Contracting With Synergies*

Alex Edmans, LBS, Wharton, NBER, and ECGI
Itay Goldstein, Wharton
John Zhu, Wharton

June 1, 2013

Abstract

This paper studies multi-agent optimal contracting with cost synergies. We model synergies as the extent to which effort by one agent reduces his colleague's marginal cost of effort. An agent's pay and effort depend not only on the synergies he exerts but also the synergies others exert on him. It may be optimal to "over-work" and "over-incentivize" a synergistic agent, due to the spillover effect on his colleagues. This result can rationalize the high pay differential between CEOs and divisional managers. An increase in the synergy between two particular agents can lead to a third agent being endogenously excluded from the team, even if their productivities are unchanged. This result has implications for optimal team composition and firm boundaries.

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, zhuyiran@wharton.upenn.edu.
We thank Patrick Bolton, Pierre Chaigneau, Jonathan Cohn, Willie Fuchs, Xavier Gabaix, Brett Green, Dirk Jenter, Ernst Maug, Salvatore Miglietta, Alan Schwartz, Anjan Thakor, and seminar participants from BI Oslo, Cambridge, LSE, Philadelphia Fed, UC Irvine, Financial Intermediation Research Society, Financial Research Association, Georgia State Conference on Incentives and Risk-Taking, the NBER Summer Institute (Law & Economics), SFS Cavalcade, and the Washington University Conference on Corporate Finance for helpful comments, and Ali Aram and Christian Goulding for good research assistance. 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, 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. Synergistic relationships are typically asymmetric: a CEO has a greater impact on a divisional manager than the other way round. Moreover, the number of synergistic relationships 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. It analyzes the effect of synergies on optimal pay and effort for not only the synergistic agent but also his colleagues, thus deriving further implications for the effect of synergies on total pay and effort across an organization, and relative pay and effort across agents. 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, and the optimal composition of a team or boundaries of a firm.

In our theory, agents contribute to the production of a joint output. 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 combination of the unidirectional influences between two 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.

We start with a model with N agents and general cost functions, to illustrate our influence concept in a broad setting. Without synergies, the effort level that the principal induces from a given agent is a trade-off between the benefits of effort (increased productivity) and its costs (increased wages). With synergies, two additional forces enter the effort determination equation. First, the cost of inducing effort is reduced due to the influence that the agent enjoys from his colleagues. Second, the benefit to the principal of inducing effort is increased: effort not only has a direct effect on output, but also reduces his colleagues' cost of effort and thus makes it easier to incentivize them. The size of this benefit is represented by the cross-partial derivative of his colleagues' cost function – the extent to which their marginal cost of effort is reduced by an increase in his own effort. Introducing synergies also causes two additional forces to enter into the wage determination equation. On the one hand, a higher optimal effort level will tend to increase the required wage. On the other hand, the influence that an agent enjoys from his colleagues will reduce his cost of effort and thus the wage.

In general, synergies inject several new forces into an optimal contracting model; these forces often work in different directions. To obtain tractable results, we specialize the model to a standard setup of a linear production function and quadratic cost function. We first consider a two-agent model before moving on to a three agent extension. In our particular setup, an agent's influence on another agent is captured by a single parameter. Moreover, we can meaningfully define the common synergy between two agents to be the sum of their (unidirectional) influence parameters. Two forces affect the relative effort levels of the agents. It would seem that the more influential agent should work harder, as his effort is more useful because it reduces his colleague's cost function. However, there is a force in the opposite direction: the principal takes advantage of this reduced cost by increasing the effort of his colleague. In our model, these two effects exactly cancel out and both agents exert the same effort level. While this cancellation is specific to our functional forms, the general point is that an agent's optimal effort level not only depends positively on his own influence on his colleague, but also positively on his colleague's influence on him. Indeed, with our functional forms, the common synergy is a sufficient statistic for the optimal effort. That is, optimal effort depends on the individual influence parameters only up to their sum.

While effort levels are equal, 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 each agent has the same direct productivity. 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 influence, and so leads him to increase his effort, as desired by the principal.

An increase in the common synergy leads to the principal implementing a higher effort level, and offering higher total wages. The result is consistent with the high level of equity incentives in start-up firms, including to rank-and-file employees. Standard principal-agent theory suggests it is never optimal to give equity to a low-level employees with little direct effect on output. 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.

An increase in agent i's influence parameter always increases i's wage, but the effect on j's wage is more nuanced. Agent j's wage increases if and only if his influence parameter is above a critical threshold, otherwise it *decreases*. The intuition is as follows. If the principal held j's wage constant, an increase in i's influence would raise 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 j's influence is sufficiently low, his effort is less useful to the team than i's. Then, the principal prefers to extract some of the surplus (created by j's lower cost of effort) by lowering agent j's wage, accepting a lower increase in his effort, and reinvesting the savings to further increase i's wage. In contrast, if j's influence is sufficiently high, the principal reinforces the increase in j's effort level by augmenting his wage, since this generates a large synergy benefit to agent i, which ultimately benefits the principal. In short, synergy determines the (common) effort level and total pay, and influence determines the agents' relative pay.

We then extend the analysis to three agents, which allows us to study differences in synergies across pairs of agents. In this setting, a "synergy component" refers to the sum of the bilateral influence parameters between a pair of agents. 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 1 will exert a higher effort level than 2. 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 whether 1 or 2 exerts higher effort will depend on who exerts more influence on 3, not who is influenced more by 3, since only the former affects the usefulness of their effort. However, if 3 has a greater influence on 1, it is less costly for the principal to induce effort from 1. 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 where one synergy component is close to zero – for example, if two divisional managers exhibit synergies with the CEO but not each other. The two non-synergistic agents can be aggregated into one and the model approximates 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. The increase in CEO in recent decades is consistent with significant developments in communication technologies, which augment a CEO's scope of influence.[1]

If one synergy component becomes sufficiently large compared to the other two, the model collapses to the two-agent setting. Thus, 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 (such as his own cost and productivity). Even if the third agent's synergy parameters do not change, if the synergy component between his two colleagues increases, this can lead to the third agent being optimally excluded as the principal focuses all of his resources on compensating the first two agents. 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 direct productivity as the first two. Similarly, if the agents are interpreted as divisions of a firm, the model has implications for firm boundaries. Conventional wisdom suggests that a division should be divested only if it does not exhibit synergies with the rest of the conglomerate. 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. It is relative, not absolute, synergies that matter.

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 uses relative performance evaluation. There are no synergies in either costs or production.[2] A rich literature, summarized by Bolton and Dewatripont (2005, Chapter 8), has built on both of these settings, analyzing questions such as collusion, mutual monitoring, and the optimal hierarchical structure, but does not consider synergies. Itoh (1991) and Ramakrishnan and Thakor (1991) study a multi-tasking problem

[1] Garicano and Rossi-Hansberg (2006) also show how improvements in communication technologies lead to increased wage inequality within an organization.

[2] 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 where efforts are perfect substitutes.

where one action increases an agent's own output, and a separate action increases his colleague's output. Here, there is a single team output, and each agent takes a single action which both improves the joint output and reduces his colleague's marginal cost. Some papers have focused on the free-rider problem under production complementarities. Che and Yoo (2001) study a repeated setting, where an agent can punish a shirking colleague by shirking himself in the future. Kremer (1993) studies maximum 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 under production complementarities.

We show that production complementarities are inherently different from the cost synergies modeled in our paper. In our paper, effort by one agent reduces his colleague's marginal cost of effort. The agent does not take into account this *externality* when making his effort decision because he is paid according to output. Synergy does not affect output. The principal prefers each agent to internalize the synergy benefits as this encourages higher effort. In order to achieve this, the principal must increase wages. On the other hand, the agent does internalize any increase in production complementarity because by definition such a change affects output. Since the agent is more productive, he will automatically exert the new, higher, effort even without any change in the contract.

Closest 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 showing 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 reservation utilities rather than cost functions. The agents' actions are observable participation decisions rather than an unobservable effort choice; there is no output or production function. Studying effort choice (out of a continuum) rather than a zero-one participation decision leads to several new results on the effect of synergies on absolute and relative effort levels. Bernstein and Winter

(2012) also focus on a participation decision, as in Segal (1999), and study 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 another's cost of effort.

The paper proceeds as follows. Section 2 illustrates our influence concept in a general model. We specialize the framework to particular functional forms in the two-agent model of Section 3 and the three-agent model of Section 4. Section 5 discusses the difference between cost and production complementarities and Section 6 concludes. Appendix A contains proofs, Appendix B analyzes negative influence, and Appendix C considers production complementarities in addition to cost synergies.

