda_i\right)-\gamma_i\sigma_q^2\lambda_i+\theta_i-\chi=$$

$$\left(\sum_{j=1}^{n}\lambda_j\left(\frac{1}{\mu_j}+\frac{1}{\mu_i}\right)+\sum_{j\neq i}\frac{\beta_j^j}{\mu_j}-\frac{1}{\mu_i}\lambda_i\right)-\gamma_i\sigma_q^2\lambda_i+\theta_i-\chi=0. \quad (37)$$

Let us rearrange the term $\sum_{j\neq i}\frac{\beta_j^j}{\mu_j}$:

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

$$\sum_{k=1}^{n}\lambda_k\sum_{j\neq k}\frac{1}{\mu_j\left(1+\mu_j\sigma_{s_j}^2\left(\gamma_j+\frac{1}{\sum_{l\neq j}\frac{1}{\gamma_l}}\right)\right)}-$$

$$\sum_{k\neq i}\lambda_k\frac{1}{\mu_i\left(1+\mu_i\sigma_{s_i}^2\left(\gamma_j+\frac{1}{\sum_{l\neq i}\frac{1}{\gamma_l}}\right)\right)}$$

$$=\sum_{k=1}^{n}\lambda_k\sum_{j\neq k,i}\frac{1}{\mu_j\left(1+\mu_j\sigma_{s_j}^2\left(\gamma_j+\frac{1}{\sum_{l\neq j}\frac{1}{\gamma_l}}\right)\right)}. \quad (38)$$

Eventually,

$$\frac{\partial\mathcal{L}}{\partial\lambda_i}=\left(\sum_{k=1}^{n}\lambda_k\left(\sum_{j\neq k,i}\frac{1}{\mu_j\left(1+\mu_j\sigma_{s_j}^2\left(\gamma_j+\frac{1}{\sum_{l\neq j}\frac{1}{\gamma_l}}\right)\right)}+\frac{1}{\mu_i}+\frac{1}{\mu_k}\mathbb{I}_{i\neq k}\right)\right)-$$

$$\gamma_i\sigma_q^2\lambda_i+\theta_i-\chi=0. \quad (39)$$

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 $B_i=\frac{1}{\mu_i\left(1+\mu_i\sigma_{s_i}^2\left(\gamma_i+\frac{1}{\sum_{j\neq i}\frac{1}{\gamma_j}}\right)\right)}$

Equating χ in expressions (39) for i and m, we get:

$$\left(\sum_{k=1}^{n}\lambda_k\left(\sum_{j\neq k,i}B_j+A_i+A_k\mathbb{I}_{i\neq k}\right)\right)-\gamma_i\sigma_q^2\lambda_i+\theta_i=$$

$$\left(\sum_{k=1}^{n}\lambda_k\left(\sum_{j\neq k,m}B_j+A_m+A_k\mathbb{I}_{i\neq k}\right)\right)-\gamma_m\sigma_q^2\lambda_m+\theta_m. \quad (40)$$

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 $\sum_{j=1}^n \lambda_j = 1$,

$$(D_i - D_m) - \lambda_i(D_i + \gamma_i\sigma_q^2) = \theta_m - \theta_i - \lambda_m(D_m + \gamma_m\sigma_q^2). \tag{41}$$

If both agents i and m hold positive shares of the company, then $\theta_i = \theta_m = 0$, and condition (10) follows. □

B.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) (9) is non-positive.[26] If λ_i was zero, then the left-hand side of condition (9) 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

[26] 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 $D_k + \gamma_k\sigma_q^2 > 0$.

equation (9), 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:

$$\lambda_0 = 1 - \sum_{i=1}^{n} \lambda_i = 1 - \sum_{i=1}^{n} \frac{D_i}{D_i + \gamma_i \sigma_q^2}. \tag{42}$$

Thus, the outsider owns shares of the company if and only if λ_0 from expression (42) 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 $k, k+1, \ldots, n$ to own some stock of the company. If in the optimal allocation they all have positive shares, then all $\theta_k = \theta_{k+1} = \cdots = \theta_n = 0$. Solving the system of first order conditions (9) 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 $k, k+1, \ldots, n$ 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 $j \in \{k, k+1, \ldots, n\}$:

$$\lambda_j = \frac{1 - \sum_{i=k}^{n} \frac{D_i - D_j}{D_i + \gamma_i \sigma_q^2}}{\sum_{i=k}^{n} \frac{D_j + \gamma_j \sigma_q^2}{D_i + \gamma_i \sigma_q^2}}. \tag{43}$$

