NBER WORKING PAPER SERIES

CONTRACTING WITH SYNERGIES

Alex Edmans
Itay Goldstein
John Y. Zhu

Working Paper 17606
http://www.nber.org/papers/w17606

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2011

We thank Patrick Bolton, Xavier Gabaix, Salvatore Miglietta, Alan Schwartz, and seminar participants from BI Oslo, Cambridge, Philadelphia Fed, and NBER Summer Institute (Law & Economics) for helpful comments, and Ali Aram for good research assistance. AE gratefully acknowledges financial support from the Dorinda and Mark Winkelman Distinguished Scholar award. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Contracting With Synergies
Alex Edmans, Itay Goldstein, and John Y. Zhu
NBER Working Paper No. 17606
November 2011
JEL No. D86,J31,J33

ABSTRACT

This paper studies optimal contracting under synergies. We define influence as the extent to which effort by one agent reduces a colleague's marginal cost of effort, and synergy to be the sum of the (unidimensional) influence parameters across a pair of agents. In a two-agent model, effort levels are equal even if influence is asymmetric. The optimal effort level depends only on total synergy and not individual influence parameters. An increase in synergy raises total effort and total pay, consistent with strong equity incentives in small firms, including among low-level employees. The influence parameters matter only for individual pay. Pay is asymmetric, with the more influential agent being paid more, even though the level and productivity of effort are both symmetric. With three agents, effort levels differ and are higher for more synergistic agents. An increase in the synergy between two agents can lead to the third agent being excluded from the team, even if his productivity is unchanged. This has implications for optimal team composition and firm boundaries. Agents that influence a greater number of colleagues receive higher wages, consistent with the salary differential between CEOs and divisional managers.

Alex Edmans
Wharton School
University of Pennsylvania
2318 Steinberg Hall - Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104
and NBER
aedmans@wharton.upenn.edu

Itay Goldstein
Wharton School
University of Pennsylvania
Philadelphia, PA 19104
itayg@wharton.upenn.edu

John Y. Zhu
Wharton School
University of Pennsylvania
Philadelphia, PA 19104
zhuyiran@wharton.upenn.edu

1 Introduction

Most work is conducted in teams. In these teams, agents' actions are typically synergistic – effort by one agent reduces the cost of effort for his colleague. For example, going on an international business trip is less costly to a manager if she has an efficient secretary; it is easier for a divisional manager to implement a new workforce practice if the CEO has developed a corporate culture that embraces change. Synergies are also important in non-corporate settings – the cost of giving an academic seminar is lower if one's coauthor has worked hard to improve the quality of the paper.

The structure of synergies within a team is complex. Within a given team, the contributions of each agent to the collective synergy are typically asymmetric. A CEO has a greater impact on the working environment of a divisional manager (through his choice of organizational structure, corporate culture, and communication policies) than the other way round; a conversation between a senior faculty member and a junior colleague usually benefits the latter more than the former. Moreover, the number of synergistic relationships that an agent will enjoy may vary across agents. A CEO likely exhibits synergies with each of his divisional managers, but a pair of divisional managers might not exhibit synergies with each other.

This paper studies an optimal contracting problem in the presence of such synergies. In our theory, agents contribute to the production of a joint project, which either succeeds or fails. We model synergies as follows. *Influence* refers to the extent to which effort by one agent reduces the marginal cost of effort of a colleague, and *synergy* between a pair of agents is the sum of the (unidirectional) influence parameters of the two agents. With more than two agents, a multidimensional *synergy profile* captures the synergies between all pairs of agents. Our framework allows for effort to be continuous, influence to be asymmetric across a given pair of agents, and agents to differ in the number of colleagues with whom they enjoy synergies and in the strength of these synergies. The model also allows for the production function to exhibit either complements or substitutes in the agents' effort levels, and shows that the effects of synergies are robust to the choice of production function.

Our analysis solves for the effect of synergies on the optimal effort level of each agent, the wage paid to a given agent if the project succeeds (both in absolute terms and relative to his colleagues), and the total wages paid out by the firm to all agents upon success. In particular, it addresses several questions that cannot be explored in a single-agent framework, such as the determinants of cross-sectional differences in pay across agents within the same firm, and the optimal composition of a team or boundaries of a firm. While standard models study how effort and pay depend on productivity and risk aversion, we study the effect of synergy.

We start with a two-agent model in which agents' efforts are perfect substitutes, i.e. the probability of project success depends on the sum of their actions. Even though the agents' influence can be asymmetric, and so one agent's effort is more "productive" in that it reduces his colleague's marginal cost more than his colleague's effort reduces his cost, the principal wants both agents to exert the same effort level. While the colleague's effort is less "productive", which would normally suggest that he should exert a lower effort level, the synergies arising from the first agent's action make it easier for the principal to induce effort from his colleague, and so she wishes the colleague to exert a similarly high level of effort. This result in the specific two-person model illustrates a more general point of the model (for any number of agents) – while influence parameters are individual and may be asymmetric across agents, the synergy is common to a group of agents. It is the common synergy, not the individual influence parameters, that determines the optimal effort level.

However, while effort levels are symmetric, wages are not. The more influential agent receives a higher wage upon success. Since the agent is paid zero upon failure, a higher success wage represents both higher incentives and a higher level of expected pay. This asymmetry in the wage occurs even though both agents exert the same effort level (and so pay is not simply a "compensating differential" for the disutility of effort), and effort by each agent has an identical effect on the production function. Instead, higher pay is optimal because it causes the agent to internalize the externalities he exerts on his colleague. When choosing his effort level, each agent takes his colleague's action as given, and so he does not take into account the impact on his colleague's cost of effort. A higher wage causes him to internalize this synergy, and so leads him to exerting the optimal level of effort.

An increase in the overall level of synergies between the agents leads to the principal implementing a higher effort level, and paying out higher total wages in the case of success. This result contrasts standard principal-agent models without synergies, in which total wages are independent of productivity parameters in the presence of risk neutrality. If an agent is more productive, the principal wishes to implement a higher level of effort (which requires steeper incentives, *ceteris paribus*), but greater productivity means that flatter incentives are required to implement a given level of effort, and these two effects exactly offset. In our model, the second effect is absent because synergy only affects cost functions and not an agent's marginal effect on production, and so an increase in synergy unambiguously leads to higher total wages. The result is consistent with the high level of equity incentives in start-up firms, including to rank-and-file employees with little direct effect on firm output. Standard principal-agent theory suggests it is never optimal to give equity incentives to a low-level employee as he has little effect on the equity price and so equity would merely subject him to risk outside his control. However, particularly in start-up firms where job descriptions are blurred and workers interact frequently with each other, agents can have a significant indirect effect on firm value through aiding their colleagues. In addition, in small firms with a shallower hierarchy, a junior employee is more likely to interact with a senior colleague.

An increase in agent i's influence parameter, holding agent j's influence parameter constant, raises total synergies and so increases total effort and total wages as explained above. Agent i's wage always increases, but the effect on agent j's wage is more nuanced. It increases if and only if his influence parameter is above a critical threshold, otherwise it decreases. The intuition is as follows. The principal could choose to hold agent j's wage constant, in which case an increase in i's influence parameter raises j's effort level because it reduces his marginal cost of effort. Thus, the principal could reduce agent j's wage slightly without his effort falling below its previous level. If agent j's influence is sufficiently low, his effort is less beneficial to the team than agent i's effort. Then, the principal prefers to extract some of the surplus (created by agent j's lower cost of effort) and reinvest it in agent i. This is achieved by lowering agent j's wage, accepting a lower increase in his effort, and reinvesting a portion of the saved cash to further increase agent i's wage. By contrast, if agent j's influence is sufficiently high, the principal chooses to reinforce the increase in j's effort level by augmenting his wage. An increase in agent i's *relative* influence – augmenting his influence parameter but decreasing agent j's to keep the total synergy constant – causes agent i's wage to increase in both absolute terms, and relative to agent j. In short, synergy determines the (common) effort level and total pay, and influence determines the agents' relative pay.

While the two-agent model fixes some basic ideas in a parsimonious manner, the core analysis of the paper is a three-agent model which allows us to study differences in the scope of synergies exerted by agents – such as the earlier example whereby a CEO influences two divisional managers, but the divisional managers do not influence each other. The "synergy component" refers to the sum of the bilateral influence parameters between a given pair of agents: i.e. agent i's influence on agent j, plus agent j's influence of agent i. There are three synergy components, one for each pair of agents. If the synergy components are sufficiently close to each other, all agents exert strictly positive effort, and the ratio of the effort (and thus wage) levels depends on the relative magnitude of all three synergy components. For example, if agent 1 exhibits more synergies with agent 3 than does agent 2, then agent 1 will exert a higher effort level than agent 2. This contrasts the two-agent case, where both agents take the same action. Note that the relative effort levels depend on the *total* synergies between each pair of agents, rather than the unidirectional influence parameters. It may seem that the optimal effort level exerted by agent 1 should depend only on his influence on agent 3, and not 3's influence on 1, since only the former affects the productivity of his effort. However, if 3 has a greater influence on 1, it is less costly for the principal to induce effort from 1, and so the optimal effort level depends on the total synergy. As in the two-agent model, the optimal effort levels depend on the collective synergy, rather than the individual influence parameters; the latter only affect relative pay.

A natural application of the three-agent model is a setting where one synergy component is close to zero – for example, if two divisional managers exhibit synergies with the CEO but less so with each other – then the two non-synergistic agents can be aggregated into a single employee and the model reduces to a close variant of the two-agent case. Thus, the CEO exerts almost the same effort level as the two divisional managers combined, and so his level of pay is also higher than each divisional manager. Bebchuk, Cremers and Peyer (2011) interpret a high level of CEO pay compared to other senior managers as inefficient rent extraction, but we show that it can be optimal given the broad scope of a CEO's activities. In addition, this result suggests that the optimal measure of firm size that determines CEO pay, in assignment models such as Gabaix and Landier (2008) and Terviö (2008), may not be an accounting measure such as assets or profits (as typically used in empirical studies), but the scope of a CEO's influence. The CEO of a large firm where divisions operate independently (such as a conglomerate) may be paid less than the manager of a small synergistic firm, such as a start-up.