2 The Influence Concept

The goal of this section is to illustrate our notion of influence in a general model, to show how it affects the principal's choice of both effort and wages. 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

$$p_i \in [0, 1] \qquad i = 1, 2, \ldots N.$$

The agents' efforts affect the firm's output, $r \in \{0, 1\}$, which is publicly observable and contractible. We will sometimes refer to $r = 1$ as "success" and $r = 0$ as "failure". The probability of success is:

$$\Pr(r = 1) = P\left(p_1, \ldots, p_N\right) = P\left(p\right), \tag{1}$$

where $p \equiv (p_1, ..., p_N)$ is a vector containing the effort levels of each agent i. As is standard, we assume that $P_i > 0$ and $P_{ii} < 0 \; \forall \; i$. We also specify $P_{ij} = 0$, i.e. no production complementarities, to ensure that our results are driven by the cost synergies that are the focus of the model. In Section 5 we discuss the difference between production complementarities and cost synergies.

The key feature of our model is that each agent's cost of effort $C^i\left(p\right)$ depends not only on his own effort level p_i, but also the effort levels exerted by his colleagues. We specify agent i's *overall* cost function as:

$$C^i(p) = g_i(p_i)\,\Pi_{j\neq i} h_{ji}(p_j) \qquad i = 1, 2, \ldots N, \tag{2}$$

The function $g_i(p_i)$ represents agent i's *individual* cost function, and depends on only his own effort level. As is standard, we assume that $g_i'(\cdot) > 0$ and $g_i''(\cdot) > 0$. The function $h_{ji}(p_j)$, where $h_{ji}'(\cdot) < 0$, represents j's *influence* on i, and captures the extent to which effort by j reduces the cost of effort by i. Note that, due to the multiplicative formulation in (2), j's effort reduces not only i's total cost of effort, but also his marginal cost and thus his effort incentives. The production function P and the cost functions C^i are common knowledge before contracting takes place.

Turning to the contract, it is automatic that each agent i will be paid zero upon failure. The principal chooses the optimal wage $w_i \geq 0$ to pay agent i upon success, that maximizes her expected output net of wages. After wages have been chosen, each agent i, observing the entire wage profile, selects his effort p_i to maximize his expected utility, given by his wage minus his cost of effort. Agents choose their efforts simultaneously, and their effort levels constitute a Nash Equilibrium.

Before we move to the analysis, a couple of points about the setup are worth making. First, agent i sets his effort p_i without observing his colleagues' effort levels (but rather only correctly expecting them in equilibrium). Since i's cost of effort depends on his colleagues' effort levels, this implies that agent i chooses his own effort without observing the implied cost (only correctly expecting it in equilibrium). The model also applies to the case in which agents choose their efforts sequentially, but each agent does not observe the efforts already taken by other agents when choosing his own. For example, a CEO may commit to a business trip in advance, but the exact cost he bears in making the trip will depend on the preparation conducted by his secretary, which is not known to the CEO until after the trip is underway.*[Itay, can you think of an example that does not involve secretaries? The issue is that they do not have a direct effect on output so they do not closely fit our model. Perhaps divisional managers.]* Alternatively, the cost function C^i may contain elements of private benefit – e.g., the extent to which the CEO enjoys his trip – which again depend on the efforts of his colleagues. Indeed, our model can be reinterpreted as each agent suffering cost $g_i(p_i)$ for choosing effort p_i, but then receiving private benefit $g_i(p_i)\left(1 - \Pi_{j\neq i} h_{ji}(p_j)\right)$.

Second, due to the combination of risk neutrality and limited liability, the agent is always paid zero upon failure. Thus, if the principal wishes to increase agent i's incentives, she will raise his success payoff w_i, but is unable to accompany this increase with a reduction in the agent's failure payoff (as this is bounded below by zero). Thus, an increase in w_i augments not only the agent's incentives (the sensitivity of pay to performance) but also expected pay (the level of pay, often referred to as the "wage" in empirical studies). Thus, in terms of empirical implications, all results pertaining to w_i are predictions for the level as well as the sensitivity of pay. The model will continue to generate predictions for both levels and sensitivities under risk aversion and unlimited liability. While the principal will now be able to accompany a rise in the success payoff w_i with a reduction in the failure payoff, the increase in the sensitivity of pay will cause the agent to demand a risk premium, again augmenting the level of pay.

2.1 Analysis

The principal solves for the optimal vector of effort levels, $p^* = (p_1^*, ..., p_N^*)$:

$$\max_{\{p_i\},\{w_i\}} P\left(p_1, \ldots, p_N\right)\left(1 - \sum_i w_i\right) = PM, \tag{3}$$

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

$$p_i \in \arg\max_{p_i} w_i P(p) - C^i\left(p\right) \qquad i = 1, 2, \ldots N. \tag{4}$$

Assuming that the first-order approach is valid, we can replace each incentive compatibility constraint (4) by the corresponding first-order condition:

$$w_i P_i(p) - C_i^i\left(p\right) = 0, \tag{5}$$

which yields the optimal wage as

$$w_i = \frac{C_i^i}{P_i}.$$

We first illustrate how the influence concept affects the principal's choice of effort levels. The effect of increasing p_i on her objective function is given by:

$$\frac{\partial}{\partial p_i}|_{p^*} PM = P_i\left[1 - \sum_i w_i\right] + P\left[-\frac{C_{ii}^i P_i - C_i^i P_{ii}}{P_i^2} - \sum_{j \neq i} \frac{C_{ij}^j}{P_j}\right] \tag{6}$$

Substituting for C^i, (6) can be written as

$$\begin{aligned} & P_i\left[1-\sum_i w_i\right] \\ & -P\frac{g_i''(p_i^*)\Pi_{j\neq i}h_{ji}(p_j^*)P_i - g_i'(p_i^*)\Pi_{j\neq i}h_{ji}(p_j^*)P_{ii}}{P_i^2} \\ & -P\sum_{j\neq i} g_j'\left(p_j^*\right)\frac{h_{ij}'(p_i^*)\Pi_{m\neq i,j}h_{mj}(p_m^*)P_j}{P_j^2}. \end{aligned} \quad (7)$$

The first-order condition with respect to p_i, (7), comprises of three components. The first is the increase in firm value from augmenting p_i, due to increased production, multiplied by the principal's share after paying wages to all agents. The second is the increased wage that the principal needs to pay agent i to induce this higher effort level, which results from the convexity of the cost function ($g_i'' > 0$) and the concavity of the production function ($P_{ii} < 0$). The first component represents a benefit to the principal and the second component a cost. Both components also exist in models without influence, and the optimal effort level is a trade-off between them. However, in a model with influence, the second component contains an additional $h_{ji}(p_j^*)$ term: the influence that agent i receives from his colleagues reduces his marginal cost of effort, and thus the cost to the principal of inducing more effort from him. Thus, in a model with influence, the optimal effort level depends not only on an agent's own productivity P_i and cost of effort g_i (functions specific to him), but also his colleagues' influence functions (functions outside his control).

The third component is the change in the wage paid to i's colleagues as a result of augmenting p_i. It arises because an increase in i's effort reduces his colleagues' costs of effort, since $h_{ij}'(p_i^*) < 0$. This in turn reduces the wage that the principal must pay to j, and represents a benefit to the principal. This entire component is also specific to a model of synergy. It demonstrates an additional benefit of effort that the principal must take into account, and will tend to increase i's optimal effort level: he becomes "over-worked" compared to a model without influence. Even though the direct productivity of his effort is unchanged, his effort has an indirect benefit of making it easier to incentivize his colleagues.

We next study how influence affects the principal's choice of wage levels. Agent i's wage is given by

$$w_i^* = \frac{C_i^i(p_i^*)}{P_i(p_i^*)} = \frac{g_i\left(p_i^*\right)}{P_i\left(p_i^*\right)}\Pi_{j\neq i}h_{ji}\left(p_j\right).$$

Without influence, the wage is given by $\frac{g_i(p_i^*)}{P_i(p_i^*)}$, the ratio of the marginal cost of effort to the marginal productivity of effort, both evaluated at the implemented effort level. Thus, influence causes the wage to change in two ways. First, as shown in equation (7), it affects the implemented effort level and thus the ratio $\frac{g_i(p_i^*)}{P_i(p_i^*)}$. Second, an additional term $\Pi_{j \neq i} h_{ji}(p_j)$ appears in the wage determination equation, since the influence of agent i's colleagues reduces his marginal cost of effort. Thus, agent i's wage, just like his effort, is affected by his colleagues' influence, in addition to functions specific to him. In general, these two effects may work in opposite directions, and so with general functional forms, the overall effect of influence on wages can be highly complex.

We thus move to specific functional forms to allow a tractable study of the determination of effort and wages in a synergistic setting. Moving to a tractable model will allow us to analyze not only how changes in agent i's influence function affect his own effort and wage levels, but also his colleagues' effort and wage levels. Solving for them will allow us to study the relative effort and wage levels between the agents, as well as the total effort and wage levels across the organization. We start with a two-agent model, as this illustrates the additional results most clearly, and then move to a three-agent model which delivers further implications.

3 The Two-Agent Model

We now specialize the production function (1) to the following:

$$\Pr(r = 1) = \frac{p_1 + p_2}{2} \tag{8}$$

and the individual cost function to:

$$g_i(p_i) = \frac{1}{4} p_i^2. \tag{9}$$

We thus specify both agents as having the same direct productivity on output and the same individual cost function. Thus, any differences in efforts and wages will be a result of differences in influence, rather than in direct productivity or individual cost.

