

Contracting With Synergies*

Alex Edmans	Itay Goldstein	John Zhu
Wharton and NBER	Wharton	Berkeley

July 13, 2011

Abstract

This paper studies optimal contracting under synergies: effort by one agent reduces a colleague's marginal cost of effort. Our framework allows for agents' contributions to synergies to be asymmetric – an agent's effect on his colleague's cost differs from his colleague's effect on him – and workers to vary in the number of synergistic relationships they enjoy. In a two-agent model, effort levels are always equal even if contributions to synergies are asymmetric. An increase in synergy raises total effort and total pay, consistent with strong equity incentives in small firms, including among low-level employees. Individual pay, however, is asymmetric, with the more influential agent receiving a greater share, even though both agents exert the same effort and have the same direct effect on output. 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. Our results are robust to whether the production function exhibits substitutes or complements.

KEYWORDS: Contract theory, complementarities, principal-agent problem, multiple agents, teams, synergies, influence.

JEL CLASSIFICATION: D86, J31, J33

*aedmans@wharton.upenn.edu, itayg@wharton.upenn.edu, johnzhu@math.berkeley.edu. We thank Salvatore Miglietta and seminar participants from BI Oslo and Cambridge for helpful comments. AE gratefully acknowledges financial support from the Dorinda and Mark Winkelman Distinguished Scholar award.

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 many non-corporate settings – the cost of presenting an academic paper at a 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 group of colleagues, 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 will vary across agents. A CEO will exhibit synergies with each of his divisional managers, but a pair of divisional managers may 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* is the sum of the (unidirectional) influence parameters across a given group of agents. We consider a general framework which 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. 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 explains a number of real-life compensation practices that contradict standard single-agent compensation models, such as the high level of equity incentives given to low-level employees with little direct effect on output. It also 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.

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 him to exert a similarly high level of effort. This specific result in the two-person model illustrates a more general point of the model (for any number of agents) – while influence parameters are individual and may exhibit asymmetries 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 higher pay. This asymmetry 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 case 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, an increase in synergies unambiguously leads to higher total wages. This is consistent with the high level of equity incentives paid in start-up firms. For example, Google implemented a broad-based option plan where rank-and-file employees (such as secretaries) received stock

options and ended up being very highly paid, even though they had 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 stock options 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 wages as explained above. 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 an interior level of 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 exerts 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 depends 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.

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. 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 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 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.

If one synergy component becomes zero – for example, if two divisional managers exhibit synergies with the CEO but not with each other – then the two non-synergistic agents can be aggregated into a single employee and the model again reduces to the two-agent case. Thus, the CEO exerts 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, it 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, 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.

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) analyzes two special features of a team-based setting: a free-rider problem when all agents contribute to a joint output, and the possibility of using relative performance evaluation to reduce noise in evaluating each agent. There are no synergies in his model; the only interaction between the agents exists through a joint production function in which the efforts are perfect substitutes rather than exhibiting complementarities. Much of the subsequent literature on team-based incentives has focused on the free-rider problem, which involves symmetric synergies in production rather than asymmetric influence in the cost function. 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. In none of these models does effort by one agent affect the cost function of a colleague.

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 participation decisions (e.g. the decision to buy a product) rather than the choice of an effort level; there is no output or production function as in this paper. 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, which is a zero-one choice in contrast to the continuous effort decision studied in this paper. Here, the magnitude of the externality an agent exerts depends not only on who is participating, but also the effort level he chooses and the effort choices of his colleagues. We show how the nature of synergies affects the optimal actions implemented by the principal, which in turn changes the optimal contract.

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 $i = 1, 2, \dots, N$. Each agent is protected with limited liability and has a reservation utility of zero. Each agent exerts an 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, $r \in \{0, 1\}$. We will sometimes refer to $r = 1$ as “success” and $r = 0$ as “failure”. The probability of success depends on effort levels of all agents as follows:

$$\Pr(r = 1) = P(p_1, p_2, \dots, p_N). \quad (1)$$

The central feature of our model is that 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:

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

where

$$\varepsilon_{ij} \geq 0 \quad 1 \leq i \neq j \leq N$$

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 . Note that the multiplicative specification in (2) means

that effort by agent i reduces the marginal cost of effort by agent j , rather than just the total cost. 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 $c_i(p)$.

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. Each agent's utility is given by his wage minus his cost of effort, i.e.:

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

The principal maximizes expected output net of wages paid to the agents, i.e. solves:

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

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

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

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.

2.1 A Useful Maximization Problem

We will make repeated use of the following maximization problem throughout this paper. 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

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

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 $\frac{b - \sqrt{b^2 - 4a^*(b)}}{2a^*(b)} = 1$ and $x^*(a, b)$ smoothly increases up to 1.

