Endogenous Contractual Externalities

Emre Ozdenoren and Kathy Yuan*

May 14, 2014

Abstract

We study effort and risk-taking behaviour in an economy with a continuum of principal-agent pairs where each agent exerts costly hidden effort. When the industry productivity is uncertain, agents have motivations to match the industry average effort, which results in contractual externalities. Contractual externalities have welfare changing effects when the information friction is correlated and the industry risk is not revealed. This is because principals do not internalize the impact of their choice on other principals' endogenous industry risk exposure. Relative to the second best, if the expected productivity is high, risk-averse principals over-incentivise their own agents, triggering a rat race in effort exertion, resulting in over-investment in effort and excessive exposure to industry risks relative to the second best. The opposite occurs when the expected productivity is low.

JEL classification codes: D86, G01, G30.
Keywords: Contractual externalities, relative and absolute performance contracts, boom-bust effort exertion and risk taking.

*Ozdenoren is from the London Business School and CEPR; e-mail: eozdenoren@london.edu. Yuan is from the London School of Economics and CEPR; e-mail: K.Yuan@lse.ac.uk. We thank Ulf Axelson, Patrick Bolton, Peter Kondor, Thomas Mariotti, Ailsa Röell, Oguzhan Ozbas, Vania Stavrakev, Gunter Strobl, and seminar participants at the Frankfurt School of Finance and Economics, HKUST, Hong Kong University, Imperial College, London School of Economics, London Business School, Shanghai Advance Institute of Finance (SAIF), European Finance Association Meeting 2013, and European Winter Finance Meeting 2014 for helpful comments, and Oleg Rubanov for excellent research assistance.

1 Introduction

It is important to understand the sources of industry boom and bust cycles, especially in light of the recent episodes in the high-tech and the finance industries. In these situations, the excited anticipation of the arrival of a "new technological era" of high productivities leads to over-investment and excessive risk-taking in the corresponding industry. This "overheating" in economic activities often contrasts with a subsequent crash where real investments and risks are substantially reduced. These pro-cyclical investment and risk-taking cycles have significant social and economic consequences (eg, the great recession of 2007-2009).

To explore the social inefficiency of industry boom-bust cycles, the theoretical literature focuses on the welfare changing effect of various pecuniary externalities in the presence of various financial frictions. In this paper, we propose a model of externality similar to pecuniary externalities but working through contractual arrangements and information frictions in the time of large positive industry productivity shocks. The model captures the pro-cyclical investment and risk-taking phenomenon, finds its properties and economic conditions under which it might arise. For example, the model predicts that social inefficient boom-busts occur in industries exposed to highly volatile common shocks and where firms are not able to coordinate in writing incentive contracts (to internalize the contractual externalities). It also predicts that excessive risk-taking during the boom is more likely to be on correlated risks rather than idiosyncratic risks. Existing empirical evidence supports these theoretical predictions. Hoberg and Phillips (2010) find that boom-bust cycles are more likely in industries with many firms and when the common industry productivity shocks are volatile and difficult to predict. Bhattacharyya and Purnananda (2011) have documented between 2000 and 2006, the period of financial industry boom, idiosyncratic risks have dropped by almost half while systematic risk have doubled among US commercial banks. That is, the potentially excessive risk-taking during the boom period is found to be correlated among firms in the same industry.

The externality in our model arises endogenously through the contractual arrangements between a continuum of principal and agent pairs in an industry. Specifically, we model a setting where the industry experiences a large productivity shock, a feature that is commonly associated with major technological innovations. Each firm in this industry has a principal and owns a project. Each principal offers a contract to a risk-averse agent who exerts costly and hidden effort. The industry productivity shock affects the return to effort of all agents in this industry. This contractual environment has two notable characteristics. First, it involves many principal-agent pairs which allows us to examine interactions among independent firms in an industry.[1] Second, when the return to effort is uncertain and is unknown to everyone in the economy (ie., the productivity shock is uncertain and unobserved), an agent's effort choice affects the level and the riskiness of his firm's output as well as the level and, potentially, the riskiness of his own compensation.

In this setup, each firm offers a contract that is based on a noisy signal of its agent's individual performance as well as a noisy signal of the average industry performance. In reality, there are a variety of circumstances where the noise in the signal about the average performance is quite large. For example, a hype around a new technology may cause a run up in the average stock price of the firms in an industry, without revealing the underlying industry productivity, eg, the dot.com boom in 1990s. The same hype would also generate a run up in individual stock prices that reflect agent-specific performances. In these circumstances, principals would be unsure whether the run up is fundamental or hype driven, and all signals could be very noisy. However, the difference between the agent-specific and the average performance signals, reflecting relative performance, can be quite informative. As a result, contracts are *effectively* weighted towards the more precise relative performance information.

Contracts that incorporate the industry average performance, ie, reflecting relative performance, incentivise agents to work harder but at the same time introduce a distinctive motive for agents to match each other's effort. This is because the industry productivity shock affects the marginal return of effort for the agents. When the industry shock gets riskier, an agent's project output becomes riskier and so does the industry average output. The difference, however, is less risky if the agent does not deviate from the industry average effort too much. By matching the industry average effort as close as possible, an agent can reduce his exposure to the industry productivity shock. This potentially introduces a feedback loop between individual and the industry average effort among agents. Therefore, given the contracts, an agent has to trade off performing well to get incentive pay versus performing to the industry average to shield himself from too much industry risk exposure.

The motivation of agents to match the industry average effort creates an externality in setting incentives among principals in the industry. The externality arises because principals take the industry average effort as given and do not take into account the impact of their choices on it. This externality has two opposing effects on the amount of performance incentives in equilibrium relative to the second best where a planner can coordinate the contracts for all principals in the industry. The first effect arises because principals do

[1] This contrasts with setups with one principal and many agents (eg, team incentives) or many principals and one agent (eg, common agency).

not take into account the industry risk exposure of other principals in the industry. When principals are risk averse, they have to trade off incentivising their agents to work harder versus exposing themselves to more industry risks in their projects. As discussed earlier, agents can hedge out the industry risk by matching the industry average effort. This means that principals are left with all the industry risk. When principals choose stronger incentives, they would expect a higher output but also a larger industry risk exposure for themselves, which is a cost they internalize. However, stronger incentives also result in larger industry risk exposures for other principals in the industry via the feedback loop between individual and industry average effort, which principals do not internalize. That is, when setting incentives, a principal optimally chooses her own risk-return tradeoff ignoring her impact on increasing other principals' industry risk exposure. By comparison, in the second best, a planner sets incentives by taking into account the feedback loop between industry average and individual effort choices. This means, relative to the second best, a principal has a tendency to provide stronger incentives for her agent.

The second effect goes in the opposite direction and arises because each principal perceives a unilateral deviation from the industry average as being too costly. To see why, suppose a principal increases her choice of incentive above the industry average unilaterally. This will incentivize her agent to work harder than the industry average which in turn exposes the agent to more industry risk. The principal has to compensate her agent for bearing this risk. By comparison, the cost of providing incentive is lower in the second best. A planner would recognize that a coordinated increase in incentives would extract more effort from agents without exposing them to higher industry risk. Hence, relative to the second best, the incentive provision could be lower in the decentralized equilibrium due to this cost consideration.

Overall, the equilibrium incentive provision, and hence efforts elicited from agents, could be either excessive or insufficient relative to the social optimum, depending on which of the two effects dominates. The first effect dominates if the expected industry productivity is high, eg, during a boom. In this case, a principal is more concerned about incentivising her agent to work harder to reap the high productivity benefit. She would like to elicit high effort from her agent by increasing performance incentives. The increased incentives triggers a rat race in effort in the whole industry. The equilibrium effort exertion by agents is excessive relative to the second best. By contrast, when the expected industry productivity is low, eg, during a recession, principals are more concerned about the cost of providing incentives and less so about the incentive provision. The logic is then reversed and the race is to the bottom: relative to the second-best, incentive provision in equilibrium is too low, leading to insufficient effort. The planner can, therefore, improve the total welfare by making performance-based incentives countercyclical: enforcing lower (higher) performance-based incentives among principals during booms (busts).

The socially suboptimal incentive provision in our model is crucially related to the fact that the productivity shock is correlated across all projects in the industry.[2] This assumption is empirically justified for industries that have experienced a major technological innovation like the dot.com or mortgage industries. To see how correlated productivity shocks are important from a theoretical perspective, consider the case when productivity shocks are purely systematic and common across projects. In this case, if an agent matches the industry average effort level, his performance relative to the industry average will not contain any exposure to the common productivity shock. Since contracts are effectively based on agents' performance relative to the industry average, agents can remove the systematic risk embedded in their performance-based compensations by matching their peers' effort-choice level.[3] This incentive to offset the systematic risk from their compensations by agents is the underlying cause for effort- and (systematic) risk-taking rat race during booms and race to the bottom during busts. By contrast, as we show in an extension of the model, when productivity shocks are idiosyncratic across projects within the industry, agents do not have motivation to match the industry average effort. In this case, since productivity shocks are uncorrelated, if an agent matches the industry average effort level, his performance relative to the industry average still contains the exposure to his firm's idiosyncratic productivity shock. Consequently, agents cannot remove the exposure to idiosyncratic productivity shocks in their compensations by matching their peers and hence have no reason to be influenced by their peers' effort choices. This unique prediction on excessive (insufficient) undertaking of systematic risks during booms (busts) is well supported by the data.[4]

In our model an agent's effort choice and the riskiness of his project are tightly linked. This is because the productivity of effort is random and correlated across firms, and thus

[2]That is, the productivity shock is systematic rather than idiosyncratic in nature. In this paper we treat the industry shock and the systematic shock as the same and use the terms interchangeably.

[3]In the less extreme case, when the risk in the productivity shock is correlated across projects, a similar result holds. In this case, agents reduce (instead of eliminate) the systematic risk embedded in the performance-based compensation by matching the average effort level in the industry. We focus on the case where the productivity shock is common for ease of exposition.

[4]To our best knowledge, only one other line of literature predicts excessive undertaking of systematic risks and the prediction is one-sided about booms. This literature studies the incentive for banks or bank managers to take on excessive risk collectively that causes the financial crisis due to bailouts (Acharya and Yorulmazer (2007); Acharya and Yorulmazer (2008); Farhi and Tirole (2011); and Acharya et al. (2011)).

when an agent increases his effort, both the expected return and the systematic risk exposure of the project are higher. We view this feature of the model desirable when studying excessive risk taking from a social perspective, especially considering that episodes of over (under) investment at the industry and/or the economy level are often observed together with excessive (insufficient) risk taking. In this way our setup and conclusions are different from those models where agents can choose effort and level of risk separately.[5] We acknowledge that in some settings agents can choose risk and return of the projects separately, however in other settings agents have to choose a portfolio of risk and return together. Put differently, agents may have to trade off investing effort in high-risk-high-return projects versus low-risk-low-return projects. By treating the risk-return as a portfolio, our framework complements and contributes to the understanding of sub-optimal risk-taking in the principal-agent framework. Importantly, we find that agents in our framework are only interested in taking a suboptimal amount of systematic rather than idiosyncratic risks since they can offset systematic risk exposures by matching the industry average effort. Additionally, in our framework sub-optimal risk taking arises due to contractual externalities as opposed to nonlinearities in payoff schedules.

The model also produces a rich set of new empirically testable hypotheses. We develop these hypotheses based on equilibrium conditions/properties and comparative-static analysis. For example, one of the key equilibrium properties is that principals are exposed to the bulk of (and sometimes the entire) industry risk but share idiosyncratic risks with agents. This indicates that when idiosyncratic risks are very high, contracting is mainly for risk sharing but the same is not true for common risks. Empirically, this predicts that for high level of idiosyncratic risk, contracts will be driven only by relative risk preferences of the principal and the agent. As the risk shifts from idiosyncratic to common, contracts will be driven by other variables such as the industry average productivity and its volatility.