If one synergy component becomes sufficiently large compared to the other two, then the model collapses to the two-agent setting. Intuitively, if the synergy between two agents is sufficiently strong, then only these two agents matter for the principal – she ignores the third agent and induces zero effort from him, even though he has the same direct effect on the production function as the first two agents. This also means that the third agent's participation depends on circumstances outside his control – in contrast to standard models in which an agent's effort level depends only on parameters specific to him. Even if his own synergy parameters do not change, if the synergy component between his two colleagues suddenly increases, this can lead to him being excluded. This is because the increased synergy between his colleagues raises the value of the firm, and thus the cost of giving the third agent equity to induce effort. This result has interesting implications for the optimal composition of a team – if two agents enjoy sufficiently high synergies with each other, there is no gain in adding a third agent, even if he has just as high a direct impact on the production function as the first two. Similarly, if the agents are interpreted as divisions of a firm, the model has implications for the boundaries of a firm and suggests which divisions should be added, divested or retained. Conventional wisdom suggests that a division should be divested only if it does not exhibit synergies with the rest of the conglomerate. However, here, even if a division enjoys strictly positive synergies, it should still be divested if its synergies are lower than those enjoyed by the other divisions – i.e. it is relative, not absolute, synergies that matter for the boundaries of the firm. Similarly, a firm should not automatically acquire a target even if it will generate strictly positive synergies in absolute terms.

We finally consider a model of perfect complements, where the success probability depends only on the minimum effort level across all agents. Even though the production function is a polar opposite, the model's core results remain robust. An increase in total synergy augments the effort levels and pay of all agents; a rise in the relative influence of a single agent raises his pay in both relative and absolute terms.

Our study builds on the literature on multi-agent principal-agent problems. Holmstrom (1982) considers two team-based settings. Where agents contribute to a joint output, a free-rider problem exists. Where each agent has his own output measure, the principal can use relative performance evaluation to reduce the noise in evaluating each agent. There are no synergies in his model: effort by one agent has no effect on the marginal productivity or marginal cost of another agent's effort. In the individual-output model, there is no interaction between the agents; in the joint-output model, the only interaction stems from a joint production function in which the efforts are perfect substitutes rather than exhibiting complementarities. A rich literature, summarized by Bolton and Dewatripont (2005, Chapter 8), has built on both of these settings, analyzing further interesting questions such as the possibility of collusion between agents, mutual monitoring between agents, and the optimal structuring of a team into hierarchies, but do not consider synergies. Itoh (1991) studies a multi-tasking problem where agents take two actions: one increases their own output, and another increases his colleague's output. This contrasts our setting where there is a single output across the team, and each agent takes a single action which both improves the joint output and reduces his colleague's marginal cost – thus, the productive action is also synergistic and the contract must change to force the agent to internalize this externality. Some of the subsequent literature on team-based incentives has focused on the free-rider problem in settings that involve complementarities in the production function under a joint output. Che and Yoo (2001) extend the free-rider problem to a repeated setting, where an agent can threaten to punish a shirking colleague by shirking himself in a future period. Kremer (1993) studies the case of extreme complementarities in production, when failure in one agent's task leads to automatic failure of the joint project, although agents do not make an effort decision. Winter (2004) extends this framework to incorporate a binary effort choice and shows that it may be optimal to give agents different incentive schemes even if they are ex ante homogenous. Extending this framework further, Winter (2006) studies how the optimal contract depends on the sequencing of agents' actions, and Winter (2010) shows how it depends on the information agents have about each other. Gervais and Goldstein (2007) analyze optimal contracting in a model with production complementarities and agents with self-perception biases. Sakovics and Steiner (2011) study optimal subsidies where there are complementarities in production.

We show in the paper that complementarities in the production function are inherently different from the synergies studied in our paper. In our paper, effort by one agent reduces the marginal cost of effort of his colleague. This can also be interpreted as an agent's effort increasing the marginal private benefit of effort by his colleague – for example, giving an academic seminar is more enjoyable if one's coauthor has worked hard on the paper. Regardless of whether we interpret an agent's effort as affecting his colleague's private cost or private benefit, the agent does not take into account this externality when making his effort decision, and has to be compensated differently to internalize it. On the other hand, in models with complementarities in the production function, the agent does internalize the effect his effort has on his colleagues' productivity, because he receives a share of the output. Thus, when the production complementarity increases, the agent will raise his effort level even if the contract is held constant – the contract does not need to change to cause him to internalize his externality. In a single-agent model, modifying the production function is isomorphic to modifying the cost function; in a multi-agent model, complementarities in production are fundamentally different from complementarities in costs.

Of closest relevance to our paper are other models of contracting with externalities. Kandel and Lazear (1992) study peer pressure, whereby an agent's effort affects the utility of other agents. Their focus is on demonstrating how to model a peer pressure situation, rather than solving for the optimal contract. In Segal (1999), agents exert externalities on each other through their impact on other agents' reservation utilities rather than cost functions. The agents' actions are observable participation decisions (e.g. the decision to buy a product) rather than the choice of an unobservable effort level; there is no output or production function as in this paper. Studying the optimal effort choice (out of a continuum) rather than a zero-one participation decision leads to several new results, such as the effect of total synergy on the optimal effort level, that efforts may be symmetric even if influence is asymmetric, and that the optimal effort level of an agent may be zero even if he enjoys strictly positive synergies. In addition, while we focus on the optimal contract, Segal's focus is on what outcomes are achievable and the bulk of the analysis concerns symmetric externalities. Bernstein and Winter (2010) also focus on a participation decision, as in Segal (1999), and study the case of heterogeneity in externalities. Dessein, Garicano, and Gertner (2010) study the optimal allocation of tasks under economies of scale, which they refer to as synergies. This is a different concept from the synergy in our paper, where effort by one agent reduces the cost of effort of another agent.

The paper proceeds as follows. Section 2 presents the most general version of the model, which we then specialize to a perfect substitutes production function in Section 3. We start with the preliminary two-agent model and then move to the core three-agent model. Section 4 analyzes a perfect complements production function and shows that the core results are robust, and Section 5 concludes. Appendix A contains all proofs not in the main text.

2 The Generic Model

This section outlines our general synergy model. Section 3 later specializes the model to the case where agents' outputs are substitutes, and Section 4 considers the case of complements.

There is a risk-neutral principal ("firm"), and N risk-neutral agents ("workers") indexed $i = 1, 2, \ldots N$. Each agent is protected with limited liability and has a reservation utility of zero. Each agent exerts an unobservable effort level

$$p_i \in [0,1] \quad i = 1, 2, \dots N.$$

The agents' efforts affect the firm's output. The firm has two possible output levels, $r \in \{0,1\}$. The output level is publicly observable and contractible. We will sometimes refer to $r = 1$ as "success" and $r = 0$ as "failure". The probability of success depends on effort levels p of all agents as follows:

$$\Pr(r = 1) = P(p_1, p_2, \dots p_N). \tag{1}$$

Each agent's cost of effort $c_i(p)$ depends not only on his own effort level p_i, but also the effort levels exerted by all other agents. We specify agent i's cost function as:

$$c_i(p) = h_i(p_i) \left(1 - \sum_{j \neq i} \varepsilon_{ji} p_j\right) \qquad i = 1, 2, \dots N, \tag{2}$$

where

$$\begin{aligned} \varepsilon_{ij} &\geq 0 \quad 1 \leq i \neq j \leq N \\ \forall i, \ \sum_{j \neq i} \varepsilon_{ji} &< 1. \end{aligned}$$

The variable ε_{ij} is an *influence parameter* that represents the influence agent i exerts on agent j. The higher the influence parameter, the greater the extent to which effort by agent i reduces the cost of effort of agent j. A central feature of our model reflected in (2) is that the effort by agent i reduces the marginal cost of effort by agent j. This is the source of the synergistic relations among agents in our model: when an agent exerts more effort, he makes it less costly for other agents to exert more effort as well. We will sometimes refer to $h_i(p_i)$ as agent i's *individual* cost function, to distinguish it from the "all-in" cost function $c_i(p)$. The influence parameters ε_{ij} and the individual cost functions $h_i(p_i)$ are common knowledge before contracting takes place. For now we consider the case of non-negative influence parameters; in Section 3.3 we extend the model to allow for $\varepsilon_{ij} < 0$.

It is automatic that each agent i will be paid zero in the case of failure. The principal wishes to solve for the optimal wage $w_i \geq 0$ to pay agent i in the case of success. The timing of the model is such that the principal chooses the wages w_i for every agent i. The wage w_i of agent i is then common knowledge to all agents. Then, given the wages, each agent i chooses his effort level p_i to maximize his expected utility, given by his wage minus his cost of effort, i.e.:

$$w_i \mathbf{1}_{r=1} - c_i\left(p\right). \tag{3}$$

Agents choose their levels of effort simultaneously, and their effort levels constitute a Nash Equilibrium.

The solution of the model is then given as follows. The principal maximizes her payoff, given by the expected output net of wages paid to the agents, i.e., she solves:

$$\max_{\{p_i\},\{w_i\}} P(p_1, p_2, \ldots p_N)\left(1 - \sum_i w_i\right), \tag{4}$$

subject to the incentive compatibility (IC) conditions for each agent i:

$$p_i \in \arg\max_p P(p_1, \ldots p_{i-1}, p, p_{i+1}, \ldots p_N) w_i - h_i(p)\left(1 - \sum_{j \neq i} \varepsilon_{ji} p_j\right) \qquad i = 1, 2, \ldots N. \tag{5}$$

Before we move to analyze specific cases of the model, a couple of points about the setup are worth making. First, as specified above, agent i sets his effort p_i without observing the effort levels of other agents (but rather only correctly expecting them in equilibrium). Since the cost of agent i's effort depends on the effort levels exerted by other agents, this implies that agent i decides on his own effort without observing the implied cost (only correctly expecting it in equilibrium). We think this is a realistic feature of the model. For example, a CEO may commit to a business trip and exert effort in advance to make it successful, but the exact cost she bears in making the trip will depend on the level of preparation conducted by her secretary, which is not known to the CEO until after the trip is completed. Alternatively, the cost function c_i may combine elements of private benefit – e.g., the extent to which the CEO enjoys her trip – which again depend on the effort by other agents in the firm.

Second, since the agent is paid zero upon failure (which is a consequence of the combination of risk neutrality, limited liability, and zero reservation utility), an increase in w_i corresponds to an increase in both incentives (the sensitivity of pay) and expected pay (the level of pay, which is often referred to as the "wage" in empirical studies). Thus, in the analysis that follows, all results pertaining to w_i are predictions for both the level and sensitivity of pay. Both move in the same direction: an increase (decrease) in w_i raises (reduces) both. These predictions do not hinge upon our assumption of risk neutrality but will continue to hold in a model with risk aversion and a binding participation constraint. An increase in the sensitivity of pay will cause the agent to demand a risk premium, augmenting the level of pay.

3 Substitute Effort

This section specializes the general production function (1) to the case in which the agents' efforts are perfect substitutes, i.e. the probability of success depends on the aggregate effort undertaken by all agents. Section 3.1 considers a preliminary two-agent model, as this version of the model is most tractable and illustrates the core ideas most clearly. Section 3.2 considers a three-agent model which is the core focus of the paper.