The influence function is given by:

$$h_{ji}(p_j) = 1 - h_{ji} p_j.$$

The constant $h_{ji} \in [0, 1)$ is an *influence parameter* that represents the extent to which

j's effort reduces i's cost of effort. For now we consider the case of non-negative influence parameters; in Appendix B we allow for $h_{ji} < 0$.

Agent i's overall cost function is thus:

$$C^i(p) = \frac{1}{4} p_i^2 (1 - h_{ji} p_j). \tag{10}$$

Our characterization of the optimal contract does depend on the form of the cost function given by (10). This is quite natural, since introducing influence injects several new forces in the model which work in different directions. Under alternative cost functions, the relative magnitudes of these counteracting forces may be different, leading to potentially different overall effects on the contract. However, while the *magnitude* of the effects is specific to our functional form, the *existence* of these effects and the economic intuition behind them will continue to hold with alternative functional forms. Thus, the general objective of our analysis is to identify the new forces that arise in a synergistic setting under any functional form. We choose a particularly tractable specification to make the economic drivers of the model as transparent as possible. In particular, this will allow us to uncover interesting phenomena that might otherwise be overlooked in less tractable settings.

The principal's program now specializes to:

$$\max_{\{p_i\},\{w_i\}} \frac{p_1 + p_2}{2} (1 - w_1 - w_2), \tag{11}$$

$$\text{s.t.} \quad p_i \in \arg\max_{p_i} \frac{p_i + p_j}{2} w_i - \frac{1}{4} p_i^2 (1 - h_{ji} p_j) \qquad i = 1, 2. \tag{12}$$

Differentiating i's utility function (12) gives his first-order condition as:

$$w_i = p_i (1 - h_{ji} p_j). \tag{13}$$

Plugging this into the principal's objective function (11) gives her reduced-form maximization problem as:

$$p_1^*, p_2^* \in \arg\max_{p_1, p_2} \frac{p_1 + p_2}{2} (1 - (p_1 + p_2) + p_1 p_2 (h_{12} + h_{21})). \tag{14}$$

We can now meaningfully define synergy:

Definition 1 ***Synergy*** *s is defined as the sum of the influence parameters $h_{12} + h_{21}$.*

Definition 1 highlights the conceptual advantage provided by the quadratic cost function we have assumed. With general cost functions, it makes no sense to define synergy as the sum of the influence parameters. Indeed, it may be difficult to designate any particular function of model parameters as the appropriate definition of *synergy*. Despite this conceptual hurdle, it is intuitive that each agent's optimal contract should depend both on his influence on the other agent and the other agent's influence on him. Moreover, it also seems quite reasonable that the larger these influence parameters are, the more the optimal contract departs from the single-agent optimal contract. Our definition of synergy plus our use of the synergy concept in the following propositions allow us to state these legitimate, but otherwise vague, intuitions in a precise manner.

[Potential alternative] Definition 1 highlights the conceptual advantage provided by the quadratic cost function we have assumed. For any cost function, it is intuitive that each agent's optimal contract will depend on both his influence on his colleague and his colleague's influence on him. Under our cost function (10), these influence parameters aggregate in the principal's reduced-form maximization problem (14) and so we can define synergy as the sum of the individual influence parameters. We can thus analyze the common synergy between agents – the sum of the individual influence parameters of each particular agent – in a precise manner.

We 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 optimal contract is characterized by Proposition 1:

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.*

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

Part (i) shows that there are two effects that determine the agents' relative effort levels. Suppose that agent i is more influential. On the one hand, i's greater influence tends to increase his optimal effort level, relative to j's, since his effort has greater cost reduction benefits. However, there is an effect in the opposite direction: the principal takes advantage of this cost reduction by increasing j's effort. With our chosen functional forms, the two forces exactly cancel out. More generally, synergy, which is a symmetric function of the influence parameters, is a "sufficient statistic" for the optimal effort levels: an influence parameter matters only through its effect on total synergy.

Part (ii) states that various measures of wage levels are increasing with synergy. While intuitive, these results are 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 – 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 cost of effort, because these two effects exactly offset each other.

Here, wages are unambiguously increasing in the synergy parameter s. The economics are as follows. Consider a rise in synergy that arises from an increase in h_{ij}. The principal wishes to induce more effort from agent i, but i's greater influence does not give him incentives to exert this greater effort level. He is paid according to output, and so his incentives to exert effort depend on the effect of his effort on output. Changing his influence parameter h_{ij} has no effect on his direct productivity: his marginal benefit from effort is $\frac{w_i}{2}$ and independent of h_{ij}. Thus, the principal increases w_i to augment agent i's marginal benefit, and cause him to internalize this externality. Agent i is thus "over-incentivized" compared to a model without synergy. 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, the complementarity would affect the agent's marginal productivity and would be internalized even under the original contract. Thus, wages may be independent of the complementarity. We will return to this point in Section 5.

That incentives rise in synergy is a potential explanation for why high equity incentives are sometimes given to junior employees, even if they have a small direct effect on output. High equity incentives are optimal if they have a significant impact on their colleagues' costs – 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. Hochberg and Lindsey (2010) document systematic evidence of broad-based option plans. Kim and Ouimet (2013) find that they increase productivity only in firms with a small number of employees, where synergy potential is likely greatest.[3]

Parts (i) and (ii) show that effort and total wages only depend on total synergy, rather than the individual influence parameters. Part (iii) shows that the individual influence parameters do affect the relative pay of each employee. The more influential agent receives the higher wage. This result holds even though both agents exert the same effort and have the same direct productivity. Instead, the wage differential is driven by two mutually reinforcing effects. Since the more influential agent exerts a greater externality, his wage must be higher to induce him to internalize his externality. Moreover, since his colleague is less influential, he suffers a higher marginal cost of effort, further increasing his required wage. Part (iii) leads to empirical predictions for within-firm differences in pay: more influential agents should receive higher wages, even if they perform the same tasks. For example, senior faculty are paid more than junior faculty even though they have the same formal job description; the former can reduce the latter's cost of effort through mentorship and guidance.

Part (iv) compares the utility of the two agents. The more influential agent receives a higher wage, but also bears a higher cost since he is helped out less by his colleague. The first effect is stronger, and so he enjoys higher utility.

Corollary 1 concerns the comparative static effects of changes in influence parameters.

Corollary 1 *(i) Suppose synergy is subcritical. An increase in either influence param-*

[3]Note that our model can only explain equity compensation to non-C-level 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. 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.

eter increases optimal effort, total wages given success, and expected total wages. (ii) Fix a subcritical synergy level. An increase in agent i's relative influence (i.e. increasing h_{ij} and lowering h_{ji} so that s is unchanged) increases both his relative and absolute wage. Specifically,

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

Part (i) of Corollary 1 follows naturally from parts (i) and (ii) of Proposition 1. Since an increase in one influence parameter, holding the other constant, raises the total synergy level s, it will raise the effort levels of both agents, total wages, and expected total wages. As discussed earlier, an increase in agent i's influence raises not only his optimal effort level, as his effort is now more useful to the principal, but also j's optimal effort level since j is now easier to incentivize. Part (ii) of Corollary 1 states that if agent i's relative influence rises (h_{ij} increases and h_{ji} decreases so that s remains constant), i's wage goes up both in absolute terms and also relative to j's wage.

The next Corollary considers the effect on absolute compensation due to a rise in h_{ij} while holding h_{ji} constant.

Corollary 2 *Suppose synergy is subcritical. An increase in h_{ij} increases w_i^*. An increase in h_{ij} increases w_j^* if and only if h_{ji} is sufficiently high. Specifically,*

$$\frac{\partial}{\partial h_{ij}} w_j^* \begin{cases} > 0 & h_{ji} \in \left(\frac{1}{6p^*(s)}, s^* - h_{ij}\right) \\ = 0 & h_{ji} = \frac{1}{6p^*(s)} \\ < 0 & h_{ji} \in \left[0, \frac{1}{6p^*(s)}\right) \end{cases} . \tag{15}$$

Finally, both $p^*w_i^*$ and $p^*w_j^*$ are increasing in h_{ij}.

It is clear that an increase in h_{ij} augments w_i^*, since total wages rise in synergy and agent i's share of total wages rises due to his greater relative influence. However, the effect on w_j^* is more subtle because there are two conflicting effects: total wages rise, but agent j's share of total wages falls. Corollary 2 provides an illuminating characterization of which force dominates when and why. As h_{ij} increases, j's cost of effort goes down. The principal can react to this in two ways. She can lower w_j^*, reasoning that a slight decrease in wage will diminish j's rent while still inducing a higher effort level from j than before. Conversely, she can increase w_j^*, arguing that the fall in j's cost of effort means that incentivizing him is more effective. Proposition 2 shows that the latter option is desirable if j's effort is particularly useful, i.e. if j's influence on i is particularly high.