Further, we decompose equilibrium performance pay into two components, relative and absolute performance pay. This decomposition also yields some unique predictions. As mentioned before, relative performance pay shields agents from industry risks, encouraging them to take on too much industry risks which are shouldered by principals completely. Yet, it still leaves agents exposed to idiosyncratic risks. Principals include absolute performance pay in the contract to achieve two different goals: controlling agents' excessive industry risk-taking requires positive sensitivity and reducing agents' exposure to idiosyncratic risks requires negative sensitivity to their absolute performance. There are several empirical pre-

[5]See Diamond (1998); Biais and Casamatta (1999); Palomino and Prat (2003); and Makarov and Plantin (2010).

dictions based on this mechanism. For example, all else equal, higher volatility in industry productivity leads contracts to contain less (more) relative (absolute) performance pay to control for endogenous industry risk taking, especially when principals are more risk-averse. The model also predicts higher absolute and relative performance pay when the expected industry productivity is large. Empirically this prediction implies that agents are more likely to be rewarded by good industry performance. When principals are less risk averse (e.g., risk-neutral), idiosyncratic risks are high and/or expected industry productivity and industry risk is low, principals would like their agents to have a negative exposure to the industry risk, and use less relative performance. We also show the robustness of the main insights of our model by varying the informativeness of the signal about the industry productivity. This robustness analysis also yields several testable hypotheses. For example, the model predicts that during the industry boom, higher volatility in the noise component of the industry signal leads contracts contain more (less) relative (absolute) performance pay, contrasting with the results on the industry productivity volatility.

The condition for the wedge between private and social optimum also gives some empirical predictions on when to expect boom or bust, offering policy guidance on managerial pay. For example, it predicts that social inefficiency is more likely in an industry where the industry-wide productivity is expected to be high and volatile and hard to contract upon. When the industry productivity signal is more informative about the underlying productivity factor, the impact of the externality is reduced and the wedge between the equilibrium and the second-best effort and risk-taking is narrowed. In fact, when the noise in the industry productivity signal approaches zero, the second best is achieved. This implies that excessive or insufficient investment is less of a concern for mature industries where there is less uncertainty about the industry productivity. However, the opposite is true for emerging new industries, where there is a need for close supervision for excessive risk-taking.

Relation to the literature. Our paper is related to the growing literature that studies the welfare effects of pecuniary externalities. This literature is based on the seminal papers of Stiglitz (1982), Greenwald and Stiglitz (1986), Geanakoplos and Polemarchakis (1985) and Arnott, Greenwald and Stiglitz (1994), which establish general conditions of welfare changing pecuniary externalities.[6] In that literature, agents do not internalize their impact

[6] There is an explosion of this literature due to the growing interests in studying social inefficiency of booms-busts. This includes but not limited to the following: Krishnamurthy (2003); Caballero and Krishnamurthy (2001; 2003); Gromb and Vayanos (2002); Korinek (2010); Bianchi (2010); Bianchi and Mendoza, (2011); Stein (2012); Gersbach and Rochet (2012); He and Kondor (2013); Farhi and Werning (2013). Davila (2011) and Stavrakeva (2013) have nice summaries of this literature.

on equilibrium prices when choosing their actions. In our paper, principals do not internalize their impact on the industry average effort, also an equilibrium outcome, when choosing their own contract terms. In a frictionless economy, variation in prices generates transfers across agents without welfare impact. Pecuniary externality creates welfare effect if there are market frictions such as borrowing constraints where an agent fails to internalize the impact of his action on other agents' borrowing constraint via prices. Similarly, in a standard contracting setup without information frictions on the correlated industry productivity risk, variation in the industry average effort should not have welfare impact since its movement generates a transfer between principals and agents. However, when there are information frictions on the correlated industry productivity, contractual externality has an impact on social welfare. This is because principals do not internalize the impact of their contracts on the risk exposure and the cost of incentive provisions of other principals in the same industry. As a result, our model generates a different type of externality that works through correlated information frictions, and provides conditions under which this externality drives a wedge between social and private equilibrium outcomes. In fact, we show that when the information friction is removed, that is, when the industry shock is revealed, contractual externalities do not have any welfare impact. By pinpointing contractual externality and its welfare effect under information friction, our model has different empirical and policy implications from the existing pecuniary externality literature about booms-busts cycles.

Since the results in our paper hinge on the fact that contracts put some weight on the industry average, our paper is closely related to the literature on relative performance (starting with Holmström 1979; 1982). We contribute to this literature theoretically in two aspects. First we decompose the equilibrium contract into two components: absolute and relative performance and highlight the role for each component. Second, we study the contractual externality among multiple principal-agent pairs and endogenize the relative benchmark that agents' performance is compared to. Our theoretical extension has many unique predictions. For example, in our model, agents are rewarded for good industry performance but not punished for bad ones. This is because, during the good time when expected industry productivity is high and volatile, a principal would like to expose her agent to some industry risk by increasing pay sensitivity to his absolute performance, not to reward her agent for good industry performance but to make the risk-averse agent internalize the risk he is taking. However, during the downturn when the expected industry productivity is low and the idiosyncratic risk is large relative to the industry risk, a principal would like to reduce her agent's idiosyncratic risk exposure by lowering pay sensitivity to his absolute performance. Empirically, these predictions match data. For example, it has been found that

CEOs are rewarded for good industry shocks (Bertrand and Mullainathan (2001)) but less so for bad ones (in Garvey and Milbourn (2006)).[7] Furthermore, early empirical evidence on relative performance evaluation using limited executive compensation data has found little evidence of its usage. For example, Gibbons and Murphy (1990); Prendergast (1999); Aggarwal and Samwick (1999) offer little empirical support for the theoretical findings in the relative performance literature. Recently, De Angelis and Grinstein (2010) and Gong et al. (2010) use more detailed disclosure data on executive compensations and find that about 25% to 34% of their respective sample firms use some forms of relative performance evaluation, indicating that the relative performance is used but its usage is limited. Our paper offers a potential resolution for the empirical and theoretical divide in this literature. In our model, principals would like the agents to be exposed to the industry risk to internalize the risks they are taking, rather than eliminate the industry risk completely from agents' contract as predicted in the traditional theoretical literature on relative performance. That is, the observed pay sensitivity to the industry index might even be positive. In fact, the comparative statics based on our theory produces a host of new testable cross-sectional predictions on the use of relative and absolute performance variables.

Many contracting situations involve rivalrous agency where principals hire agents who compete on behalf of principals. The literature on rivalrous agency (for example, Myerson (1982), Vickers (1985), Freshtman and Judd (1986; 1987), Sklivas (1987), Katz (1991), among others)) has examined the strategic aspect of incentive provision among principals to explore implications for oligopolistic conducts. Our paper differs in two aspects. First, the interactions among agents in our model arises endogenously via contracts that are based on correlated information, even though principals do not engage in direct competition. Second, our focus is different. We explore implications of contractual externalities for aggregate inefficiencies.[8]

The structure of the paper is as follows. In section 2, we present the model. In section

[7] This line of empirical literature has suggested various alternative explanations to this puzzling phenomenon such as CEO setting their pay. Our model not only provides an alternative (and rational) explanation but also generates new empirical predictions. For example, it predicts this phenomenon is more pronounced in industries where industry shocks are large and with more risk-averse principals.

[8] Our model differs from the study of 'common agency' (Pauly (1974); Bernheim and Whinston (1986)) where multiple principals share the same agency since our principals do not share any agents. Furthermore, this literature studies contracting when externality is given. In our model, externality arises endogenously through contracting. Our paper is also different from the contract theory literature on (rank order) tournaments (Akerlof (1976); Lazear and Rosen (1981); Green and Stokey (1983); Nalebuff and Stiglitz (1983); and Bhattacharya and Mookherjee (1986)) where they study one principal and many agents.

3, we lay out agents' and principals' optimization problems. In section 4, we study the case with only relative performance pay. We solve for the optimal linear contract under the equilibrium and the second best, compare the two and present some comparative static results. In section 5, we study the case with both relative and absolute performance pay, and present some numerical results. Section 6 concludes.

2 Model

In this section, we describe our setup, information environment, and equilibrium definition.

2.1 The Setup

There is a continuum of principals in an industry. Each principal owns a firm which in turn owns a project. There is also a continuum of agents who are able to obtain a fixed reservation utility in a competitive labor market. The principal hires an agent to work on the project and offers the agent a contract. Each principal, agent and project triplet is indexed by $i \in [0,1]$. The principal's objective is to maximize her expected utility which is based on the expected final value of the project. Principals are potentially risk averse, and their utility is given by $u_s(w) = -\exp(-r_s w)$ where $r_s \geq 0$.[9]

There are three dates $t = 0, 1, 2$. At $t = 0$, principal i offers agent $i \in [0,1]$ a contract. We assume that contracts are offered simultaneously. Agent i observes his contract and decides whether to accept or reject it. If he accepts the contract, he chooses hidden effort denoted by e_i on project i. Agent i's effort is costly and the cost is specified as $C(e_i) = e_i^2/2$. We assume that all agents have identical CARA preferences so that $u(w_i, e_i) = -\exp(-r(w - C(e_i)))$ where $r \geq 0$. At $t = 1$, two payoff-relevant public signals about project i are revealed. One is about agent i's performance and the other is about the average performance of all projects in the industry. We assume that these signals are contractible and determine agent i's compensation. All agents are paid at time 1. At $t = 2$, the final values of all projects are realized and principals receive their payoffs. For simplicity we assume no discounting.

[9]Note that we allow for risk-averse principals. In presence of contractual externalities risk-neutrality of principals is not an innocuous assumption. Later in the paper, we discuss differences in the results when $r_s = 0$ and $r_s \neq 0$. In reality, there are a number of reasons why principals might be risk-averse or act as if they are risk-averse. Banal-Estañol and Ottavinani (2006) have discussed these in detail, which include concentrated ownership, limited hedging, managerial control, limited debt capacity and liquidity constraints, and stochastic productions.

2.2 Production Technology

We assume that project i generates output V_i, which is a random function of agent i's unobservable effort and two stochastic shocks,

$$V_i = V(e_i, \tilde{h}, \tilde{\epsilon}_i). \tag{1}$$

The randomness arises from a common random variable $\tilde{h}$, and a project-specific random variable $\tilde{\epsilon}_i$. We interpret $\tilde{h}$ as a common productivity shock to all projects and $\tilde{\epsilon}_i$ as an output shock specific to the individual project. In the rest of the paper, we call $\tilde{h}$ the industry shock or the systematic shock as it cannot be diversified away. The important assumption is that $\partial^2 V_i/(\partial h \partial e_i) \neq 0$, ie, the state of nature that is common across agents, affects the productivity of effort. This specification is meant to capture the uncertainty about industry productivity after a technological innovation.

Our results are based on a linear specification where $V_i = \tilde{h}e_i + \tilde{\epsilon}_i$. In our model, the random variable $\tilde{h}$ is normally distributed with mean $\bar{h} > 0$ and variance σ_h^2 (ie, precision $\tau_h = 1/\sigma_h^2$). The random variable $\tilde{\epsilon}_i$ is normally distributed with mean zero and variance σ_ϵ^2 (ie, precision $\tau_\epsilon = 1/\sigma_\epsilon^2$). To show that the result on contractual externality is crucially dependent on the productivity shock being common across projects, we also analyze an alternative specification where the productivity function is $V_i = \tilde{k}_i e_i + \tilde{\epsilon}_i$. Here $\tilde{k}_i$ is a project-specific random term.

Note that in our specifications, the productivity shock enters multiplicatively with effort. When $\sigma_h = 0$, the specification for output in our model is standard. In the more general case where $\sigma_h > 0$, higher average effort generates a higher return, but since the productivity of effort is random it also leads to higher volatility. Here, we have in mind a broad interpretation of effort as choosing the scale of the project, eg, by devoting more resources (time, personnel, etc.) to it.

2.3 Information Structure

In our model principals receive contractible signals about the output of their projects, as well as the average output of the industry. We assume that the industry average reveals the industry productivity shock $\tilde{h}$ with noise. The idea is that after major technological innovation there is uncertainty about industry productivity and it is difficult to assess the realization of this uncertainty through public signals such as industry stock price indices, which themselves are very noisy.

Specifically, we assume that there are two noisy signals on outcomes of the projects in the industry which principals can contract on. The first is a noisy signal of project i's outcome, ie, agent i's performance, given by

$$s_i = \tilde{h}e_i + \tilde{\epsilon}_i + \tilde{\zeta}, \tag{2}$$

where the signal noise term $\tilde{\zeta}$ is normally distributed with mean zero and variance σ_ζ^2 (ie, precision $\tau_\zeta = 1/\sigma_\zeta^2$). It is common to all projects, ie, an industry-wide noise, reflecting that the difficulty is in assessing the industry-wide (rather than project-specific) technology shock.