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 $k, k+1, \ldots, n$ 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 (11) 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 (11) for $k = 2$ (or n agents to be owners of the company and check condition (11) 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.[27] 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 $j = 0, \ldots, k-1$ have no shares. □

B.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[28]

$$\sum_{i=k+1}^{n} \frac{\hat{D}_i - \hat{D}_k}{\hat{D}_i + \hat{\gamma}_i \sigma_q^2} = \sum_{i=N \cdot \lceil \frac{k}{N} \rceil + 1}^{n} \frac{\hat{D}_i - \hat{D}_k}{\hat{D}_i + \hat{\gamma}_i \sigma_q^2} < 1. \tag{44}$$

[27] The process will end, because condition (11) is satisfied for $k = n - 1$.

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

We can order limit types by

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

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:

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

For a big enough N the types are sorted the same way as in the limit.[29] Since there are N identical agents of each type l, the left hand side of inequality (44) equals:[30]

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

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

$$\frac{N}{\sigma_q^2} \sum_{i=l+1}^{b} \frac{D_i^N - D_l^N}{\gamma_i}. \tag{45}$$

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:

$$\lambda_0 = 1 - N \sum_{i=1}^{b} \frac{D_i^N}{D_i^N + \gamma_i^N \sigma_q^2} \to 1. \tag{46}$$

If σ_q^2 grows slower than N ($\sigma_q^2 = o(N)$), then expression (45) 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.

[29] 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.

[30] Here l is the type corresponding to agent i in inequality (44), that is $l = \lceil \frac{k}{N} \rceil$.

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 (44) 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 $R_i < (\leq) R_j$ (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 (12)) 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

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

□

B.5 Proof of Proposition 5.1

Proof. The left hand side of expression (11) positively depends on D_i for $i > k$ ($\frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} = \frac{1 - \frac{D_k}{D_i}}{1 + \frac{\gamma_i \sigma_q^2}{D_i}}$), does not depend on D_i for $i < k$ and negatively depends on D_k (each term only has $-D_k$ in the numerator). Also, D_i negatively depends on μ_i, and D_k does not depend on μ_i for $k \neq i$. Suppose that agents $m, m+1, \ldots, n$ are owners of the firm.

Suppose μ_i decreases, then D_i increases. It means that the participation of agents k for $k > i$ is not affected, they are still owners of the firm till the point when D_i becomes equal to D_k. For agents $k < i$, the left hand side of expression (11) increases, so if it was not satisfied (agent k was not an owner), it is still not satisfied. If it was satisfied, at some point it might stop being be satisfied. So, some owners $k \neq i$ might stop being owners even if they were owners for lower μ_i, but all existing non-owners still have no shares. The left hand side of condition (11) for agent $i = k$ himself decreases with D_k, so if he was an owner, he will still be an owner. And even if he was not an owner, he might become one.

The effect of an increase in μ_i (decrease in D_i) is the opposite. All conditions (11) for $k > i$ are not affected, and D_i decreases, so they will never be affected. For $k < i$, the left hand side of expression (11) decreases when D_i decreases, so if it was satisfied, it will still be satisfied. It means that all existing owners remain owners, but some non-owners might become owners. For $k = i$ the left hand side of expression (11) increases when D_k decreases, so if it was not satisfied, it is still not satisfied. But even if it was satisfied (agent i was the owner), it might stop being satisfied. So, if agent i was not an owner before μ_i increased, he is still not an owner. □

B.6 Proof of Proposition 5.2

Proof. In this section we will show that if $\lambda_j > 0$, then $\frac{d\lambda_j}{d\mu_j} < 0$. If additionally $\lambda_k > 0$, then $\frac{d\lambda_k}{d\mu_j} > 0$. In other words, the share of agent m negatively depends on μ_j, and the shares of all other agents positively depend on μ_j.

Notice that D_i does not depend on μ_j for $j \neq i$.