The intuition is as follows. The total effect the influence parameters have on the agents' absolute effort levels unfolds through an "echo" mechanism between the two agents. The influence of agent i on agent j raises i's optimal effort level, which reduces j's cost of effort, which raises j's optimal effort level, which, due to the influence of agent j on agent i, reduces i's cost of effort, which raises i's optimal effort level, and so on. In this process, both influence parameters help increase each agent's optimal effort. This "echo" intuition explains why, as synergy increases, the common effort level increases (part (i) of Proposition 1). For the echo mechanism to work well, both sides of the echo – i.e. both influence parameters – need to be strong. Thus, if and only if h_{ji} is sufficiently high, the principal wants to further increase w_j^* to take advantage of the echo. The threshold level for h_{ji} is $\frac{1}{6p^*(s)}$, which is decreasing in the common effort level and thus the common synergy. The larger the synergy, the greater the echo potential, and therefore the more willing the principal is to increase w_j^*. Put differently, if the echo is strong enough, the increase in h_{ij} causes such a large increase in total wages that it outweighs the fall in j's share of the wage pool. Separately, while the change in j's absolute wage depends on h_{ji}, the expected wage $p^*w_j^*$ unambiguously rises (regardless of h_{ji}), due to the increase in the optimal effort level p^*.

Given the importance of influence for pay, a natural question to ask is whether influence is inherent to a person, or to his position in the organization. Either may be true depending on the setting. The former will apply to a manager with strong leadership skills that encourage his colleagues to work hard, either through increasing their private benefit from working (employees enjoy working for an inspirational leader) or, equivalently, reducing their cost of doing so. Under this interpretation, influence can be considered a dimension of managerial talent. In talent assignment models with moral hazard (e.g. Edmans, Gabaix, and Landier (2009)), talent affects productivity and is thus internalized by the agent without the need to modify the contract; here, influence affects colleagues' cost functions and so affects the optimal wage. The latter will apply to an employee who occupies a central position in an organization that allows him to exert influence. He may be in this position through historical accident or entrenchment, but even though his high wages are a result of luck (occupying a central position) rather than rents to talent, they are efficient.

4 The Three-Agent Model

We now extend the model to three agents. Under this setup, synergy will no longer be a single parameter shared by all agents, but several dimensions that vary across agents: in particular, one agent may enjoy synergies with both of his colleagues, but his colleagues may not enjoy synergies with each other. We show that these differences across synergy levels will lead not only to differences in wages as in the two-agent model, but also differences in effort.

The production function (1) now becomes:

$$\Pr(r=1) = \frac{p_1 + p_2 + p_3}{3}. \tag{16}$$

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 (5) gives his first-order condition as:

$$w_i(p_i) = p_i \left(1 - \sum_{j \neq i} h_{ji} p_j \right), \tag{17}$$

and plugging this into the principal's objective function (3) gives her reduced-form maximization problem as:

$$p_1^*, p_2^*, p_3^* \in \arg \max_{p_1, p_2, p_3 \in [0,1]} \frac{(p_1 + p_2 + p_3)}{3} \left(1 - (p_1 + p_2 + p_3) + Ap_1p_2 + Bp_1p_3 + Cp_2p_3\right), \tag{18}$$

where

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

We now generalize our synergy concept:

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* $s = ||(A, B, C)||$.

Each synergy component is an analog of the synergy scalar s in the two-agent model. In the three-agent model, there are three relevant synergy components for each of the three pairs of agents, which together form the synergy profile $\boldsymbol{s}$.

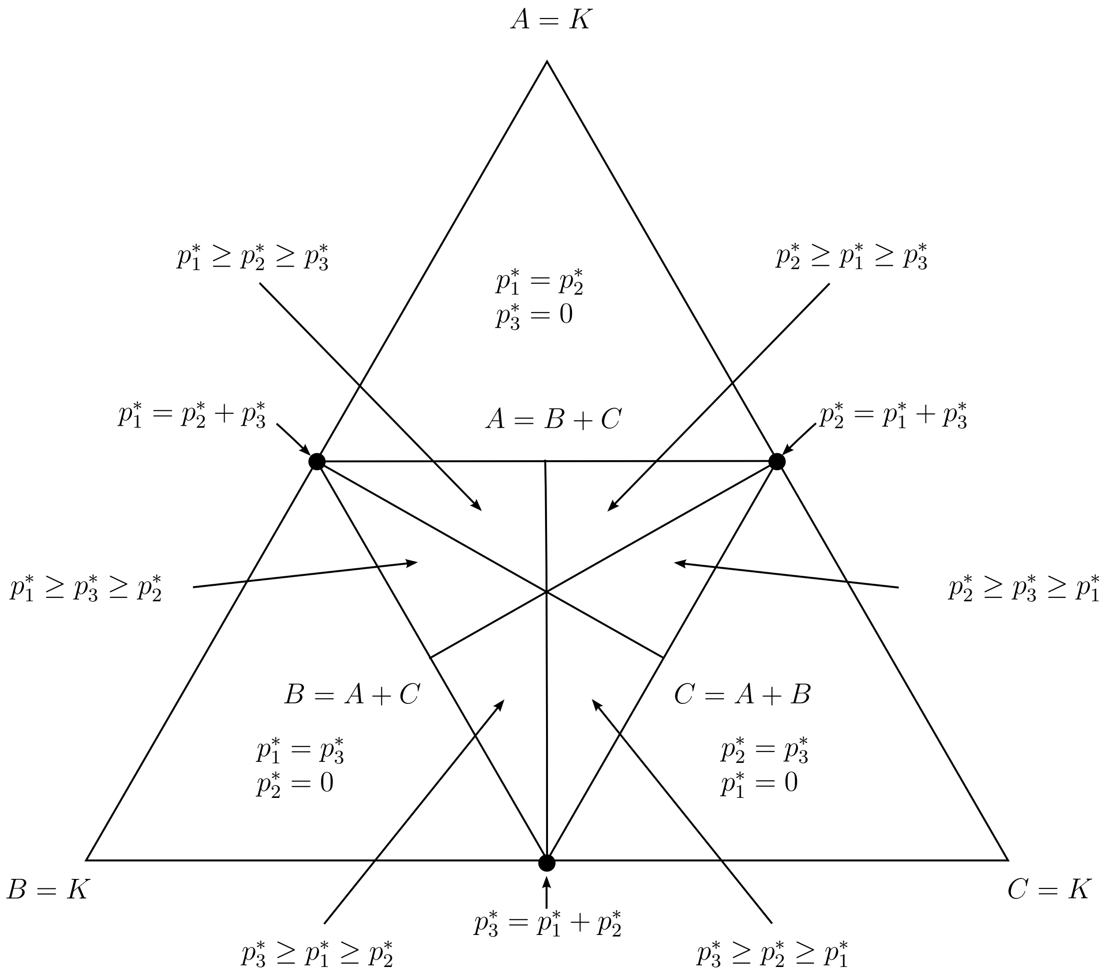

Figure 1: The optimal efforts as a function of the synergy profile (A, B, C). A cross section of the synergy profile space: $A + B + C = K$ where $K > 0$ is some constant.

The solution to the model is given by Proposition 2 below.

Proposition 2 *The optimal efforts are functions of the synergy profile:* $(p_1^*(\mathbf{s}), p_2^*(\mathbf{s}), p_3^*(\mathbf{s}))$.
(i) If each synergy component is weakly smaller than the sum of the other two, then every agent exerts effort and the direction of the synergy profile determines the direction of the optimal effort profile. The optimal efforts satisfy

$$\frac{p_1^*(\mathbf{s})}{p_2^*(\mathbf{s})} = \frac{C}{B}\frac{A+B-C}{A+C-B} \quad ; \quad \frac{p_2^*(\mathbf{s})}{p_3^*(\mathbf{s})} = \frac{B}{A}\frac{A+C-B}{B+C-A} \quad ; \quad \frac{p_3^*(\mathbf{s})}{p_1^*(\mathbf{s})} = \frac{A}{C}\frac{B+C-A}{A+B-C}. \tag{19}$$

(ii) If a single synergy component is strictly larger than the sum of the other two, then the model degenerates into the two-agent model. The third agent who does not enjoy the largest synergy component exerts no effort and receives no wage. The other two agents' effort and wage levels are determined as in Proposition 1.

The simplex in Figure 1 illustrates the characterization in Proposition 2. It fixes the sum of the synergy components $A + B + C$ at a constant K and studies the effect of changing their relative level. Part (i) of Proposition 2 states that, if the synergy components are balanced so that no single component exceeds the sum of the other two, the optimal effort profile is interior and given by equation (19). This equilibrium is illustrated by the middle triangle, bounded by the three dots. 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, 1 exerts more effort than 2.

