Discussion Paper Series

RIEB

Kobe University

DP2021-20

Socially Responsible Investment: Ex-ante Contracting or Ex-post Bargaining?

Meg ADACHI-SATO

September 28, 2021

Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Socially responsible investment: ex-ante contracting or ex-post bargaining?*

Meg Adachi-Sato†

Initial version: May 10, 2020
Revised: August 20, 2021

Abstract

This paper shows how a socially and environmentally aware firm principal can motivate a profit-oriented manager to pursue environmental, social and governance (ESG) outcomes. In the model, the manager produces a verifiable output that is detrimental to ESG, but also engages in an unverifiable output that promotes ESG. I show that an ex-post bargaining contract is preferred to an ex-ante commitment contract if the unverifiable output substantially improves ESG or if there is a large negative externality. The paper also demonstrates how social impact bonds can be more effective than short-term debt to finance social programs.

Keywords: Socially responsible investment, ESG, multitask, hold-up, incomplete contracts, social impact bonds, sustainability-linked bonds.

JEL Codes: D86, G11, G23, M12, M14

*This paper was originally titled "Contract duration and socially responsible investment." I sincerely acknowledge the valuable comments and assistance of two anonymous referees and the associate editor. I am also grateful to Clara-Aimo A-S-Osano, Bruce Chapman, Neal Galpin, Hideshi Itoh, Kamiya Kazuya, Andrew Macrae, Hiroshi Osano, David Stern, and Jason Vladusic for their useful advice about this work. This research is financially supported by Yu-cho Foundation (grant-in-aid for research 2020) and Research grant from the Japan Society for the Promotion of Science (JSPS), research category (C), Grant Number 21K01559.

†Research Institute for Economics and Business Administration, Kobe University, Kobe 657-8501, Japan. E-mail: megasato@diamond.kobe-u.ac.jp. Faculty of Business Administration and Accountancy, Khon Kaen University, Khon Kaen 40002, Thailand.

1 Introduction

Socially responsible investment (SRI) has attracted the interest of investors, employees, and consumers alike. Many investment funds, particularly socially responsible funds, consider asset allocation from an environmental, social, and governance (ESG) perspective. Despite this movement, firm managers often still pursue profit maximization. This article shows how principals can use ex-post bargaining and ex-ante commitments to motivate profit-maximizing managers to pursue socially responsible investment.

To this end, I examine an effort (in the form of firm-specific capital investment) allocation problem in the two-period multitask agency model following Holmstrom and Milgrom (1991). In this model, a manager of a firm must produce an observable and verifiable output x that incurs social costs or negative externality z, and an observable but unverifiable output y that reduces social costs z. Examples of x include normal day-to-day outputs or production. Examples of y include embedding a firm culture that values environmental and social issues, or that has policies and practices to combat bullying. A firm's culture is difficult to evaluate by outsiders (unverifiable), even though one can see or hear from the employees (observable).[1] If y effectively reduces z, or in other words, if the sensitivity of y reducing z denoted by ζ is high, it is worth producing y. If the sensitivity ζ is low, there is no point in engaging in the production.

To achieve these outputs, the manager must, during the first period, make an observable but unverifiable effort/investment in each task, denoted by I_x and I_y, to increase the productivity needed to produce each output. Examples of I_x and I_y include efforts to develop internal employment relationships, build up a network of connection, and attain firm-specific knowledge that helps the firm produce outputs more efficiently. I_y also includes an effort that promotes a culture that values lowering the firm's social costs. The information in I_x and I_y is collected over time through frequent and personal contact between the principal and the manager, but it is difficult to transmit to outsiders because of a lack of a knowledge of the full context.

[1] In reality, there are both verifiable and unverifiable outputs that help reduce social costs. However, as we can see from the debate at the Davos conference, the line between verifiable and unverifiable outputs is rather grey when it comes to reducing social costs. Therefore, this paper focuses only on the unverifiable outputs that can reduce social costs. See for example Hughes, K., Sakano, A., Gore, A., Lacqua, F., Niinami, T. and Wijsen M. 2020, "Breaking Free from Single-Use Plastics", World Economic Forum Annual Meeting 2020, <https://jp.weforum.org/events/world-economic-forum-annual-meeting-2020/sessions/breaking-free-from-single-use >, accessed May 5, 2020.

In reality, managers may pursue their own interest when choosing an investment, rather than seeking to reduce social costs. In this paper, however, I consider a principal who values verifiable outputs and is also concerned about ESG, and thus wishes to reduce negative externalities through the production of unverifiable outputs.

The paper uses a two-period model that consists of a profit-maximizing manager and an ESG-oriented principal. The principal offers the manager either a contract that binds both parties to determine the second-period wage at the beginning of the first period (hereafter the ex-ante commitment contract) or a contract that binds both parties to negotiate the second-period wage at the beginning of the second period (hereafter the ex-post bargaining contract).[2] In both contracts, the first-period wage is fixed and the difference between them is how the second-period wage is determined. The first-period wage does not affect the principal's choice of offering an ex-post bargaining or an ex-ante commitment contract, because it is determined to satisfy the IR constraint of the manager before he undertakes investment under either wage contract. The ex-ante commitment contract allows the principal to keep the entire surplus by making a take-it-or-leave-it offer to the manager, knowing the manager will not be motivated to work on the unverifiable output. The ex-post bargaining contract forces the principal to split the surplus in the second period, knowing this induces the manager to make efforts to produce the unverifiable output.

This model shows that the principal's decision to offer an ex-post bargaining or ex-ante commitment contract depends both on the effectiveness of the unverifiable output to reduce social costs (denoted as parameter ζ), and on the amount of social costs (denoted by z).

If the bargaining power of the manager is assumed to be the same as that of the principal at the beginning of the second period, then the more effective the unverifiable output is at reducing social costs (i.e. the larger the ζ), or the larger the social costs (i.e. the larger the z), the more likely it is that the principal offers an ex-post bargaining contract with a fixed wage.[3] [4] On the

[2]The manager remains employed under both wage contracts on the equilibrium path. Moreover, the ex-ante commitment contract is renegotiation-proof because neither the principal nor the manager chooses any action in the second period. This is different from Inderst and Mueller (2010) and Adachi-Sato (2018), who examine optimal managerial compensation and replacement contracts.

[3]The incumbent manager is no longer competitive at the beginning of the second period if he has invested during the first period, because he has gained firm-specific skills no newly hired manager could obtain in the second period.

[4]I show in section 3 that if the manager is risk neutral, it can be either a fixed wage or an incentive payment. I further show in Appendix D that if the manager is risk averse, the principal offers a fixed-wage contract.

contrary, the smaller the ζ or the smaller the z, the more likely it is that the principal offers an ex-ante commitment contract with incentive pay. These results hold regardless of whether the manager is risk neutral or risk averse.

Given this result, further interesting findings emerge when the bargaining power between the parties changes. That is, the greater the manager's bargaining power, the more likely it is the principal will choose ex-post bargaining contract. To this end, I use a generalized Nash bargaining game in a risk-neutral setting for both z and ζ.

The intuition behind the main result, where the bargaining power stays the same for both the principal and the manger, is as follows. If the principal offers the manager a second-period wage after the manager makes investments for both verifiable and unverifiable outputs (an ex-post bargaining contract), the manager has an incentive to make investments for both the verifiable and unverifiable outputs during the first period. This is because the manager will seek a larger bargaining surplus (Nash product) by doing so. The more the manager invests in the production of both the verifiable and unverifiable outputs, the larger the whole bargaining surplus. However, as the bargaining surplus must be split between both parties, the hold-up problem arises. Thus, neither the investment for the verifiable or unverifiable output is determined optimally by the manager under the ex-post bargaining contract.

If, however, the principal offers a second-period wage before the manager has made an effort towards either output, and if she is bound to this contract (an ex-ante commitment contract), the manager has no incentive to invest in the unverifiable output at all. Instead, the manager will be motivated to invest in the optimal level of verifiable output, because his wage depends only on the verifiable output in the second period. Thus, the principal offers a contract that will induce the manager to invest in the verifiable output, which maximizes the firm's expected total net payoff without social costs.[5]

In short, the choice between incentive contracting (an ex-ante commitment contract) and hold-up (an ex-post bargaining contract) depends on the trade-off between the manager's motivation to maximize the firm's expected total net payoff without considering any social costs, and the manager's motivation to reduce social costs by making socially responsible investment. Consequently, if the unverifiable output effectively reduces social costs, or if the social costs are

[5] One example of a principal caring about an observable and verifiable primary output to detriment of social cost is the former US President's pursuit of an 'America first' policy while opting out of the Paris Agreement.

large, the principal is more likely to offer an ex-post bargaining contract with a fixed wage.

Traditionally, classical hold-up problems were solved by selling the entire project to the agent – in this case, the manager. However, when a manager must produce an output for which they do not care for and which is unverifiable, such as a social benefit, the traditional solution of "selling the entire project to the agent" does not induce a profit-maximizing agent to produce a socially beneficial output. Thus, this paper provides new insight into how to motivate such an agent to produce an unverifiable output that does not maximize profit. That is, in order to motivate the agent to invest in ESG outcomes that do not necessarily reflect in his profit, a socially or environmentally aware firm principal should make him a residual claimant to some of the principal's value in the social good by creating opportunity for hold-up against the principal herself, such as offering the type of ex-post bargaining contract studied in this paper.