First, consider the case when agents $m, m+1, \dots, n$ own shares of the firm. Equations (12) and (14) define how λ_j depends on the parameters of the model. Differentiating it with respect to the productivity parameter μ_j of agent j, we obtain:

$$\frac{d\lambda_j}{d\mu_j} = \frac{d\frac{1-\sum_{i=m}^{n}\frac{D_i-D_j}{D_i+\gamma_i\sigma_q^2}}{\sum_{i=m}^{n}\frac{D_j+\gamma_j\sigma_q^2}{D_i+\gamma_i\sigma_q^2}}}{d\mu_j} = \frac{\frac{dD_j}{d\mu_j}\sum_{i=m}^{n}\frac{1}{D_i+\gamma_i\sigma_q^2}}{\sum_{i=m}^{n}\frac{D_j+\gamma_j\sigma_q^2}{A_i-B_i+\gamma_i\sigma_q^2}} + \frac{\frac{dD_j}{d\mu_j}\sum_{i=1}^{n}\frac{1}{D_i+\gamma_i\sigma_q^2}\left(1-\sum_{i=1}^{n}\frac{D_i-D_j}{D_i+\gamma_i\sigma_q^2}\right)}{\left(\sum_{i=1}^{n}\frac{D_j+\gamma_j\sigma_q^2}{D_i+\gamma_i\sigma_q^2}\right)^2}. \tag{47}$$

Both terms of equation (47) are $\frac{dD_j}{d\mu_j}$ multiplied with a positive number. Therefore, the sign of $\frac{d\lambda_m}{d\mu_m}$ is the same as the sign of $\frac{dD_j}{d\mu_j}$. Looking at equation (8), we see that μ_m is in the denominator, hence, $\frac{dD_j}{d\mu_j}$ is negative. Thus, $\frac{d\lambda_m}{d\mu_m}$ is negative, which means that when the agent becomes less productive, she gets a lower share of the firm.

Let us see how agent ks share changes when agent j becomes less productive (μ_j

increases).

$$\lambda_k = \frac{1 - \sum_{i=m}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2}}{\sum_{i=m}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}} =$$

$$\frac{1 - \sum_{i=m, i\neq j}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} - \frac{D_j - D_k}{D_j + \gamma_j \sigma_q^2}}{\frac{D_k + \gamma_k \sigma_q^2}{D_j + \gamma_j \sigma_q^2} + \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}} = \frac{\frac{D_k + \gamma_j \sigma_q^2}{D_j + \gamma_j \sigma_q^2} - \sum_{i=m, i\neq j}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2}}{\frac{D_k + \gamma_k \sigma_q^2}{D_j + \gamma_j \sigma_q^2} + \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}} =$$

$$\frac{D_k + \gamma_j \sigma_q^2 - (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2}}{D_k + \gamma_k \sigma_q^2 + (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}}. \quad (48)$$

$$\frac{d\lambda_k}{d\mu_j} =$$

$$-\frac{dD_j}{d\mu_j} \frac{1}{\left(D_k + \gamma_k \sigma_q^2 + (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}\right)^2} \times$$

$$\left(\sum_{i=m, i\neq j}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} \left(D_k + \gamma_k \sigma_q^2 + (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2} \right) + \right.$$

$$\left. \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2} \left(D_k + \gamma_j \sigma_q^2 - (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} \right) \right) =$$

$$-\frac{dD_j}{d\mu_j} \frac{1}{\left(D_k + \gamma_k \sigma_q^2 + (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}\right)^2} \times$$

$$\left(\sum_{i=m, i\neq j}^{n} \frac{D_i - D_k}{D_i + \gamma_i \sigma_q^2} \left(D_k + \gamma_k \sigma_q^2\right) + \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2} \left(D_k + \gamma_j \sigma_q^2\right) \right)$$

$$= -\frac{dD_j}{d\mu_j} \frac{\sum_{i=m, i\neq j}^{n} \frac{D_i + \gamma_j \sigma_q^2}{D_i + \gamma_i \sigma_q^2} \left(D_k + \gamma_k \sigma_q^2\right)}{\left(D_k + \gamma_k \sigma_q^2 + (D_j + \gamma_j \sigma_q^2) \sum_{i=m, i\neq j}^{n} \frac{D_k + \gamma_k \sigma_q^2}{D_i + \gamma_i \sigma_q^2}\right)^2}. \quad (49)$$