The second is a noisy signal of the industry average project outcome, ie, the average performance of all agents, given by

$$t = \tilde{h}\overline{e} + \tilde{\zeta}, \tag{3}$$

where $\overline{e} = \int_0^1 e_i di$ is the average effort of all agents. Note that since the outputs of all projects are observed with correlated noises, the industry average output is also observed with noise $\tilde{\zeta}$. This industry signal reveals the industry productivity $\tilde{h}$ with noise.

Technically, this environment is equivalent to the one in which the principal observes a signal about the agent's relative performance relative to his peers, ie,

$$p_i = s_i - t = \tilde{h}\left(e_i - \overline{e}\right) + \tilde{\epsilon}_i \tag{4}$$

and a noisy signal of his absolute performance which is s_i. When the common noise $\tilde{\zeta}$ is extremely volatile, especially relative to the project-specific noise $\tilde{\epsilon}_i$, we show later that the contract puts more weight on relative information p_i which is more precise, rather than the absolute performance signal. Intuitively, when there is a great uncertainty about the industry productivity, it is relatively easy to assess an agent's performance relative to his peers. That is, the information on the ranking of agents is more precise than the information on an agent's absolute performance level. Empirically, we observe that stock analysts are better at ranking stocks than pricing stocks (Da and Schaumburg (2011)). The finance literature is more successful in explaining cross-sectional equity returns while the equity premium remains a puzzle. Moreover, this information structure parsimoniously captures the tournament-like incentives that agents face in the real world. For example, the ranking of businesses, university programs, fund managers, doctors in different specialities, and even economists of different vintages, is prevalent when there is also (possibly quite noisy) information on their individual performance.

2.4 Equilibrium Definition

We restrict attention to linear compensation contracts which is common in the theoretical literature on principal-agent models. We assume that agent i's compensation contract has three components. The first component is a fixed wage W_i. The other two components, $l_i p_i$ and $m_i s_i$, condition the agent's payment on the realization of the two signals. Therefore, agent i's total compensation I_i is given by

$$I_i\left(l_i, m_i, W_i\right) = l_i p_i + m_i s_i + W_i, \tag{5}$$

where l_i measures the power of relative performance-based pay and m_i measures the power of absolute performance-based pay. Agent i's utility is given by $u\left(I_i - C\left(e_i\right)\right)$ where I_i is his income and $C\left(e_i\right)$ is the cost of exerting effort e_i.[10]

Now we are ready to specify agent i's optimization problem. We assume that agents' reservation utility is $u\left(\bar{I}\right)$. Agent i accepts contract $\left(l_i, m_i, W_i\right)$ if his expected utility from accepting the contract exceeds his reservation utility

$$E\left[u\left(I_i\left(l_i, m_i, W_i\right) - C\left(e_i\left(l_i, m_i, W_i\right)\right)\right)\right] = E\left[u\left(l_i p_i + m_i s_i + W_i - C\left(e_i\left(l_i, m_i, W_i\right)\right)\right)\right] \geq u\left(\bar{I}\right),$$

where $e_i\left(l_i, m_i, W_i\right)$ is the optimal effort choice conditional on accepting the contract. That is,

$$e_i\left(l_i, m_i, W_i\right) = \arg\max_{e_i \geq 0} E\left[u\left(l_i p_i + m_i s_i + W_i - C\left(e_i\right)\right)\right]. \tag{6}$$

We define an equilibrium of the model as follows.

Definition 1: An equilibrium consists of contracts $\left(l_i^*, m_i^*, W_i^*\right)$, effort choices $e_i^* = e_i\left(l_i^*, m_i^*, W_i^*\right)$ for each $i \in [0,1]$ and average effort $\bar{e} = \int_0^1 e_i^* di$ such that given $\bar{e}$ the

[10] We can rewrite (5) as

$$I_i\left(a_i, b_i, W_i\right) = a_i\left(s_i - t\right) + b_i t_i + W_i,$$

where

$$a_i = l_i + m_i, \text{ and } m_i = b_i.$$

Written in this way, a_i denote the total performance combining both relative (l_i) and absolute (m_i) performance pay while b_i denotes the exposure to the industry risk. This specification focuses on the tradeoff between performance pay and industry risk exposure in the contract. By comparison, the specification in the text (5) focuses on the tradeoff between relative and absolute performance pays in the contract. These are equivalent specifications. We choose to illustrate the model intuition using the absolute and relative performance pay specification aiming to generate some empirical testable implications in a later section of the paper.

contract solves principal's problem who chooses (l_i, m_i, W_i) to maximize $E\left[u_s\left(V_i - I_i\right)\right]$ subject to $E\left[u\left(I_i - C\left(e_i\right)\right)\right] \geq u\left(\bar{I}\right)$, where $e_i = e_i\left(l_i, m_i, W_i\right)$ (given in (6)) and $V_i = \tilde{h}e_i\left(l_i, m_i, W_i\right) + \tilde{\epsilon}_i$.

We begin our analysis in section 3 by first solving the agents' and principals' problems in the contractual environment discussed above. In section 4, we obtain analytical results when the absolute performance signal is not informative ie, $\tau_\zeta = 0$. We discuss how the equilibrium effort level compares with the second-best optimum and present results on comparative statics. In section 5, we present some numerical results when the absolute performance signal is informative and show the intuition obtained in section 4 is robust.

3 Agents' and Principals' Problems

In this section we first solve agents' equilibrium effort choices for a given contract. We then characterize principals' choice of optimal contract.

3.1 Agents' Effort Choice

Given contract (l_i, m_i, W_i) agent i's compensation is:

$$I_i = l_i p_i + m_i s_i + W_i = l_i\left(\tilde{h}\left(e_i - \overline{e}\right) + \tilde{\epsilon}_i\right) + m_i\left(\tilde{h}e_i + \tilde{\epsilon}_i\tilde{\zeta}\right) + W_i, \tag{7}$$

and agent i chooses e_i to maximize:

$$E\left[u\left(l_i\left(\tilde{h}\left(e_i - \overline{e}\right) + \tilde{\epsilon}_i\right) + m_i\left(\tilde{h}e_i + \tilde{\epsilon}_i + \tilde{\zeta}\right) + W_i - C\left(e_i\right)\right)\right]. \tag{8}$$

Computing the expectation in the above expression, agent i's problem in (8) can be restated as choosing e_i to maximize:

$$\left(l_i + m_i\right)\bar{h}e_i - l_i\bar{h}\overline{e} + W_i - C\left(e_i\right) - \frac{1}{2}r\left(\left(l_i\left(e_i - \bar{e}\right) + m_i e_i\right)^2\frac{1}{\tau_h} + \left(l_i + m_i\right)^2\frac{1}{\tau_\epsilon} + m_i^2\frac{1}{\tau_\zeta}\right). \tag{9}$$

From (9) we see how a given incentive package shapes agent i's exposure to various sources of risks. His risk exposure to the common productivity shock $(\tilde{h})$ depends on (i) the power of the relative performance-based pay l_i times the difference between his effort and the average effort $(e_i - \bar{e})$, and (ii) the power of absolute performance pay m_i times his effort e_i. His risk exposure to the common noise $\tilde{\zeta}$ depends solely on the power of absolute performance-based pay while his risk exposure to the idiosyncratic noise $\tilde{\epsilon}_i$ depends on the power of *total* performance-based pay. From this we can see that to lower his compensation risk coming from the industry productivity shock agent i has incentive to match the industry average effort $\bar{e}$. Put differently, agent i is able to completely hedge his exposure to the industry risk that comes through his relative performance pay, although he might still be exposed to some industry risk that comes through his absolute performance pay. At the same time, this hedging possibility creates a matching motivation among effort choices of agents. Taking the first-order condition and solving for e_i, we obtain agent i's effort choice as

$$e_i = \frac{(l_i + m_i)\,\bar{h} + \frac{r}{\tau_h} l_i\,(l_i + m_i)\,\bar{e}}{1 + \frac{r}{\tau_h}\,(l_i + m_i)^2}. \tag{10}$$

Note that agent i's effort is increasing in $\bar{e}$, the average effort exerted by all the other agents with a positive relative performance pay sensitivity. Thus, when the average effort increases, agent i's best response is to increase his effort.

Typically, the more risk averse an agent is (ie., the higher r is) and/or the more volatile the industry shock becomes (ie., the lower τ_h is), the lower effort level he will choose. This is because by lowering his effort, the agent reduces his exposure to the industry risk. This effect is captured by the term r/τ_h in the denominator of (10). In our setup, the term r/τ_h is also in the numerator capturing the fact that when r is higher or τ_h is lower, an agent has a stronger incentive to match the average effort to hedge the industry risk. Through this second effect, for a given contract (l_i, m_i), the agent's effort may increase when his risk aversion is higher or the industry productivity shock becomes more volatile.

Crucially the total performance-based pay, $(l_i + m_i)$, has two opposing effects on agent i's effort choice. Increasing it makes agent i increase his effort because his pay becomes more sensitive to average productivity, $\bar{h}$, and the magnitude of his performance relative to the industry average $\bar{e}$. This is captured by the numerator in (10). At the same time, increasing $(l_i + m_i)$ causes agent i to bear more industry risk by making the agent deviate more from the industry average. This increase in risk exposure causes agent i induces him to lower his effort. This effect is reflected by the denominator in (10). Moreover, both effects become stronger as the industry risk $1/\tau_h$ increases. As we show later, these two effects underly the externalities that principals have to face when writing the compensation contracts. Note that in the limit, as the industry risk approaches zero, our model delivers the standard result where agent i's effort is determined by his performance pay and the productivity of his effort, ie, $(l_i + m_i)\bar{h}$.

3.2 Principals' Choice of Optimal Contract

Now we turn to the principals' problem. Principal i chooses the contract terms (l_i, m_i, W_i) to maximize her expected utility:

$$E\left[u_s\left(V_i - I_i\right)\right] \tag{11}$$

subject to agent i's participation constraint: $E\left[u\left(I_i - C\left(e_i\right)\right)\right] \geq u\left(\bar{I}\right)$ where e_i is given by (10).

We proceed to solve the equilibrium contract terms (l_i, m_i, W_i). Using (7) we obtain principal i's final payoff as

$$V_i - I_i = \tilde{h}e_i + \tilde{\epsilon}_i - l_i\left(\tilde{h}\left(e_i - \bar{e}\right) + \tilde{\epsilon}_i\right) - m_i\left(\tilde{h}e_i + \tilde{\epsilon}_i + \tilde{\zeta}\right) - W_i.$$

Computing expectations, principal i's problem in (11) can be stated as

$$\begin{aligned}&\max_{l_i, m_i, W_i} \left(1 - l_i - m_i\right)\bar{h}e_i + l_i\bar{h}\bar{e} - W_i \\ &-\frac{1}{2}r_s\left(\left(e_i - l_i\left(e_i - \bar{e}\right) - m_i e_i\right)^2\frac{1}{\tau_h} + \left(1 - l_i - m_i\right)^2\frac{1}{\tau_\epsilon} + m_i^2\frac{1}{\tau_\zeta}\right),\end{aligned} \tag{12}$$

where e_i is given by (10). Using (9) and agent i's individual rationality constraint we obtain

$$-\left(l_i + m_i\right)\bar{h}e_i + l_i\bar{h}\bar{e} - W_i = -C\left(e_i\right) - \frac{1}{2}r\left(\left(\left(l_i + m_i\right)e_i - l_i\bar{e}\right)^2\frac{1}{\tau_h} + \left(l_i + m_i\right)^2\frac{1}{\tau_\epsilon} + m_i^2\frac{1}{\tau_\zeta}\right) - \bar{I}.$$

We substitute the above equation into (12) and simplify it further as

$$\begin{aligned}&\max_{l_i, m_i} \bar{h}e_i - C\left(e_i\right) - \frac{1}{2}r_s\left(\left(e_i - l_i\left(e_i - \bar{e}\right) - m_i e_i\right)^2\frac{1}{\tau_h} + \left(1 - l_i - m_i\right)^2\frac{1}{\tau_\epsilon} + m_i^2\frac{1}{\tau_\zeta}\right) \\ &-\frac{1}{2}r\left(\left(l_i\left(e_i - \bar{e}\right) + m_i e_i\right)^2\frac{1}{\tau_h} + \left(l_i + m_i\right)^2\frac{1}{\tau_\epsilon} + m_i^2\frac{1}{\tau_\zeta}\right) - \bar{I}.\end{aligned} \tag{13}$$

The above expression has an intuitive interpretation as it is principal i's and agent i's combined surplus. The first term is the expected output of the project, the second term is the cost of agent i's effort, and the next two terms are the disutilities from the risk exposures of the agent and the principal respectively. The following proposition characterizes the equilibrium contract.