(iii) If $b > 1$ then $\frac{b - \sqrt{b^2 - 4a^*(b)}}{2a^*(b)} < 1$ and $x^*(a, b)$ **explodes** up to 1 upon reaching the critical threshold $a^*(b)$.

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

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:

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

We assume a quadratic individual cost function:

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

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

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

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

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

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 is 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*

$$p^*(s) = \begin{cases} \frac{4 - \sqrt{16 - 12s}}{6s} & s \in (0, s^*) \\ 1 & s \geq s^*. \end{cases}$$

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,

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

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. However, greater effort by the more influential agent decreases the cost of effort of his colleague, inducing the latter to exert more effort.

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. To glean the intuition, the synergy can be thought of as an “echo” between the two agents. When $s < s^*$, the echo 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.

Here, wages are unambiguously increasing in the synergy parameter s . Synergies have no direct effect on output; they only affect output indirectly through changing the other agent's cost of effort and in turn affecting his effort choice. When choosing his effort level, agent i does not take into account this indirect effect via agent j 's cost function, because he takes agent j 's effort level as given. Thus he does not internalize his externality, and so the principal chooses to give him a sharper contract to cause

him to do so.

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 secretary reduces the cost of a CEO going to a meeting. 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. Oyer (2004) proposes an alternative explanation for broad-based option plans: 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.

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. In addition, it also increases agent i 's wage in both 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.

3.2 The Main Three-Agent Model

We now present the three-agent model which is the core analysis of this section, as it allows us to study the effect of an agent's scope of influence on his wage. The production function (1) now specializes to:

$$\Pr(r = 1) = \frac{p_1 + p_2 + p_3}{3}. \quad (7)$$

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:

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

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. 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 \mathbf{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:

$$Ap_2^*(\mathbf{s}) + Bp_3^*(\mathbf{s}) = Ap_1^*(\mathbf{s}) + Cp_3^*(\mathbf{s}) = Bp_1^*(\mathbf{s}) + Cp_2^*(\mathbf{s}) \quad (8)$$

which implies

$$\frac{p_1^*(\mathbf{s})}{p_2^*(\mathbf{s})} = \frac{C}{B} \frac{A + B - C}{A + C - B} \quad \frac{p_2^*(\mathbf{s})}{p_3^*(\mathbf{s})} = \frac{B}{A} \frac{A + C - B}{B + C - A} \quad \frac{p_3^*(\mathbf{s})}{p_1^*(\mathbf{s})} = \frac{A}{C} \frac{B + C - A}{A + B - C}. \quad (9)$$

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.¹ 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 direction of the synergy profile such that the optimal effort profile is interior. An increase in agent i 's relative influence increases both his relative and

¹Recall that part (i) implies that, in this interval, the direction of the optimal effort profile is fixed.

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.*) (i) 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.

(ii) Suppose there is a single synergy component that is zero. Then the sum of the efforts exerted by the two agents who do not have synergy on each other equals the effort of the other agent. By viewing the two non-synergistic agents as a single agent, 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:

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 wage coincides with the ratios of optimal effort.*