3.1 The Preliminary Two-Agent Model

The production function (1) now specializes to*:*

$$\Pr(r = 1) = \frac{p_1 + p_2}{2}. \tag{6}$$

We assume a quadratic individual cost function:

$$h_i(p_i) = \frac{1}{4}p_i^2.$$

Differentiating agent i's expected utility function (5) gives his first-order condition as:

$$w_i = p_i(1 - \varepsilon_{ji}p_j).$$

Plugging this into the principal's objective function (4) gives her reduced-form maximization problem as:

$$p_1^*, p_2^* \in \arg\max_{p_1,p_2} \frac{p_1 + p_2}{2} \left(1 - (p_1 + p_2) + p_1 p_2(\varepsilon_{12} + \varepsilon_{21})\right). \tag{7}$$

We define the following term:

Definition 1 ***Synergy*** *is defined to be the sum of the influence parameters* $s = \varepsilon_{12} + \varepsilon_{21}$.

We also make the following assumption to resolve cases in which the principal is indifferent between two contracts:

Assumption 1 *When computing the optimal contract, if the principal is indifferent between two arrangements A and B, and A is preferred by all agents over B, then A is chosen.*

The solution to the model and its properties are given by Proposition 1 below.

Proposition 1 *(Substitute production function, two agents.) (i) For all nonzero synergy, optimal efforts are equal: $p_1^*(s) = p_2^*(s) \equiv p^*(s)$. There exists a critical synergy level $s^* > 0$ such that*

$$p^*(s) = \begin{cases} \frac{2-\sqrt{4-3s}}{3s} & s \in (0, s^*) \\ 1 & s \geq s^*. \end{cases}$$

Optimal effort $p^*(s)$ is strictly increasing on $(0, s^*]$ and explodes to 1 at s^*. When there is no synergy, any combination of efforts that sum to $\frac{1}{2}$ is optimal.

(ii) Total wages given success, $w_1^* + w_2^*$, and expected total wages $\frac{p_1^*+p_2^*}{2}(w_1^* + w_2^*) = p^*(w_1^* + w_2^*)$ are both increasing in s on $(0, s^*]$.

(iii) Suppose synergy is subcritical. An increase in either influence parameter will lead to increases in optimal effort, total wages given success, and expected total wages.

(iv) Suppose synergy is subcritical. The more influential agent receives the higher wages upon success, i.e. $w_1^* > w_2^*$ if and only if $\varepsilon_{12} > \varepsilon_{21}$.

(v) Fix a subcritical synergy level. An increase in agent i's relative influence (i.e. increasing ε_{ij} and lowering ε_{ji} so that s is unchanged) increases both his relative and absolute wealth. Specifically,

$$\frac{w_i^*}{w_j^*}, \ \frac{w_i^*}{w_i^* + w_j^*}, \ w_i^* \text{ and } p^* w_i^* \quad \text{all strictly increase.}$$

(vi) Suppose synergy is subcritical. An increase in ε_{ij} increases w_i^* and $p^* w_i^*$. The effect on w_j^* depends on ε_{ji} as follows: An increase in ε_{ij} leads to an increase in w_j^* if and only if ε_{ji} is sufficiently high. Specifically,

$$\frac{d}{d\varepsilon_{ij}} w_j^* \begin{cases} > 0 & \varepsilon_{ji} \in (\frac{1}{6p^*(s)}, s^* - \varepsilon_{ij}) \\ = 0 & \varepsilon_{ji} = \frac{1}{6p^*(s)} \\ < 0 & \varepsilon_{ji} \in [0, \frac{1}{6p^*(s)}) \end{cases} . \tag{8}$$

Finally, $\frac{d}{d\varepsilon_{ij}}p^* w_j^*$ is always positive.
(vii) The more influential agent receives the higher utility.

We now discuss the intuition behind and implications of each part of the Proposition. Part (i) states that each agent exerts the same effort level. This result may appear surprising as it seems efficient for the more influential agent to exert the greater effort level. It is tempting to consider influence as a component of an agent's "productivity", and conclude that the more influential agent is more "productive" and so should work harder. However, this is not the case, because greater effort by the more influential agent decreases the cost of effort of his colleague, inducing the latter to exert more effort. Mathematically, from the principal's reduced-form objective function (7), we can see that the cost saving due to synergy is given by $p_1 p_2 (\varepsilon_{12} + \varepsilon_{21})$. It thus depends on the product $p_1 p_2$ and is highest when $p_1 = p_2$. Assume without loss of generality that agent i is the more influential agent. If $p_i > p_j$, then the principal is not benefiting much from agent i's influence on agent j since agent j is exerting little effort; thus, it is optimal for her to increase p_j. If $p_i < p_j$, then the principal should increase p_i to allow agent j (who is exerting high effort) to benefit from agent i's influence. In sum, for the principal to gain from the synergy, she needs p_i to be high to reduce j's cost of effort, and p_j also to be high so that j benefits from this reduced cost of effort.

The expression for the cost saving also shows that it is only the sum of influence parameters – i.e., the synergy s – that matters for the determination of equilibrium effort levels $p^*(s)$. The components ε_{12} and ε_{21} themselves do not matter beyond their sum. To glean the intuition, the synergy can be thought of as an "echo" between the two agents – the influence of agent i on agent j raises the optimal effort level for agent i, which reduces the cost of effort for agent j, which raises the optimal effort level for agent j, which, due to the influence of agent j on agent i, reduces the cost of effort for agent i, which raises the optimal effort level for agent i, and so on. In this process, it is the combination of influence parameters (here, their sum) that determines the optimal efforts of the two agents and, in the current specification, pulls them closer to each other.

Intuitively, we see that as the synergy parameter s increases, effort by each agent is more productive – in addition to its (unchanged) direct effect on firm output, it now has a greater effect on the other agent's cost function, and so it is efficient for the principal to implement a higher effort level. When the synergy crosses a threshold s^*, the optimal effort level jumps discontinuously to its maximum value of 1. Essentially, when $s < s^*$, the echo between the two agents is dampening and the solution is interior. When

$s > s^*$, the synergy is so strong that the echo is amplifying and the model "explodes", leading to the maximum effort being optimal.

Part (ii) states that wages increase with synergy. While intuitive, this result is far from automatic. With greater synergies, it is efficient to implement a higher effort level, which requires a higher wage holding all else equal. However, it seems that there is a counteracting effect in the opposite direction – when synergies are higher, each agent's cost of effort is lower, and so a lower wage is required to implement a given effort level. Indeed, in a single-agent moral hazard model under risk neutrality and limited liability, the optimal contract involves paying the agent one-half of the firm's output, regardless of the agent's productivity or cost of effort, because these two effects exactly offset each other.

Here, wages are unambiguously increasing in the synergy parameter s. The key is that synergies have no direct effect on output. The synergy parameter does not appear in the production function and so does not affect the direct marginal productivity of an agent's effort. It only affects output indirectly through changing the other agent's cost of effort and in turn affecting his effort choice. In a Nash equilibrium, when choosing his effort level, agent i takes agent j's effort choice as given and does not take into account the effect he has on agent j's cost function. Thus, he does not internalize his externality, and so the principal chooses to give him a sharper contract to cause him to do so. Importantly, this result illustrates the difference between our approach of modeling the complementarity between the agents in the cost function (or private benefit function), and an alternative approach of modeling it in the production function. Under the alternative approach, wages would be independent of the complementarity. We will return to this point in Section 3.4.

Part (ii) implies that total wages, as a fraction of output, will be higher in firms in which synergies are greater. Moreover, these higher wages come in the form of performance-sensitive pay. This is a potential explanation for why high equity incentives are given to low-level employees, even though they may have a small direct effect on output. High equity incentives are optimal if they have a large indirect effect by changing another agent's cost function – for example, an efficient analyst in a private equity firm reduces the cost of a director going to a meeting by producing accurate briefing materials. Synergies are likely particularly high in small and young firms, where job descriptions are often blurred and interactions are frequent. This may explain why incentive-based compensation is particularly high in start-ups, even among low-level employees – as was the case in firms such as Google. Hochberg and Lindsey (2010) document systematic evidence of broad-based option plans. Note that our model can only explain equity compensation to rank-and-file employees that exert significant synergies on a sufficiently large number of people. If firms grant equity to non-synergistic employees, this is likely for alternative reasons already in the literature.[1]

Part (iii) follows naturally from parts (i) and (ii). Since an increase in a single influence parameter, holding the other influence parameter constant, raises the total synergy level s, it will raise the effort levels of both agents, total wages, and expected total wages. However, a single influence parameter has no independent effect on effort and total wages other than through its impact on the total synergy. Total synergy is a "sufficient statistic" for effort and total wages – how the synergy is divided between the two influence parameters does not matter. The influence parameters do have an independent effect on the relative pay of each employee, as shown in part (iv). The more influential agent receives the higher wage. This result holds even though both agents exert the same level of effort, so the higher wage is not merely a "compensating differential" for the disutility of exerting a higher level of effort. It also holds even though the agents have the same direct productivity in the production function (6): each agent's task is equally important to firm value. Instead, the wage differential is driven purely by the indirect effect each agent has on his colleague. Part (iv) leads to empirical predictions for within-firm differences in pay: more influential agents should receive higher wages, even if all the tasks they perform are the same. For example, senior faculty within academic departments are paid more than junior faculty even though they all have the same formal job description (teaching courses and writing papers); the former can reduce the latter's cost of effort through mentorship and guidance.

Part (v) analyzes the effect of an increase in agent i's relative influence: it increases agent i's wage both in absolute terms and also relative to agent j's wage. Since agent i is exerting a greater externality, it is efficient to pay him more to cause him to internalize this externality.

While part (iii) shows that an increase in i's influence parameter augments total wages, part (vi) studies the effect on the individual wages of each agent. It is clear that agent i's wage rises, since total wages rise (part (iii)) and i's share of total wages

[1] Oyer (2004) justifies broad-based option plans from a retention perspective: options are worth more when employees' outside options are higher, persuading them to remain within the firm. Oyer and Schaefer (2005) find support for both this explanation and the idea that option compensation screens for employees with desirable characteristics. They do not test our synergy explanation, which has not been previously proposed to our knowledge. Bergman and Jenter (2007) present theory and evidence that option plans are used to take advantage of employees' irrational overvaluation of their firm's options.

rises due to his greater relative influence (part (v)). However, there are two conflicting effects on agent j's absolute wage: total wages rise, but j's share of total wages falls. Part (vi) characterizes which force dominates when. If the principal held w_j^* constant, the rise in ε_{ij} would increase agent j's effort because it reduces his marginal cost of effort. However, the principal need not hold w_j^* constant. She could choose to decrease j's wage and thus extract part of the "surplus" created by the rise in ε_{ij} by paying j less; in return she accepts a smaller (but still positive) increase in agent j's effort. Put differently, since j's marginal cost of effort has fallen, it is cheaper to induce effort from him and she takes advantage of this by lowering his wage. Alternatively, she could increase j's wage and reinforce the increase in j's effort brought about by the rise in ε_{ij}. Put differently, since it is cheaper to induce effort from j, she can take advantage of this by increasing j's effort even further (above and beyond the increase already occurring from the rise in ε_{ij}) via a higher wage. The latter option is desirable if j's effort is particularly beneficial to the team, i.e., if j's influence on i is particularly high.