The intuition for the results derived from comparative statics on the manager's bargaining power is also interesting to examine in detail. First, an increase in the manager's bargaining power increases the investment he makes, and thus the expected utility of the principal under the ex-post bargaining contract, while the investment level and the principal's utility remain unchanged under the ex-ante commitment contract. This is because an increase in the manager's bargaining power endogenously increases the investment under the ex-post bargaining contract, as both parties know they will split the expected surplus at the beginning of the second period. This is in contrast to the ex-ante commitment contract, in which the second-period wage is determined at the beginning of the first stage regardless of the bargaining power.[6] This means that the greater the manager's bargaining power, the less he fears the hold-up, and thus the more he invests, contributing to a larger expected output.

Second, the expected marginal cost of the incremental investment is the same for both the manager and the principal, as the principal has to compensate for the manager's expected cost to induce the manager to participate in the contract. Thus, given the principal's expected marginal revenue of investment exceeds the manager's expected marginal cost of investment, the incremental investment will mean the principal's expected marginal revenue exceeds her expected marginal cost.[7] As a result, the principal's net expected utility increases under the ex-

[6] Furthermore, the manager's reservation utility is fixed at the beginning of the first stage.

[7] By substituting the IR constraint into the principal's expected utility, I show that her expected marginal revenue of investment is larger than the manager's expected marginal cost of investment. See Appendix A.

post bargaining contract, while it remains unchanged under the ex-ante commitment contract as the agent's bargaining power increases.

This shift in the principal's expected utility under the ex-post bargaining contract shifts the threshold at which the principal is indifferent about ex-ante and ex-post contracts. The threshold shifts towards zero or smaller ζ (or z or both), where the ex-ante contract is more likely to be chosen, as shown in the first result of this paper. This means there is greater likelihood the principal will choose the ex-post bargaining contract, because the distance between the threshold and zero represents the possibility of the ex-ante commitment contract – but this is narrower due to the new threshold. Thus, the more the threshold shifts towards zero or small z (or ζ or both), the greater the chance the principal will choose the ex-post bargaining contract.

The implications of my results are as follows. First, suppose a firm in which the unverifiable output can effectively reduce social costs caused by the verifiable output, or in which the social costs are substantial. In this case, it is better for the firm to hold wage negotiations frequently in order to promote socially responsible investment or activity if the firm's founder or the majority of the shareholders of the firm value ESG principles. On the other hand, for a firm in which the unverifiable output is not effectively reducing social costs, or in which the social costs are not substantial, it is better not to hold wage negotiations too often. Furthermore, if verifiable and unverifiable outputs are managed by two different managers, future wages for the manager who will produce verifiable outputs should be agreed at the beginning of the initial contract, whereas future wages for the manager who is expected to produce unverifiable outputs should be negotiated more often. For example, managers who are paid fixed pay rather than incentive pay are generally motivated by promotion or wage renewal by promotion, as is the case for bureaucrats or officers and employees of public sector companies. Therefore, in equilibrium, governments or government-owned companies can reduce negative externalities by investing more in unverifiable outputs.

Second, a number of companies, such as ALCOA and Royal Dutch Shell, have recently started to tie executive pay to specific ESG targets. In terms of my model, this is like offering the ex-ante commitment contract in which the wage is linked to some sort of signal of social costs, z. This implies that if the specific ESG targets are imprecise or vague, these firms actually resemble a firm offering an ex-ante commitment contract. Thus, if the unverifiable output can substantially

reduce social costs or if social costs are large, such firms might consider adopting a contract that is similar to the ex-post bargaining contract developed in this model.

Third, the manager's bargaining power in this paper can be interpreted as the skill gap between an incumbent manager and a newly hired manager. My result implies that if the skill gap increases, the company is more likely to offer an ex-post bargaining contract, because an incumbent manager cannot be substituted or replaced easily. Testable implications regarding this are discussed at the end of section 3.

Finally, my research can be applied to the analysis of social impact or sustainable-linked bonds. Under these arrangements, investors receive financial returns based on the accomplishment of predefined social objectives. Indeed, the optimal ex-post bargaining contract characterized in this paper can be implemented using traditional short-term debts, whereas the optimal ex-ante commitment contract can be implemented using social impact or sustainable-linked bonds. If the unverifiable output contributes to reducing social disutility, or if the production of verifiable outputs involves more social disutility, traditional short-term debts are better than social impact (sustainable-linked) bonds. Conversely, if the unverifiable output does not substantially contribute to reducing social disutility, or if the production of verifiable outputs involves less social disutility, social impact (sustainable-linked) bonds are better than traditional short-term debts. In addition, traditional short-term debts are more likely to be preferred, as the private investors' bargaining power increases.

The remainder of the paper is organized as follows. Section 2 discusses the related literature. Section 3 analyzes the basic model of managerial compensation. Section 4 studies the social impact and sustainability-linked bonds in the context of the models developed in section 3. The final section concludes. All of the proofs of the propositions and corollaries in the text are provided in Appendices A and B. Appendix C discusses limited liability constraints. Appendix D also investigates the case of a risk-averse agent.

2 Literature

The theoretical literature on socially responsible investing is limited. Heinkel, Kraus, and Zechner (2001) discuss a problem in which risk sharing is reduced when the firm is excluded by socially responsible investors. Hart and Zingales (2017) investigate a firm with prosocial investors who

dislike social costs if they feel directly responsible for them.[8] Chowdhry, Davies, and Waters (2019) deal with the financing of a profit-maximizing firm and examine how socially minded investors induce the firm to commit to pursuing social goals. Morgan and Tumlinson (2019) study firm behavior, when shareholders care about public goods as well as profits, and when managerial compensation reflects these concerns. They show that managers can redirect more profits toward public good than shareholders would when acting separately. Furthermore, if public good is sufficiently desirable, they also indicate that the manager will select the socially optimal level of output. Oehmke and Opp (2020) examine the ability of socially responsible investors to influence firms by relaxing financial constraints for clean production, when firm production generates social costs and socially responsible investors care about externalities regardless of whether they are directly responsible for the social costs. Broccardo, Hart, and Zingales (2020) examine the relative effectiveness of exit (divestment and boycott) and voice (engagement) strategies in promoting socially desirable outcomes in companies. However, none of these papers mentioned above discuss the agency problem with hold-up between the socially responsible principal and the manager in an incomplete contract framework, as shown in my paper. I consider how the socially responsible principal induces the manager to pursue socially responsible investment or activity by selecting the timing and commitment of contracts in the incomplete contract setting with a multitask principal-agent relationship.

Fudenberg, Holmstrom, and Milgrom (1990) and Ray and Salanie (1990) discuss how and when the principal can achieve the utility level of a long-term contract by repeating short-term contracts.[9] However, my focus in the present paper considers the condition under which the principal is better off offering ex-post bargaining contracts than an ex-ante commitment contract. Kamiya and Adachi-Sato (2013) present a general model of long-, short-, and medium-term wage contracts. Adachi-Sato and Kamiya (2013) develop a multi-task and job allocation model in which the agent has to produce not only verifiable but also unverifiable outputs where both outputs contribute to the firm's revenues in a framework as observed in actual labor markets.

[8] Extending the model of Heinkel, Kraus, and Zechner (2001), Pastor, Stambaugh, and Taylor (2020) derive an ESG factor in an asset-pricing equilibrium model. Also see Pederson, Fitzgibbons, Pomorski (2020).

[9] Dutta and Reichelstein (1996) show that short-term contracts can be better than a long-term contract in a different context. That is, in their model, agents get fired on the equilibrium path, and hence they allow agents to change sequentially. Their model-setting is different from mine in which the principal wishes to motivate one agent in a dynamic framework.

The present paper is quite different from the above two papers in that I investigate an ESG investment problem with social costs generated from the verifiable output. Furthermore, my paper examines completely different production processes from these two papers: the verifiable output contributes to the firm's revenue but generates the social cost of production, while the unverifiable output reduces the social cost of production. Indeed, the concept of the social costs, and the sensitivity to which the unverifiable output reduces the social cost is original in my model. As a result, unlike the above-mentioned two papers, my paper shows that ex-post bargaining contracts may be strictly better than offering an ex-ante commitment contract in the context of socially responsible investment. In addition, this paper also allows interpretation and examination of the use of social impact bonds using the discussion of security design.

Farrell and Shapiro (1989) and Bernheim and Whinston (1998) present models with verifiable and unverifiable attributes, where it is better not to contract, or to contract incompletely, over even verifiable attributes. My paper may seem somewhat similar to theirs. However, their logics are quite different from mine. Indeed, in Proposition 1 in Farrell and Shapiro, the seller does not prefer to sign a contract on verifiable attributes, because doing so becomes a constraint on optimizing unverifiable attributes. This argument has nothing to do with ex-post bargaining, and cannot be applied to my case. This is because in my model, the principal does not choose any variables so as to optimize her utility after signing a contract. Bernheim and Whinston (1998) demonstrate that if some aspects of performance are noncontractible, it may be optimal to leave other verifiable aspects of performance unspecified. This is quite different from my argument about the trade-off between ex-ante commitment and ex-post bargaining in inducing investments of the manager for verifiable and unverifiable outputs. Again, none of these papers examine the agency problem with hold-up between the socially responsible principal and the manager.

This paper contributes to the small but emerging body of literature on social impact or sustainable-linked bonds. Pauly and Swanson (2017) consider whether social impact bonds can finance projects that might not otherwise be undertaken using traditional bonds. They argue that social impact bonds will achieve greater program success if investors' efforts depend on incentives and can positively affect project outcomes. Tortorice, Bloom, Kirby, and Regan (2020) discuss a model of social impact bonds where there is asymmetric information about the probability of

project success. They indicate that social impact bonds expand the set of implementable projects if the government is pessimistic about the likelihood of a project success, or if the government is averse to paying costs associated with a project in excess of benefits. However, these studies leave the question of whether social impact bonds will be effective under social programs with complex outputs. Neither of these studies investigate the agency problem with hold-up between the ESG-oriented principal and the manager in the incomplete contract setting explored in my paper. Furthermore, my model considers the condition where social impact or sustainable-linked bonds are preferred to traditional short-term debts.