Note that it is the *total* synergy between agents 1 and 3 (relative to the total synergy between 2 and 3) that determines the relative values of p_1 and p_2, not 1's unidirectional influence on 3, h_{13} (relative to 2's unidirectional influence on 3, h_{23}). It may seem that p_1 should only depend on h_{13} (and not h_{31}) as only the former affects the usefulness of 1's effort. However, when h_{31} rises, 1's cost function is lower and so it is cheaper to implement a higher p_1. The intuition is similar to the two-agent model: it is the total synergy that matters, not one's individual influence parameter. Hence, we have $p_1 > p_2$ when $B > C$, rather than when $h_{13} > h_{23}$.

On the one hand, this result extends the principle in the two-agent case, that the optimal effort level depends on the common synergy, not the individual influence parameters. On the other hand, the result contrasts the two-agent case, since all agents no longer exert the same effort level. In the two-agent case, there is only one synergy component and so one common effort level. Here, the existence of three synergy components allows for asymmetry in effort levels between the three agents.

A natural application of the model is where one synergy component (say C) is close to zero. For example, agent 1 is a CEO who shares synergies with two division managers, agents 2 and 3, but they share few synergies with each other. Figure 1 shows that agent 1 exerts the highest effort. Essentially, agents 2 and 3 can be aggregated, and their combined effort level is close to the effort exerted by agent 1.

Part (ii) of Proposition 2 shows that, when the synergy between two agents is sufficiently strong relative to the other two synergies, then the model collapses to the two-agent model of Proposition 1. Intuitively, if the synergy between two agents is sufficiently strong, then only these two 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 has the same direct effect on the production function (16) as the other two, yet is completely ignored. Moreover, it means that even if there is no change to the synergies between agent 3 and his colleagues, an increase in the synergies between agents 1 and

2 can lead to him being excluded. Thus, 3's participation depends not only on his own synergies with others, but also on the synergy components that do not involve him.

That the third agent is excluded seems surprising, since we have a convex cost function and the marginal cost of effort for agent 3 is zero. Thus, it may seem cheaper to increase agent 3's effort from 0 to ε than to increase the effort of agents 1 and 2 from $p-\varepsilon$ to p. The key to the intuition is that it is not the marginal cost of effort that is relevant but the marginal increase in wage that the principal needs to offer to induce effort. From (17), the wage $w_i(p_i) = p_i\left(1 - \sum_{j\neq i} h_{ji}p_j\right)$, is linear, rather than convex, in the effort level. It is true that the principal has to pay agents 1 and 2 more than 3 in *absolute* terms, due to the convex cost of effort, but the *marginal* increase in the wage is the same for all agents. Even though the marginal cost of effort is high for agents 1 and 2, the marginal benefit for them is also high since they are already receiving a wage (and thus a share of output); thus only a small increase in the wage is required. For agent 3, the marginal cost of additional effort is low, but the marginal benefit is also low since he currently has no share of output. Thus, without synergies, it is just as costly to increase effort by agent 3 from 0 to ε as it is to increase effort by agents 1 and 2 from $p-\varepsilon$ to p. With synergies, it is more beneficial to induce effort from agents 1 and 2 rather than 3. Another way to view the intuition is that increased synergy between 1 and 2 raises the value of the firm, and thus the cost to the principal of giving 3 equity to induce effort from him. It becomes so expensive to induce effort from 3 that the principal chooses not to do so.

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 he has just as high direct productivity as the existing two agents and has strictly positive synergies with them. Moreover, the agents can be interpreted as divisions of a firm, in which case Proposition 2 has implications for firm boundaries. If two divisions exhibit sufficiently strong synergies with each other, it may be optimal to divest a third division even if it exhibits strictly positive synergies with the first two. Conventional wisdom is that any division that enjoys positive synergies should be included within a firm. Here, it is relative, not absolute, synergies that determine firm boundaries. 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.[4]

[4] This implication assumes that it is not possible to compensate the division based only on its own performance, but only according to the performance of the overall conglomerate. While divisional

When one of the agents is excluded, Proposition 1 completely characterizes optimal effort as a function of synergy. When all three agents exert effort, the effect of the synergy profile on the optimal effort profile is similar to the two-agent results.

Corollary 3 *Suppose every agent exerts effort. Fix a direction of the synergy profile. 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. At the critical synergy size s^*, the optimal effort profile explodes so that at least one agent is now applying effort 1.*

Analogous to part (i) of Proposition 1, Corollary 3 states that, when synergy size increases, effort by each agent becomes both more useful and easier to incentivize, and so it is optimal for the principal to implement a higher effort profile. When synergies become sufficiently strong, the principal implements the maximum effort level of 1 for at least one agent.

Corollary 4 turns from the optimal effort profile to the optimal wage profile. It shows that, when all three agents exert effort, the comparative statics of the relative and absolute sizes of the wages are in the same spirit as the two-agent results.

Corollary 4 *(i) Suppose every agent exerts effort and synergy is subcritical. Total wages given success and expected total wages are strictly increasing in s up to the critical synergy size s^*.*

(ii) 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 $\{h_{ij}\}_{j \neq i}$ and decreasing some elements of $\{h_{ji}\}_{j \neq i}$ so that $\mathbf{s}$ is unchanged) increases both his relative and absolute wage. Specifically,

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

and

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

(iii) Suppose further that the influence between any pair of agents is symmetric. Then the ratios of optimal wages coincide with the ratios of optimal efforts.

profit measures are typically available, there is only a stock price for the overall conglomerate. The stock price incorporates many additional pieces of information than profits, such as growth prospects; consequently, managerial pay is typically much more sensitive to the stock price than to profits (see, e.g., Murphy (1999)).

Part (i) of Corollary 4 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 (ii) of Corollary 4 is analogous to part (iv) of Proposition 1. An increase in one agent's influence augments his wage in both absolute and relative terms. Moreover, if the influence parameters are symmetric across all pairs of agents, the entire wage profile can be fully solved: part (iii) states that the ratios of optimal wages coincides with the ratios of optimal effort.

Given these results, the model can potentially 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 optimal since the centrality of the CEO leads to him exhibiting greatest synergies. High wages for the CEO lead his colleagues to anticipate that he will work harder, in turn inducing them to exert greater effort themselves.[5] More generally, the 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. The increase in communication technology over the past few decades has plausibly increased the CEO's influence, consistent with the rise in CEO pay over the same period. Talent assignment models argue that CEO pay depends on firm size (e.g. Gabaix and Landier (2008), Terviö (2008)), which 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 than the manager of a small firm where there are strong synergies (e.g. a start-up).

5 Cost Synergies vs Production Complementarities

A key feature of our model is that an agent's effort reduces the marginal cost of effort of his colleague – or alternatively increases the colleague's marginal private benefit from his own effort. This feature generates the synergies in our model. To what extent is this different from modeling complementarities in the production function so that an

[5]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 as per the model of Lazear and Rosen (1981)). They find that the pay differential between the CEO and other senior managers is positively related to firm performance. Aggarwal, Fu, and Pan (2011) use the pay differential as a proxy for non-CEO executives to monitor the CEO, as in the theory of Acharya, Myers, and Rajan (2011).

agent's effort increases the marginal *productivity* of his colleagues' effort?

One might be tempted to posit an equivalence between cost synergies and production complementarities. 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 factor. However, in a multi-agent world, we show there is a subtle but significant economic difference between the two: cost synergies are fundamentally different from production complementarities because the former represent a true externality, but the latter do not. Loosely speaking, *since 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, consider a two-agent model with production complementarities but no cost synergies. The production function is given by

$$\Pr(r=1) = a\frac{p_1+p_2}{2} + b\sqrt{p_1 p_2},$$

where b parameterizes the complementarity. The principal maximizes:

$$\left(a\frac{p_1+p_2}{2} + b\sqrt{p_1 p_2}\right)(1 - w_1 - w_2), \tag{20}$$

and agent i's objective function is:

$$\left(a\frac{p_i+p_j^*}{2} + b\sqrt{p_i p_j^*}\right) w_i - \frac{1}{4}p_i^2. \tag{21}$$

The solution is given in Proposition 3 below.

Proposition 3 *(Complementaries in production function, two agents.)*
(i) The optimal effort profile is $p_1^* = p_2^* = \frac{a+b}{4}$.
(ii) To implement an arbitrary effort profile (p_1, p_2)*, the principal offers wages* $w_1 = \frac{p_1}{a+b\sqrt{\frac{p_2}{p_1}}}$ *and* $w_2 = \frac{p_2}{a+b\sqrt{\frac{p_1}{p_2}}}$.
(iii) At the optimal effort profile (p_1^*, p_2^*)*, the principal offers wages* $w_1^* = w_2^* = \frac{1}{4}$.

Part (i) shows that the optimal effort profile is increasing in the production complementarity b, just as in the core model where it was increasing in the cost synergy s. Part (ii) demonstrates a counteracting effect absent from the core model. For a given effort level p_i, the wage required to implement this effort level is decreasing in

b. Since the production complementarity term $b\sqrt{p_1 p_2}$ is shared across agents, an increase in the production complementarity P_{ij} (through raising b) also augments agent i's marginal productivity P_i, as can be seen from i's objective (21). Since agent i takes into account his increased productivity, he does not need as high a wage to induce a given effort level. Indeed, part (iii) shows that, under the current functional forms, these two effects exactly cancel out: while a rise in b increases the optimal effort level, it also reduces the wage required to induce a given effort level, so the overall wage is unchanged. Since i fully internalizes his increased productivity, he will exert the new, higher, optimal effort level even under the original contract.