Moreover, the threshold level of ε_{ji}, $\frac{1}{6p^*(s)}$, is decreasing in the common effort level and thus the common synergy – i.e., the higher the synergy, the greater the range of parameters ε_{ji} under which agent j's wage increases. As explained earlier, the synergy creates an "echo" between the agents which amplifies the effect of changes in a parameter on the equilibrium. If the echo is strong enough, the increase in ε_{ij} causes such a large increase in total wages that it outweighs the fall in j's share of the total wage pool. Thus, j's wage rises in absolute terms. While the change in j's absolute wage depends on ε_{ji}, the expected wage $p^* w_j^*$ unambiguously rises (regardless of ε_{ji}), due to the increase in the optimal effort level p^* from part (iii).

Finally, part (vii) compares the utility of the two agents. The more influential agent receives a higher wage, but also has a higher cost function since he is helped out less by his colleague. The Proposition shows that the first effect is stronger, and so the more influential agent receives the higher utility.

3.2 The Main Three-Agent Model

We now present the three-agent model which is the core analysis of this section. The production function (1) now specializes to:

$$\Pr(r = 1) = \frac{p_1 + p_2 + p_3}{3}. \tag{9}$$

and we continue to assume a quadratic individual cost function, which is now given by:

$$h_i(p_i) = \frac{1}{6}p_i^2.$$

Differentiating agent i's utility function (3) gives his first-order condition as:

$$w_i(p_i) = p_i \left(1 - \sum_{j \neq i} \varepsilon_{ji} p_j\right),$$

and plugging this into the principal's objective function (4) gives her reduced-form maximization problem as:

$$p_1^*, p_2^*, p_3^* \in \arg \max_{p_1, p_2, p_3 \in [0,1]} \frac{(p_1 + p_2 + p_3)}{3} \left(1 - (p_1 + p_2 + p_3) + Ap_1p_2 + Bp_1p_3 + Cp_2p_3\right),$$

where

$$A = \varepsilon_{12} + \varepsilon_{21} \quad B = \varepsilon_{13} + \varepsilon_{31} \quad C = \varepsilon_{23} + \varepsilon_{32}.$$

We define the following terms:

Definition 2 *The **synergy profile** $\mathbf{s}$ is defined to be the vector (A, B, C). The quantities A, B and C are the **synergy components** of the synergy profile. The **size** of $\mathbf{s}$ is defined to be $s = ||(A, B, C)||$.*

Quantity A is the analog of the synergy scalar s in the two-agent model: it measures the sum of the influence that agents 1 and 2 exert on each other, and B and C are defined analogously for agents 1 and 3 and agents 2 and 3, respectively. In a three-agent model, there are three relevant synergy components between each of the three pairs of agents, which together form the synergy profile $\boldsymbol{s}$.

The solution to the model is given by Proposition 2 below for the case of an interior solution, and Proposition 3 for the case of a boundary solution.

Proposition 2 *(Substitute production function, three agents, interior solution.) (i) Suppose the synergy profile $\mathbf{s}$ is strictly nonzero and the optimal effort profile $\mathbf{p}^*(\mathbf{s}) = (p_1^*(\mathbf{s}), p_2^*(\mathbf{s}), p_3^*(\mathbf{s}))$ is interior. Then we have:*

$$Ap_2^*(\mathbf{s}) + Bp_3^*(\mathbf{s}) = Ap_1^*(\mathbf{s}) + Cp_3^*(\mathbf{s}) = Bp_1^*(\mathbf{s}) + Cp_2^*(\mathbf{s}) \tag{10}$$

which implies

$$\frac{p_1^*(\mathbf{s})}{p_2^*(\mathbf{s})} = \frac{C}{B}\frac{A+B-C}{A+C-B} \quad ; \quad \frac{p_2^*(\mathbf{s})}{p_3^*(\mathbf{s})} = \frac{B}{A}\frac{A+C-B}{B+C-A} \quad ; \quad \frac{p_3^*(\mathbf{s})}{p_1^*(\mathbf{s})} = \frac{A}{C}\frac{B+C-A}{A+B-C}. \tag{11}$$

In particular, interior optimal effort profiles occur only when each synergy component is strictly smaller than the sum of the other two. Moreover, the optimal effort ratios are homogenous of degree 0 in A, B and C. Therefore the direction of the synergy profile is sufficient to determine the direction of the optimal effort profile provided it is interior.

(ii) Fix a direction of the synergy profile such that each component is strictly smaller than the sum of the other two. There exists a critical synergy size threshold s^* such that, if s is subcritical then the optimal effort profile is interior, and the size of the optimal effort profile is a strictly increasing function of synergy size.[2] *At the critical synergy size s^*, the optimal effort profile explodes so that at least one agent is now applying effort 1.*

(iii) Total wages given success and expected total wages are strictly increasing in s up to the critical synergy size s^*.

(iv) Fix a synergy profile such that the optimal effort profile is interior. An increase in agent i's relative influence (i.e., increasing at least one element of $\{\varepsilon_{ij}\}_{j\neq i}$ and decreasing some elements of $\{\varepsilon_{ji}\}_{j\neq i}$ so that $\mathbf{s}$ is unchanged) increases both his relative and absolute wealth. Specifically,

$$\frac{w_i^*}{\sum_j w_j^*},\ w_i^* \text{ and } p^* w_i^* \text{ all strictly increase,}$$

and

$$\frac{w_i^*}{w_j^*} \text{ weakly increases for all } j \text{ and strictly increases at least one } j.$$

Proposition 3 *(Substitute production function, three agents, boundary solution.) Suppose there is a single synergy component that is greater than the sum of the other two. Then the efforts exerted by the two agents who have the largest synergy with each other are equal and the other agent does not exert effort. The size of the other two synergy components has no effect on the optimal effort profile and the model is isomorphic to the 2-agent model.*

Combining the results of Propositions 2 and 3 gives the full solution to the model as Theorem 2, the key result of this section:

[2] Recall that part (i) implies that, in this interval, the direction of the optimal effort profile is fixed.

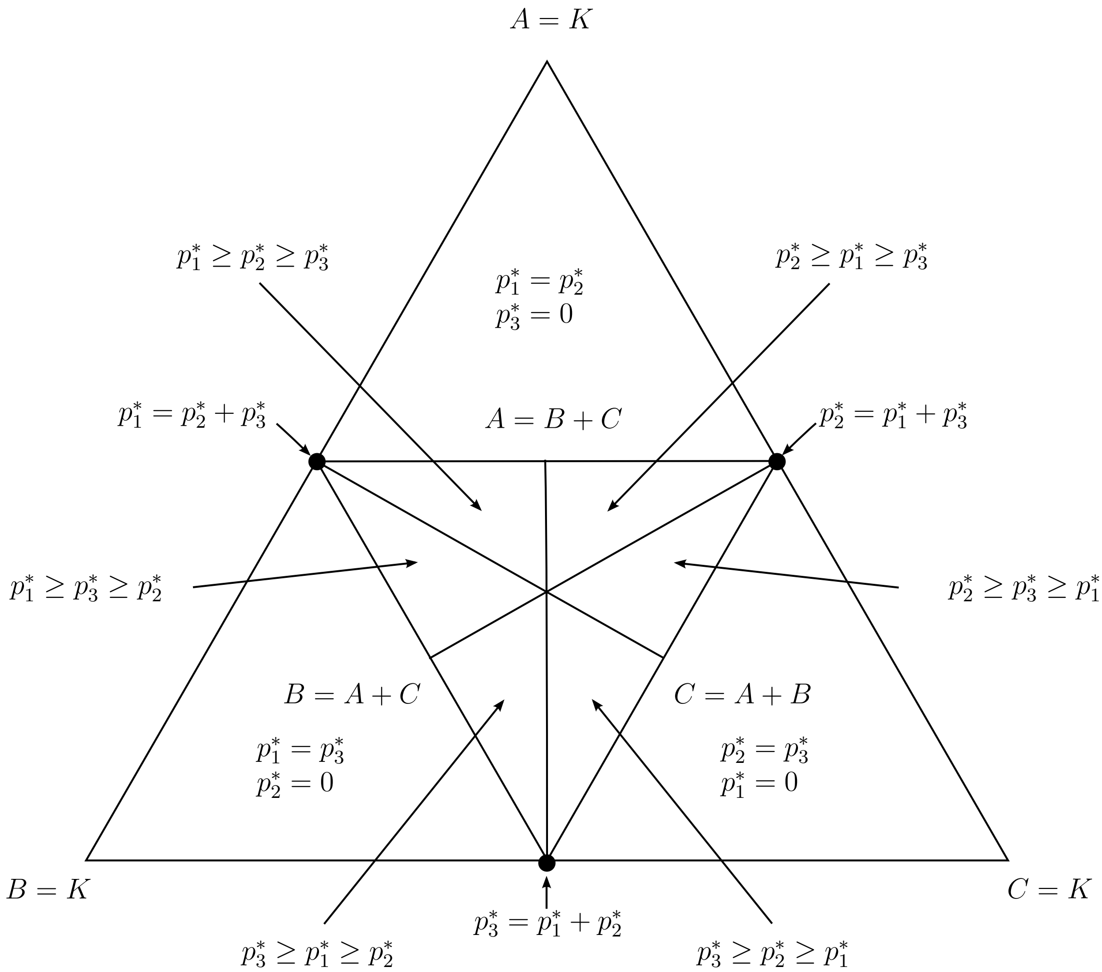

Figure 1: The $A + B + C = K$ simplex where $K > 0$ is some constant.

Theorem 2 *The optimal effort profile is summarized in Figure 1.*

Corollary 1 *Suppose the influence between any pair of agents is symmetric. That is for each $i \neq j$, $\varepsilon_{ij} = \varepsilon_{ji}$. Then when the optimal effort profile is in the interior, the ratios of optimal wages coincide with the ratios of optimal efforts.*

We now discuss the intuition behind and implications of each of the above results. Part (i) of Proposition 2 states that the ratio of the optimal effort levels only depends on the relative size of the different synergy components A, B and C, and not their absolute magnitude. Thus, a proportional increase in each synergy component will augment each effort level to the same degree, leaving the ratios unchanged.