3 The basic model of managerial compensation

3.1 Model setting

3.1.1 Basic environment

There is a principal and a manger. The principal delegates the management of her firm to the manager. I assume that both of them are risk neutral. The manager wishes to maximize only her expected revenues less investment costs. However, the principal wishes to maximize the firm's expected revenues less the social cost of the production. If the principal is a founder family of the firm, this can be justified by assuming that the founder family has an intrinsic motive not to cause social harm. If the principal is a fund, the fund is a socially responsible investor that follows ESG criteria.

In order to motivate such a manager, the principal either offers an ex-ante commitment contract or an ex-post bargaining contract in a two-period model. The former contract binds both parties to the first- and second period wages for the manager. The latter contract only binds both parties to the first-period wage and lets them decide the second period wage via Nash bargaining at the beginning of the second period.

There are two types of output produced by the firm. One is an observable and verifiable output $x > 0$ that generates not only the firm's revenue but also disutility of a nonpecuniary negative externality, expressed as a constant $z > 0$, which is interpreted as social cost by the principal, who is aware of ESG. The other is an observable but unverifiable output $y > 0$ that reduces the principal's disutility by ζyz. The parameter $\zeta \geq 0$ is the sensitivity at which the unverifiable output reduces social costs. To focus on the role of y, in the subsequent analysis, I

assume that z is an observable but unverifiable constant value. There are two verifiable output levels, x^H and x^L, where $x^H > x^L > 0$. The probabilities of x^H and x^L are denoted by $P^H \in [0,1]$ and $P^L = 1 - P^H$. There are two unverifiable output levels, y^H and y^L, where $y^H > y^L > 0$. The probabilities of y^H and y^L are denoted by $Q^H \in [0,1]$ and $Q^L = 1 - Q^H$. During the first period, the manager makes two types of effort (hereafter firm-specific investments) to generate outputs, $I_x \geq 0$ and $I_y \geq 0$, to increase the productivity for producing x and y, respectively. I assume that both I_x and I_y are observable but unverifiable, and that P^H and Q^H in the second period are functions of I_x and I_y, denoted by $P^H(I_x)$ and $Q^H(I_y)$, respectively. As is formally stated in Assumptions 1 and 2 imposed below, I assume that the random variables x and y are stochastically independent and that $P^H = Q^H = 0$ in the first period. That is, I assume that the investments made in the first period will increase the manager's productivity from the second period onwards. The manager incurs disutilities in making the investments, denoted $D_x(I_x)$ and $D_y(I_y)$. Note that there is no complementarity or substitutability between I_x and I_y, as x and y are stochastically independent and the total cost of the investments is additively separable, i.e., $D_x(I_x) + D_y(I_y)$.

The manager's investments, I_x and I_y, develop an internal employment relation or organization, build up a network of connection, and attain firm-specific know-how to efficiently implement firm production. The information about these investments is collected over time through frequent and personal contacts between the principal and the manager, but is difficult to transmit to outsiders because of a lack of context. Hence, these investments are observable but unverifiable. In addition, if the incumbent manager is replaced with a new manager from outside, the newly hired manager cannot utilize the predecessor's investment made in the first-period to produce x or y because he lacks in the firm-specific skills. Hence, $P^H = Q^H = 0$ in the second period if the incumbent manager is replaced with a new manager from the outside.

The wage for each period is paid at the end of each period, or after the realization of the outputs in each period.[10] As x is the only verifiable variable, the wage depends on the realization of x only: the wages for x^H and x^L are denoted by w^H and w^L, respectively. Let w_t^i denote w^i, $i = H, L$, in period $t = 1, 2$. Because of risk neutrality, w_2^i need not depend on the realization of x in the first period. I first investigate the model without limited liability constraints. In

[10] As the agent is risk neutral, I can consider a model in which the wages for both periods are paid together at the end of the second period. This is, however, a special case of an ex-ante commitment contract.

Appendix C, I show that similar results still hold even after I impose these constraints.

Throughout this paper, I make the following assumption.

Assumption 0 $x^L - (1 - \zeta y^L)z - \underline{u} \geq 0,$

where $\underline{u} > 0$ is the reservation payoff of the manager in the competitive labor market in each period. This assumption is justified if x^L is sufficiently large while neither z nor $\underline{u}$ is sufficiently large. It ensures the payoffs of the principal and the manager during the second period under the ex-post bargaining contract are nonnegative.

Next, the following assumptions on the functions D_x, D_y, P^H, and Q^H are standard.

Assumption 1

1. $\frac{dD_i}{dI_i} > 0$, $\frac{d^2D_i}{dI_i^2} > 0$, $D_i(0) = 0$, and $\frac{dD_i(0)}{dI_i} = 0$, $i = x, y$.
2. $P^H(0) = 0$, $\frac{dP^H}{dI_x} > 0$, and $\frac{d^2P^H}{dI_x^2} < 0$.
3. $Q^H(0) = 0$, $\frac{dQ^H}{dI_y} > 0$ and $\frac{d^2Q^H}{dI_y^2} < 0$.
4. The random variables x and y are stochastically independent.

In addition, for simplicity, I make the following assumption.[11]

Assumption 2 The probabilities of x^H and y^H are zero in the first period.

Under this assumption, the principal need not determine w_1^H in the first period.

I assume that there are a lot of competitive managers when the principal hires a new manager. As a result, the principal is able to extract the full surplus of the firm. Hence, when the principal hires a new manager at any stage, she posts a take-it-or-leave-it wage offer to him.

I also assume that the manager obtains some firm-specific skills if he makes firm-specific investments. Thus, if the manager makes firm-specific investments, he can retain part of the surplus at the subsequent contracting stage because he has knowledge that gives him an advantage over outsiders. Accordingly, if the principal and the manager have agreed to the ex-post bargaining contract, they can bargain over the wage at the beginning of the subsequent period.

[11] I can obtain what I would like to achieve in this analysis without this assumption. The assumption is imposed purely for the sake of simplifying the analysis.

For simplicity, I use Nash bargaining for the negotiation between the two parties. I assume that the principal's and manager's bargaining powers are equal and that if they fail in bargaining, they need to find new partners, i.e., they can access the labor market and match with a new partner. More specifically, if bargaining fails, the principal makes a take-it-or-leave-it offer to a new manager from the outside market, whereas the incumbent manager is hired by another firm as a new manager through the firm's take-it-or-leave-it offer. However, given that the incumbent manager's investments I_x and I_y cannot be utilized by the newly hired manager because of the lack of firm-specific skills and the fact that these investments are firm-specific, the principal obtains only $x^L - (1 - \zeta y^L)z - \underline{u}$ by hiring the new manager while the incumbent manager receives only $\underline{u}$. Note that $\underline{u}$ is the manager's reservation utility determined in the competitive market. Thus, the principal's and manager's outside option values in the second period are $x^L - (1 - \zeta y^L)z - \underline{u}$ and $\underline{u}$, respectively. Hence, the threat point of the Nash bargaining is $(x^L - (1 - \zeta y^L)z - \underline{u}, \underline{u})$. It is worth noting that I obtain similar results even if their bargaining power is different in the second period.

Assumption 3 The principal posts a take-it-or-leave-it offer when a contract is signed with a new manager. The principal and the manager Nash bargain over the wages with the threat point held at $(x^L - (1 - \zeta y^L)z - \underline{u}, \underline{u})$ at the beginning of the second stage when both parties are bound by the ex-post bargaining contract.

3.1.2 Timing

At the start of the first period: The principal decides to offer an agent either an ex-ante wage commitment contract or an ex-post wage bargaining contract through a take-it-or-leave-it offer: both contracts bind the first-period wage to be w_1^L[12]; the second-period wage w_2 is contingent on the verifiable output x for the ex-ante commitment contract, while it is determined by the negotiation at the beginning of the second period for the ex-post bargaining contract. The agent accepts the contract that has been offered to him. If the agent is offered an ex-post bargaining contract, he accepts it considering the outcome of the Nash bargaining that will take place.

[12]The first-period wage is determined at the beginning of the first period to satisfy the IR constraint for the manager in both types of contract. It does not affect the principal's choice of offering an ex-post bargaining or ex-ante commitment contract, because it is determined before the manager undertakes investment under either contract.

Finally, the agent decides the investment levels for x and y, i.e., I_x and I_y.

At the end of the first period: The output is realized but it is only realized as $(x_1^L,\, y_1^L)$, as the investment made in the first period takes effect in the second period. Accordingly, the wage w_1^L is paid to the agent.

At the start of the second period: The principal and the agent Nash bargain over the agent's wage w_2^i if an ex-post bargaining contract was agreed at the start of the first period. If the negotiation breaks down, the manager is fired and a new manager is hired. Nothing happens if they had agreed on the ex-ante commitment contract.

At the end of the second period: The output $(x_2^i,\, y_2^i)$ where $i = H, L$, is realized for the ex-ante commitment contract and the manager receives a prespecified w_2^i contingent on x_2^i. The output $(x_2^i,\, y_2^i)$ where $i = H, L$, is realized for the ex-post bargaining contract with the manager retained from the start of the first period. The manager negotiates w_2^i and receives a share of the bargaining surplus (plus his threat point) determined in the negotiation. With the new manager who started at the beginning of the second period, the output is realized as $(x_2^L,\, y_2^L)$ and he receives $\underline{u}$. The manager fired at the beginning of the second period receives $\underline{u}$ in the new firm.