Since $D_i \geq 0$ for all i and $\frac{dD_j}{d\mu_j} < 0$, $\frac{d\lambda_k}{d\mu_j} > 0$. Notice, that if the external investor owns some shares of the firm, then according to equation (14), the share each other agent holds depends only on D_k, but not on the productivity parameters of other agents, so it does not change with μ_j for $j \neq k$. Agent j's share decreases with μ_j in this case, because

$$\lambda_j = \frac{D_j}{D_j + \gamma_j \sigma_q^2} = \frac{1}{1 + \frac{\gamma_j \sigma_q^2}{D_j}},$$

so it positively depends on D_j, and D_j negatively depends on μ_j. All the shares the agent whose productivity decreases loses go to the outsider. □

B.7 Proof of Proposition 5.3

Proof. The proof is analogous to the proof of Proposition 5.1. The only difference is that D_i depends positively, not negatively on $\sigma^2_{s_i}$. □

B.8 Proof of Proposition 5.4

Proof. The proof of this proposition is analogous to the proof of Proposition 5.2. Indeed, notice that only variable D_i depends on $\sigma^2_{s_i}$, but none of the other variables. Then (both for $k \neq j$ and $k = j$):

$$\frac{d\lambda_k}{d\sigma^2_{s_j}} = \frac{d\lambda_k}{dD_j}\frac{dD_j}{d\sigma^2_{s_j}}.$$

So, the only difference when we differentiate λ_k with respect to $\sigma^2_{s_j}$, instead of μ_j, is the term $\frac{dD_j}{d\sigma^2_{s_j}}$ instead of term $\frac{dD_j}{d\mu_j}$. While the former term is negative, the latter term is positive,[31] so all claims of Proposition 5.2 are reverted. □

B.9 Proof of Proposition 5.5

Proof. By setting $k = 0$ in (11) we see that the outside investor owns a positive number of shares if and only if:

$$\sum_{i=1}^{n} \frac{D_i}{D_i + \gamma_i \sigma^2_q} < 1. \tag{50}$$

The left hand side of condition (50) strictly decreases with σ^2_q, and when $\sigma^2_q = 0$, it is equal to n, so the condition is not satisfied. When $\sigma^2_q \to \infty$, the left-hand side of condition (50) monotonically converges to zero, therefore there is a unique value of $\sigma^2_q = \bar{\sigma}^2_q$, such that if $\sigma^2_q \leq \bar{\sigma}^2_q$ the outsider does not own any shares of the company, and if $\sigma^2_q > \bar{\sigma}^2_q$, then the outsider owns a positive share. If the outside investor owns a positive share of the firm, then substituting $m = 0$ in equation (9) yields (15). □

[31]The variance $\sigma^2_{s_j}$ is in the denominator of the denominator, so the dependency of D_j on $\sigma^2_{s_j}$ in equation (8) is positive

C (For Online Appendix) Pareto optimal allocations as sums of certainty equivalents

In this section of the appendix we show that all Pareto optimal distributions in the CARA-normal model maximize the sum of certainty equivalents of the agents.

Lemma C.1. *Assume that the economy consists of n agents with von Neumann–Morgenstern CARA utility functions $u_i(X_i) = 1 - e^{-\gamma_i X_i}$ or linear utility functions $u_i(X_i) = w_i$, where X_i is the amount of money agent i gets, and money is transferable from one agent to the other.*[32] *Let $\mathcal{X}$ be the set of all feasible allocations $X(\varepsilon) = (X_1, \ldots, X_n)$ in the problem consists of linear combinations of normally distributed variables ε and all non-state-dependent transfers, then all Pareto optimal amongst these feasible allocations are solutions of the following problem:*

$$\max_{\mathcal{X}} \sum_{i=1}^{n} \left(\mathbb{E}X_i - \frac{\gamma_i}{2} \mathbb{V}X_i \right).$$

Before moving to the proof, let us notice, that typically the social planner's problem is given by

$$\max_{\mathcal{X}} \sum_{i=1}^{n} \nu_i \mathbb{E}u_i(X_i)$$

with *any positive coefficients* ν_i and the *utility functions* as summation terms. Here, instead, we add up *certainty equivalents* with *equal coefficients*.