Part (ii) states that, if the size of the synergy profile s is sufficiently small, and the synergy components are balanced so that no single component exceeds the sum of the other two, the optimal effort profile is strictly interior. Analogous to part (i) of Proposition 1, when synergy size increases, effort by each agent becomes more productive as it now has a greater impact on the other agents' cost functions, and so it is efficient for the principal to implement a higher effort profile. When synergies become sufficiently strong, it becomes optimal for the principal to implement the maximum effort level of 1 for at least one agent.

The simplex in Figure 1 fixes the sum of the synergy components $A + B + C$ at a constant K and studies the effect of changing the relative level of the synergy components. The middle triangle (bounded by the three dots) in Figure 1 illustrates the case of an interior effort profile summarized by Proposition 2. For an interior effort profile, all three synergy components matter for the relative size of the individual effort levels. For example, if and only if $B > C$ (i.e. the left-hand side of the triangle), we have $p_1 > p_2$: since agent 1 generates more synergies with agent 3 than does agent 2, it is efficient for agent 1 to exert a higher effort level than the level of effort exerted by agent 2; from Corollary 1, if pairwise influences are symmetric, agent 1 will also enjoy higher pay.

Note that it is the *total* synergy between agent 1 and agent 3 (relative to the total synergy between agent 2 and agent 3) that determines the relative values of p_1 and p_2, not agent 1's unidirectional influence on agent 3, ε_{13} (relative to agent 2's unidirectional influence on agent 3, ε_{23}). It may seem that p_1 should only depend on ε_{13} (and not ε_{31}) as only the former affects the productivity of agent 1's effort. However, when ε_{31} rises, agent 1's cost function is lower and so it is cheaper to implement a higher level of effort. The intuition is similar to that in the two-agent model, whereby synergy can be thought of as an echo between two agents, and it is the combination of their influences on each other that matters, not each influence parameter separately. Hence, when the synergistic relation between agents 1 and 3 is stronger than that between agents 2 and 3, it is optimal for agent 1 to exert higher effort than agent 2. Similarly, if and only if $A > C$, then $p_1 > p_3$, and if and only if $A > B$, then $p_2 > p_3$. In sum, the relative size of the total synergies between each of the three pairs of agents determines their relative effort levels. The agent that exhibits the greatest total synergies with both of his colleagues will work the hardest (and earn the highest pay, if influence is symmetric).

On the one hand, this result extends the principle in the two-agent case, that the optimal effort level depends on the common synergy, and not the individual influence parameters. The synergy profile is a "sufficient statistic" for the effort profile; how it is divided into the individual influence parameters does not matter. On the other hand, the result also contrasts the two-agent setting, since it is no longer the case that all agents exert the same effort level. In the two-agent case, there is only one synergy component (agent 1's synergy component with agent 2 is identical to agent 2's synergy component with agent 1) and so one common effort level. Here, the existence of three synergy components allows for asymmetry in effort levels between the three agents. However, while there are individual effort levels, they still only depend on the common synergy components, not the individual influence parameters.

A natural application of the model is the case where one agent has strong synergies with both other agents, but the other agents do not have strong synergies between them. Suppose, for example, that A and B are large, but C is close to zero. The synergistic agent, agent 1, is the CEO, who shares synergies with two division managers, agents 2 and 3, but they do not share strong synergies between them. As we can see in Figure 1, in this case, the effort exerted by agent 1 will be highest. Essentially, agents 2 and 3 can be aggregated, and their combined effort level is close to the effort exerted by agent 1.

Proposition 3 considers the case of a boundary effort profile. It states that, if one synergy component exceeds the sum of the other two, then the model collapses to the two-agent model of Proposition 1. Intuitively, if the synergy between two agents is sufficiently strong, then only those two agents matter for the principal – she ignores the third agent and induces zero effort from him. This "corner" result (captured by the three triangles that surround the middle triangle in Figure 1) is striking because the third agent still has the same direct effect on the production function (9) as the other two agents, yet is being completely ignored. Moreover, it means that even if there is no change to the synergies exerted by the third agent on his colleagues, an increase in the synergies between agents 1 and 2 can lead to him being excluded. Thus, the third agent's participation depends not only on his own synergy parameters, but also on parameters that have no direct relevance to him. Since the synergies between agents 1 and 2 are so strong, it is always more efficient to increase their effort level from $p - \varepsilon$ to p rather than to increase the third agent's effort level from 0 from ε. Note that this result holds even though we have a convex function and so it is more costly to increase the effort levels of agents 1 and 2 than agent 3. The convex cost function is why, even if $A > B$ and $A > C$, agent 3 typically exerts a strictly positive effort level even though he exhibits fewer synergies. Only if $A > B + C$ are the synergies between the first two agents sufficiently strong to outweigh the effect of the convex cost function and lead to agent 3's effort level being zero. Due to the strong synergy, raising the effort levels of agents 1 and 2 "echoes" many times and is thus more effective than raising the effort level of agent 3. Another way to view the intuition is that increased synergy between agents 1 and 2 raises the value of the firm, and thus the cost to the principal of giving agent 3 equity to induce effort from him.

The above result has interesting implications for the optimal composition of a team. If two agents exhibit sufficiently high synergies with each other, there is no benefit in adding a third agent to the team, even if the third agent has just as high a direct impact on firm value as the existing two agents and has strictly positive synergies with the first two agents. If the third agent was added, he would become a redundant "third wheel" and be asked to implement zero effort, so there is no loss in excluding him from the team. Moreover, the three agents can be interpreted as three different divisions of a firm, in which case Proposition 3 has implications for the boundaries of the firm. If two divisions exhibit sufficiently strong synergies with each other (e.g. there are spillovers in marketing campaigns), it may be optimal to divest a third division even if that third division makes a strong direct contribution to overall firm value and the first two divisions exhibit no direct synergies in the production function. Conversely, it may be optimal for a two-division firm not to acquire a third division even if it would generate strictly positive synergies, if those synergies are low relative to those enjoyed by the two existing divisions. Conventional wisdom is that any division that enjoys positive synergies should be included within a firm. Here, even though the third division enjoys strictly positive synergies with the first two, it is relative, not absolute, synergies that determine the optimal boundaries of the firm. The empirical implication is that a decision to divest (or not acquire) a division might not be driven by the low synergies generated (or potentially generated) by this division, but rather by the strong synergies between other divisions.

While in Proposition 2, all three synergy components matter for the optimal effort profile, in Proposition 3 only the largest synergy component matters and the other two are irrelevant. For example, within the middle triangle, the relative size of B and C affects the relative size of p_1 and p_2, as discussed earlier. In the top triangle (where $A > B + C$), we have $p_1 = p_2$ regardless of the relative size of B and C. Intuitively, the synergy between agents 1 and 2 is so important that their individual synergies with agent 3 become irrelevant. Wages are then determined as in the two-agent model and depend on the relative influence of each agent.

Having considered the optimal effort profile, we now turn to the implications for the optimal wage profile. Part (iii) of Proposition 2 is analogous to part (ii) of Proposition 1: total wages depend on the total synergy across all agents. While total synergy determines total wages, the influence parameters determine relative wages: part (iv) of Proposition 2 is analogous to part (iv) of Proposition 1. An increase in one agent's influence parameter augments his wage in both absolute and relative terms; the intuition is as earlier. Moreover, if the influence parameters are symmetric across all pairs of agents, the entire wage profile can be fully solved: Corollary 1 states that the ratios of optimal wages coincides with the ratios of optimal effort.

The model can thus explain why CEOs earn significantly more than other senior managers. Bebchuk, Cremers, and Peyer (2011) argue that this is due to inefficient rent extraction by the CEO, but our theory suggests that it may be efficient: the centrality of the CEO leads to him exhibiting greatest synergies, increasing his optimal effort level and thus pay.[3] Thus, the three-agent model shows that a CEO's wage depends on the scope of the firm under his control, i.e. the number of agents (or divisions) with which he exhibits synergies and the strength of these synergies. Talent assignment models argue that CEO pay depends on the size of the firm under his control (e.g. Gabaix and Landier (2008), Terviö (2008)), where firm size is typically measured by an accounting variable such as total assets or profits. Our theory suggests that the relevant measure of firm size is the scope and depth of the CEO's synergies. Thus, the CEO of a large firm in which the divisions operate independently (e.g. a holding company) may be paid less highly than the manager of a small firm where there are strong synergies (e.g. a start-up).

3.3 Negative Influence Parameters

This subsection extends the model to the case where the influence parameters ε_{ij} can be negative. We start with the two-agent model and then move to the three-agent model.

3.3.1 The Preliminary Two-Agent Model

Recall the principal's reduced-form maximization problem is given by:

$$p_1^*, p_2^* \in \arg\max_{p_1, p_2} \frac{p_1 + p_2}{2} \left(1 - (p_1 + p_2) + p_1 p_2 (\varepsilon_{12} + \varepsilon_{21})\right).$$

There are thus two cases to consider.

[3] Kale, Reis, and Venkateswaran (2009) study another reason for why high pay for the CEO may be efficient – to provide tournament incentives for other senior managers. They find that the pay differential between the CEO and other senior managers is positively related to firm performance.

Case 1. $\varepsilon_{12} > 0 > \varepsilon_{21}$, and $\varepsilon_{12} + \varepsilon_{21} > 0$.

By inspecting the maximization problem, we can see that the solution only depends on the total synergy s and not the individual influence parameters ε_{ij}. Since we have $s > 0$, we are in the case of the core model and so Proposition 1 holds.

Case 2. $\varepsilon_{12} + \varepsilon_{21} < 0$.

Since we now have $s < 0$, by inspecting the maximization problem we can see that the solution requires $p_1^* p_2^* = 0$ and so one agent exerts zero effort. Since both agents have the same direct productivity, it does not matter which agent this is. Without loss of generality, assume that $p_2^* = 0$. Then the principal solves:

$$p_1^* \in \arg\max_{p_1} \frac{p_1}{2} (1 - p_1) .$$

This is a single-agent model. The solution is standard, and is given by Proposition 4 below:

Proposition 4 *(Substitute production function, two agents, negative synergy.) Suppose that the total synergy s is negative. Then only one agent exerts strictly positive effort; without loss of generality, assume this is agent 1. The analog of Proposition 1 is as follows:*

(i) The optimal effort levels are given by $p_1^*(s) = \frac{1}{2}$, $p_2^*(s) = 0$.

(ii) The wage levels are given by $w_1^* = \frac{1}{2}$ and $w_2^* = 0$, and are independent of s as long as $s < 0$.

(iii) An increase in either influence parameter has no effect on effort and wages as long as $s < 0$.

(iv) Since the principal is indifferent over which agent has the zero effort and wage level, it is possible to have $w_1 > w_2$ for $\varepsilon_{12} < \varepsilon_{21}$.