PROPOSITION 1: Given $\bar{h}$, r, r_s, there exists $\bar{\tau}_h$ such that for all $\tau_h > \bar{\tau}_h$ there exists a unique equilibrium contract which is symmetric.

In this paper we will restrict attention to situations where the equilibrium is unique. Proposition 1 guarantees the existence of a unique equilibrium as long as the industry risk is not too large.[11] Note that once the values of $\bar{h}$, r, r_s are fixed, Proposition 1 guarantees that there is a unique equilibrium for large enough τ_h regardless of the values of τ_ϵ and τ_ζ which allows us to fix τ_h and perform comparative statics with respect to τ_ϵ and τ_ζ (without losing existence and uniqueness of the equilibrium).[12]

4 Equilibrium with Only Relative Performance

Often major technological innovations introduce extreme difficulties for economic agents to assess the industry productivity. We capture this feature by allowing the absolute signals associated with the industry productivity shock to be extremely noisy, that is, by letting precision τ_ζ be zero. As a result, $m_i^* = 0$, that is, contracts do not include an absolute performance-based pay component.

In this case, principals do not have any information about $\tilde{h}$ directly, and both signals s_i and t are uninformative by themselves. However, their difference p_i is informative because it is unaffected by the common noise $\tilde{\zeta}$. Consequently, principals can only assess how much better or worse their agents are performing relative to the average and have to base agents' compensation on this information alone. In section 5, we relax this assumption and study the equilibrium results when τ_ζ is larger than zero.

To solve her problem, principal i takes $\bar{e}$ as given and chooses the optimal linear contract which we denote by l_i^*. The following proposition characterizes the equilibrium contract and effort levels.

PROPOSITION 2: When $\tau_\zeta = 0$, under the conditions in Proposition 1, a unique symmetric equilibrium contract exists and satisfies

$$\frac{\bar{h}^2}{\frac{r}{\tau_h}(l^*)^2 + 1}(1 - l^*)\left(1 - \frac{r_s}{\tau_h} l^*\right) - \frac{1}{\tau_\epsilon}(rl^* - r_s(1 - l^*)) = 0. \tag{14}$$

Moreover, $l^* \in (0,1)$ and the equilibrium effort level is $e^* = \bar{e} = l^* \bar{h}$.

4.1 Equilibrium Properties

There are several notable features of this equilibrium contract. To illustrate, we dissect the equilibrium condition (14) into terms that are associated with the usual tradeoff between

[11] When the industry risk is large, it is possible to construct examples of multiple equilibria. The multiplicity of equilibrium contracts is an interesting possibility that is worth studying further in a future work.

[12] This is because terms that are higher order with respect to τ_h do not depend on these variables.

incentives and risk-sharing. To do so we define the *incentive provision* as the level of compensation when the sole purpose of the contract is to incentivise the agents to exert effort, and the *risk-sharing provision* as the level of compensation when the purpose of the contract is to allow risk sharing between principals and agents. The following corollary characterizes the optimal contract in two limiting cases.

COROLLARY 1: When τ_ϵ goes to infinity, the optimal linear contract reflects only the incentive provision and is given by $l_i^* = \min\{1, \tau_h/r_s\}$. When τ_ϵ goes to zero, the optimal linear contract reflects only the risk-sharing provision and is given by $l_i^* = r_s/(r_s + r)$.

Corollary 1 allows us to identify the terms in the equilibrium condition (14) that correspond to incentive and risk sharing provisions:

$$\bar{h}^2 \frac{1}{\underbrace{\frac{r}{\tau_h}\left(l_i^*\right)^2 + 1}_{\substack{\text{Cost of Unilateral Deviations} \\ \text{in Incentive Provision}}}} \underbrace{\left(1 - l_i^*\right)\left(1 - \frac{r_s}{\tau_h} l_i^*\right)}_{\text{Incentive Provision}} - \frac{1}{\tau_\epsilon} \underbrace{\left(r l_i^* - r_s\left(1 - l_i^*\right)\right)}_{\text{Risk-Sharing Provision}} = 0. \tag{15}$$

A few comments are in order. It is intuitive that the magnitude of risk-sharing provision depends solely on the relative risk-aversions of both principals and agents, $r_s/(r_s + r)$. The magnitude of the incentive provision needs some explanation. In setting up the incentive pay, principals would like to give agents the maximum performance pay ($l^* = 1$) if they are risk-neutral. However, in our model, principals are risk-averse and shoulder all industry risk in equilibrium.[13] Moreover, the amount of industry risk exposures in the projects are endogenous and depend on agents' effort choices. Since a higher-powered incentive leads an agent to take on more industry risk which is then passed onto his principal, a risk-averse principal would be setting incentives according to the level of the industry risk and her appetite for it (τ_h/r_s). This also means that the higher the industry risk, the larger the risk-aversion of the principal, the lower the incentive provision is in equilibrium. That is, $l^* = \tau_h/r_s$. Note that this term is associated with endogenous risk-taking by agent and is unique to our model.

While the incentive and the risk-sharing provisions are captured by the (labeled) two terms in (15), the relative importance of these two provisions in the contract is indicated by the coefficients of these two terms. To measure the importance of incentive provision

[13]To see why this is this case, recall in equilibrium $e_i^* = \bar{e}$. This implies that each agent's industry risk exposure in his compensation contract is zero in equilibrium (from (9)).

relative to risk sharing, we take the ratio of the coefficients of these two terms and obtain: $\bar{h}^2\tau_\epsilon / \left(r/\tau_h \left(l_i^*\right)^2 + 1\right)$. Hereafter we refer to this term simply as relative importance of incentive provision.

It is intuitive that relative importance of incentive provision is increasing in expected productivity $\bar{h}$ (since principals are keen to get the agents to work harder) and decreasing in idiosyncratic risk $1/\tau_\epsilon$ (since risk sharing becomes more important as the amount of idiosyncratic risk increases). Note that the industry risk $1/\tau_h$ does not appear in the same way as the idiosyncratic risk $1/\tau_\epsilon$ in the relative importance. This is because principals bear the industry risk entirely and share only the idiosyncratic risk with agents in equilibrium. However, the industry risk affects the relative importance of incentive provision through the term $(r/\tau_h(l_i^*)^2 + 1)$, the intuition for which is not immediately apparent and needs some explanation. Recall that this term appears in equation (10) when we solve agent i's optimal effort (except that here $m = 0$). It captures agent i's disutility from taking on additional industry risk when incentivised to work (potentially) more than the industry average.[14] Higher the industry risk, higher this disutility. This captures the cost of providing incentives unilaterally since a principal has to compensate her agent for this disutility. As a result when this term increases, the cost of incentive provision increases, and consequently relative importance of incentive provision decreases.

4.2 Comparative Statics

In the spirit of developing testable implications of the model, we derive some comparative statics of this equilibrium. Empirically, it is possible to obtain proxies or estimates for model parameters such as industry (marginal) productivity, industry risks and idiosyncratic risks, as well as risk aversions of the principals and agents. For example, one can use the proportion of institutional investors in the shareholder base of a firm as a proxy for (the inverse of) risk aversion of the firm. The next proposition characterizes the comparative statics of the equilibrium with respect to parameters $\bar{h}$ and τ_ϵ.

Proposition 3: If $\tau_h/r_s > r_s/(r_s + r)$, the equilibrium contract l^* increases in $\bar{h}$ and τ_ϵ. If $\tau_h/r_s < r_s/(r_s + r)$, l^* decreases in $\bar{h}$ and τ_ϵ.

To understand this proposition, recall that in Corollary 1 we find that when either $\bar{h}$

[14]Of course, in equilibrium, agents industry risk exposure in his compensation contract is zero since agents hedge industry risk by choosing $e_i^* = \bar{e}$. Since each principal takes other principals and agents behaviours as given, in her view, deviating from equilibrium choice and providing stronger incentives unilaterally would impose her agent to bear more risk and hence result in this additional cost of incentive provision.

or τ_ϵ is zero, contracting in equilibrium reflects only the risk-sharing concern, requiring the performance-pay sensitivity to be $r_s/(r_s+r)$. When $\bar{h}$ or τ_ϵ approaches infinity, the contracting reflects only the incentive concern, requiring the performance-pay sensitivity to be $\min\{1, \tau_h/r_s\}$ in equilibrium. Thus, the above proposition says that when the incentive concern requires a larger (smaller) performance-pay sensitivity than the risk-sharing concern, the power of the equilibrium contract l^* is monotone increasing (decreasing) in $\bar{h}$ and τ_ϵ. Intuitively, when $\bar{h}$ and τ_ϵ are increasing, that is, the expected productivity of effort is higher and the idiosyncratic risk of the project is lower, principals are more concerned about incentivising agents than risk-sharing, that is, the purpose of the contracting is weighted more towards incentive provision rather than risk sharing. Since the former requires a larger (smaller) pay sensitivity, l^* is increasing (decreasing) in $\bar{h}$ and τ_ϵ. Empirically, principals tend to be less risk-averse than agents. Hence testable implications should be that incentive provision is increasing in expected productivity and decreasing in the idiosyncratic risks unless in rare cases with extremely risk-averse principals. The monotonicity result established in Proposition 3 leads to the following corollary.

Corollary 2: The equilibrium contract l^* always takes a value between $\min\{1, \tau_h/r_s\}$ and $r_s/(r_s+r)$.

The comparative statics on τ_h is also straightforward. Algebraically, we observe that when τ_h is bigger, that is, the systematic risk is lower, the magnitude of the incentive provision is larger since principals are less concerned about endogenous risk-taking. Further, when τ_h is larger, the relative importance of the incentive provision in (15) become larger, since the unilateral incentive provision cost is lower. Thus, when the systematic risk is smaller (higher τ_h), principals are more willing to give larger performance-pay compensations (larger l_i). The following proposition states this result formally.

Proposition 4: The equilibrium contract l^* is strictly increasing in τ_h.

Comparative statics of the equilibrium contract l^* with respect to agents' level of risk aversion parameter, r, is quite complex because of two effects through the risk-sharing provision and the incentive provision. First, since contracts impose risk on agents, as agents get more risk averse (i.e, r becomes larger), they require reducing performance-pay sensitivities to improve risk sharing. This is the direct effect and is shown in (15) where $l^* = r_s/(r+r_s)$. Second, a larger r also increases the unilateral incentive provision cost and as a result, the risk-sharing provision becomes relatively more important (see (15)) which is an indirect effect. As the next proposition shows, when $\tau_h/r_s > r_s/(r_s+r)$, the indirect effect reinforces the direct effect and l^* is decreasing in r.[15]

PROPOSITION 5: If $\tau_h/r_s > r_s/(r_s+r)$ then the equilibrium contract l^* decreases in r.

The risk aversion of principals, r_s, on the other hand, affects the magnitude of both incentive and risk-sharing provisions but in the opposite way as shown in (15). Although comparative statics with respect to r_s are generally ambiguous, it is instructive to look at the limiting case where principals are risk-neutral ($r_s = 0$):

$$\bar{h}^2 \underbrace{\frac{1}{\frac{r}{\tau_h}\left(l_i^*\right)^2+1}}_{\substack{\text{Cost of Unilateral Deviations}\\ \text{in Incentive Provision}}} \underbrace{\left(1-l_i^*\right)}_{\text{Incentive Provision}} - \frac{1}{\tau_\epsilon} \underbrace{r l_i^*}_{\text{Risk-Sharing Provision}} = 0. \tag{16}$$

With risk-neutral principals, the incentive provision calls for giving agents the maximum performance pay ($l^* = 1$) since they are not concerned about their own exposure to the industry risk and hence about the endogenous risk-taking by agents. The risk-sharing provision dictates only fixed wages for agents ($l^* = 0$) from (16). The equilibrium contract reflects the balance of these two corner solutions and the balance hinges on the expected productivity $\bar{h}$, the volatilities of the industry and the idiosyncratic shocks ($1/\tau_h$ and $1/\tau_\epsilon$), and the agents' risk aversion r. When the industry productivity is known, ie, when τ_h is infinite, the optimal contract has the simple textbook solution $l_i^* = \bar{h}^2/\left(\bar{h}^2 + r/\tau_\epsilon\right)$.