With other functional forms, it may be that the two effects do not exactly cancel out. The point that is robust to functional forms is that the second effect always exists (not that it exactly cancels out the first effect): changing production complementarities will, in general, elicit effort responses from the agents even if the contract is unchanged. The same cannot be said about influence parameters in our cost synergies model. It is this missing internalization by the agents in our cost synergies model that drives the non-trivial comparatives statics with respect to synergy.

Indeed, in the cost synergies model, the principal maximizes:

$$\frac{p_i + p_j}{2} \left(1 - p_i \left(1 - h_{ji} p_j\right) - p_j \left(1 - h_{ij} p_i\right)\right) \tag{22}$$

and agent i's objective function is:

$$\frac{p_i + p_j^*}{2} w_i - \frac{1}{4} p_i^2 \left(1 - h_{ji} p_j^*\right). \tag{23}$$

To implement an arbitrary effort profile (p_i, p_j), the principal offers agent i a wage of

$$w_i = p_i (1 - h_{ji} p_j). \tag{24}$$

Differentiating the principal's objective (22) with respect to p_i shows that, when i's influence h_{ij} increases, the principal wishes to implement a higher effort level p_i. However, inspecting the agent's objective (23) shows that agent i does not consider his influence on agent j, h_{ij}, when choosing his effort level: this term does not appear in his objective function (unlike the production complementarity b in (21)). An increase in cost influence has no effect on agent i's direct productivity: the only effect is on agent j's cost of effort which is noncontractible, and so i does not consider it when making his own effort choice. Unlike the production complementarity term which is shared across agents, the cost function $\frac{1}{4}p_j^2(1 - h_{ij}p_i)$ is specific to agent j, so a change in the cost synergy C_{ij} (through raising h_{ij}) does not affect agent i's marginal cost C_i. Therefore, the wage required to implement a given effort profile does not depend on the cost influence h_{ij} (see equation (24)).

The analysis of Section 2 showed that the effect of increasing agent 1's influence on his optimal effort level is independent of production complementarities. In Appendix C, we analyze the model of Sections 3 and 4 under the case of production complementarities as well as cost synergies. We show that the addition of production complementarities does not change the implications generated by cost synergies.

6 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 notion of synergy is fundamentally different than complementarities in the production function and leads to a number of new results.

In a general model, the existence of synergies generates additional forces that the principal must consider in the contracting problem. The optimal effort level will, in general, be higher due to two forces: effort from a given agent is cheaper to induce since he enjoys synergies from his colleagues, and more valuable to the principal due to the positive influence his effort has on others. The optimal wage level also changes, both due to the different optimal effort level, and also because synergy reduces the agent's marginal cost of effort.

The standard setting of a linear production function and quadratic cost function allows us to derive additional results. With two agents, each agent's optimal effort level depends not only on his productivity, cost, and influence, but also his colleague's influence. In fact, optimal efforts depends on the two influence parameters only up to their sum, which we define as the common synergy shared between the two agents. Since synergy is symmetric over influence, both agents exert the same effort. Wages, however, differ across agents with the more influential agent receiving higher pay. Total wages increase with the total level of synergy: since synergy is a true externality, agents do not take it into account when choosing their effort level, and so must be "over-incentivized" to cause them to internalize it. This result suggests that it may be optimal to grant rank-and-file employees strong equity incentives, even if their direct effect on output is low. An increase in one agent's influence parameter augments his own effort and pay. Interestingly, such an increase raises his colleague's pay if and only if his colleague is sufficiently influential.

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

$$\frac{\partial}{\partial x} U(x, a, b)|_{x=1} = 1 - 2bx + 3ax^2|_{x=1} = 1 - 2b + 3a > 0 \tag{25}$$

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 $\frac{\partial}{\partial x} U(x, a, b) > 0$ for $x = 1$ (from (25)), 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 $\frac{\partial}{\partial x} U(x, a, b) > 0$ 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 1 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,

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

Fact 3: On the other hand,

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

Fact 4: For all sufficiently low 1, $x^*(a,b) = x^{loc}(a,b)$. To see this, notice since $\lim_{a\downarrow 0} x^{loc}(a,b) = \frac{1}{2b} < 1$, so for all sufficiently low 1, 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 1 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 1'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 1. An increase in 1 “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 1 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 1 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 1's such that these constraints are satisfied. By contrast, if $U(x,a,b)$ has a single critical point, 1 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 $\frac{\partial U\left(x^{loc}(a,b),a,b\right)}{\partial a} < \frac{\partial U(1,a,b)}{\partial a}$, 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 1 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 $U\left(1, a^*\left(b\right), b\right) > U\left(1, x^{loc}\left(a^*\left(b\right)\right), b\right)$ (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 $[0, a^*(b))$

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

$$\Rightarrow \frac{\partial}{\partial a}U(x^*(a, b), a, b) = -2bx_1^*(a, b) + 6ax^*(a, b)x_1^*(a, b) + 3x^{*2}(a, b) = 0 \tag{26}$$

Now

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

Equation (26) then implies

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

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 $p_1 = p_2 = p$. This allows us to apply Lemmas 1 and 2 with $x = \frac{p_1+p_2}{2} = p$; parts (i) and (ii) are essentially transcriptions of these two 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.

For part (iii), note if i is more influential than j then $h_{ij} > h_{ji}$. This implies:

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

For part (iv), the first-order condition yields: $w_1 = p_1 (1 - h_{21}p_2)$. Thus, agent 1's utility is given by:

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

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

Proof of Corollary 1

Part (i) is a direct application of Lemma 3. For part (ii), holding synergy fixed, an increase in agent i's relative influence means both increasing h_{ij} and decreasing h_{ji}. This causes both an increase in w_i^* and a decrease in w_j^*.

Proof of Corollary 2

We use a dot to denote the derivative with respect to h_{ij}.

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

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

A linear combination of the two gives us

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

Since $\dot{p^*} > 0$, this means that, when $s < \bar{s}$, $\dot{w_j^*}$ and $6h_{ij}p^* - 1$ have the same sign. Equation (15) follows immediately. Turning to the expected wage, we have:

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

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

A linear combination of the two gives us

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

Proof of Proposition 2

Holding total effort constant,

$$p_1^*(\mathbf{s}), p_2^*(\mathbf{s}), p_3^*(\mathbf{s}) \in \arg \max_{p_1, p_2, p_3 \in [0,1]} Ap_1p_2 + Bp_1p_3 + Cp_2p_3. \tag{27}$$

The first-order conditions which characterize interior solutions to this convex problem are captured by

$$Ap_2^*(\mathbf{s}) + Bp_3^*(\mathbf{s}) = Ap_1^*(\mathbf{s}) + Cp_3^*(\mathbf{s}) = Bp_1^*(\mathbf{s}) + Cp_2^*(\mathbf{s}). \tag{28}$$

Part (i) is then easily deduced.

Without loss of generality, suppose $A > B \geq C$ and $A > B + C$. Looking at the convex problem of equation (27), it is clear that $p_3^* = 0$. The principal's maximization problem then becomes symmetric in p_1 and p_2 and there is nontrivial synergy between agents 1 and 2. Part (ii) then follows from the two-agent case.

Proof of Corollary 3

Since the maximization problem of equation (27) is convex, the optimal effort profile will satisfy the ratios of equation (19) 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 corollary and condition 2 holds if synergy is sufficiently small. Suppose then that synergy is small. Let p denote the highest effort of the optimal effort profile. Then there exists $1 \geq \alpha \geq \beta > 0$ such that the other two effort levels 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

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

The corollary now follows from Lemma 1.

Proof of Corollary 4

Statement (i) 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 $\{h_{ij}\}_{j \neq i}$ and a corresponding decrease of some elements in $\{h_{ji}\}_{j \neq i}$. 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. 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 (ii) now follows.

For (iii), recall the optimal wage for agent i is

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

Equation (28) 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 3

The principal solves:

$$\max_{p_1, p_2, w_1 \geq 0, w_2 \geq 0} \left(a \frac{p_1 + p_2}{2} + b\sqrt{p_1 p_2}\right)(1 - w_1 - w_2),$$

subject to

$$p_i \in \arg\max_{p \in [0,1]} \left(a \frac{p + p_{-i}}{2} + b\sqrt{p p_{-i}}\right) w_i - \frac{1}{4} p^2, \quad i = 1, 2.$$

We prove the Proposition using a series of lemmas.