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

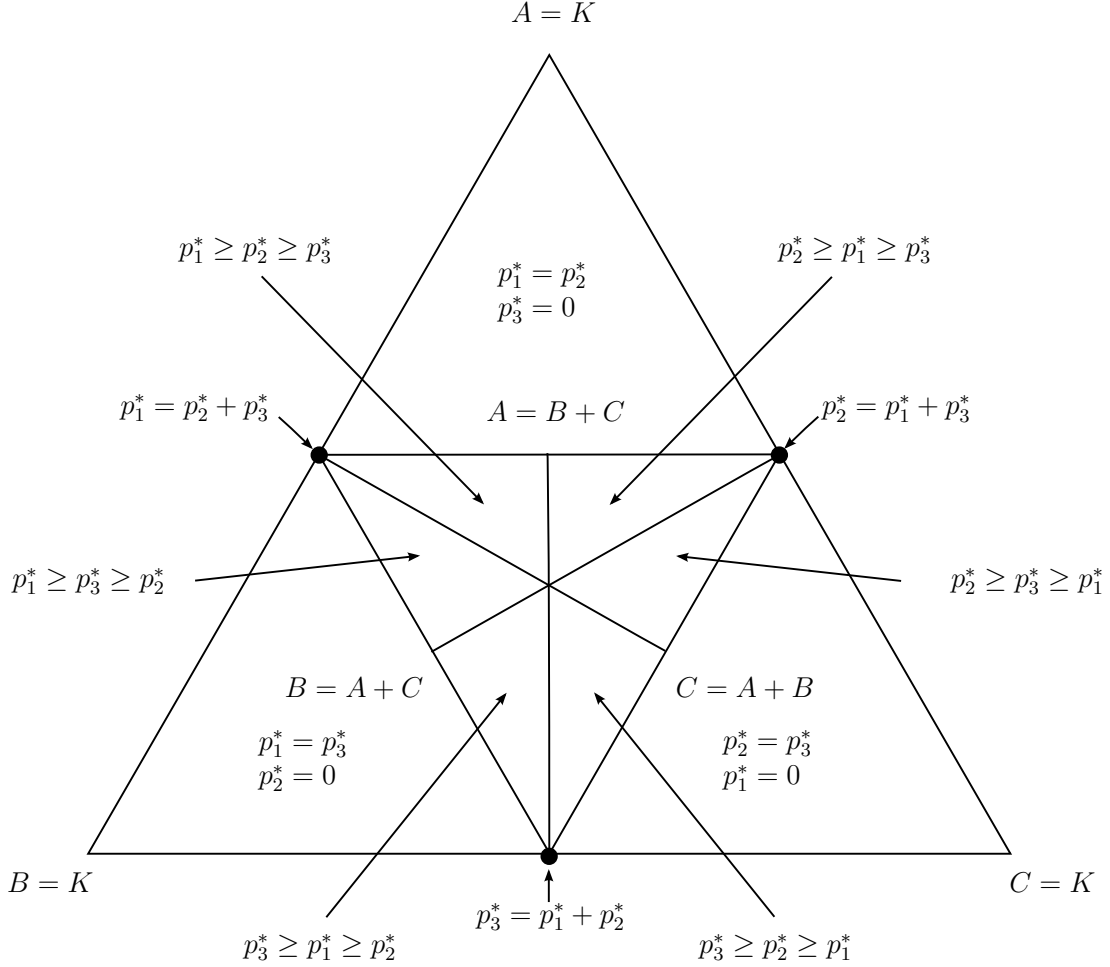


Figure 1: The $A + B + C = K$ simplex where $K > 0$ is some constant.

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 1. 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; 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. 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 synergies are 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 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.

Proposition 3 considers the case of a boundary effort profile. Part (i) 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 is striking because the third agent still has the same direct effect on the production function (7) as the other two agents, yet is being completely ignored. Moreover, it means that even if there is no change at all 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 to 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 the agent 3's optimal effort level being zero.

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 2 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. Conventional wisdom is that any division that enjoys strictly 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.

While in Proposition 2, all three synergy components matter for the optimal effort profile, here 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 becomes irrelevant. Wages are then determined as in the two-agent model and depend on the relative influence of each agent.

Part (ii) of Proposition 3 considers the case where a single synergy component is zero. In this case, the two non-synergistic agents can be considered a single agent in aggregate and the model again reduces to the 2-agent model of Proposition 1. The third agent, who exhibits synergies with both of his colleagues, exerts the same effort level as the two other agents combined. This may correspond to the case of a CEO and two divisional managers. For example, assume agent 1 is a CEO and agents 2 and 3 are divisional managers. Agent 1 exhibits synergies with each of the divisional

managers, but agents 2 and 3 exhibit no synergies with each other. This corresponds to the case of $C = 0$ and we are on the straight line on the left side of the outer triangle, between the points $A = K$ and $B = K$. The CEO exerts the highest effort level; the relative effort levels of the two divisional managers depends on the relative size of their synergies with the CEO (A and B).

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. 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).

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 a pair of agents, then relative wages within this pair are the same. This allows the entire wage profile to be fully solved: Corollary 1 states that the ratios of optimal wages coincides with the ratios of optimal effort. Harder-working agents are paid more.

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:

$$\Pr(r = 1) = \min(p_1, p_2, \dots, p_N). \quad (10)$$

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:

$$p_1 = p_2 = \dots = p_N \equiv p, \quad (11)$$

and

$$w_i(p) = \kappa_i p \left(1 - \sum_{j \neq i} \varepsilon_{ji} p \right). \quad (12)$$

These first-order conditions already give us some preliminary results. Equation (11) shows that all agents will exert the same effort level, as is intuitive given the perfect complementarities production function (10). Equation (12) 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 (11) and (12) 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 4 below.

Proposition 4 (*Complementary production function.*) (i) *There exists a unique critical synergy threshold $s^*(\kappa) > 0$ such that optimal effort is given by:*

$$p^*(s) = \begin{cases} \frac{2\kappa - \sqrt{4\kappa - 12s}}{6s} & s \in [0, s^*(\kappa)) \\ 1 & s \geq s^*(\kappa). \end{cases}$$

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 nondistortionary² then there is an increase in agent i 's 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 4 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.

²In other words, the decrease in the other agents' total influence is achieved by simply multiplying their influence parameters with a common scalar $c < 1$.