3.2 The first-best solution

I first determine the first-best optimal allocation without agency problems. For simplicity, I assume that no agents discount their payoffs or utility. With no moral hazard problems, the principal can determine the investment amounts I_x and I_y by herself as follows:

$$\max_{I_x, I_y} x^L - (1-\zeta y^L)z - D_x(I_x) - D_y(I_y) + \left[\sum_{j=H,L} P^j(I_x)x^j - \sum_{j=H,L} Q^i(I_y)(1-\zeta y^i)z\right].$$

The first-order conditions with respect to I_x and I_y for the above problem are given by

$$\frac{dD_x(I_x)}{dI_x} = \frac{dP^H(I_x)}{dI_x}(x^H - x^L),$$

and

$$\frac{dD_y(I_y)}{dI_y} = \frac{dQ^H(I_y)}{dI_y}\zeta(y^H - y^L).$$

Hence, the first-best investment levels are characterized by the above two equations.

3.3 An ex-post bargaining contract

Under the ex-post bargaining contract, the principal offers the first-period wage at the beginning of the first period, and they bargain over the second-period wage at the beginning of the second period. The manager can make investments during the first period to maximize his own expected payoff. However, the principal cares about both her expected revenues and the social costs of production. Then, by Assumption 3, the principal's problem in the first period under the ex-post bargaining contract is to offer a take-it-or-leave-it offer on the first-period wage in order to induce the manager to implement the principal's preferred investment levels, subject to the individual rationality constraint and the incentive compatibility constraint on investments:

$$\max_{w_1^L, I_x, I_y} x^L - w_1^L - (1 - \zeta y^L)z + V_2^p(I_x, I_y), \tag{1}$$

$$\text{s.t.} \quad w_1^L - D_x(I_x) - D_y(I_y) + V_2^m(I_x, I_y) \geq 2\underline{u}, \tag{2}$$

$$\begin{aligned} & w_1^L - D_x(I_x) - D_y(I_y) + V_2^m(I_x, I_y) \\ & \geq w_1^L - D_x(I_x') - D_y(I_y') + V_2^m(I_x', I_y'), \ {}^\forall I_x', I_y', \end{aligned} \tag{3}$$

where $V_2^p(I_x, I_y)$ and $V_2^m(I_x, I_y)$ are the principal's and manager's utilities in the second period when the investments are I_x and I_y, where $V_2^p(I_x, I_y)$ and $V_2^m(I_x, I_y)$ are determined by backward induction explained below. The individual rationality constraint is given by (2) and the incentive compatibility constraint is represented by (3).

The manager has bargaining power at the beginning of the second period. Applying Assumption 3, the principal and the manager Nash bargain over wages: for a given (I_x, I_y),

$$\max_{w_2^H, w_2^L} \left\{ \sum_{j=H,L} P^j(I_x)(x^j - w_2^j) - \sum_{i=H,L} Q^i(I_y)(1 - \zeta y^i)z - [x^L - (1 - \zeta y^L)z - \underline{u}] \right\}$$

$$\times \left\{ \sum_{j=H,L} P^j(I_x) w_2^j - \underline{u} \right\}.$$

As both players are risk neutral, they obtain the same surplus from the Nash bargaining solution,

which is half of the total surplus. Formally, their utilities are expressed as

$$\begin{aligned} V_2^p(I_x, I_y) &= \frac{1}{2}\left\{\sum_{j=H,L} P^j(I_x)x^j - \sum_{i=H,L} Q^i(I_y)(1-\zeta y^i)z - [x^L - (1-\zeta y^L)z]\right\} \\ &\quad + x^L - (1-\zeta y^L)z - \underline{u} \\ &\geq 0, \end{aligned} \tag{4a}$$

$$V_2^m(I_x, I_y) = \frac{1}{2}\left\{\sum_{j=H,L} P^j(I_x)x^j - \sum_{i=H,L} Q^i(I_y)(1-\zeta y^i)z - [x^L - (1-\zeta y^L)z]\right\} + \underline{u} \geq 0, \tag{4b}$$

where the last inequalities of (4a) and (4b) are evident from Assumption 0, $x^H > x^L$, and $y^H > y^L$.

3.4 An ex-ante commitment contract

Under the ex-ante commitment contract, the principal and the manager agree on and bind themselves to the wages for both periods at the beginning of the first period. The manager makes investments during the first period. In line with Assumption 3, the principal's contracting problem is to make a take-it-or-leave-it offer on the first- and second-period wages in order to induce the manager to implement the principal's preferred investment levels, subject to the individual rationality and incentive compatibility constraints on investments:

$$\max_{w_1^L, I_x, I_y, w_2^H, w_2^L} x^L - w_1^L - (1-\zeta y^L)z + \left[\sum_{j=H,L} P^j(I_x)(x^j - w_2^j) - \sum_{i=H,L} Q^i(I_y)(1-\zeta y^i)z\right], \tag{5}$$

$$\text{s.t.} \quad w_1^L - D_x(I_x) - D_y(I_y) + \sum_{j=H,L} P^j(I_x)w_2^j \geq 2\underline{u}, \tag{6}$$

$$w_1^L - D_x(I_x) - D_y(I_y) + \sum_{j=H,L} P^j(I_x)w_2^j \geq w_1^L - D_x(I_x') - D_y(I_y') + \sum_{j=H,L} P^j(I_x')w_2^j, {}^\forall I_x', I_y'. \tag{7}$$

The principal's utility is given by (5). Inequalities (6) and (7) are the individual rationality and incentive compatibility constraints of the manager.

3.5 A comparison of the two types of wage contract

I explain below the mechanism through which the principal decides between the two contracts. The following proposition and its corollary show that the result depends on both the effectiveness of y in reducing social costs, ζ, and the size of the social costs, z.

Proposition 1 *1. Under the ex-ante commitment contract, there exists an optimal level of investment for the verifiable output which maximizes the expected total net payoff generated by the firm if there were no social costs or if the principal or society entirely ignore the social cost. This optimal investment for the verifiable output is larger than that under the ex-post bargaining contract.*

2. Under the ex-ante commitment contract, w_2^H is strictly larger than w_2^L. Under the ex-post bargaining contract, the fixed wage, i.e., $w_2^H = w_2^L$, can be offered.

3. There exists a threshold $\bar{\zeta} > 0$ such that the principal prefers an ex-ante commitment to an ex-post bargaining contract at the beginning of the first period for $\zeta \in [0, \bar{\zeta})$, and prefers an ex-post bargaining to an ex-ante commitment contract for $\zeta \in (\bar{\zeta}, \infty)$.

Proof: See Appendix A. ∎

Corollary 1 *There exists a threshold $\overline{z} > 0$ such that the principal prefers an ex-ante commitment to an ex-post bargaining contract at the beginning of the first period for $z \in [0, \overline{z})$, and prefers an ex-post bargaining to an ex-ante commitment contract for $z \in (\overline{z}, \infty)$.*

Proof: See Appendix A. ∎

Several remarks about Proposition 1 and its corollary are in order. Under the ex-post bargaining contract, the bargaining position/surplus of the manager at the beginning of the second period depends on his productivity in producing y as well as on his productivity in producing x. Thus, the principal can induce the agent to invest in I_y. However, the investment level for both outputs is reduced due to hold-up. Furthermore, a fixed wage can be used to motivate the manager, which is also true under the risk-averse setting in Appendix D.

Under the ex-ante commitment contract, at the beginning of the first period, the principal can offer a second-period wage depending on the output x the manager is going to produce in the second period. However, she cannot offer a second-period wage that reflects the amount of y the manager is going to produce in this period, as y is observable but unverifiable. As a result, the ex-ante commitment contract cannot motivate the manager to invest in I_y at all, which is the investment to reduce social costs generated by x. However, the principal can motivate the

manager to invest more in I_x by making w_2^H much bigger than w_2^L. Indeed, the equilibrium level of I_x produces the optimal level for the verifiable output, which maximizes the expected total net payoff generated by the firm for both the first and second periods, exclusive of the social costs represented by z.

Under the ex-ante commitment contract, the first-best allocation can be achieved if unverifiable output y is useless in reducing the social cost z, that is if $\zeta = 0$, and therefore there is no need to produce y, or formally $I_y = 0$. Given that the principal can set the manager's utility equal to $\underline{u}$ under the optimal ex-ante commitment contract, and that the ex-post bargaining contract cannot achieve the first-best allocation, the principal strictly prefers the ex-ante commitment contract when $\zeta = 0$. However, when ζ increases from 0, (meaning ζ is no longer 0) the investment allocation between I_x and I_y is distorted under the ex-ante commitment contract. This is because in this situation the principal does not have any incentive schemes to control I_y, although I_y can reduce the social costs created by I_x. As a result, the principal is forced to induce the manager to choose an inefficiently high level of I_x under the ex-ante commitment contract. Because the principal has an incentive scheme to control I_y under the ex-post bargaining contract, it is possible that she prefers the ex-post bargaining contract to the ex-ante commitment contract if ζ is sufficiently large. This logic can also apply to the variation of z.

Hence, the principal's choice between an ex-post bargaining and an ex-ante commitment contract depends on both the sensitivity/effectiveness of y in reducing social costs and the size of the social costs z. That is, if the ESG-oriented principal wishes the manager to invest in both I_x and I_y, she will offer him the ex-post bargaining contract if y effectively reduces social costs or if the social costs are sufficiently large. See Fig. 1.

The practical implications of Proposition 1 and its corollary are as follows. First, in an industry or a firm where unverifiable outputs substantially contribute to a reduction in social costs caused by verifiable outputs or where the social costs are substantially large, it is better to hold ex-post wage negotiations frequently; otherwise, it is better not to hold ex-post wage negotiations too often. In addition, even in the same firm, if one manager is mainly involved in producing verifiable outputs with social costs whereas the other manager is mainly involved in producing unverifiable outputs for reducing the social costs, then the firm should commit to future wages agreed at the beginning of the initial contract for the former manager, whereas the

firm should negotiate future wages for the latter manager more often.