Lemma 4 *Given wages w_1 and w_2, the agents' effort levels satisfy*

$$p_i = \begin{cases} 0 & \textit{if } w_i = 0, \\ \left(a + b\sqrt{\frac{p_{-i}}{p_i}}\right) w_i & \textit{if } w_i > 0 \textit{ and } (a + b\sqrt{p_{-i}}) w_i < 1, \\ 1 & \textit{if } (a + b\sqrt{p_{-i}}) w_i \geq 1, \end{cases}$$

for $i = 1, 2$.

Proof. Let

$$f_i(p) = \left(a \frac{p + p_{-i}}{2} + b\sqrt{p p_{-i}}\right) w_i - \frac{1}{4} p^2, \quad i = 1, 2,$$

denote agent i's expected utility function. Then, the first and second order derivatives are

$$f_i'(p) = \left(\frac{a}{2} + \frac{b}{2}\sqrt{\frac{p_{-i}}{p}}\right) w_1 - \frac{1}{2} p, \quad i = 1, 2,$$

$$f_i''(p) = -\frac{1}{4}\left(\frac{b}{p}\sqrt{\frac{p_{-i}}{p}}w_1\right) - \frac{1}{2} < 0, \quad i = 1, 2.$$

Since f_i is strictly concave and $[0,1]$ is a compact set, there is a unique maximizer p. Moreover, if $f_i'(0) \leq 0$, then $p = 0$. If $f_i'(1) \geq 0$, then $p = 1$. Otherwise, if $f_i'(0) > 0 > f_i'(1)$, then there is a unique $p \in (0,1)$ given by the first-order condition. Specifically, note that $f_i'(0) = 0$ if $w_i = 0$, and for $w_i > 0$,

$$0 \geq f_i'(0) = \left(\frac{a}{2} + \frac{b}{2}\lim_{p\downarrow 0}\sqrt{\frac{p_{-i}}{p}}\right) w_i = \begin{cases} \frac{1}{2}aw_i & \text{if } p_{-i} = 0, \\ +\infty & \text{if } p_{-i} > 0, \end{cases}$$

so $p = 0$ if $w_i = 0$;

$$0 \leq f_i'(1) = \left(\frac{a}{2} + \frac{b}{2}\sqrt{p_{-i}}\right) w_i - \frac{1}{2},$$

so $p = 1$ if $w_i \geq \frac{1}{a+b\sqrt{p_{-i}}}$; and, finally, in all other cases, the first-order condition implies that $p \in (0,1)$ satisfies

$$p = \left(a + b\sqrt{\frac{p_{-i}}{p}}\right) w_i.$$

■

We will now substitute the wages into the principal's problem, so that it becomes a function of the effort levels p_i only.

Lemma 5 *The principal's problem reduces to*

$$(p_i^*, p_2^*) \in \arg\max_{\substack{0\leq p_1\leq 1,\\ 0\leq p_2\leq 1}} \left(a\frac{p_1+p_2}{2} + b\sqrt{p_1p_2}\right)\left[1 - \left(\frac{p_1}{a+b\sqrt{\frac{p_2}{p_1}}} + \frac{p_2}{a+b\sqrt{\frac{p_1}{p_2}}}\right)\right].$$

with optimal success wages given by

$$w_i^* = \begin{cases} 0 & p_i^* = 0, \\ \frac{p_i^*}{a+b\sqrt{\frac{p_{-i}^*}{p_i^*}}}, & 0 < p_i^* \leq 1. \end{cases} \tag{29}$$

Proof. Let $W_i^p = \{w_i \geq 0 : p_i = p\}$. By Lemma 4, $W_i^1 = [\frac{1}{(a+b\sqrt{p_{-i}})}, \infty)$. Since $\Pr(r = 1)$ is a nonnegative constant for all $w_i \in W_i^1$, and $1 - w_1 - w_2$ is decreasing in w_i, it is enough to limit consideration to the smallest element in W_i^1, which is $\frac{p_i}{(a+b\sqrt{p_{-i}})}$ evaluated at $p_i = 1$. By Lemma 4, $W_i^0 = \{0\}$, so $w_i = 0$ for $p_i = 0$, and $w_i = \frac{p_i}{(a+b\sqrt{\frac{p_{-i}}{p_i}})}$ for each $p \in (0,1)$. Hence, we can replace w_i in the objective function with $w_i = \frac{p_i}{(a+b\sqrt{\frac{p-i}{p_i}})}$, where we take $w_i = 0$ for $p_i = 0$ since $\lim_{p_i \downarrow 0} \frac{p_i}{(a+b\sqrt{\frac{p-i}{p_i}})} = 0$. ∎

Equation (29) is part (ii) of Proposition 3. It now remains to prove parts (i) and (iii). To solve the problem of Lemma 5, we introduce the following functions. Let

$$f(p,q) = g(p,q)(1 - h(p,q)), \tag{30}$$

where the functions C^1 and C^2 are defined by

$$g(p,q) = a\frac{p+q}{2} + b\sqrt{pq}, \tag{31}$$

$$h(p,q) = \frac{p}{a + b\sqrt{\frac{q}{p}}} + \frac{q}{a + b\sqrt{\frac{p}{q}}}, \tag{32}$$

with $h(0,z) = h(z,0) = \frac{z}{a}$.

Then, the reduced problem is equivalent to solving

$$\max_{\substack{0 \le p_1 \le 1, \\ 0 \le p_2 \le 1}} f(p_1, p_2). \tag{33}$$

Remark 1 *Note that this problem* (33) *is not concave. For example, let* $a = \frac{1}{10}$, $b = \frac{9}{10}$, *and consider* $p \in \{\frac{1}{500}, \frac{1}{50}, \frac{1}{5}\}$ *for* $q = \frac{9}{10}$. *Then,* $f(\frac{1}{500}, \frac{9}{10}) \approx -0.443$, $f(\frac{1}{50}, \frac{9}{10}) \approx -0.475$, *and* $f(\frac{1}{5}, \frac{9}{10}) \approx -0.357$. *So,* $\frac{1}{500} < \frac{1}{50} < \frac{1}{5}$, *but* $f(\frac{1}{500}, \frac{9}{10}) > f(\frac{1}{50}, \frac{9}{10}) < f(\frac{1}{5}, \frac{9}{10})$, *which violates concavity. Nevertheless, this problem does have a unique global maximum, which we demonstrate below.*

The next few lemmas help establish that the solution involves symmetric effort levels, $p_1 = p_2$.

Lemma 6 *For all* $p, q \ge 0$, *there exist* $x, y \ge 0$ *such that the functions* C^1 *and* C^2 *in* (31) *and* (32) *satisfy*

$$g(p,q) = g(x,x) \quad \text{and} \quad h(p,q) = h(y,y).$$

Namely,

$$x = \frac{a}{a+b}\frac{p+q}{2} + \frac{b}{a+b}\sqrt{pq}, \tag{34}$$

$$y = \frac{a+b}{2}\left(\frac{p}{a + b\sqrt{\frac{q}{p}}} + \frac{q}{a + b\sqrt{\frac{p}{q}}}\right), \tag{35}$$

$$g(x,x) = (a+b)x, \tag{36}$$

and

$$h(y,y) = \frac{2y}{a+b}. \tag{37}$$

Proof. The result follows by substitution of (34) into (31) and (35) into (32). ■

Lemma 7 *The functions C^1 and C^2 are homogeneous of degree one. Namely, for all $p,q \geq 0$, and for all $\alpha \geq 0$, we have $g(\alpha p, \alpha q) = \alpha g(p,q)$ and $h(\alpha p, \alpha q) = \alpha h(p,q)$.*

Proof. The result follows immediately from (31) and (32). ■

Lemma 8 *For all $p,q \geq 0$, the values x and y defined in Lemma 6 satisfy $x \leq y$. Moreover, $x < y$ if $p \neq q$.*

Proof. Let $\beta = \frac{b}{a+b}$. Then, $x \leq y$ if and only if

$$(1-\beta)\frac{p+q}{2} + \beta\sqrt{pq} \leq \frac{1}{2}\left(\frac{p}{(1-\beta)+\beta\sqrt{\frac{q}{p}}} + \frac{q}{(1-\beta)+\beta\sqrt{\frac{p}{q}}}\right). \tag{38}$$

Note that both sides of (38) are zero for $p = q = 0$. Moreover, if $p = 0$, then condition (38) reduces to

$$\frac{1}{2}(1-\beta)q \leq \frac{1}{2}\frac{1}{1-\beta}q,$$

which always holds since $0 < 1-\beta < 1 < \frac{1}{1-\beta}$. A similar property holds for p when $q = 0$, so $x \leq y$ when either $p = 0$ or $q = 0$.