(v) For a fixed $s < 0$, changes in agent i's relative influence have no effect.

(vi) As long as $s < 0$, changes in agent i's absolute influence have no effect.

(vii) The agent who is exerting effort has the higher utility. Since the principal is indifferent over which agent has the zero effort and wage level, it is possible that this is the less influential agent.

We can summarize the above results as follows. Case 1 shows that, as long as the total synergy is positive, the core model's result of equal effort levels (irrespective of individual influence parameters) continues to hold in the case where one influence parameter is negative. It may seem surprising that the principal chooses to hire (i.e., induce strictly positive effort from) an agent that exert negative influence, but this is optimal if it is outweighed by the other agent exerting a sufficiently positive influence so that the total synergy is positive. Case 2 shows that, if total synergy is negative, the principal only wishes to hire one agent, and the individual influence parameters are irrelevant for the choice of agent. Again, it is the total synergy that matters for whether both agents exert effort, so it does not matter if one influence parameter is negative as long as the total synergy is positive.

3.3.2 The Main Three-Agent Model

In the two-agent model, the solution depended on whether the total synergy (rather than the individual influence parameters) was positive or negative. In the three-agent model, the solution depends on whether the synergy components are positive or negative. Without loss of generality, we will assume that A is the largest synergy component, followed by B and then C. There are four cases to consider:

Case 1. $A > B > C > 0$.

If each synergy component is positive, we are in the case of the core model and Propositions 2 and 3 continue to hold.

Case 2. $A > B > 0 > C$.

Here, one of the synergy components is negative. This ensures that there is a single synergy component that is greater than the sum of the other two: $A > B + C$. We thus obtain the corner solution of Proposition 3. Only the two agents who have the largest synergy with each other exert effort, and the problem reduces to the 2 agent model.

Case 3. $A > 0 > B > C$

This case is similar to Case 2 in that we have $A > B + C$. We thus obtain the corner solution of Proposition 3.

Case 4. $0 > A > B > C$.

In this case, only one agent exerts effort. Since all three agents have the same direct productivity, it does not matter which agent this is. Without loss of generality, assume that $p_2^* = p_3^* = 0$. We are in a single agent model where $p_1^* = \frac{1}{2}$ and the analogy of Proposition 4 applies.

3.4 Discussion: The Synergy Concept

A key feature of our model is that an agent's effort reduces the marginal cost of effort of his colleague. Alternatively, as mentioned before, this can be interpreted as an agent's effort increasing the marginal private benefit that the colleague derives from his own effort. This feature generates the synergies among agents in our model. To what extent is this different from instead assuming that there are complementarities in the production function, i.e., that an agent's effort increases the marginal productivity of the other agent's effort in the production function?

In a single-agent model with separable utility, changing the agent's marginal productivity by multiplying the production function by a constant factor is indeed isomorphic to changing his marginal cost by dividing the cost function by the same multiple. However, in a multi-agent world, synergies in the cost function are fundamentally different from complementarities in the production function. The most conceptually important difference is that cost synergies are a true externality, but production complementarities are not. To illustrate this distinction, suppose that agents do not affect other agents' cost of effort, i.e., $c_i = h_i(p_i)$ ($= \frac{1}{4}p_i^2$ in the two-agent model), but that the production function exhibits complementarities, e.g., in the two-agent model $\Pr(r = 1) = \frac{p_1+p_2}{2} + s'\sqrt{p_1 p_2}$. Complementarity is captured by the positive cross partial $s = \frac{s'}{4\sqrt{p_1 p_2}}$. An increase in agent i's effort will increase the productivity of agent j, but does not take this into account because he holds agent j's effort fixed when calculating his own optimal action. One might be tempted to conclude that the complementarity therefore represents an externality. To show that it does not, consider the case of a single agent who internalizes everything – he owns production and exerts both efforts p_1 and p_2. If the complementarity captured by the positive cross partial s is an externality, it should be taken into account by the single agent since he internalizes all externalities. However, the single agent's optimization problem for his choice of p_i is

$$\max_{p_i} \frac{p_i + p_j^*}{2} + s'\sqrt{p_i p_j^*} - \frac{1}{4}p_i^2 - \frac{1}{4}p_j^{*2},$$

and analogously for the choice of p_j. Thus, even in a single-agent model, the complementarity between p_i and p_j is ignored. Even the single agent holds p_j fixed when choosing p_i, which is why p_j enters as p_j^* in the above objective function rather than as a function of p_i. Thus, the optimal p_i is independent of the cross-partial with respect to p_i and p_j. To the extent that first-order conditions are sufficient, then, by definition, second-order effects such as production complementarities do not matter.

To illustrate further that the multi- and single-agent models are similar under production complementarities, and thus that such complementarities are not true externalities, consider the case in which there are no cost synergies (but there may be production complementarities), there is unlimited liability, and utility is quasi-linear in money. Then, the optimality conditions of the aggregate agent in a single-agent model equal the aggregate of the optimality conditions of each individual agent in a multi-agent model. Thus, if the optimal contract for the single agent is $w(r)$, then the optimal contracts in the two-agent world are simply two copies of $w(r)$ plus some lump-sum transfers. This is essentially a consequence of the the analysis of the free-rider problem in Holmstrom (1982). This will not be the case if there are cost synergies.

More generally, because contracts are contingent upon output but cannot be made contingent on effort costs, agents naturally internalize the effects of their efforts on production but not on costs. To illustrate, in a multi-agent model with production complementarities but no cost synergies, agent i's objective function is:

$$\left(\frac{p_i + p_j^*}{2} + s' \sqrt{p_i p_j^*}\right) w_i - \frac{1}{4} p_i^2. \tag{12}$$

In the current model, with cost synergies only, an aggregate agent's optimization problem for his choice of p_i is:

$$\frac{p_i + p_j^*}{2} - \frac{1}{4} p_i^2 \left(1 - \varepsilon_{ji} p_j^*\right) - \frac{1}{4} p_j^{*2} \left(1 - \varepsilon_{ij} p_i\right). \tag{13}$$

In a multi-agent model, agent i's objective function is:

$$\frac{p_i + p_j^*}{2} w_i - \frac{1}{4} p_i^2 \left(1 - \varepsilon_{ji} p_j^*\right). \tag{14}$$

Equation (12) shows that agent i does internalize the production complementarity s when choosing his effort level. This is because the complementarity s affects output, and he receives a share of output since output is contractible. By contrast, equation (14) shows that agent i does not consider his influence on agent j, ε_{ij}, when considering his effort level: this term does not appear in his objective function. This is because his influence affects agent j's cost of effort (which is non-contractible, and so he does not share in this effect) but has no effect on output, and so he does not internalize it. Thus, synergies in the cost function represent true externalities that are not internalized by the agents. Even with unlimited liability and quasilinear utility, the optimality conditions of the aggregate agent in the single-agent model do not equal the aggregate of the optimality conditions of each individual agent in a multi-agent model. In the multi-agent model, the principal would like the agents to internalize the cost synergies as they affect total surplus, as shown by equation (13), and thus varies the contract to cause them to do so. Indeed, the paper's main objective is to analyze how the principal increases incentives to induce the agents to internalize their cost externalities, although such internalization is only partial since it is costly to the principal – due to limited liability, increased incentives can only be achieved by an increase in the success payoff and not a reduction in the failure payoff, and the principal trades off this cost with the benefits of internalization.

Note that, even though cost and production synergies are fundamentally different from a modeling standpoint, in that the latter but not the former are internalized by an agent, they are similar in the economic idea that they represent. In the presence of production synergies, effort by one agent increases the marginal productivity of his colleague, for a given unit cost. In the presence of cost synergies, effort by one agent reduces the marginal cost of his colleague, for a given unit productivity. Thus, although we are modeling synergies differently from a framework in which they appear in the production function, our model continues to capture the same economic idea that synergies improve a colleague's productivity-to-cost ratio.

In the next section, we analyze our model of cost synergies in addition to complementarities in the production function, and show that the presence of complementarities in the production function over and above cost synergies does not change the implications generated by cost synergies so far in the paper.

4 Complementary Effort

This section specializes the general production function (1) to the case in which the agents' efforts are perfect complements, i.e. the probability of success depends on the minimum effort level undertaken by all agents. The production function (1) now specializes to*:*

$$\Pr(r = 1) = \min(p_1, p_2, ..., p_N). \tag{15}$$

We continue to assume a quadratic individual cost function:

$$h_i(p_i) = \frac{\kappa_i}{2} p_i^2.$$

Differentiating agent i's utility function (3) gives his first-order conditions as:

$$p_1 = p_2 = \ldots = p_N \equiv p, \tag{16}$$

and

$$w_i(p) = \kappa_i p \left(1 - \sum_{j \neq i} \varepsilon_{ji} p\right). \tag{17}$$

These first-order conditions already give us some preliminary results. Equation (16) shows that all agents will exert the same effort level, as is intuitive given the perfect complementarities production function (15). Equation (17) shows that agent i's wage is linear in his cost parameter κ_i, i.e. agents with more difficult tasks (higher κ_i) will receive higher wages.

Plugging the first-order conditions (16) and (17) into the principal's objective function (4) gives her reduced-form maximization problem as:

$$p^* \in \arg\max_p \; p\left(1 - \sum_i w_i(p)\right) = \arg\max_p \; p\left(1 - p\sum_i \kappa_i + p^2 \sum_i \left(\sum_{j\neq i} \varepsilon_{ij}\kappa_j\right)\right).$$

We define the following terms:

Definition 3 ***Synergy*** *is defined to be the sum of each agent's total influence:*

$$s = \sum_i \left(\sum_{j\neq i} \varepsilon_{ij}\kappa_j\right)$$

Difficulty *is defined to be the sum of the cost parameters,* $\kappa \equiv \sum_i \kappa_i$.

Assumption 3 *Difficulty* $\kappa > \frac{1}{2}$.

This is a nontriviality assumption about the difficulty of the project being not too low. It ensures that the problem has nontrivial solutions in agent efforts for at least some realized levels of synergy.

The solution to the model is given by Proposition 5 below.

Proposition 5 *(Complementary production function.) (i) There exists a unique critical synergy threshold* $s^*(\kappa) > 0$ *such that optimal effort is given by:*

$$p^*(s) = \begin{cases} \frac{\kappa - \sqrt{\kappa - 3s}}{3s} & s \in [0, s^*(\kappa)) \\ 1 & s \geq s^*(\kappa). \end{cases}$$

Optimal effort $p^*(s)$ *is strictly increasing on* $[0, s^*(\kappa)]$. *Furthermore, if difficulty* $\kappa > 1$, *then* $p^*(s)$ *explodes to 1 when the critical synergy level* $s^*(\kappa)$ *is reached.*

(ii) Total wages given success, $w^*(s) = \sum_i w_i^*(s)$, *and expected total wages* $p^*(s)w^*(s)$ *are both strictly increasing on* $[0, s^*(\kappa)]$.