Second, managers who are involved in producing unverifiable outputs and hence receive more fixed pay may be seen as motivated by promotion or wage renewal by promotion. This tendency towards promotion is significantly observed among managers in companies owned by central or local government. Thus, if these firms incur social costs, and they can reduce these costs overall with socially responsible investment, their government owners are more likely to be successful in persuading them to do so.

Finally, a number of companies have recently started to embed ESG more deeply by relating executive pay to specific ESG targets.[13] However, if the specific ESG targets are imprecise or not relevant to the firm's ESG objective, these firms actually resemble a firm offering the ex-ante commitment contract in my model. In this case, if the unverifiable output substantially contributes to reducing social costs of verifiable outputs or if the social costs are substantially large, it may be better for the firm to hold ex-post wage negotiation frequently, like a firm offering the ex-post bargaining contract in my model.

Next, I consider the effect of firm specificity—that is, the manager's bargaining power—on the choice of an ex-post bargaining or an ex-ante commitment contract.

Proposition 2 *If the manager's bargaining power comparatively increases, the principal is more likely to offer the manager an ex-post bargaining contract.*

Proof: See Appendix B. ∎

Intuitively, an increase in the manager's bargaining power β increases the investments I_x and I_y he makes, and thus the expected utility of the principal under the ex-post bargaining contract, while the investment level and the principal's utility remain unchanged under the ex-ante commitment contract. This is because an increase in the manager's bargaining power endogenously increases the investment under the ex-post bargaining contract, as both parties know they will split the expected surplus at the beginning of the second period. This is in contrast to the ex-ante commitment contract, in which w_2 is determined at the beginning of

[13] For example, Royal Dutch Shell announced plans to tie executive pay to three-to-five year targets for net carbon footprints from 2020 (see King, 2020). In ALCOA, 20 percent of executive cash compensation is tied to safety, environmental stewardship (including greenhouse gas emissions reductions and energy efficiency), and diversity goals (see https://corpgov.law.harvard.edu/2019/09/10/executive-compensation-and-esg/).

the first stage regardless of the bargaining power. This means that the greater the manager's bargaining power, the less he fears the hold-up, and thus the more he invests, contributing to a larger expected output.

Second, the expected marginal cost of the incremental investment is the same for both the manager and the principal, as the principal has to compensate for the expected cost to induce the manager to participate in the contract. Moreover, by substituting the manager's IR constraint into the principal's expected utility, I show that her expected marginal revenue of investment is larger than the manager's expected marginal cost of investment. See Appendix A. Thus, the incremental investment will induce the principal's expected marginal revenue to exceed her expected marginal cost.

As a result, as β increases, the principal's net expected utility increases under the ex-post bargaining contract, while it remains unchanged under the ex-ante commitment contract.

This shift in the principal's expected utility under the ex-post bargaining contract shifts the threshold at which the principal is indifferent about ex-ante and ex-post contracts. The threshold shifts towards zero, where the ex-ante contract is more likely to be chosen, as was shown in Fig. 1. This means there is greater likelihood the principal will choose the ex-post bargaining contract, because the distance between the threshold and zero represents the possibility of the ex-ante commitment contract – but this is narrower due to the new threshold. Thus, the more the threshold shifts towards zero, the greater the chance the principal will choose the ex-post bargaining contract.

As the manager's bargaining power in this paper depends on the firm-specificity of investments, it can be interpreted as the skill gap between an incumbent manager and a newly hired manager.[14] Then, the implication of this is that if the skill gap between the incumbent manager and a newly hired manager increases, the more likely an ex-post bargaining contract is chosen. Thus, the company is more likely to choose an ex-post bargaining contract as the skill gap between the incumbent manager and a newly hired manager increases so that the incumbent manager cannot be substituted or replaced easily.

[14] On the other hand, a shrinkage of the manager's supply can be viewed as an increase in $\underline{u}$. Because an increase in $\underline{u}$ increases the principal's utility under the ex-post bargaining and the ex-ante commitment contract by the same extent, it has no effects on the choice of the ex-post bargaining and the ex-ante commitment contracts.

Berry, Bizjak, Lemmon, and Naveen (2006) empirically find that diversified firms need a manager of greater ability because managing a diversified firm is a more difficult task than managing a focused firm. As it is more costly for diversified firms to find a suitable manager, they suggest that the cost for replacing an incumbent manger is higher in diversified than in focused firms. This implies that the skill gap between the incumbent manager and a newly hired manager is higher in diversified firms than focused firms. Hence, diversified companies are more likely to choose an ex-post bargaining contract if they face the social cost problem discussed in this paper.

As argued in Laux (2012), promoting an insider to the manager is not only less time-consuming but also less costly because the insider has already acquired firm-specific human capital. As a result, firms that have a well-organized insider succession plan are more likely to offer an ex-ante commitment contract.

Furthermore, Laux (2012) also suggests that for firms in which the incumbent manager has already established that he is the right person for the position, replacing him is very costly. This conversely means that replacing the incumbent manager is not costly for firms in which the incumbent manager is a relatively new hire (maybe from outside) with uncertain talent or fit, and for firms that have recently changed their business strategy so that it is unclear if the incumbent remains a good match. These firms are more likely to choose an ex-ante commitment contract.

4 Extensions: social impact and sustainability-linked bonds

The analysis of this article can also be applied to social impact bonds. That is, bonds issued by public entities to finance social services or programs. This section examines the effectiveness of social impact bonds in comparison with traditional short-term borrowing. It also discusses sustainability-linked bonds, which are normally issued by a private company or a public entity.[15]

The general structure of the social impact bond is as follows. An issuer borrows funds from a private for-profit investor to execute a social program. The issuer is most often a public

[15] For simplicity, in the subsequent discussion, I assume there is no default, regardless of whether the issuer uses traditional short-term debts, or social impact bonds or sustainability-linked bonds. This implies that the issuer has enough funds to repay debt or bond payments, even though for political reasons it cannot make enough funds available for the project prior to proven success.

entity with altruistic preferences, for example, local government. The issuer then furnishes the funds to a nonprofit service provider that needs to finance up-front costs to execute the program. For simplicity, I assume the issuer and the nonprofit service provider are the same, that is, the principal.[16] The issuer and the investor, such as an investment bank, then agree to a performance-contingent debt contract that allows the issuer to pay only if a pre-defined performance target is met. If the program successfully attains the target, the issuer pays both principal and interest; otherwise, the issuer pays nothing in most cases.[17]

Furthermore, the social impact bond induces the private investor to exert an effort to positively influence program performance. Indeed, the private investor not only expresses his concern about the social program and the current inability of the government to deal with it, but also can offer specific ideas about methods and techniques to solve the problem. Pauly and Swanson (2017) present evidence that existing social impact bonds engage private investors with program-specific expertise to improve program performance (see Section 6 and Appendix A in their study).[18]

Alternatively, the issuer can finance the social program using traditional short-term debt at each period. In this case, the issuer needs to pay both principal and interest to the private investor, shouldering all financial risk. In addition to the social impact bond and traditional short-term debt, I assume that the issuer can also obtain a part of the funds from government transfer.

The timing of the model is as follows. At the beginning of period 1, the issuer offers the social impact bond or traditional short-term debt to the private investor to finance the social program. During periods 1 and 2, the program is executed. If the issuer uses traditional short-term debt, it rolls over the short-term debt at the beginning of period 2. At the end of period 2, the program's final success or failure is realized. Under the social impact bond, the issuer pays

[16] Tortorice, Bloom, Kirby, and Regan (2020) make the same assumption.

[17] For example, the first US-based social impact bond program is the NYC ABLE Project for Incarcerated Youth, launched in 2012. In this social impact bond, the issuer (government payer) is New York City's Department of Corrections, while the investor is Goldman Sachs. The social impact bond implements a recidivism-reduction program targeted to Rikers Island adolescent inmates. The issuer pays back only if the recidivism rate falls. If the recidivism rate falls by 10 percent, Goldman Sachs receives its capital back;, if it exceeds 11 percent, Goldman Sachs also receives a financial return consistent with typical community development. See Pauly and Swanson (2017).

[18] Managers of nonprofit service providers may also exert productive effort. However, to focus on the role of the private investor, I assume here that their productive effort is fixed and invariant irrespective of the financing method.

both principal and interest only if the program successfully attained its targets; whereas under traditional short-term debt, the issuer pays both principal and interest regardless of outcomes.

The issuer needs to finance up-front capital expenditures $\underline{u}$ in each period to execute the social program. If the social program is executed, the performance outcome of the social program for the issuer is measured by the observable and contractible output $x > 0$ in each period. However, this program may generate disutility for program participants or running costs (exclusive of $\underline{u}$), $z > 0$, in each period that reduces the issuer's utility, where z is observable but noncontractible.[19] However, if the observable but noncontractible output $y > 0$ is produced, the principal's disutility is reduced by ζyz in each period. The observable but noncontractible effort $I_x \geq 0$ and $I_y \geq 0$ can be viewed as the private investor's effort to increase productivity for the production of x and y, respectively.[20]

The ex-post bargaining contract given in the preceding section can be transformed into traditional short-term debt, and the ex-ante commitment contract into social impact bond. For traditional short-term debt, the issuer borrows $\underline{u}$ from the private investor at the beginning of period 1, and pays back w_1^L at the end of period 1. Then, if the issuer rolls over the short-term debt at the beginning of period 2, she can make a fixed payment to the private investor at the end of period 2. However, if the issuer fails to roll over the short-term debt, she must finance $\underline{u}$ from a new private investor in the outside loan market. On the other hand, for the social impact bond, the issuer offers a performance-contingent bond at the beginning of period 1: he borrows $2\underline{u} - w_1^L$ from the private investor at the beginning of period 1 and pays w_2^H (w_2^L) to the private investor at the end of period 2 if the pre-specified performance outcome is (is not) met, that is ,$x = x^H$ (x^L). In this case, w_1^L needs to be funded by the government transfer at the beginning of period 1. This interpretation particularly holds true if x^L is sufficiently small.