5 Conclusion

This paper has studied the effect of synergies on optimal effort levels and wages in a team-based setting. 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 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. 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

Proof of Lemma 1

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

$$\frac{d}{dx}U(x, a, b)|_{x=1} = 1 - 2bx + 3ax^2|_{x=1} = 1 - 2b + 3a > 0 \quad (13)$$

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 $b^2 - 3a = 3a(b \cdot \frac{b}{3a} - 1) > 0$. This implies $\frac{b}{3a} > 2$. But $\frac{b}{3a}$ is the inflection point of $U(x, a, b)$. Combined with (13), we conclude that if a local maximum exists, the corresponding argument value is greater than 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,

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

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 $x^*(a, b) = 1$. So 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 $\tilde{a} \leq \frac{b^2}{3}$. Fact 2 and Fact 3 imply that $U(x^{loc}(a, b), a, b)$ and $U(1, a, b)$ satisfy the single crossing property on the interval $[0, \tilde{a}]$. Also, it is clear they must cross at some point $a^*(b)$ and Fact 4 implies that on $[0, a^*(b))$, $x^*(a, b) = x^{loc}(a, b)$.

Proof of Lemma 2

The first claim follows from Fact 1 in the proof of Lemma 1. For the second 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 third 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 $a^*(b) < 1$ is if both $x^{loc}(a^*(b), b)$ and the inflection point are both strictly smaller than 1.

Proof of Lemma 3

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

$$\begin{aligned} \frac{d}{dx}U(x, a, b)|_{x^*(a, b)} &= 1 - 2bx^*(a, b) + 3ax^{*2}(a, b) = 0 \\ \Rightarrow \frac{d}{da}U(x^*(a, b), a, b) &= -2bx_1^*(a, b) + 6ax^*(a, b)x_1^*(a, b) + 3x^{*2}(a, b) = 0 \end{aligned} \quad (14)$$

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 (14) 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

Statements (i), (ii) and (iii) are essentially transcriptions of Lemmas 1, 2 and 3. 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).

Proof of Proposition 2

Holding total effort constant,

$$p_1^*(\mathbf{s}), p_2^*(\mathbf{s}), p_3^*(\mathbf{s}) \in \arg \max_{p_1, p_2, p_3 \in [0,1]} Ap_1p_2 + Bp_1p_3 + Cp_2p_3 \quad (15)$$

The first-order conditions which characterize interior solutions to this convex problem are captured by equation (8). This proves (i).

Since the maximization problem of equation (15) is convex, the optimal effort profile will satisfy the ratios of equation (9) 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 agent 1's effort is highest, agent 2's effort is agent 1's effort and agent 3's effort is agent 1's effort. Then the principal's maximization problem becomes

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

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. 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 (15), 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. Statement (i) then follows from the preliminary two-agent case.

Without loss of generality, suppose $C = 0$. If $A \neq B$ then the result is already covered by Statement (i). Thus we can assume $A = B$. Looking at the principal's reduced-form maximization problem, it is clear that the optimal choice of p_2 and p_3

is defined only up to their sum. Moreover, the maximization problem is symmetric between agent 1 and the sum of agents 2 and 3 and there is nontrivial synergy between them. Statement (ii) 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 (8) 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 4

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] Che, Yeon-Koo and Seung-Weon Yoo (2001): “Optimal Incentives for Teams.” *American Economic Review* 91, 525-541
- [5] Gabaix, Xavier and Augustin Landier (2008): “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics* 123, 49-100
- [6] 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
- [7] Holmstrom, Bengt (1982): “Moral Hazard in Teams.” *Bell Journal of Economics* 13, 324-340
- [8] Kandel, Eugene and Edward P. Lazear (1992): “Peer Pressure and Partnerships.” *Journal of Political Economy* 100, 801-817
- [9] Kremer, Michael (1993): “The O-Ring Theory of Economic Development.” *Quarterly Journal of Economics* 108, 551-575
- [10] Oyer, Paul (2004): “Why Do Firms Use Incentives That Have No Incentive Effects?” *Journal of Finance* 59, 1619-1649
- [11] 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
- [12] Segal, Ilya (1999): “Contracting With Externalities.” *Quarterly Journal of Economics* 64, 337-388
- [13] Terviö, Marko (2008): “The Difference that CEOs Make: An Assignment Model Approach.” *American Economic Review* 98, 642-668

- [14] Winter, Eyal (2004): “Incentives and Discrimination.” *American Economic Review* 94, 764-773
- [15] Winter, Eyal (2006): “Optimal Incentives for Sequential Production Processes.” *RAND Journal of Economics* 37, 376-390
- [16] Winter, Eyal (2010): “Transparency and Incentives Among Peers.” *RAND Journal of Economics* 41, 504-523