Proof. Any Pareto optimal allocation is a solution of the following maximization problem for some i and $\{\bar{u}_j\}_{j \neq i}$:

$$\max_{X \in \mathcal{X}} \mathbb{E}u_i(X_i), \text{ s.t. } \forall j \neq i : \mathbb{E}u_j(X_j) \geq \bar{u}_j.$$

Since all utility functions are CARA, and $\mathcal{X}$ contains only linear combinations of normally distributed variables, both the maximized expression and the inequality constraints can

[32] If the normally distributed random variables are $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ and the transfers between agent i and agent j are t_{ij} (what agent i pays to agent j), then $X_i = \sum_{j=1}^{m} \zeta_i^j \varepsilon_j + \sum_{k=1}^{n} t_{ik}$ for some constant ζ_j^i and t_{ik}.

be equivalently rewritten in terms of certainty equivalents:[33]

$$\max_{X\in\mathcal{X}} \mathbb{E}X_i - \frac{\gamma_i}{2}\mathbb{V}X_i, \text{ s.t. } \forall j \neq i : \mathbb{E}X_j - \frac{\gamma_j}{2}\mathbb{V}X_j \geq -\frac{\log(1-\bar{u}_j)}{\gamma_j}.$$

The Lagrangian of this problem is:

$$\mathcal{L} = \mathbb{E}X_i - \frac{\gamma_i}{2}\mathbb{V}X_i + \sum_{j\neq i}\nu_j\left(\mathbb{E}X_j - \frac{\gamma_j}{2}\mathbb{V}X_j + \frac{\log(1-\bar{u}_j)}{\gamma_j}\right),$$

where ν_j are non-negative Lagrange multipliers of the corresponding constraints. We assumed that all non-state-dependent transfers are feasible, so denote a transfer from agent i to agent j as t_{ij}. The first order Kuhn-Tucker condition with respect to transfer t_{ij} yields:

$$\frac{\partial\mathcal{L}}{\partial t_{ij}} = 1 - \nu_j = 0.$$

Hence, all solutions maximizing the Lagrangian will have $\nu_j = 1$ for all j. The constrained problem becomes equivalent to the unconstrained problem of maximizing the Lagrangian with $\nu_j = 1$, which is in turn equivalent to maximizing

$$\max_{\mathcal{X}} \sum_{i=1}^{n}\left(\mathbb{E}X_i - \frac{\gamma_i}{2}\mathbb{V}X_i\right).$$

□

D (For Online Appendix) Check of concavity

In this section we will study concavity of the utility function.[34] Equation (36) shows first derivatives of the central planner's objective function with respect to λ_is (by the envelope theorem, the derivative if we substitute optimal β_i^j is equal to the full derivative). We will find the Hessian in λ's by substituting optimal βs into the first derivative (equation (39)), and then taking all second and cross derivatives.

[33]If agent j's utility function is linear, rather than CARA, substitute the corresponding $\gamma_j = 0$ and $-\frac{\log(1-\bar{u}_j)}{\gamma_j}$ is replaced with just $\bar{u}_j$. In any case, this is a constant which does not affect the proof below.

[34]Second derivatives of the Lagrangian do not contain any Lagrange multipliers, because all the constraints are linear in choice variables, so concavity of the Lagrangian is equivalent to the concavity of the objective function.

$$\frac{\partial^2 \mathcal{L}}{\partial \lambda_i^2} = \left(\sum_{j \neq i} \frac{1}{\mu_j \left(1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{l \neq j} \frac{1}{\gamma_l}}\right)\right)} + \frac{1}{\mu_i} \right) - \gamma_i \sigma_{q_i}^2. \tag{51}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \lambda_i \partial \lambda_j} = \left(\sum_{j \neq i} \frac{1}{\mu_j \left(1 + \mu_j \sigma_{s_j}^2 \left(\gamma_j + \frac{1}{\sum_{l \neq j} \frac{1}{\gamma_l}}\right)\right)} + \frac{1}{\mu_i} + \frac{1}{\mu_j} \right). \tag{52}$$

Notice that we can think of the objective function as a sum of two functions: $-\sum_{i=1}^{n} \gamma_i \sigma_{q_i}^2 \lambda_i^2$ and the remainder. The first first function is definitely concave, so if the second function is also concave, their sum is also concave. Therefore, from now on we will only consider the remainder of the function (eliminate the term $\gamma_i \sigma_{q_i}^2$).