Suppose the issuer uses short-term debt to finance the social program. Then, at the beginning of period 1, the issuer offers short-term debt to maximize her expected utility represented by (1), subject to the following constraints: the private investor's participation constraint, (2), which ensures that his net expected payoff at the beginning of period 1 is equal to the total lending

[19] For example, in prisoner rehabilitation program, the better performance outcome of the program may increase effort disutility of prisoners or additional running costs of prisons.

[20] The private investor's effort can also be viewed as his effort to apply his specific ideas about methods and technique in order to solve the design and management problem.

amount $2\underline{u}$, and his incentive compatibility constraint, (3), which implies that he chooses his efforts during the first period to maximize his own net expected payoff at the beginning of period 1. Because the private investor obtains some program-specific skills in period 1, he has bargaining power at the beginning of period 2.[21] Hence, the issuer and the private investor Nash bargain over the period 2 debt payment. If bargaining fails, the issuer borrows $\underline{u}$ by making a take-it-or-leave-it offer to a new private investor in the outside market, while the initial private investor lends $\underline{u}$ by accepting a take-it-or-leave-it offer from another borrower. Because I assume that the risk-free interest rate is equal to zero and that no new private investor can utilize the initial lender's program-specific expertise obtained in period 1, the issuer's and initial lender's outside options are represented by $x_L - z - \underline{u}$ and $\underline{u}$, as indicated by the bargaining problem characterized in Section 3.1.

Next, suppose that the issuer uses the social impact bond to finance the social program. In this case, at the beginning of period 1, the issuer and the private investor agree on the debt payment contingent on the observable performance outcome x at the end of period 2. Hence, the issuer offers the social impact bond to maximize her expected utility represented by (5) subject to the individual rationality constraint for the private investor, (6), and the incentive compatibility constraint on the private investor's efforts, (7).

These arguments show that the optimal contract derived in the preceding section can be implemented as follows: the optimal ex-post bargaining contract can be implemented using traditional short-term debt, whereas the optimal ex-ante commitment contract can be implemented using the social impact bond.

Accordingly, applying Propositions 1 and 2, I obtain the following proposition and corollary.

Proposition 3 *1. There exists a $\bar{\zeta} > 0$ such that the issuer prefers the social impact bond to traditional short-term debt at the beginning of the first period for $\zeta \in [0, \bar{\zeta})$, and prefers traditional short-term debt to the social impact bond for $\zeta \in (\bar{\zeta}, \infty)$.*

2. If the private investor has more bargaining power, the issuer is more likely to offer the private investor the traditional short-term debt.

[21] This discussion is reminiscent of the literature regarding the bank's bargaining power in the bank loan contract. See Rajan (1992).

Corollary 2 *There exists a* $\overline{z} > 0$ *such that the issuer prefers the social impact bond to traditional short-term debt at the beginning of the first period for* $z \in [0, \overline{z})$*, and prefers traditional short-term debt to the social impact bond for* $z \in (\overline{z}, \infty)$*.*

The implications of this proposition and its corollary are provided as follows. First, the social impact bond gives more incentive for the private investor to make efforts to achieve the higher performance outcome with social disutility by offering him contingent debt payments. Moreover, under the social impact bond, the equilibrium effort level for the higher performance outcome maximizes the expected total net utility enjoyed by the issuer who does not consider social disutility.

Second, traditional short-term debt motivates the private investor to make efforts both to achieve the higher performance outcome with social disutility and to reduce the social disutility; however, the effort level for the higher performance outcome is not the one that maximizes the expected total net utility generated by the issuer (who, again, does not consider social disutility).

Third, if the effectiveness of y in reducing social disutility, ζ, improves or if the social disutility is significantly large, the issuer is more likely to prefer traditional short-term debt to the social impact bond. In other words, if the unverifiable output reduces more social disutility or if the production of verifiable outputs involves more social disutility, the issuer is more likely to choose traditional short-term debt.

Finally, if the private investor's bargaining power increases, the less likely it is the social impact bond is chosen. Thus, when the current private investor cannot be substituted or replaced easily, the issuer is less likely to choose the social impact bond if the private investor's bargaining power increases.

In the case of sustainability-linked bonds, which have financial features that vary according to whether the issuer achieves predefined ESG key performance indicators, these bonds can be issued by any for-profit company or public entity with access to capital markets.[22] If the likelihood of the issuer meeting the target for key performance indicators highly depends on the effort or monitoring of investors, the analysis of this section is applicable to the case of sustainability-linked bonds.

[22] For sustainability-linked bonds, see Uzsoki (2020).

5 Conclusion

In this article, I explore how a profit-maximizing manager can be motivated to pursue socially responsible investment with wage contracts. I demonstrate that incentive contracting (an ex-ante commitment contract) and hold-up (an ex-post bargaining contract) are alternative ways to motivate a manager to make socially responsible investments. That is, an ex-ante commitment contract does not allow for hold-up for the investment for the verifiable output that accompanies social costs. Rather, it deprives the manager of the incentive to invest in the unverifiable output that reduces social costs. An ex-post bargaining contract allows for a hold-up for the investment in the verifiable output, but provides the manager with investment incentive for the unverifiable output. Hence, an appropriate use of contracts of different types can mitigate the inefficiency caused by the trade-off.

If the unverifiable output substantially contributes to reducing social costs or if the social costs are substantially large, the principal offers an ex-post bargaining contract with a fixed wage. Otherwise, the principal offers an ex-ante commitment contract with incentive pay. These results hold regardless of whether the manager is risk neutral or risk averse. In addition, under the risk-neutral setting, if the manager's bargaining power increases, the principal is more likely to offer the manager an ex-post bargaining contract.

This paper provides new insight into how to motivate a profit-maximizing manager to produce an unverifiable output that does not maximize profit. That is, in order to motivate such a manager to invest in ESG outcomes that do not necessarily reflect in his expected profit, a socially or environmentally aware firm principal should make him a residual claimant to some of her value in the social good. This can be achieved by creating opportunity for hold-up against the principal herself, such as offering the type of ex-post bargaining contract studied in this paper. In normal situations, when the manager faces a hold-up problem, the classical solution of "sell the entire firm to the manager" resolves the agency problem and the first-best solution is obtained. However, when the manager has to engage in the activity that does not bring him a direct profit such as doing social good (and hence he is not interested), the classical solution of "selling the entire project to the manager" does not motivate him to engage in such production as he simply does not care for such issues. Consequently, this paper demonstrates that the

under-effort problem of the manager can be mitigated by transferring the part of the principal's surplus through Nash bargaining process.

An useful implication of this study is to investigate whether a social impact or sustainability-linked bond is preferred to traditional short-term debt when financial investing involves ESG impact.

References

Adachi-Sato, Meg, 2018. Stock vesting conditions, control benefits and managerial replacement. Canadian Journal of Economics/Revue Canadienne D'economique, 51(1), 275-313. doi:10.1111/caje.12323

Adachi-Sato, Meg, Kamiya, Kazuya, 2013. A dynamic multitask model: Fixed wages, no externalities, and holdup problems. CIRJE Discussion Paper F-825.

Bernheim, B Douglas, Whinston, Michael D., 1998. Incomplete contracts and strategic ambiguity. The American Economic Review, 88(4), 902-932.

Berry, Tammy K., Bizjak, John M., Lemmon, Michael L., Naveen, Lalitha, 2006. Organizational complexity and CEO labor markets: Evidence from diversified firms. Journal of Corporate Finance, 12(4), 797-817. doi:10.1016/j.jcorpfin.2005.04.001

Broccardo, Eleonora, Hart, Oliver D., Zingales, Luigi, 2020. Exit vs. voice. NBER Working Paper, 27710. doi:10.3386/w27710

Chowdhry, Bhagwan, Davies, Shaun William, Waters, Brian, 2019. Investing for impact. The Review of Financial Studies, 32(3), 864-904. doi:10.1093/rfs/hhy068

Dutta, Sunil, Reichelstein, Stefan, 2003. Leading indicator variables, performance measurement, and long-term versus short-term contracts. Journal of Accounting Research, 41(5), 837-866. doi:10.1046/j.1475-679X.2003.00125.x

Farrell, Joseph, Shapiro, Carl, 1989. Optimal contracts with lock-in. The American Economic Review, 79(1), 51-68.

Fudenberg, Drew, Holmstrom, Bengt, Milgrom, Paul, 1990. Short-term contracts and long-term agency relationships. Journal of Economic Theory, 51(1), 1-31. doi:10.1016/0022-0531(90)90048-O

Hart, Oliver, Zingales, Luigi, 2017. Companies should maximize shareholder welfare not market value. Journal of Law, Finance, and Accounting, 2(2), 247-275. doi:10.1561/108.00000022

Heinkel, Robert, Kraus, Alan, Zechner, Josef, 2001. The effect of green investment on corporate behavior. Journal of Financial and Quantitative Analysis, 36(4), 431-449. doi:10.2307/2676219

Hildenbrand, Werner, 1974. Core and equilibria of a large economy. Princeton University Press(Princeton, New Jersey), viii - 251.