Now, suppose that $p,q > 0$. Multiplying both sides of (38) by $\frac{2}{\sqrt{pq}}$ yields the equivalent condition

$$(1-\beta)\left(\sqrt{\frac{p}{q}} + \sqrt{\frac{q}{p}}\right) + 2\beta \leq \frac{\sqrt{\frac{p}{q}}}{(1-\beta)+\beta\sqrt{\frac{q}{p}}} + \frac{\sqrt{\frac{q}{p}}}{(1-\beta)+\beta\sqrt{\frac{p}{q}}} \tag{39}$$

Let

$$w = \sqrt{\frac{p}{q}} + \sqrt{\frac{q}{p}}.$$

Note that $w \geq 2$ for all $p,q > 0$. This lower bound on w follows by substitution of $\alpha = \sqrt{\frac{p}{q}}$ and minimization:

$$\min_{p,q>0} \sqrt{\frac{p}{q}} + \sqrt{\frac{q}{p}} = \min_{\alpha>0} \alpha + \frac{1}{\alpha} = 2,$$

which follows because $\alpha + \frac{1}{\alpha}$ is a strictly convex function on $\alpha > 0$ that is minimized at $\alpha = 1$ on $\alpha > 0$.

Now by substitution and simplification of the right-hand side of (39) into a single ratio, we have $x \leq y$ if and only if

$$(1-\beta)w + 2\beta \leq w\frac{(1-\beta) - 2\beta\frac{1}{w} + \beta w}{(1-\beta)^2 + \beta(1-\beta)w + \beta^2}. \tag{40}$$

Note that the left-hand side of (40) is a convex-combination of 2 and w, and because $w \geq 2$, is at most w. The right-hand side of (40) is at least w. To see this, note that because $w \geq 2$ and $\beta > 0$,

$$1 + \beta(w-2) \geq 1 \geq \frac{2}{w},$$

which after multiplying by β and rearranging gives

$$\beta^2 w + \beta(1-2\beta) \geq \frac{2\beta}{w}.$$

Now, using the identities $\beta^2 = \beta - \beta(1-\beta)$ and $\beta(1-2\beta) = (1-\beta) - [(1-\beta)^2 + \beta^2]$, we have

$$[\beta - \beta(1-\beta)]w + (1-\beta) - [(1-\beta)^2 + \beta^2] \geq \frac{2\beta}{w},$$

which can be rearranged as

$$(1-\beta) - 2\beta\frac{1}{w} + \beta w \geq (1-\beta)^2 + \beta(1-\beta)w + \beta^2.$$

This inequality implies that

$$\frac{(1-\beta) - 2\beta\frac{1}{w} + \beta w}{(1-\beta)^2 + \beta(1-\beta)w + \beta^2} \geq 1;$$

hence, the term multiplying w on the right-hand side of (40) is at least one. Thus, the right-hand side of (40) is at least w and the left-hand side of (40) is at most w. Therefore, $x \leq y$ for all $p, q > 0$. Furthermore, all of the above inequalities are strict if $p \neq q$ since $w > 2$, which leads to $x < y$. ■

Lemma 9 *For all $p, q \geq 0$, the function f in* (30) *satisfies*

$$f(p,q) \leq \frac{(a+b)^2}{8}.$$

Moreover, the inequality is strict for $p \neq q$.

Proof. First, note that $f(0,0) = 0$, and that for all $z \geq 0$,

$$f(0,z) = f(z,0) = \frac{1}{2}az\left(1 - \frac{z}{a}\right),$$

which is strictly concave in z on $z \geq 0$. Hence, the first-order condition yields the optimal $z = \frac{a}{2}$ and

$$f(0,z) = f(z,0) \leq f\left(\frac{a}{2}, 0\right) = \frac{a^2}{8} < \frac{(a+b)^2}{8}.$$

So, the result holds if $p = 0$ or $q = 0$. Now, suppose $p, q > 0$. The maximum value of f along the ray $\{(\alpha p, \alpha q) : \alpha \geq 0\}$ is given by

$$\max_{\alpha \geq 0} f(\alpha p, \alpha q) = \max_{\alpha \geq 0} \alpha g(p,q)(1 - \alpha h(p,q)) = \frac{1}{4}\frac{g(p,q)}{h(p,q)}.$$

The first equality follows by application of Lemma 7. The second equality follows from the first-order condition, which gives $\alpha = \frac{1}{2h(p,q)} > 0$ as the unique maximizer since the objective is strictly concave on $\alpha \geq 0$.

By Lemma 6, there exist $x, y \geq 0$ such that

$$\max_{\alpha \geq 0} f(\alpha p, \alpha q) = \frac{1}{4}\frac{g(p,q)}{h(p,q)} = \frac{1}{4}\frac{g(x,x)}{h(y,y)} = \frac{(a+b)^2}{8}\frac{x}{y}.$$

By Lemma 8, we have $\frac{x}{y} \leq 1$, with the inequality strict for $p \neq q$. Thus,

$$f(p,q) \leq \frac{(a+b)^2}{8},$$

with the inequality strict for $p \neq q$. ■

Lemma 10 *The solution to the model in which the synergy is in the production function, not the cost function, is the following. The agents exert the same level of effort*

$$p_1^* = p_2^* = \frac{a+b}{4},$$

and receive the same success wage

$$w_1^* = w_2^* = \frac{1}{4},$$

which results in optimal expected utility $\frac{a+b}{16}(1-\frac{a+b}{4})$ *for each agent and optimal profit* $\frac{(a+b)^2}{8}$ *for the principal.*

Proof. The reduced problem of (33) when $p_1 = p_2 = p$ is

$$\max_{p\geq 0} f(p,p) = \max_{p\geq 0} g(p,p)(1-h(p,p)) = \max_{p\geq 0}(a+b)p\left(1-\frac{2p}{a+b}\right) = \frac{(a+b)^2}{8}.$$

The second equality follows by Lemma 6. The third equality follows by the first order condition of the function being maximized, which is strictly concave in p on $p \geq 0$ and has the unique maximizer $p = \frac{a+b}{4}$. Hence, $p_1 = p_2 = \frac{a+b}{4}$ is the unique maximizer among all symmetric effort levels, and by Lemma 9 it is the unique global maximizer since it achieves an objective value that is strictly larger than that achieved by any non-symmetric effort levels. The optimal success wages are obtained by (29). ■

Lemma 10 proves parts (i) and (iii), and so the Proposition is proven.

B Negative Influence Parameters

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

B.1 The Two-Agent Model

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

$$p_1^*, p_2^* \in \arg\max_{p_1,p_2} \frac{p_1+p_2}{2}\left(1-(p_1+p_2)+p_1p_2(h_{12}+h_{21})\right).$$

There are thus two cases to consider.

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

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

Case 2. $h_{12} + h_{21} < 0$.

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 $h_{12} < h_{21}$.

(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 results continue 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. However, 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. 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.

B.2 The 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 Proposition 2 continues 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 2. 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 2.

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.

C Complementary Effort

In Sections 3 and 4, efforts were perfect substitutes in the production function. This section considers a model in which agents' efforts are perfect complements, i.e. the probability of success depends on the minimum effort level undertaken by all agents, and shows that the results are robust. The production function now becomes:

$$\Pr(r = 1) = \min(p_1, p_2, ..., p_N). \tag{41}$$

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 (5) gives his first-order conditions as:

$$p_1 = p_2 = \ldots = p_N \equiv p, \tag{42}$$

and

$$w_i(p) = \kappa_i p \left(1 - \sum_{j \neq i} h_{ji} p\right). \tag{43}$$

These first-order conditions already give us some preliminary results. Equation (42) shows that all agents will exert the same effort level, as is intuitive given the perfect complementarities production function (41). Equation (43) 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 (42) and (43) into the principal's objective function (3) gives her reduced-form maximization problem as:

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

We define the following terms:

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

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

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

Assumption 2 *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; the proof is essentially the same as in Proposition 1.

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[6] *then there is an increase in agent i's relative and absolute wage. 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 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 Sections 3 and 4, 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 Sections 3 and 4, 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.

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

References

[1] Acharya, Viral V., Stewart C. Myers, and Raghuram G. Rajan (2011): "The Internal Governance of Firms." *Journal of Finance* 66, 689–720

[2] Aggarwal, Rajesh K. Huijing Fu, and Yuhui Pan (2011): "An Empirical Investigation of Internal Governance." Working Paper, University of Minnesota

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

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

[5] Bernstein, Shai and Eyal Winter (2012): "Contracting with Heterogeneous Externalities." *American Economic Journal: Microeconomics*, forthcoming

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

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

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

[9] Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009): "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium." *Review of Financial Studies* 22, 4881–4917

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

[11] Garicano, Luis and Esteban Rossi-Hansberg (2006): "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics* 121, 1383–1435

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

[13] 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

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

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

[16] 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

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

[18] Kim, E. Han and Paige Ouimet (2013): "Broad-based Employee Stock Ownership: Motives and Outcomes." Working Paper, University of Michigan

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

[20] Lazear, Edward P., and Sherwin Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89, 841–864.

[21] Murphy, Kevin (1999): "Executive Compensation" in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3b. New York and Oxford: Elsevier/North-Holland, 2485–2563

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

[23] 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

[24] Ramakrishnan, Ram T.S. and Anjan V. Thakor (1991): "Cooperation Versus Competition in Agency." *Journal of Law, Economics, and Organization* 7, 248–283

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

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

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

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

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

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