4.3 Comparison with the Second Best

In this section, we compare the equilibrium effort and risk-taking with the second-best level. We define the second best as the solution to the planner's problem where the planner maximizes the sum of the utilities of all principals conditional on the incentive and individual rationality constraints for the agents. Formally,

DEFINITION 2: A second-best solution consists of a contract $\left(l^{SB}, W^{SB}\right)$ and effort choice e^{SB} where $e^{SB} = e_i\left(l^{SB}, W^{SB}\right)$ and the contract solves the planner's problem, ie, the planner chooses (l, W) to maximize $\int_0^1 E\left[u_s\left(V_i - I_i\right)\right] di$, subject to $E\left[u\left(I_i - C\left(e_i\right)\right)\right] \geq u\left(\bar{I}\right)$, where $I_i = lp_i + W$ and $V_i = \tilde{h} e_i\left(l, W\right) + \tilde{\epsilon}_i$.

[15] When $\tau_h/r_s < r_s/(r_s+r)$, that is when the incentive provision requires a lower performance sensitivity, it is possible to have a region where the equilibrium incentive is increasing in r. However, this requires extreme and unrealistic values for risk aversion parameters.

This definition calls for several comments. The planner's role is limited to coordinating the contracts written by principals. In particular, the planner must give agents incentives to exert effort. Also, our definition requires the planner to give agents their reservation utility, which implies that including agents' utility in the planner's objective would not change the solution.

From Definition 2 we see that the planner chooses the contract term l to maximize the sum of principals' utilities subject to incentive and participation constraints. Since principals' optimization problems are identical, the planner's problem can be seen equivalently as maximizing the utility of one of the principals taking into account that $e_i^* = \bar{e} = l\bar{h}$. That is, the planner internalizes the impact of the contract term l on the industry average effort level $\bar{e}$ when choosing l. Thus, the planner's problem is

$$\max_{l \geq 0} \left[-\frac{1}{2}\bar{h}^2 l \left(l \left(1 + \left(\frac{r_s}{\tau_h} \right) \right) - 2 \right) - \frac{1}{2} l^2 \frac{r}{\tau_\epsilon} - \frac{1}{2} (1-l)^2 \frac{r_s}{\tau_\epsilon} \right]. \tag{17}$$

The first-order condition of the problem is

$$\bar{h}^2 \underbrace{\left(1 - l^{SB} \left(\frac{r_s}{\tau_h} + 1 \right) \right)}_{\text{Incentive Provision}} - \frac{1}{\tau_\epsilon} \underbrace{\left(r l^{SB} - r_s \left(1 - l^{SB} \right) \right)}_{\text{Risk-Sharing Provision}} = 0, \tag{18}$$

and the solution to the planner's problem is:

$$l^{SB} = \frac{\frac{r_s}{\tau_\epsilon} + \bar{h}^2}{\frac{r}{\tau_\epsilon} + \frac{r_s}{\tau_\epsilon} + \bar{h}^2 \left(\frac{r_s}{\tau_h} + 1 \right)}. \tag{19}$$

Like the optimal equilibrium contract, the second-best solution also reflects the incentive and risk-sharing provisions. The following corollary summarizes the limiting results for the second-best contract.

Corollary 3: When τ_ϵ goes to infinity, the second-best contract reflects only incentive provision and is given by $l^{SB} = 1/(r_s/\tau_h + 1)$. When τ_ϵ goes to zero, the second-best contract reflects only risk-sharing provision and is given by $l^{SB} = r_s/(r_s + r)$.

Although the second-best solution of (18) is similar to the equilibrium solution of (15) in reflecting both incentive and risk-sharing provisions, there are two important differences. First, the second best requires a lower magnitude of incentive provision, that is, $l^{SB} = 1/(r_s/\tau_h + 1)$. Second, there is no cost of unilateral deviations in incentive provision. That is, contractual externality has two opposing effects. Intuitively, the first effect arises because principals do not take into account the industry risk exposure of other principals in the industry. When principals are risk averse, they have to trade off incentivising their agents to work harder versus exposing themselves to more industry risks in their projects. Stronger the incentive they choose, higher the output they would expect, and larger the industry risk they are exposed. Their industry risk exposure is endogenously linked to the strength of the incentives they provide. When setting incentives, a principal optimally chooses her own risk-return tradeoff ignoring her impact on increasing other principals' industry risk exposure. In the second best, a planner sets incentives by taking into account the feedback loop between industry average and individual effort choices and consequences of industry risk exposure for other principals in the industry. This means, the second best requires weaker incentives for agents.

The second effect goes in the opposite direction and arises because each principal perceives a unilateral deviation from the industry average as being too costly. Recall, the cost of unilateral deviations in incentive provision is incurred in equilibrium when a principal, who takes the industry average effort $\bar{e}$ as given, considers increasing incentives and making her agent work harder unilaterally. The principal realizes that by doing so, her agent's effort would be above $\bar{e}$ which imposes costly industry risk on the agent, and she has to compensate the agent for this risk. In the second best, this unilateral deviation cost disappears because planner can coordinate (dictate) incentive provision across all principals in the industry. Therefore, the relative importance of incentive provision is higher in second best.

To summarize, the externality in the model has two opposing effects on the performance-pay sensitivity in the contract. Compared with the second best, *the magnitude* of equilibrium incentive provision is larger because principals do not internalize the impact of their incentive provision on the average effort level and the industry risk exposure of other principals, consequently, provide too much incentive. However, *the relative importance* of equilibrium incentive provision is lower because principals perceive unilateral increases in incentive provision as too costly and want to free-ride. Naturally, there is excessive performance sensitivity in equilibrium relative to the second best when the free-riding effect dominates. The next proposition gives the exact condition that characterizes when the equilibrium contract is more sensitive to performance than the second-best contract.

PROPOSITION 6: The equilibrium contract is more sensitive to performance than the second-best contract (i.e, $l^{SB} < l^*$), and consequently agents put more effort in equilibrium than the second best, $e^{SB} < e^*$, if and only if

$$\bar{h}^2 \frac{r_s}{\tau_h}\left(\frac{r_s}{\tau_h}+1\right)+\frac{r_s}{\tau_h}\frac{r_s}{\tau_\epsilon}\left(1+\frac{r}{\tau_h}\right)+\frac{r}{\tau_\epsilon}\left(\frac{r_s}{\tau_h}-\frac{r}{\tau_h}\right)>0. \tag{20}$$

Thus, if this inequality is reversed, the second-best contract is more performance sensitive than the equilibrium contract.

The above proposition shows a pro-cyclical pattern of incentive provision, effort choice and risk-taking in the economy. To see this, note that condition (20) is satisfied for a sufficiently large $\bar{h}$ if principals are risk-averse. When $\bar{h}$ is large, the incentive provision term gets a larger weight in equilibrium than in the second-best (shown as the coefficient in front of the incentive-concern term in equations (15) and (18)). This means that when $\bar{h}$ is large, ie, during the productivity boom, the contracting between principals and agents is more motivated by the incentive concern. Intuitively, during this time, the expected productivity of effort is very high, principals are more concerned with providing incentives than cutting down incentive costs, and hence they would like to offer their own agents a contract with a high performance sensitivity. By doing so, they do not internalize the impact of their own incentive-provision on increasing the industry average effort, hence over-incentivise their own agent and trigger a rat race in effort-making. Since marginal productivity of effort is random in our model, an immediate consequence of this result is that there is excess risk-taking behavior among agents in equilibrium. The planner, in this case, can improve the total welfare by enforcing lower performance-based pay sensitivities in agents' compensation contracts.

By contrast, when $\bar{h}$ is low, eg, during downturns, condition (20) is not satisfied. In this case, since the expected productivity of effort is low, the incentive provision term gets a lower weight in equilibrium than in the second-best. The cost of providing incentives unilaterally becomes a major consideration for principals. Principals would like to free-ride each other in incentive provision, offering their agents a contract with a low performance-pay sensitivity. By doing so, principals again do not internalize the impact of their own incentive-provision on increasing the industry average effort, and hence under-incentivise the agents relative to the second-best. This again triggers a race but this time causes a race to the bottom. There is insufficient effort- and risk-taking. In this case, the planner can improve the total welfare by enforcing contracts with higher performance based pay-sensitivities.

When principals are risk-neutral, the relative importance of the incentive provision term in equilibrium is different from that in the second best, but the magnitude of the incentive concern is at the same corner solution in both cases: giving the maximum incentive pay to agents by setting $l = 1$. That is, externality can only affect the performance-pay sensitivity in the contract through influencing the relative importance of the incentive provision but not its magnitude. This is shown by comparing (16) with (21)

$$\bar{h}^2 \underbrace{\left(1 - l^{SB}\right)}_{\text{Incentive Provision}} - \frac{1}{\tau_\epsilon} \underbrace{r l^{SB}}_{\text{Risk-Sharing Provision}} = 0. \tag{21}$$

Algebraically, it is clear from the above equation that when principals are risk-neutral, the externality creates a wedge between the equilibrium and the second best contracts only through affecting the relative importance of the incentive provision term but not through its magnitude. Intuitively, when principals are risk-neutral, they are not concerned by agents' endogenous risk-taking incentivised by high performance sensitivity in pay. Therefore, risk-neutrality of the principal is not an innocuous assumption and shuts down an important channel through which externality has an impact. The next corollary shows that in equilibrium risk-neutral principals always under-provide incentives relative to the second-best.

COROLLARY 4: When the principals are risk neutral, the second-best contract is more performance sensitive than the equilibrium contract.

4.4 Industry-wide vs. Idiosyncratic Variations in Productivity

In this section, we show that the excessive (insufficient) effort provision is related to the common/systematic rather than project-specific/idiosyncratic risk. To highlight the source of externality we consider the case where the productivity shock is idiosyncratic rather than common to the industry. Specifically, we let $V_i = \tilde{k}_i e_i + \tilde{\epsilon}_i$ where $\tilde{k}_i$ is a project-specific random term which is independently and normally distributed across agents with mean $\overline{k}$ and variance $1/\tau_k$.

Analogous to the main model, we assume that the two contractible signals are

$$s_i = \tilde{k}_i e_i + \tilde{\epsilon}_i + \tilde{\zeta}.$$

and

$$t = \overline{k}\bar{e} + \tilde{\zeta}$$

where $\bar{e}$ is the average effort exerted by the agents in the industry. Clearly this is equivalent to simply observing $p_i = \tilde{k}_i e_i - \overline{k}\bar{e} + \tilde{\epsilon}_i$.

We can now write agent i's compensation when absolute performance signals are not contractible (eg., $\tau_\zeta = 0$) as

$$I_i = l_i p_i + W_i = l_i \left(\tilde{k}_i e_i - \overline{k}\bar{e} + \tilde{\epsilon}_i\right) + W_i. \tag{22}$$

Using (22), given a contract (l_i, W_i) and average effort $\bar{e}$, agent i chooses e_i to maximize

$$E\left(u\left(I_i - C\left(e_i\right)\right)\right).$$

Plugging in p_i and computing the expectation in the above equation, the agent's problem can be restated as choosing e_i to maximize

$$l_i\bar{k}e_i - l_i\bar{k}\bar{e} + W_i - C\left(e_i\right) - \frac{1}{2}r\left(\left(l_ie_i\right)^2\frac{1}{\tau_k} + l_i^2\frac{1}{\tau_\epsilon}\right).$$

Taking the first-order condition and solving for e_i, we obtain agent i's effort choice as

$$e_i = \frac{l_i\bar{k}}{1 + \frac{r}{\tau_k}l_i^2}. \tag{23}$$

The above equation shows that, as one would expect, a volatile project-specific risk $(1/\tau_k)$ lowers the effort level. More importantly, it shows that when the productivity shock is idiosyncratic, there are no feedback loop between the industry average effort and an individual agent's effort. Therefore, the results we obtained earlier on excessive (insufficient) effort provision can only arise in an environment where the productivity shock has a systematic component across projects in the industry.