Holmstrom, Bengt, Milgrom, Paul, 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. The Journal of Law, Economics, and Organization, 7, 24-52.

Hughes, Kristin, Sakano, Akira, Gore, Al, Lacqua, Francine, Niinami, Tak, Wijsen, Melati, 2020. Breaking free from single-use plastics. World Economic Forum Annual Meeting 2020, <https://jp.weforum.org/events/world-economic-forum-annual-meeting-2020/sessions/ breaking-free-from-single-use>, accessed May 5, 2020.

Inderst, Roman, Mueller, Holger M., 2010. CEO replacement under private information. The Review of Financial Studies, 23(8), 2935-2969. doi:10.1093/rfs/hhq018

Kamiya, Kazuya, Adachi-Sato, Meg, 2013. Multiperiod contract problems with verifiable and unverifiable outputs. Mimeo.

King, Timothy, 2020. The many questions of tying ESG to executive compensation. Corporate Board Member, <https://boardmember.com/questions-tying-esg-executive-compensation/>, accessed Aug 1, 2021.

Laux, Volker, 2012. Stock option vesting conditions, CEO turnover, and myopic investment. Journal of Financial Economics, 106(3), 513-526. doi:10.1016/j.jfineco.2012.06.003

Morgan, John, Tumlinson, Justin, 2019. Corporate provision of public goods. Management Science, 65(10), 4489-4504. doi:10.1287/mnsc.2018.3137

Oehmke, Martin, Opp, Marcus M., 2020. A theory of socially responsible investment. CEPR Discussion Paper No. DP14351.

Pastor, Luboš, Stambaugh, Robert F., Taylor, Lucian A., 2020. Sustainable investing in equilibrium. Journal of Financial Economics, in press. doi:10.1016/j.jfineco.2020.12.011

Pauly, Mark V., Swanson, Ashley, 2017. Social impact bonds: New product or new package? The Journal of Law, Economics, and Organization, 33(4), 718-760. doi:org/10.1093/jleo/ewx012

Pedersen, Lasse Heje, Fitzgibbons, Shaun, Pomorski, Lukasz, 2020. Responsible investing: The ESG-efficient frontier. Journal of Financial Economics, in press. doi:10.1016/j.jfineco.2020.11.001

Rajan, Raghuram G., 1992. Insiders and outsiders: The choice between informed and arm's-length debt. The Journal of Finance, 47(4), 1367-1400. doi:10.1111/j.1540-6261.1992.tb04662.x

Rey, Patrick, Salanie, Bernard, 1990. Long-term, short-term and renegotiation: On the value of commitment in contracting. Econometrica, 58(3), 597-619.

Tortorice, Daniel L., Bloom, David E., Kirby, Paige, Regan, John, 2020. A theory of social impact bonds. NBER Working Paper, 27527.

Uzsoki, David, 2020. Sustainability-linked bonds: A new way to finance COVID-19 stimulus. International Institute for Sustainable Development, <https://www.iisd.org/articles/sustainability-linked-bonds>, accessed July 22, 2020.

Appendices

A. Proof of Proposition 1 and Corollary 1

An ex-post bargaining contract

In the first period, given (4) and $\sum_{i=H,L} Q^i(I_y) = 1$, (3) is rearranged so that the manager chooses I_x and I_y satisfying the following incentive compatibility constraint:

$$\max\ w_1^L - D_x(I_x) - D_y(I_y) + \frac{1}{2}\left\{ \sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z\right] \right\}$$
$$+x^L - (1-\zeta y^L)z - \underline{u}. \tag{A1}$$

The first-order conditions then yield

$$\frac{dD_x(I_x)}{dI_x} = \frac{1}{2}\frac{dP^H(I_x)}{dI_x}(x^H - x^L), \tag{A2a}$$

and

$$\frac{dD_y(I_y)}{dI_y} = \frac{1}{2}\frac{dQ^H(I_y)}{dI_y}\zeta(y^H - y^L)z. \tag{A2b}$$

Note that by Assumption 1 the second-order conditions are satisfied. Let the solutions of the above equations be I_x^* and I_y^*. On the other hand, it follows from (2) that the principal must set

$$w_1^L = D_x(I_x^*) + D_y(I_y^*) - V_2^m(I_x^*, I_y^*) + 2\underline{u}. \tag{A3}$$

As discussed in the text, the Nash bargaining solution is (4). That is,

$$\begin{aligned} V_2^p(I_x^*, I_y^*) &= \frac{1}{2}\left\{ \sum_{j=H,L} P^j(I_x^*)x^j - z + \sum_{i=H,L} Q^i(I_y^*)\zeta y^i z - \left[x^L - (1-\zeta y^L)z\right] \right\} \\ &\quad +x^L - (1-\zeta y^L)z - \underline{u}, \\ V_2^m(I_x^*, I_y^*) &= \frac{1}{2}\left\{ \sum_{j=H,L} P^j(I_x^*)x^j - z + \sum_{i=H,L} Q^i(I_y^*)\zeta y^i z - \left[x^L - (1-\zeta y^L)z\right] \right\} + \underline{u}. \end{aligned}$$

Then, it follows from (A3) that the principal's expected utility, (1), is obtained:

$$\begin{aligned} & x^L - w_1^L - (1-\zeta y^L)z + V_2^p(I_x^*, I_y^*) \\ = \ & x^L - (1-\zeta y^L)z - D_x(I_x^*) - D_y(I_y^*) + V_2^m(I_x^*, I_y^*) + V_2^p(I_x^*, I_y^*) - 2\underline{u} \\ = \ & x^L - (1-\zeta y^L)z - D_x(I_x^*) - D_y(I_y^*) \\ & + \left[\sum_{j=H,L} P^j(I_x^*)x^j - z + \sum_{i=H,L} Q^i(I_y^*)\zeta y^i z \right] - 2\underline{u}. \end{aligned} \tag{A4}$$

Finally, as discussed in the text, the Nash bargaining solution shows that the principal can choose a fixed wage, i.e.,

$$w_2^H = w_2^L = V_2^m(I_x^*, I_y^*) = \frac{1}{2}\left[\sum_{j=H,L} P^j(I_x^*)x^j + \sum_{i=H,L} Q^i(I_y^*)\zeta y^i z - x^L - \zeta y^L z\right] + \underline{u}.$$

∎

An ex-ante commitment contract

Let I_y^{**} be the optimal investment level that satisfies (7). Then, from Assumption 1.1, $I_y^{**} = 0$. Suppose that $w_2^j = x^j - r$, $j = H, L$, where r is the principal's utility in period two. Then, substituting $w_2^j = x^j - r$, $j = H, L$, into (7), I obtain the following first-order condition with respect to I_x:

$$\frac{dD_x(I_x)}{dI_x} = \frac{dP^H(I_x)}{dI_x}(x^H - x^L). \tag{A5}$$

Let I_x^{**} be the solution.

On the other hand, I can consider the following maximization problem of the joint utility of the principal and the manager for $I_x^{**} = 0$:

$$x^L - (1 - \zeta y^L)z - D_x(I_x) + \sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(0)\zeta y^i z. \tag{A6}$$

Note that $D_y(0) = 0$. Then, it is evident that the first-order condition with respect to I_x is again obtained by (A5). Under Assumptions 1.1 and 1.2, this implies that I_x^{**} also maximizes the joint utility of the principal and the manager when $I_y^{**} = 0$.

Using (6) with $I_y^{**} = 0$, the principal must set

$$w_1^L = D_x(I_x^{**}) - \sum_{j=H,L} P^j(I_x^{**})w_2^j + 2\underline{u}. \tag{A7}$$

Then, the principal's utility, (5), for $I_y^{**} = 0$ is expressed as follows:

$$\begin{aligned} &x^L - w_1^L - (1 - \zeta y^L)z + \textstyle\sum_{j=H,L} P^j(I_x^{**})(x^j - w_2^j) - z + \sum_{i=H,L} Q^i(0)\zeta y^i z \\ &= x^L - (1 - \zeta y^L)z - D_x(I_x^{**}) + \textstyle\sum_{j=H,L} P^j(I_x^{**})x^j - z + \sum_{i=H,L} Q^i(0)\zeta y^i z - 2\underline{u}. \end{aligned} \tag{A8}$$

As has been shown above, when $I_y^{**} = 0$, I_x^{**} maximizes the joint utility of the principal and the manager, and satisfies (6) and (7) for $w_2^j = x^j - r$, $j = H, L$. Given that the manager's reservation utility is set equal to a constant level $2\underline{u}$, these findings show that the optimal ex-ante

commitment contract consists of $(I_x,\ I_y) = (I_x^{**},\ I_y^{**}) = (I_x^{**},\ 0)$ and $w_2^j = x^j - r,\ j = H, L$. Finally, it follows from $w_2^j = x^j - r, j = H, L$, that w_2^H is larger than w_2^L. ∎

A comparison of two types of contract

First, comparing (A2a) and (A5), the manager undertakes more investment in I_x under the ex-ante commitment contract than under the ex-post bargaining contract, i.e., $I_x^* < I_x^{**}$.

When $\zeta = 0$, the principal prefers the ex-ante commitment contract to the ex-post bargaining contract, i.e., (A8) is larger than (A4). Indeed, when $\zeta = 0$, it follows from (A2b) with $\frac{dD_y(0)}{dI_y} = 0$ that $I_y^* = 0$ is chosen even in the ex-post bargaining contract. Thus, using $\zeta = I_y^* = I_y^{**} = 0$,

$$\text{(A8)} - \text{(A4)} = -D_x(I_x^{**}) + \sum_{j=H,L} P^j(I_x^{**})x^j - \left(-D_x(I_x^*) + \sum_{j=H,L} P^j(I_x^*)x^j \right) > 0.$$

Given Assumptions 1 and 2, the last inequality holds because (A5) implies that I_x^{**} maximizes $-D_x(I_x) + \sum_{j=H,L} P^j(I_x)x^j$.