To simplify future derivations, let us denote $A_i = \frac{1}{\mu_i}$ and $B_i = \frac{1}{\mu_i \left(1 + \mu_i \sigma_{s_i}^2 \left(\gamma_i + \frac{1}{\sum_{j \neq i} \frac{1}{\gamma_j}}\right)\right)}$. Notice that $B_i < A_i$, because the denominator is the same μ_i, but multiplied with $1 + \mu_i \sigma_{s_i}^2 \left(\gamma_i + \frac{1}{\sum_{j \neq i} \frac{1}{\gamma_j}}\right) > 1$. Let us show that for λ, such that $\sum_{i=1}^{n} \lambda_i = 1$ the function is concave. Concavity of function f with a constraint g is studied using the border Hessian:

$$\mathcal{H} = \begin{pmatrix} 0 & \frac{\partial g}{\partial \lambda_1} & \frac{\partial g}{\partial \lambda_2} & \cdots & \frac{\partial g}{\partial \lambda_n} \\ \frac{\partial g}{\partial \lambda_1} & \frac{\partial^2 f}{\partial \lambda_1^2} & \frac{\partial^2 f}{\partial \lambda_1 \partial \lambda_2} & \cdots & \frac{\partial^2 f}{\partial \lambda_1 \partial \lambda_n} \\ \frac{\partial g}{\partial \lambda_2} & \frac{\partial^2 f}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 f}{\partial \lambda_2^2} & \cdots & \frac{\partial^2 f}{\partial \lambda_2 \partial \lambda_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial \lambda_n} & \frac{\partial^2 f}{\partial \lambda_n \partial \lambda_1} & \frac{\partial^2 f}{\partial \lambda_n \partial \lambda_2} & \cdots & \frac{\partial^2 f}{\partial \lambda_n^2} \end{pmatrix}.$$

For all i the derivative $\frac{\partial g}{\partial \lambda_i} = 1$. Substituting $\frac{\partial^2 f}{\partial \lambda_i \partial \lambda_j}$ from equations (51) and (52), we can write down the Hessian:

$$\mathcal{H} = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & A_1 + \sum_{i \neq 1} B_i & A_1 + A_2 + \sum_{i \neq 1,2} B_i & \cdots & A_1 + A_n + \sum_{i \neq 1,n} B_i \\ 1 & A_1 + A_2 + \sum_{i \neq 1,2} B_i & A_2 + \sum_{i \neq 2} B_i & \cdots & A_2 + A_n + \sum_{i \neq 2,n} B_i \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_1 + A_n + \sum_{i \neq 1,n} B_i & A_2 + A_n + \sum_{i \neq 2,n} B_i & \cdots & A_n + \sum_{i \neq n} B_i \end{pmatrix}.$$

In order to show that the function is concave on $\sum_{i=1}^{n} \lambda_i = 1$, we need to show that the n principal minors (except for the first one) change signs (the first 2×2 minor is negative, the next 3×3 is negative, etc.).

Consider the $l+1$-st minor:

$$\mathcal{H}_l = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & A_1 + \sum_{i \neq 1} B_i & A_1 + A_2 + \sum_{i \neq 1,2} B_i & \dots & A_1 + A_l + \sum_{i \neq 1,l} B_i \\ 1 & A_1 + A_2 + \sum_{i \neq 1,2} B_i & A_2 + \sum_{i \neq 2} B_i & \dots & A_2 + A_l + \sum_{i \neq 2,l} B_i \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_1 + A_l + \sum_{i \neq 1,l} B_i & A_2 + A_n + \sum_{i \neq 2,n} B_i & \dots & A_l + \sum_{i \neq l} B_i \end{pmatrix}.$$

Subtract the first column multiplied by $\sum_{i=1}^{n} B_i$ from all other columns (it does not change the determinant). We obtain:

$$\begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & A_1 - B_1 & A_1 + A_2 - B_1 - B_2 & \dots & A_1 + A_l - B_1 - B_l \\ 1 & A_1 + A_2 - B_1 - B_2 & A_2 - B_2 & \dots & A_2 + A_l - B_2 - B_l \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_1 + A_l - B_1 - B_l & A_2 + A_n - B_2 - B_l & \dots & A_l - B_l \end{pmatrix}.$$

Denote $D_i = A_i - B_i$.

$$\begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & D_1 & D_1 + D_2 & \dots & D_1 + D_l \\ 1 & D_1 + D_2 & D_2 & \dots & D_2 + D_l \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & D_1 + D_l & D_2 + D_l & \dots & D_l \end{pmatrix}.$$

Now, subtract the last column from all columns, except the first column:

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & -D_l & D_2 - D_l & \dots & D_1 + D_l \\ 1 & D_1 - D_l & -D_l & \dots & D_2 + D_l \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & D_1 & D_2 & \dots & D_l \end{pmatrix}.$$