5 Relative and Absolute Performance

In the previous section, we derive a closed-form solution and explore the properties of the model when contracts contain only the relative performance pay. This could happen when the signal about agent's absolute performance is measured very noisily, ie, τ_ζ is zero. In this section, we relax this assumption and let τ_ζ to be larger than zero, that is, principals receive an informative signal about absolute performances. We start with a complete characterization for the case where τ_ζ approaches infinity. As mentioned in section 2, in the intermediate cases where $0 < \tau_\zeta < \infty$ a general closed form solution to the agents' and principals' maximization problems (given in equations (10) and (13) respectively) is not possible. Hence we illustrate the solution numerically for intermediate values of τ_ζ. The numerical analysis shows that when the noise $\tilde{\zeta}$ becomes more precise, the impact of the externality weakens and the gap between the equilibrium and the second best narrows.

We begin with the limiting case when the noise $\tilde{\zeta}$ disappears.

PROPOSITION 7: When τ_ζ approaches infinity, the effort choices and contracts coincide in equilibrium and in the planner's solution.

In other words, if the industry productivity shock is perfectly revealed, principals are able to completely counteract the impact of externalities among agents' effort-taking through optimal contracting. To see this algebraically, let τ_ζ go to infinity, set $a_i = l_i + m_i$ and $c_i = l_i \bar{e}$. We can restate principal i's problem as choosing (a_i, c_i) to maximize:

$$\bar{h}e_i - \frac{1}{2}r_s\left((e_i - a_ie_i + c_i)^2\frac{1}{\tau_h} + (1-a_i)^2\frac{1}{\tau_\epsilon}\right) - C(e_i)$$
$$-\frac{1}{2}r\left((a_ie_i - c_i)^2\frac{1}{\tau_h} + a_i^2\frac{1}{\tau_\epsilon}\right) - \bar{I}$$

where agent i's effort is given by

$$e_i = \frac{a_i\bar{h} + \frac{r}{\tau_h}a_ic_i}{1 + \frac{r}{\tau_h}a_i^2}$$

Note that stated this way principals' problems are completely separated and $\bar{e}$ no longer plays a role. This is because principal i can completely eliminate the impact of the industry average effort $\bar{e}$ by adjusting c_i. By redefining the principals' optimization problem this way, we see that it coincides with the planner's problem and Proposition 7 is obvious.

It is apparent that contractual externality has no welfare changing effects when information friction on the industry risk is removed, that is when $\tau_\zeta = \infty$. Here we can observe the parallel between the workings of contractual and pecuniary externalities. In general, pecuniary externality also does not have welfare changing effects except for conditions as established in Stiglitz (1982), Greenwald and Stiglitz (1986), Geanakoplos and Polemarchakis (1985), Arnott, Greenwald and Stiglitz (1994), and more recently Farhi and Werning (2013). Intuitively, in our setup, this is because the principals have two contracting instruments: the relative performance p_i and the absolute performance s_i, and face two types of risk: systematic and idiosyncratic. Using both instruments a principal can choose optimally her agent's exposure to each type of risks. Since she can choose her exposure to the industry risk regardless of the industry average effort, she is able to undo the welfare effect of contractual externality. The planner, therefore, has no role to play in this environment. In the next proposition, we explicitly characterize the equilibrium solution in this case.

Proposition 8: When τ_ζ approaches infinity, the optimal contract (l^*, m^*, W^*) is symmetric and unique. The total performance sensitivity $a^* = l^* + m^*$ is the unique positive root to the following equation:

$$\begin{aligned} H(a) \;=\; & -\bar{h}^2\left(\frac{r}{\tau_h}a + 1\right)\left(\frac{r}{\tau_h} + \frac{r_s}{\tau_h}\right)^2(a-1) \\ & -\left(\frac{r}{\tau_h} + \frac{r_s}{\tau_h} + \left(\frac{r}{\tau_h}\right)^2 a^2\left(\frac{r_s}{\tau_h} + 1\right) + 2\frac{r_s}{\tau_h}\frac{r}{\tau_h}a\right)^2\left(a\frac{r}{\tau_\epsilon} + (a-1)\frac{r_s}{\tau_\epsilon}\right) = 0. \end{aligned} \tag{24}$$

Given a^* the contract term m^* is given by:

$$m^* = a^* - \frac{\frac{r}{\tau_h}\left(1 + \frac{r}{\tau_h}(a^*)^2\right) + \frac{r_s}{\tau_h}(a^* - 1)\left(\frac{r}{\tau_h}a^* + 1\right)}{\left(\frac{r}{\tau_h}a^* + 1\right)\left(\frac{r}{\tau_h} + \frac{r_s}{\tau_h}\right)}. \tag{25}$$

It is interesting to note that when m is strictly positive, agents are rewarded for better industry performance. In contrast, in the single-agent relative performance model, under corresponding assumptions, agents would not be rewarded by what seems to be luck rather than effort.[16] The difference in result is due to the fact that the principal would like to expose her agent to some industry risk, not to reward the agent for good industry performance but to make the risk-averse agent internalize the risk he is taking. We next use the characterisation of m^* in the above proposition to provide an exact condition when the absolute performance signal gets a positive or negative weight in equilibrium.

Corollary 5: The weight on the absolute performance signal m^*, is positive if and only if (20) holds.

Note the condition for a positive absolute performance pay is the same condition in Proposition 6, under which the equilibrium performance sensitivity would exceed the second-best when the contracts only contain the relative performance pay. This gives a different perspective on the results of excessive/insufficient risk taking in the setting where absolute performance pays are not contractible. When (20) is satisfied, a principal would like to offer her agent some absolute performance pay, ie, a positive m, incentivising her agent to work harder since $\bar{h}$ is high but also internalize the industry risk. However, when the principal is constrained from giving her agent any absolute performance pay, she increases the use of relative performance pay instead, ie, a larger l, triggering feedback loops in the industry average and individual agent's effort choice, causing excessive effort and risk level in equilibrium relative to the second best. Similarly, when a principal is constrained from giving a negative exposure to the absolute performance pay to control effort, the opposite

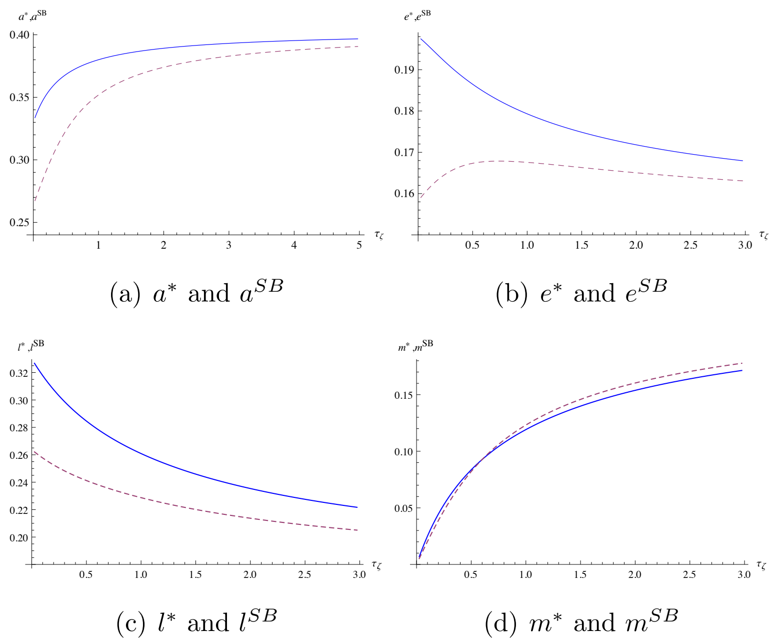
[16] In the standard relative performance model principal observes two signals: a noisy signal of the agent's performance and a second signal that is uninformative about the agent's performance but correlated with the noise term of the first signal. The second signal could be the performance of other agents working on the project but could also be any other information correlated with the signal about the agent's performance. When the two signals are positively correlated, the second signal gets a negative weight. This is because when the second signal is higher, the principal learns that the noise in the first signal is likely to be high. Putting a negative weight on the second signal, allows the principal not to reward the agent for luck.

would happen. She lowers the use of relative performance pay instead, triggering a race to the bottom, resulting in insufficient equilibrium effort and risk-taking relative to the second best.

This analysis shows that in our framework both relative and absolute performance pays are essential in compensation contracts. Similar to Holmström (1979), principals in our model use relative performance to filter the correlated noise, ie., the luck component, when agents' performance signals are correlated. However, in our setup, the fact that relative performance exposes agents to the idiosyncratic firm risk but shields them from the industry productivity risk means that agents do not consider the impact of their effort choice on their firm's total industry risk, potentially exposing their principals to excessive industry risk. Therefore, different from the classical relative performance literature, our model finds that principals use absolute performance to control and share risks with agents which relative performance alone cannot achieve.

Specifically, in our model, absolute performance pay plays two roles: either exposing agents to the industry risk to reduce their incentive to take correlated industry risk or offsetting agents' idiosyncratic risk exposure. Principals, especially the risk-averse ones or those in a high industry risk environment, would like to give agents a positive sensitivity to their absolute performance (ie., $m > 0$) so that agents would internalize their tendency to take on too much industry risk. By contrast, principals, especially those who are close to risk-neutral or those in a low industry risk environment, would like to give agents a negative sensitivity to their absolute performance (ie., $m < 0$) to reduce their idiosyncratic risk exposure. Corollary 5 establishes the condition under which absolute performance pay sensitivity is positive. It shows that when agents' incentive to take correlated industry risks is relatively large, that is when industry productivity shock has a high expected mean and volatility and when principals are risk-averse, the first force prevails; otherwise, the second force dominates.

This finding regarding the purposes of absolute versus relative performance in compensation contracts offers a unique explanation to various empirical puzzles regarding the usage of relative performance contracts. For example, the empirical phenomenon of "paying for luck" might be due to the fact that principals want to control agents' excessive risk-taking tendency. In fact, there might be a large amount of relative performance in the compensation contracts (high l) even when the pay is positively correlated with industry risk (high m). Our model calls for a decomposition of the pay package to undercover this usage of relative versus absolute performance instruments. Furthermore, based on the condition in Corollary 5, our model predicts this phenomenon occurs more often in industries with volatile and



(a) a^* and a^{SB} (b) e^* and e^{SB}

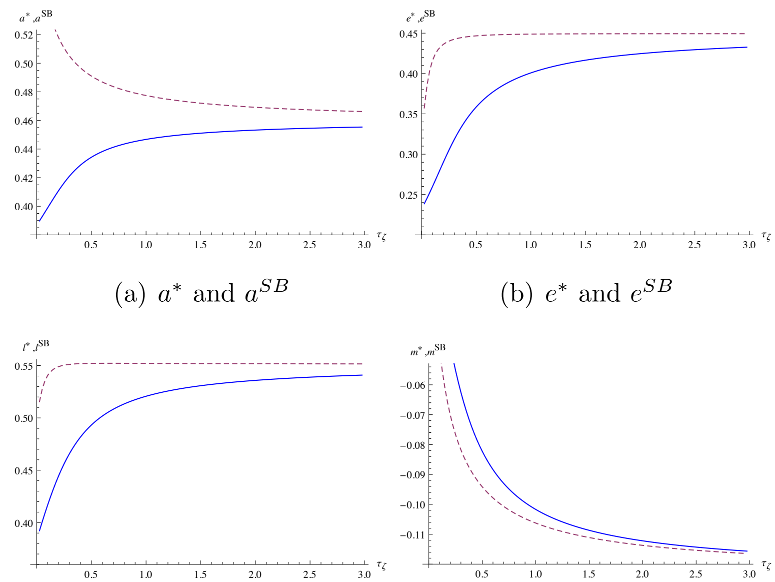
(c) l^* and l^{SB} (d) m^* and m^{SB}

Figure 1: Noisy Industry Signal (τ_ζ) and Excessive Effort: The solid and the long-dashed lines represent how the total performance sensitivity a, relative performance l, absolute performance m, and effort (e) change with respect to the noise of the average industry performance signal (τ_ζ) in equilibrium and in the planner's optimum, respectively. The parameters are fixed at $r = 0.3$, $r_s = 0.16$, $\tau_\epsilon = 1$, $\bar{h} = 0.6$, and $\tau_h = 0.05$.

high productivity, while in industries where expected sector productivity is low, and firm-specific risks are larger, our model finds that the sensitivity to industry risks is much lower, even turns negative, predicting an asymmetry in "paying for luck." These are new testable implications.