(iii) Suppose synergy is subcritical. An increase in any influence parameter of any agent will lead to increases in optimal effort, total payment given success and total expected success payment.

(iv) Fix a subcritical synergy level. Suppose agent i's relative influence increases, i.e. his total influence increases while holding synergy constant. If the resulting decrease in the total influence of the other agents is nondistortionary[4] *then there is an increase in agent i's relative and absolute wealth. Specifically,*

$$\frac{w_i^*}{\sum_j w_j^*},\ w_i^* \ \textit{and}\ \ p^* w_i^* \ \ \textit{all strictly increase,}$$

and

$$\frac{w_i^*}{w_j^*} \ \ \textit{weakly increases for all } j \textit{ and strictly increases at least one } j.$$

Proposition 5 shows that our model's key results are robust to the nature of the production function. Even though the perfect complements production function of this section is the polar opposite of the perfect substitutes production function of Section 3, the main insights regarding the effort and wage profiles remain unchanged. In addition to demonstrating robustness to the specification of the production function, this section also shows that the results naturally extend to the case of N agents.

As in Section 3, an increase in total synergy leads to an increase in the implemented effort levels, total pay and expected total pay; the intuition is the same. An increase in a single agent's influence parameters augments total synergy (thus leading to the above effects) and his own pay in both relative and absolute terms.

5 Conclusion

This paper has studied the effect of synergies on optimal effort levels and wages in a team-based setting. We model synergies as effort by one agent reducing the cost, or increasing the private benefit, of effort by a colleague. This is a fundamentally different notion of synergy to complementarities in the production function and leads to a number of new results. In a two-agent framework, effort levels are equal even though influence may be asymmetric. Wages differ across agents, even though both agents exert the same effort level and have the same direct impact on output, with the

[4]In other words, the decrease in the other agents' total influence is achieved by simply multiplying their influence parameters with a common scalar $c < 1$.

more influential agent receiving higher pay. Total wages increase with the total level of synergy, consistent with the high equity incentives in small start-up firms. In short, total synergy determines total effort and total wages; individual influence parameters only affect individual wages. The model also shows that it may be optimal to grant rank-and-file employees strong equity incentives, even if their direct effect on output is low, if they exert sufficiently high synergies. This prediction is consistent with the frequency of broad-based stock option plans.

With three agents, optimal effort levels differ and depend on the total synergies an agent enjoys with his colleagues rather than his unidirectional influence. If synergies between two agents are sufficiently strong, it is optimal for the principal to focus entirely on these agents and ignore the third. This result has implications for the optimal composition of a team and optimal firm boundaries – if synergies between two agents (divisions) become sufficiently strong, it is efficient to discard the third agent (division) even if his (its) own parameters do not change. Agents that exert synergies over a greater number of colleagues receive higher pay, consistent with the wage premia CEOs enjoy over divisional managers.

A Proofs

We first start with a maximization problem which we will make repeated use of in these proofs. Consider the following maximization problem where $a, b \geq 0$:

$$\max_{x \in [0,1]} x(1 - bx + ax^2).$$

Let $x^*(a, b)$ denote the set of argument solutions.

Lemma 1 *(i) If $b \leq \frac{1}{2}$, then $x^*(a, b) = 1$.*
(ii) If $b > \frac{1}{2}$, then there exists a threshold $a^*(b) > 0$ such that

$$x^*(a,b) = \begin{cases} \frac{b - \sqrt{b^2 - 3a}}{3a} & a < a^*(b) \\ \{\frac{b - \sqrt{b^2 - 3a}}{3a}, 1\} & a = a^*(b) \\ 1 & a > a^*(b) \end{cases}.$$

Proof. We first define some notation. Let $U(x, a, b) = x(1 - bx + ax^2)$ and $x^{loc}(a, b) = \frac{b - \sqrt{b^2 - 3a}}{3a}$.

First let $b \leq \frac{1}{2}$. If $a = 0$, it is clear that $x^*(0, b) = 1$. If $a > 0$, then

$$\frac{d}{dx} U(x, a, b)|_{x=1} = 1 - 2bx + 3ax^2|_{x=1} = 1 - 2b + 3a > 0 \tag{18}$$

To show $x^*(a, b) = 1$, it suffices to show there is no local maximum of $U(x, a, b)$ on $(0, 1)$. By the quadratic formula, a local maximum exists (anywhere) if and only if $b^2 - 3a = 3a(b \cdot \frac{b}{3a} - 1) > 0$. Since $b \leq \frac{1}{2}$, this implies $\frac{b}{3a} > 2$. In addition, $\frac{b}{3a}$ is the inflection point of $U(x, a, b)$. Since $U(x, a, b)$ is a positive cubic, the inflection point lies above the local maximum. Thus, since $\frac{d}{dx} U(x, a, b) > 0$ for $x = 1$ (from (18)), and the inflection point is not reached until $x = \frac{b}{3a} > 2$, the local maximum must be between $x = 1$ and $x = \frac{b}{3a}$. Thus, we must also have $\frac{d}{dx} U(x, a, b) > 0$ for all $x < 1$. Thus, there is no local maximum of $U(x, a, b)$ on $(0, 1)$.

Now consider $b > \frac{1}{2}$. We have the following facts:

Fact 1: $x^{loc}(a, b)$ is strictly increasing in a on $[0, \frac{b^2}{3}]$. This follows from the fact that $b - \sqrt{b^2 - 3a}$ is convex while $3a$ is linear and both are equal to zero when $a = 0$.

Fact 2: By the envelope theorem,

$$\frac{d}{da} U\left(x^{loc}(a, b), a, b\right) = \left[x^{loc}(a, b)\right]^3 < 1 \ \text{ when } \ x^{loc}(a, b) < 1$$

Fact 3: On the other hand,

$$\frac{d}{da}U(1,a,b) = 1$$

Fact 4: For all sufficiently low a, $x^*(a,b) = x^{loc}(a,b)$. To see this, notice since $\lim_{a\downarrow 0} x^{loc}(a,b) = \frac{1}{2b} < 1$, so for all sufficiently low a, the local maximum is in the interval $(0,1)$. Of course when $a = 0$, the local maximum is the global maximum. By continuity, the fact is true.

Clearly, whenever $x^{loc}(a,b) > 1$ or does not exist, then $x^*(a,b) = 1$. Therefore, suppose $x^{loc}(a,b) \leq 1$ and exists. Fact 1 implies that the set of a that satisfy these two conditions is of the form $[0,\tilde{a}]$ where $\tilde{a} \leq \frac{b^2}{3}$. We wish to show that $U\left(x^{loc}(a,b),a,b\right)$ and $U(1,a,b)$ satisfy the single crossing property on the interval $[0,\tilde{a}]$. $\tilde{a}$ is the upper bound on the interval of a's such that $x^{loc}(a,b) \leq 1$ and exists. Thus, there are two cases to consider. First, we could have $x^{loc}(\tilde{a},b) = 1$, in which case the functions $U\left(x^{loc}(a,b),a,b\right)$ and $U(1,a,b)$ cross at $a = \tilde{a}$. Second, we could have $x^{loc}(\tilde{a},b) < 1$. Note that at $a = \tilde{a}$, the function $U(x,a,b)$ must have a single critical point. If it had two critical points, we could increase a. An increase in a "flattens" out the cubic by bringing the value of the local minimum and local maximum closer, but since there are two critical points to begin with, this can be done without violating the requirement that at least one critical point, $x^{loc}(a,b)$, exists. An increase in a also raises $x^{loc}(\tilde{a},b)$ (from Fact 1), but since $x^{loc}(\tilde{a},b) < 1$, this can be done without violating the constraint that $x^{loc}(a,b) \leq 1$. Since a can be increased without violating the constraints that $x^{loc}(a,b) \leq 1$ and exists, $\tilde{a}$ would not meet the requirement of being the upper bound on the interval of a's such that these constraints are satisfied. By contrast, if $U(x,a,b)$ has a single critical point, a cannot be increased further as the function would then have no critical points. Since $U(x,a,b)$ has a single critical point, it is non-decreasing in x. Thus, $x^{loc}(\tilde{a},b) < 1$ implies $U\left(x^{loc}(\tilde{a},b),\tilde{a},b\right) < U(1,\tilde{a},b)$. Facts 2 and 3 imply that $\frac{dU\left(x^{loc}(a,b),a,b\right)}{da} < \frac{dU(1,a,b)}{da}$, and we also have $U\left(x^{loc}(0,b),0,b\right) < U(1,0,b)$. Thus, the functions $U\left(x^{loc}(a,b),a,b\right)$ and $U(1,a,b)$ must cross at some point $a^*(b) \in [0,a]$. Finally, Fact 4 implies that on $[0,a^*(b))$, $x^*(a,b) = x^{loc}(a,b)$. ■

Lemma 2 *(i) If $b > \frac{1}{2}$ then $x^*(a,b)$ is strictly increasing on $[0,a^*(b))$.*

(ii) If $b \in (\frac{1}{2},1]$ then $\frac{b-\sqrt{b^2-3a^*(b)}}{3a^*(b)} = 1$ and $x^*(a,b)$ smoothly increases up to 1.