Now, subtract the last row from all other rows, except the first one:

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 1 \\ 0 & -D_1 - D_l & -D_l & \dots & D_1 \\ 0 & -D_l & -D_2 - D_l & \dots & D_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & D_1 & D_2 & \dots & D_l \end{pmatrix}.$$

The determinant of this matrix is equal to minus determinant of the matrix, which is obtained by eliminating the first and the last columns and rows:

$$\begin{pmatrix} -D_1 - D_l & -D_l & \dots & -D_l \\ -D_l & -D_2 - D_l & \dots & -D_l \\ \vdots & \vdots & \ddots & \vdots \\ -D_l & -D_l & \dots & -D_{l-1} - D_l \end{pmatrix}.$$

Subtract the last row from all other rows:

$$\begin{pmatrix} -D_1 & 0 & \dots & D_{l-1} \\ 0 & -D_2 & \dots & D_{l-1} \\ \vdots & \vdots & \ddots & \vdots \\ -D_l & -D_l & \dots & -D_{l-1} - D_l \end{pmatrix}.$$

This determinant is equal to:

$$(D_{l-1}) \sum_{i=1}^{l-1} \prod_{j \neq i} (-D_j) + (-D_{l-1} - D_l) \prod_{j \neq l, l-1} (-D_j).$$

This is equal to (expanding the last column):

$$\sum_{i=1}^{l} \prod_{i \neq j} (-D_j).$$

But for each i: $\text{sign}(\prod_{i \neq j}(-D_j)) = (-1)^{l-1}$, because $D_j > 0$. So, the original determinant has the same sign as $(-1)^l$. Hence, the signs of the principal minors alternate, and the function on $\sum_{i=1}^{n} \lambda_i = 1$ is concave.

E (For Online Appendix) Sample firms description

In this section we describe the firms used in examples in Section 5.

The agents in all examples are heterogenous and their parameters are mostly results of a random number generator (except for Figure 3).

The following example is used for Figures 1, 6 and 7.

Agent number	Production cost m_i	Variance of signal $\sigma_{s_i}^2$	Risk aversion γ_i
1	0.057	0.912	0.288
2	0.522	0.104	0.415
3	0.336	0.746	0.465
4	0.176	0.736	0.764
5	0.209	0.562	0.818
6	0.905	0.184	0.1
7	0.675	0.597	0.178
8	0.468	0.3	0.36

Table 3: Description of the agents for Figures 1, 6 and 7.

Output noise variance is $\sigma_q^2 = 4$ and it is close to the sum of variances of individual signals $\sum \sigma_{s_i}^2 = 4.141$.

Agent number	Production cost m_i	Variance of signal $\sigma_{s_i}^2$	Risk aversion γ_i
1	changes	0.083	0.677
2	0.412	0.720	0.988
3	0.603	0.996	0.767
4	0.751	0.355	0.337
5	0.584	0.971	0.662
6	0.552	0.346	0.244
7	0.584	0.887	0.296
8	0.512	0.455	0.680

Table 4: Description of the agents for Figure 2

Output noise variance is $\sigma_q^2 = 1.6$.

Agent number	Production cost m_i	Variance of signal $\sigma_{s_i}^2$	Risk aversion γ_i
1	changes	0.924	0.525
2	0.500	0.224	1.000
3	0.500	0.100	1.000
4	1.000	0.200	1.000

Table 5: Description of the agents for Figure 3

Output noise variance is $\sigma_q^2 = 0.3$.

Agent number	Production cost m_i	Variance of signal $\sigma_{s_i}^2$	Risk aversion γ_i
1	changes	0.544	0.318
2	0.639	0.721	0.119
3	0.545	0.522	0.940
4	0.647	0.994	0.646

Table 6: Description of the agents for Figure 4

Output noise variance is $\sigma_q^2 = 1.6$.

Agent number	Production cost m_i	Variance of signal $\sigma_{s_i}^2$	Risk aversion γ_i
1	changes	0.870	0.770
2	0.256	0.265	0.350
3	0.613	0.318	0.662
4	0.582	0.119	0.416
5	0.541	0.940	0.842

Table 7: Description of the agents for Figure 5

Output noise variance is $\sigma_q^2 = 0.5$.

References

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

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

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.

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. (2017). Moral hazard and subjective evaluation with multiple agents. *Working Paper*.

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

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

Levin, J. 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.

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