Lastly, we study the intermediate case cases where $0 < \tau_\zeta < \infty$. In these cases, the information friction restricts principals' ability to use absolute performance signals in contracts, causing externalities to prevail. We provide two numerical examples for intermediate value of τ_ζ, showing how the equilibrium and the second best incentives change with the noise in this signal. One example is a case when contractual externalities cause excessive effort/risk taking relative to the second best; and the other is the opposite. In both examples, when τ_ζ increases, the impact of the endogenous contractual externality becomes smaller as principals' ability to span the risk space of the agents strengthens.

The graphs in Figure 1 illustrate the intuition in the case where equilibrium effort level exceeds the second best. This is the case when the condition in Corollary 5 is likely to be met. Note that the condition in Corollary 5 also characterizes the condition when principals would like to give a positive exposure to the absolute performance signal ($m^* > 0$) but con-



(a) a^* and a^{SB}

(b) e^* and e^{SB}

(c) l^* and l^{SB}

(d) m^* and m^{SB}

Figure 2: Noisy Industry Signal (τ_ζ) and Insufficient Effort: The solid and the long-dashed lines represent how the total performance sensitivity l, relative performance l, absolute performance m, and effort (e) change with respect to the noise of the average industry performance signal (τ_ζ) in equilibrium and in the planner's optimum, respectively. The parameters are fixed at $r = 0.3$, $r_s = 0.01$, $\tau_\epsilon = 1$, $\bar{h} = 0.6$, and $\tau_h = 0.05$.

strained to do so due to the noise in the industry performance signal. As this noise becomes smaller (ie., τ_ζ gets larger), the information friction on using absolute performance pay is less constraining, principals increase the sensitivity to the absolute performance in equilibrium (m^*) to give agents a positive exposure to the industry risk and better control agents' excessive correlated risk-taking. This is shown in Figure 1(d). Correspondingly, this switch to the usage of the absolute performance instrument in contracts leads to the sensitivity to relative performance (l^*) to drop in equilibrium, as shown in Figure 1(c). However, the total performance sensitivity in the equilibrium contract ($a^* = l^* + m^*$) increases since principals are able to use both contractual instruments on absolute and relative performances more effectively as the noise in the industry output signal becomes smaller. Interestingly, agents reduce effort in equilibrium because they are now exposed to the industry risk through absolute performance pay, as shown in Figure 1(b). Further, as expected, when τ_ζ gets larger, the information friction gets smaller, the impact of the externality weakens, the gap between the equilibrium and the second best narrows.

The graphs in Figure 2 illustrate the intuition in the case where equilibrium effort level falls below the second best. This is the case when the condition in Corollary 5 is likely to be violated, and as a result, principals would like to give agents a negative exposure to the absolute performance ($m^* < 0$) but constrained to do so due to the noise in this signal. This happens when principals are less risk-averse and/or idiosyncratic risks are relatively larger than the industry risks. In these occasions, principals would not mind taking over a large portion of idiosyncratic risks from their agents to lower contracting costs, which implies a negative absolute performance pay sensitivity. The noise in the industry output signal, $\tilde{\zeta}$, constraints principals' ability to do so. When τ_ζ gets larger, this constraint gets less binding, which explains the increased use of absolute performance pay (that is, the contract is more negatively related to absolute performance) in Figure 2(d). This in turns makes it less costly for principals to use the relative performance pay since agents bear less firm-specific risks, resulting in higher usage of relative performance pay as in Figure 2(c). The total performance sensitivity in the equilibrium contract (a^*), just as in the previous case, also increases in τ_ζ. Again when the noise in absolute performance signals becomes smaller, principals are less constrained to use both contractual instruments on industry and firm-specific performance to span the risk space agents are facing, as shown in Figure 2(a). Interestingly, agents increase their effort in equilibrium because they are less exposed to the idiosyncratic risk through the lowered absolute performance sensitivity and more incentivized to take on correlated industry risk through increased relative performance sensitivity, which is shown in Figure 2(b). Further, as expected, when τ_ζ gets larger, the information friction gets smaller, the impact of the externality weakens, the gap between the equilibrium and the second best narrows. These numerical examples show that the insights of the models are robust and endogenous contractual externalities exist when there exists some form of the information friction in the contracting environment.

6 Conclusion

In this paper, we study the welfare changing effect of endogenous contractual externality. Contractual externality in our model arises when there is a great productivity uncertainty, for example, during major technological innovations. In times like this, the information about the common productivity shock is extremely noisy, relative performance information is, however, relatively precise and used in contracting between principal-agent pairs. Consequently, industry average effort affect individual agent's utility. Principals, when setting the optimal contract for their own agents, do not internalize the impact of their incentive provision on the average effort in the industry. This may set off the ratchet effect in effort choices among agents. For example, risk-averse principals are eager to provide more powerful incentives during booms, causing the industry average effort to be high, triggering a rat race among agents to exert excessive effort, which results in excessive systematic risk exposure in the economy, relative to the second-best. During recessions, the opposite might happen: The incentive provision is too weak and the equilibrium level of effort is lower than the second best. We find that contractual externality, similar to pecuniary externality, has welfare changing effect only in the presence of information friction on the industry risk factor. Besides theoretical contributions, our results may offer an explanation to the boom-bust cycle of investment and risk-taking observed in industries that experience new but uncertain productivity shocks. These episodes are abundant in recent years. For example, following the introduction of the Internet, the dot.com industry has been flooded with investment which is subsequently reduced. Similarly, following the innovations in the financial products such as asset-based securities, the financial industry has expanded significantly followed by a sharp contraction. Compensation regulations such as enforcing counter-cyclical performance pay could improve the total welfare.

References

Acharya, V., M. Pagano, and P. Volpin (2011): "Seeking alpha: Excess risk taking and competition for managerial talent," New York University, University of Naples Federico II, London Business School and CEPR Working Paper.

Acharya, V. and T. Yorulmazer (2007): "Too many to fail: An analysis of time-inconsistency in bank closure policies," *Journal of Financial Intermediation*, 16, 1–31.

——— (2008): "Cash in the market pricing and optimal resolution of bank failures," *Review of Financial Studies*, 21, 2705–2742.

Aggarwal, R. K. and A. A. Samwick (1999): "Executive compensation, strategic competition, relative performance evaluation: Theory and evidence," *The Journal of Finance*, 54, 1999–2043.

Akerlof, G. (1976): "The economics of caste and of the rat race and other woeful tales," *Quarterly Journal of Economics*, 90, 599–617.

Arnott, R., B. C. Greenwald, and J. E. Stiglitz (1994): "Information and economic efficiency," *Information Economics and Policy*, 6, 77.

Banal-Estanol, A. and M. Ottaviani (2006): "Mergers with product market risk," *Journal of Economics & Management Strategy*, 15, 577–608.

Bernheim, B. D. and M. Whinston (1986): “Common agency,” *Econometrica*, 54, 923–942.

Bertrand, M. and S. Mullainathan (2001): “Are CEOS rewarded for luck? The ones without principals are,” *The Quarterly Journal of Economics*, 116, 901–932.

Bhattacharya, S. and D. Mookherjee (1986): “Portfolio choice in research and development,” *Rand Journal of Economics*, 17, 594–605.

Bhattacharyya, S. and A. Purnanandam (2011): “Risk-taking by banks: What did we know and when did we know it?” Working paper.

Biais, B. R. and C. Casamatta (1999): “Optimal leverage and aggregate investment,” *Journal of Finance*, 54, 1291–1323.

Bianchi, J. (2010): “Credit Externalities: Macroeconomic Effects and Policy Implications,” *American Economic Review*, 100, 398–402.

Bianchi, J. and E. G. Mendoza (2011): “Overborrowing, Financial Crises and Macro-prudential Policy,” IMF Working paper 11/24.

Caballero, R. J. and A. Krishnamurthy (2001): “?International and domestic collateral constraints in a model of emerging market crises,” *Journal of Monetary Economics*, 48, 513–548.

——— (2003): “?Excessive Dollar Debt: Financial Development and Underinsurance,” *Journal of Finance*, 58, 867–894.

Da, Z. and E. Schaumburg (2011): “Relative valuation and analyst target price forecasts,” *Journal of Financial Markets*, 14, 161–192.

Davila, E. (2011): “?Dissecting Fire Sales Externalities,” Harvard University Working Papers.

De Angelis, D. and Y. Grinstein (2010): “Relative performance evaluation in CEO Compensation: Evidence from the 2006 disclosure rules,” Cornell University Working Paper.

Diamond, P. A. (1998): “Managerial incentives: On the near linearity of optimal compensation,” *Journal of Political Economy*, 34, 931–957.

Farhi, E. and J. Tirole (2011): "Collective moral hazard, maturity mismatch and systemic bailouts," *American Economic Review*, forthcoming.

Farhi, E. and I. Werning (2013): "?On the Inefficiency of Financial Market Equilibria in Macroeconomics Models with Nominal Rigidities," Harvard Working paper.

Fershtman, C. and K. L. Judd (1986): "Strategic incentive manipulation in rivalrous agency," Working paper.

——— (1987): "Incentive equilibrium in oligopoly," *American Economic Review*, 77, 927–940.

Garvey, G. T. and T. T. Milbourn (2006): "Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad," *Journal of Financial Economics*, 82, 197–225.

Geanakoplos, J. and H. M. Polemarchakis (1985): "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete.?" Cowles Foundation for Research in Economics Discussion Papers 764.

Gersbach, H. and J.-C. Rochet (2012): "Aggregate Investment Externalities and Macroprudential Regulation?" CEPR Discussion Papers 8764.

Gibbons, R. and K. J. Murphy (1990): "Relative performance evaluation for chief executive officers," *Industrial and Labor Relations Review*, 43, 30–51.

Gong, G., L. Y. Li, and J. Y. Shin (2010): "Relative performance evaluation and related peer groups in executive compensation contracts," Pennsylvania State University and University of Illinois at Urbana-Champaign Working Paper.

Green, J. R. and N. L. Stokey (1983): "A comparison of tournaments and contracts," *The Journal of Political Economy*, 91, 349–364.

Greenwald, B. C. and J. E. Stiglitz (1986): "Externalities in Economies with Imperfect Information and Incomplete Markets," *The Quarterly Journal of Economics*, 101, 229–264.

Gromb, D. and D. Vayanos (2002): "?Equilibrium and welfare in markets with financially constrained arbitrageurs," *Journal of Financial Economics*, 66, 361–407.

He, Z. and P. Kondor (2013): "Inefficient investment waves?" NBER Working Papers.

Hoberg, G. and G. Phillips (2010): "Real and Financial Industry Booms and Busts," *Journal of Finance*, 65, 45–86.

Holmström, B. (1979): "Moral hazard and observability," *Bell Journal of Economics*, 10, 74–91.

——— (1982): "Moral hazard in teams," *Bell Journal of Economics*, 13, 324–340.

Katz, M. L. (1991): "Game-playing agents: Unobservable contracts as precommitments," *RAND Journal of Economics*, 22, 307–328.

Korinek, A. (2010): "Hot money and serial financial crises," University of Maryland Working Papers.

Krishnamurthy, A. (2003): "?Collateral constraints and the amplication mechanism," *Journal of Economic Theory*, 111, 277?–292.

Lazear, E. P. and S. Rosen (1981): "Rank-order tournaments as optimum labor contracts," *The Journal of Political Economy*, 89, 841–864.

Makarov, I. and G. Plantin (2010): "Rewarding trading skills without inducing gambling," London Business School Working Paper.

Myerson, R. (1982): "Optimal coordination mechanisms in generalized principal-agent problems," *Journal of Mathematical Economics*, 10, 67–81.

Nalebuff, B. J. and J. E. Stiglitz (1983): "Prizes and incentives: Towards a general theory of compensation and competition," *The Bell Journal of Economics*, 14, 21–34.