*(iii) If $b > 1$ then $\frac{b-\sqrt{b^2-3a^*(b)}}{3a^*(b)} < 1$ and $x^*(a,b)$ **explodes** up to 1 upon reaching the critical threshold $a^*(b)$.*

Proof. The first claim follows from Fact 1 in the proof of Lemma 1. For the third claim, note $x^{loc}(a,b)$ is only defined when $a \leq \frac{b^2}{3}$ and $x^{loc}(\frac{b^2}{3}, b) = \frac{1}{b}$. Fact 1 then implies the $b > 1$ claim. For the second claim, now suppose $b \leq 1$. Then $x^{loc}(\frac{b^2}{3}, b) = \frac{1}{b} \geq 1$ and it is also the inflection point. In general the inflection point is $\frac{b}{3a}$. Thus as a decreases from $\frac{b^2}{3}$, the inflection point is increasing. In particular, it remains above 1. However, the only way that we can have $U(1, a^*(b), b) > U\left(1, x^{loc}(a^*(b)), b\right)$ (i.e. an explosion) is if both $x^{loc}(a^*(b), b)$ and the inflection point are both strictly smaller than 1. Thus, there is no explosion. ■

Lemma 3 *If $b > \frac{1}{2}$ then the quantities $bx^*(a,b) - ax^{*2}(a,b)$ and $x^*(a,b)(bx^*(a,b) - ax^{*2}(a,b))$ are both increasing on $[0, a^*(b))$.*

Proof. On $[0, a^*(b))$

$$\frac{d}{dx}U(x,a,b)|_{x^*(a,b)} = 1 - 2bx^*(a,b) + 3ax^{*2}(a,b) = 0$$

$$\Rightarrow \frac{d}{da}U(x^*(a,b),a,b) = -2bx_1^*(a,b) + 6ax^*(a,b)x_1^*(a,b) + 3x^{*2}(a,b) = 0 \tag{19}$$

Now

$$\frac{d}{da}bx^*(a,b) - ax^{*2}(a,b) = bx_1^*(a,b) - 2ax^*(a,b)x_1^*(a,b) - x^{*2}(a,b)$$

Equation (19) then implies

$$\frac{d}{da}bx^*(a,b) - ax^{*2}(a,b) = \frac{b}{3}x_1^*(a,b) > 0$$

This shows $bx^*(a,b) - ax^{*2}(a,b)$ is increasing. Since $x^*(a,b)$ is positive and increasing as well, so $x^*(a,b)(bx^*(a,b) - ax^{*2}(a,b))$ is also increasing. ■

Proof of Proposition 1

The principal's objective function is $\frac{p_1+p_2}{2}(1 - (p_1+p_2) + p_1p_2s)$. We first wish to prove that $p_1 = p_2$. Fix a given $X = p_1 + p_2$. The term p_1p_2s is maximized, for a given X, by setting $p_1 = p_2$. The other terms in the objective function are all terms in X. Thus, we have $p_1 = p_2$. This allows us to apply Lemmas 1, 2 and 3 with $x = \frac{p_1+p_2}{2}$; statements (i), (ii) and (iii) are essentially transcriptions of these three Lemmas, respectively. The only difference is that at the critical synergy level, we now discriminate between the two optimal efforts in accordance with Assumption 1.

To see (iv), note if i is more influential than j then $\varepsilon_{ij} > \varepsilon_{ji}$. This implies:

$$w_i^*(s) = p^*(s)(1 - \varepsilon_{ji}p^*(s)) > p^*(s)(1 - \varepsilon_{ij}p^*(s)) = w_j^*(s).$$

More generally, holding synergy fixed, an increase in agent i's relative influence means both increasing ε_{ij} and decreasing ε_{ji}. This causes both an increase in w_i^* and a decrease in w_j^*, which proves (v).

The proof of part (vi) is as follows. We use a dot to denote the derivative with respect to ε_{ij}.

$$\dot{w_j^*} = \dot{p^*} - 2\varepsilon_{ij}p^*\dot{p^*} - p^{*2}$$

$$p^* \in \arg\max_p p(1 - 2p + s^*p^2) \Rightarrow 1 - 4p^* + 3sp^{*2} = 0 \Rightarrow -4\dot{p^*} + 6sp^*\dot{p^*} + 3p^{*2} = 0$$

A linear combination of the two gives us

$$\dot{w_j^*} = \frac{1}{3}\dot{p^*}\left(6\varepsilon_{ij}p^* - 1\right)$$

Since $\dot{p^*} > 0$, this means that, when $s < \bar{s}$, $\dot{w_j^*}$ and $6\varepsilon_{ij}p^* - 1$ have the same sign. Equation (8) follows immediately. Turning to the expected wage, we have:

$$\dot{p^*w_j^*} = 2p^*\dot{p^*} - 3\varepsilon_{ij}p^{*2}\dot{p^*} - p^{*3}$$

$$-4\dot{p^*} + 6sp^*\dot{p^*} + 3p^{*2} = 0 \Rightarrow p^*\left(-4\dot{p^*} + 6sp^*\dot{p^*} + 3p^{*2}\right) = 0$$

A linear combination of the two gives us

$$\dot{p^*w_j^*} = 3\varepsilon_{ij}p^{*2}\dot{p^*} + \frac{1}{2}p^{*3} > 0.$$

Finally, for part (vii), the first-order condition yields: $w_1 = p_1\left(1 - \varepsilon_{21}p_2\right)$. Thus, agent 1's utility is given by:

$$\begin{aligned} U_1 &= p_1\left(1 - \varepsilon_{21}p_2\right)\left(\frac{p_1 + p_2}{2}\right) - \frac{1}{4}p_1^2\left(1 - \varepsilon_{21}p_2\right) \\ &= \frac{3}{4}p^2\left(1 - \varepsilon_{21}p\right) \end{aligned}$$

where $p = p_1 + p_2$, and similarly $U_2 = \frac{3}{4}p^2\left(1 - \varepsilon_{12}p\right)$. Hence $U_1 > U_2$ if and only if $\varepsilon_{12} > \varepsilon_{21}$.

Proof of Proposition 2

Holding total effort constant,

$$p_1^*(\mathbf{s}), p_2^*(\mathbf{s}), p_3^*(\mathbf{s}) \in \arg \max_{p_1, p_2, p_3 \in [0,1]} Ap_1p_2 + Bp_1p_3 + Cp_2p_3 \tag{20}$$

The first-order conditions which characterize interior solutions to this convex problem are captured by equation (10). This proves (i).

Since the maximization problem of equation (20) is convex, the optimal effort profile will satisfy the ratios of equation (11) so long as:

1. Each synergy component is strictly smaller than the sum of the other two.

2. The restriction of each effort being no greater than 1 is nonbinding.

Condition 1 is assumed in this lemma and condition 2 holds if synergy is sufficiently small. Suppose then that synergy is small. Call by p the highest effort of the optimal effort profile. Then there exists $1 \geq \alpha \geq \beta > 0$ such that the other two efforts are αp and βp. Assume without loss of generality that agent 1's effort is highest, agent 2's effort is α times agent 1's effort and agent 3's effort is β times agent 1's effort. Then the principal's maximization problem becomes

$$p^* \in \arg \max_{p \in [0,1]} (1 + \alpha + \beta)p \left(1 - (1 + \alpha + \beta)\, p + (A\alpha + B\beta + C\alpha\beta)\, p^2\right).$$

Statement (ii) now follows from Lemma 1. Statement (iii) follows from Lemma 3.

Holding the synergy profile fixed, an increase in agent i's relative influence means both an increase of at least one element of $\{\varepsilon_{ij}\}_{j \neq i}$ and a corresponding decrease of some elements in $\{\varepsilon_{ji}\}_{j \neq i}$. This causes an increase in w_i^* and a decrease in at least one element of $\{w_j^*\}_{j \neq i}$ provided the effort profile is interior. p28, final paragraph of the proof of proposition 2. Moreover, since (p_1, p_2, p_3) is a function of the synergy profile only, it is unaffected by changes in relative influence and so p^* is unchanged. Statement (iv) now follows.

Proof of Proposition 3

Without loss of generality, suppose $A > B \geq C$ and $A \geq B + C$. Looking at the convex problem of equation (20), it is clear that $p_3^* = 0$. But then the principal's maximization problem becomes symmetric in p_1 and p_2 and there is nontrivial synergy between agents 1 and 2. The statement in the proposition then follows from the preliminary two-agent case.

Proof of Corollary 1

Recall the optimal wage for agent i is

$$w_i^*(p_i^*) = p_i^* \left(1 - \sum_{j \neq i} \varepsilon_{ji} p_j^*\right).$$

Equation (10) and the corollary's assumption about the influence parameters imply that the quantity inside the parentheses is the same for all i. The result now follows immediately.

Proof of Proposition 5

The proof is essentially the same as in Proposition 1.

References

[1] Bebchuk, Lucian A., Martijn Cremers, and Urs C. Peyer (2011): "The CEO Pay Slice." *Journal of Financial Economics*, forthcoming

[2] Bergman, Nittai K. and Dirk Jenter (2007): "Employee Sentiment and Stock Option Compensation." *Journal of Financial Economics* 84, 667–712

[3] Bernstein, Shai and Eyal Winter (2010): "Contracting with Heterogeneous Externalities." Working paper, Harvard University

[4] Bolton, Patrick and Mathias Dewatripoint (2005): "Contract Theory." *MIT Press, Cambridge.*

[5] Che, Yeon-Koo and Seung-Weon Yoo (2001): "Optimal Incentives for Teams." *American Economic Review* 91, 525–541

[6] Dessein, Wouter, Luis Garicano, and Robert Gertner (2010): "Organizing for Synergies." *American Economic Journal: Microeconomics* 2, 77–114

[7] Gabaix, Xavier and Augustin Landier (2008): "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics* 123, 49–100

[8] Gervais, Simon and Itay Goldstein (2007): "The Positive Effects of Biased Self-Perceptions in Firms." *Review of Finance* 11, 453–496

[9] Hermalin, Benjamin E. (1998): "Toward an Economic Theory of Leadership: Leading By Example." *American Economic Review* 88, 1188–1206

[10] Hochberg, Yael V. and Laura Lindsey (2010): "Incentives, Targeting, and Firm Performance: An Analysis of Non-executive Stock Options." *Review of Financial Studies* 23, 4148–4186

[11] Holmstrom, Bengt (1982): "Moral Hazard in Teams." *Bell Journal of Economics* 13, 324–340

[12] Itoh, Hideshi (1991): "Incentives to Help in Multiagent Situations." *Econometrica* 59, 611–636

[13] Kale, Jayant R., Ebru Reis, and Anand Venkateswaran (2009): "Rank-Order Tournaments and Incentive Alignment: The Effect on Firm Performance." *Journal of Finance* 64, 1479–1512

[14] Kandel, Eugene and Edward P. Lazear (1992): "Peer Pressure and Partnerships." *Journal of Political Economy* 100, 801–817

[15] Kremer, Michael (1993): "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics* 108, 551–575

[16] Oyer, Paul (2004): "Why Do Firms Use Incentives That Have No Incentive Effects?" *Journal of Finance* 59, 1619–1649

[17] Oyer, Paul and Scott Schaefer (2005): "Why Do Some Firms Give Stock Options To All Employees?: An Empirical Examination of Alternative Theories." *Journal of Financial Economics* 76, 99–133

[18] Sakovics, Jozsef and Jakub Steiner (2011): "Who Matters in Coordination Problems?" *American Economic Review*, forthcoming

[19] Segal, Ilya (1999): "Contracting With Externalities." *Quarterly Journal of Economics* 64, 337–388

[20] Terviö, Marko (2008): "The Difference that CEOs Make: An Assignment Model Approach." *American Economic Review* 98, 642–668

[21] Winter, Eyal (2004): "Incentives and Discrimination." *American Economic Review* 94, 764–773

[22] Winter, Eyal (2006): "Optimal Incentives for Sequential Production Processes." *RAND Journal of Economics* 37, 376–390

[23] Winter, Eyal (2010): "Transparency and Incentives Among Peers." *RAND Journal of Economics* 41, 504–523