Palomino, F. and A. Prat (2003): "Risk taking and optimal contracts for money managers," *Rand Journal of Economics*, 34, 113–137.

Pauly, M. (1974): "Overprovision and public provision of insurance," *Quarterly Journal of Economics*, 93, 541–562.

Prendergast, C. (1999): "The provision of incentives in firms," *Journal of Economic Literature*, 37, 7–63.

Sklivas, S. D. (1987): "The strategic choice of managerial incentives," *RAND Journal of Economics*, 18, 452–458.

Stavrakeva, V. (2013): "Welfare Implications of the Structure of the Banking Sector in a Small Open Economy?" Harvard University Working Papers.

Stein, J. (2012): "Monetary policy as financial stability regulation," *The Quarterly Journal of Economics*, 127, 57.

Stiglitz, J. E. (1982): "The Inefficiency of the stock market equilibrium," *Review of Economic Studies*, 49, 241–261.

Vickers, J. (1985): "Delegation and the theory of the firm," *Economic Journal*, 95, 138–147.

Appendix

In the proof we drop the subscript i when there is no room for confusion and we use the following notation:

$$t = \frac{r}{\tau_h},\ v = \frac{r}{\tau_\epsilon},\ u = \frac{r}{\tau_\zeta},\ p = \frac{r_s}{\tau_h},\ q = \frac{r_s}{\tau_\epsilon},\ s = \frac{r_s}{\tau_\zeta}.$$

Proof of Proposition 1

Note that

$$e_i = \frac{(l_i + m_i)\,\bar{h} + tl_i\,(l_i + m_i)\,\bar{e}}{1 + t\,(l_i + m_i)^2} = (l_i + m_i)\,\bar{h} + o\,(1/\tau_h)\,.$$

Let $a_i = l_i + m_i$. Substituting this into (13) we obtain:

$$\max_{a_i, b_i} a_i\left(\bar{h}\right)^2 - \frac{\left(a_i\bar{h}\right)^2}{2} - \frac{1}{2}\left(a_i^2 v + m_i^2 u\right) - \frac{1}{2}\left((1 - a_i)^2 q + m_i^2 s\right) + o\,(1/\tau_h) - \bar{I}.$$

For sufficiently large τ_h this function is concave in (a_i, m_i) and has a unique maximum (a^*, m^*) which is identical for all principals. From this we solve for the unique $l^* = a^* - m^*$.

Proof of Proposition 2

We know the existence part holds for Proposition 1.

We use (10) to plug in for e in the principals' problem (13) (where m is set to zero given that the signal t is uninformative) and take the derivative of the objective function with respect to l to find the first-order condition as a function of $\bar{e}$.

In equilibrium, $\bar{e} = l\bar{h}$. Therefore any equilibrium must solve for the first-order condition and $\bar{e} = l\bar{h}$. To find an equilibrium we plug $\bar{e} = l\bar{h}$ in the first order condition. After simplifying we find the equilibrium condition:

$$-\bar{h}^2\frac{(l-1)\,(lp-1)}{tl^2+1} - (vl - q\,(1-l)) = 0.$$

The next lemma is useful in proving the comparative statics results:

Lemma 1: Let

$$\Psi\,(l) = -\bar{h}^2\frac{(l-1)\,(lp-1)}{tl^2+1}. \tag{26}$$

Suppose there is a unique equilibrium, then the following are true. (i) $\Psi(l) - (vl - q(1-l))$ crosses zero from above at l^* where Ψ is given in (26). (ii) If $1/p > q/(v+q)$ then $l^* < 1/p$, otherwise $l^* > 1/p$.

Proof of Lemma 1

Part (i) follows from $\Psi(0) + q > 0$, $\Psi(1) - vl < 0$ and uniqueness. The proof of (ii) is immediate if $1/p > 1$. So suppose $1/p < 1$. Otherwise, $\Psi(l)$ crosses zero at $1/p < 1$. $\Psi(l) - (vl - q(1-l))$ is above $\Psi(l)$ for $l < q/(v+q)$ and is below $\Psi(l)$ for $l > q/(v+q)$. This and the fact that there is a unique equilibrium l^* prove part (ii).

Proof of Proposition 3

Equilibrium l^* solves

$$\Psi(l^*) - (vl^* - q(1-l^*)) = 0$$

where Ψ is given in (26). We write $\Psi\left(l^*\left(\bar{h}\right), \bar{h}\right)$ to make the dependence of Ψ and l^* on $\bar{h}$ explicit. We use similar notation for other parameters, eg, $\Psi(l^*(\tau_h), \tau_h)$.

Taking the total derivative of the equilibrium condition with respect to $\bar{h}$ we obtain

$$\frac{\partial l^*\left(\bar{h}\right)}{\partial \bar{h}} = \frac{-\frac{\partial \Psi}{\partial \bar{h}}}{\frac{\partial \Psi\left(l^*\left(\bar{h}\right), \bar{h}\right)}{\partial l} - (v+q)}.$$

Denominator is negative by Lemma 1 (i). By Lemma 1 (ii),

$$\frac{\partial \Psi\left(l^*, \bar{h}\right)}{\partial \bar{h}} = -2\bar{h}\frac{(1-l^*)(pl^*-1)}{t(l^*)^2+1} \gtrless 0$$

if $1/p \gtrless q/(v+q)$ which proves part (i) for $\bar{h}$. Proof for the result on τ_ϵ is entirely analogous.

Proof of Proposition 4

Taking the total derivative of the equilibrium condition with respect to τ_h, we obtain

$$\frac{\partial l^*(\tau_h)}{\partial \tau_h} = \frac{-\frac{\partial \Psi}{\partial \tau_h}}{\frac{\partial \Psi(l^*(\tau_h), \tau_h)}{\partial l} - (v+q)}.$$

Denominator is negative by Lemma 1 (i). Moreover,

$$\frac{\partial \Psi(l^*, \tau_h)}{\partial \tau_h} = \bar{h}^2 (1-l^*) \frac{\left(\frac{r_s}{(\tau_h)^2} l^*\right) + \left(\frac{r}{(\tau_h)^2} (l^*)^2\right)}{\left(\frac{r}{\tau_h}(l^*)^2 + 1\right)^2} > 0.$$

Thus l^* is increasing in τ_h.

Proof of Proposition 5

Taking the total derivative of the equilibrium condition with respect to r, we obtain:

$$\frac{\partial l^*(r)}{\partial r} = \frac{-\frac{\partial \Psi}{\partial r} + \frac{1}{\tau_\epsilon} l^*}{\frac{\partial \Psi(l^*(r),r)}{\partial l} - (v+q)}.$$

Denominator is negative by Lemma 1 (i). By Lemma 1 (ii), if $1/p > q/(v+q)$ then

$$\frac{\partial \Psi(l^*, r)}{\partial r} - \frac{1}{\tau_\epsilon} l^* = \bar{h}^2 \frac{(1-l^*)(pl^*-1)}{\left(t(l^*)^2+1\right)^2} \frac{(l^*)^2}{\tau_h} - \frac{1}{\tau_\epsilon} l^* < 0.$$

Thus, $\frac{\partial l^*(r)}{\partial r} < 0$ if $1/p > q/(v+q)$.

Proof of Proposition 6

Proof follows from plugging l^{SB} in the equilibrium condition (14) and checking whether its value is positive (in which case $l^{SB} < l^*$) or negative (in which case $l^{SB} > l^*$).

Proof of Proposition 8:

Let $a = l + m$. The principals' objective can be written as

$$\bar{h}e - \frac{1}{2}e^2 - \frac{1}{2}\left((ae - l\bar{e})^2 t + a^2 v\right) - \frac{1}{2}\left(((1-a)e + l\bar{e})^2 p + (1-a)^2 q\right). \tag{27}$$

and the first order condition yields the optimal level of effort as a function of contract terms and the average effort

$$e = \frac{a\bar{h} + tla\bar{e}}{1+ta^2}.$$

Next substituting for the effort level in the objective function of (27) we obtain

$$\begin{aligned}&\left(\bar{h}\left(\frac{a\bar{h}+tla\bar{e}}{1+ta^2}\right) - \frac{1}{2}\left(\frac{a\bar{h}+tla\bar{e}}{1+ta^2}\right)^2 - \frac{1}{2}\left(\left(a\left(\frac{a\bar{h}+tla\bar{e}}{1+ta^2}\right) - l\bar{e}\right)^2 t + a^2 v\right)\right) \\ &-\frac{1}{2}\left(\left((1-a)\left(\frac{a\bar{h}+tla\bar{e}}{1+ta^2}\right) + l\bar{e}\right)^2 p + (1-a)^2 q\right).\end{aligned} \tag{28}$$

For a given a the above function is negative quadratic in l. Thus for a given a principals' objective function is maximized at $l(a)$ which is given by

$$l(a) = \frac{\bar{h}a\left(t\left(ta^2+1\right) - p(1-a)(ta+1)\right)}{\bar{e}\left(t\left(ta^2+1\right) + p(ta+1)^2\right)}. \tag{29}$$

Substituting (29) for $l(a)$ in (28) we reduce principals' problem to choosing a to maximize

$$\frac{1}{2}\bar{h}^2 a(t+p)\frac{a(t-1)+2}{t+p+t^2a^2+pt^2a^2+2pta}-\frac{1}{2}a^2v-\frac{1}{2}(1-a)^2q.$$

Taking the derivative with respect to a, we obtain

$$\frac{-\bar{h}^2(ta+1)(t+p)^2(a-1)}{(t+p+t^2a^2+pt^2a^2+2pta)^2}-av-(a-1)q.$$

Note that the above function starts as positive and crosses to negative once. Thus the objective function is maximized at a^* that solves

$$H(a)=-\bar{h}^2(ta+1)(t+p)^2(a-1)-\left(t+p+t^2a^2(p+1)+2pta\right)^2(av+(a-1)q)=0.$$

Note that $a^*\in(0,1)$. In equilibrium

$$\bar{e}=\frac{a^*\bar{h}+ta^*l(a^*)\bar{e}}{1+t(a^*)^2}.$$

Plugging for $l(a^*)$, we obtain

$$\bar{e}=\frac{\bar{h}a^*(ta^*+1)(t+p)}{t\left(t(a^*)^2+1\right)+p(ta^*+1)^2}.$$

Using the above to substitute for $\bar{e}$ in (29), we obtain

$$l(a^*)=\frac{t\left(1+t(a^*)^2\right)+p(a^*-1)(ta^*+1)}{(ta^*+1)(t+p)}.$$

Thus

$$\begin{aligned} m^* &= a^*-l(a^*) \\ &= a_i^*-\frac{\frac{r}{\tau_h}\left(1+\frac{r}{\tau_h}(a_i^*)^2\right)+\frac{r_s}{\tau_h}(a_i^*-1)\left(\frac{r}{\tau_h}a_i^*+1\right)}{\left(\frac{r}{\tau_h}a_i^*+1\right)\left(\frac{r}{\tau_h}+\frac{r_s}{\tau_h}\right)}. \end{aligned}$$

Proof of Corollary 5:

We continue to use the notation in the proof of Proposition 8. First, note that

$$m^*=a^*-l(a^*)>0\Leftrightarrow l(a^*)<a^*.$$

Using the expression for x_{a^*} we see that

$$l(a^*)<a^*\Leftrightarrow p(1+ta^*)>t(1-a^*),$$

or if and only if

$$a^* > \left(\frac{t-p}{t\,(1+p)}\right).$$

Note

$$H\left(\frac{t-p}{t\,(1+p)}\right) = \frac{1}{t}\frac{1}{(p+1)^3}\,(t+p)^2\,(t+1)^2\left(pv - tv + \bar{h}^2p^2 + \bar{h}^2p + pq + pqt\right).$$

Thus $l\,(a^*) \leq a^*$ if and only if

$$\left(\bar{h}^2\frac{r_s}{\tau_h}\left(\frac{r_s}{\tau_h}+1\right)+\frac{r_s}{\tau_\epsilon}\frac{r_s}{\tau_h}\left(\frac{r}{\tau_h}+1\right)\right) \geq \frac{r}{\tau_\epsilon}\left(\frac{r}{\tau_h}-\frac{r_s}{\tau_h}\right).$$