To investigate the effect of an increase in ζ on the choice of contracts, using (A4) and (A8), I only need to investigate

$$\kappa(\zeta) = -D_y(I_y^*) + \sum_{i=H,L} \left[Q^i(I_y^*) - Q^i(0) \right] \zeta y^i z,$$

because (A2a) and (A5) imply that neither I_x^* nor I_x^{**} depends on ζ.

Then, it follows from (A2b) with Assumptions 1.1 and 1.3 that

$$\kappa'(\theta) = \frac{1}{2} \frac{dQ^H(I_y^*)}{dI_y^*} \zeta (y^H - y^L) z \cdot \frac{dI_y^*}{d\zeta} + \sum_{i=H,L} \left[Q^i(I_y^*)) - Q^i(0) \right] y^i z,$$

where

$$\frac{dI_y^*}{d\zeta} = \frac{\frac{1}{2} \frac{dQ^H(I_y^*)}{dI_y^*} (y^H - y^L) z}{\frac{d^2 D_y(I_y^*)}{dI_y^{*2}} - \frac{1}{2} \frac{d^2 Q^H(I_y^*)}{dI_y^{*2}} \zeta (y^H - y^L) z} > 0. \tag{A9}$$

Note that κ is a strictly increasing function of ζ, and goes towards $+\infty$ as ζ goes towards $+\infty$. This implies that the principal's utility under the ex-post bargaining contract, (A4), is larger than that under the ex-ante commitment contract, (A8), when ζ is sufficiently large.[23] In contrast, when $\zeta = 0$, the principal strictly prefers the ex-ante commitment contract to the ex-post

[23] Note that Assumption 0 does not prevent this logic because it is more likely to hold when ζ is larger.

bargaining contract. Thus, there exists a $\overline{\zeta} > 0$ such that the principal prefers the ex-ante commitment contract to the ex-post bargaining contract for $\zeta \in [0, \overline{\zeta})$, and prefers the ex-post bargaining contract to the ex-ante commitment contract for $\zeta \in (\overline{\zeta}, \infty)$. ∎

Proof of Corollary 1:

Social cost z appears only as ζz in (A2b) and the deduction of [(A8)−(A4)]. Accordingly, applying the procedure used in "A comparison of Two Types of Contract" analyzed above, I can verify the statement of this corollary. ∎

B. Proof of Proposition 2

I begin with examining the effect of an increase in the manager's bargaining power on the principal's utility under the ex-post bargaining contract. Let the bargaining power of the principal and the manager be $1-\beta$ and β. Then the second period bargaining becomes:

$$\max_{w_2^H, w_2^L} \left\{ \sum_{j=H,L} P^j(I_x)(x^j - w_2^j) - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z - \underline{u}\right] \right\}^{1-\beta}$$
$$\times \left[\sum_{j=H,L} P^j(I_x) w_2^j - \underline{u} \right]^{\beta}. \tag{B1}$$

Applying the generalized Nash bargaining solution to (B1), I obtain

$$V_2^p(I_x, I_y) = (1-\beta) \left\{ \sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z\right] \right\}$$
$$+ x^L - (1-\zeta y^L)z - \underline{u}, \tag{B2}$$

$$V_2^m(I_x, I_y) = \beta \left\{ \sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z\right] \right\} + \underline{u}. \tag{B3}$$

Repeating a procedure similar to that used in the proof of Proposition 1, I can also show that the first-order conditions with respect to I_x and I_y under the ex-post bargaining contract are as follows:

$$\frac{dD_x(I_x)}{dI_x} = \beta \frac{dP^H(I_x)}{dI_x}(x^H - x^L), \tag{B4}$$

$$\frac{dD_y(I_y)}{dI_y} = \beta\zeta \frac{dQ^H(I_y)}{dI_y}(y^H - y^L)z. \tag{B5}$$

Define $\widehat{I}_x^*$ and $\widehat{I}_y^*$ as I_x and I_y that satisfy (B4) and (B5). Let $\Psi(I_x, I_y)$ denote the principal's utility attained in period 1 under the ex-post bargaining contract.[24] Repeating a similar procedure used in the proof of Proposition 1, I can derive

$$\begin{aligned}\Psi(\widehat{I}_x^*, \widehat{I}_y^*) &= x^L - (1-\zeta y^L)z - D_x(\widehat{I}_x^*) - D_y(\widehat{I}_y^*) \\ &\quad + \left[\sum_{j=H,L} P^j(\widehat{I}_x^*)x^j - z + \sum_{i=H,L} Q^i(\widehat{I}_y^*)\zeta y^i z\right] - 2\underline{u}. \end{aligned} \tag{B6}$$

Now, differentiating $\Psi(I_x, I_y)$ with respect to β and evaluating it at $(I_x, I_y) = (\widehat{I}_x^*, \widehat{I}_y^*)$ yields

$$\begin{aligned}\frac{\partial \Psi(\widehat{I}_x^*, \widehat{I}_y^*)}{\partial \beta} &= \left[-\frac{dD_x(\widehat{I}_x^*)}{d\widehat{I}_x^*} + \frac{dP^H(\widehat{I}_x^*)}{d\widehat{I}_x^*}(x^H - x^L)\right]\frac{\partial \widehat{I}_x^*}{\partial \beta} \\ &\quad + \left[-\frac{dD_y(\widehat{I}_y^*)}{d\widehat{I}_y^*} + \frac{dQ^H(\widehat{I}_y^*)}{d\widehat{I}_y^*}\zeta(y^H - y^L)z\right]\frac{\partial \widehat{I}_y^*}{\partial \beta}. \end{aligned} \tag{B7}$$

Given Assumptions 1.1–1.3 and repeating a similar procedure used in deriving (A9), it follows from (B4) and (B5) that $\frac{\partial \widehat{I}_x^*}{\partial \beta} > 0$ and $\frac{\partial \widehat{I}_y^*}{\partial \beta} > 0$. Accordingly, it is found from (B4), (B5), and (B7) with $x^H > x^L$ that $\frac{\partial \Psi}{\partial \beta} > 0$. Because the principal's utility in period 1 under the ex-ante commitment contract is independent of β, this implies that the ex-post bargaining contract is more likely to be preferred as β is larger. ∎

C. Limited liability constraints

I discuss below the role of limited liability constraint. I consider two types of constraint: (i) all wages are nonnegative,[25] and (ii) $w_1^L + w_2^i \geq 0, i = H, L$.

For the ex-post bargaining contract, I can set $w_2^H = w_2^L = V_2^m(I_x^*, I_y^*) \geq 0$, where I_x^* and I_y^* are the optimal investment levels chosen under the ex-post bargaining contract (see (A2a) and (A2b) in Appendix A). Then, it follows from (2) that the principal must set

$$w_1^L + w_2^i = D_x(I_x^*) + D_y(I_y^*) + 2\underline{u} \geq 0, \qquad i = H, L.$$

Thus, the limited liability constraint of type (ii) is always satisfied. Moreover, if

$$D_x(I_x^*) + D_y(I_y^*) - V_2^m(I_x^*, I_y^*) + 2\underline{u} > 0, \tag{C1}$$

[24] Even though the principal can set w_1^L to be arbitrarily negative in the absence of limited liability, she must then increase V_2^m to satisfy (A3). Hence, an increase in β does not always lead to an increase in the principal's utility $\Psi(I_x, I_y)$ under the ex-post bargaining contract even without limited liability constraints.

[25] This case can be interpreted as a minimum wage because the zero wage can be viewed as the minimum wage.

then w_1^L can be nonnegative, that is, (i) is satisfied.

For the ex-ante commitment contract, I can set $w_2^H = x^H - r$ and $w_2^L = x^L - r$, where r is the principal's utility in period two (see Appendix A). Then, it follows from (6) with $I_y^{**} = 0$ and Assumption 1.1 that the principal must set

$$r = w_1^L - D_x(I_x^{**}) + \sum_{j=H,L} P^j(I_x^{**})x^j - 2\underline{u}, \tag{C2}$$

where I_x^{**} is the optimal investment level chosen under the ex-ante commitment contract (see (A5)). Hence, using (C2) with $x^H > x^L$ and $w_2^j = x^j - r$, $j = 1, 2$, I obtain

$$w_1^L + w_2^H > w_1^L + w_2^L = x^L + D_x(I_x^{**}) - \sum_{j=H,L} P^j(I_x^{**})x^j + 2\underline{u}.$$

The right-hand side is positive for a sufficiently large $\underline{u}$, as I_x^{**} does not depend on $\underline{u}$. Thus, the limited liability constraint of type (ii) is not binding for a sufficiently large $\underline{u}$. Note that I can also find a sufficiently large $\underline{u}$ such that (i) is also satisfied. If I consider the case in which $\underline{u}$ is not sufficiently large, these limited liability constraints are binding under the ex-ante commitment contract. Thus, the principal's utility under the ex-ante commitment contract in the presence of these limited liability constraints is smaller than in their absence.

I now provide the following proposition and corollary that have the equivalent results with Proposition 1.2, 1.3, and Corollary 1 for the limited liability constraints of type (i) and (ii).

Proposition 4 *If a limited liability constraint is imposed, optimal contracts satisfy the following properties.*

1. Under the ex-ante commitment contract, w_2^H is larger than w_2^L. Under the ex-post bargaining contract, a fixed wage, i.e., $w_2^H = w_2^L$, can be offered.

2. For the limited liability constraint of type (ii), there exists a $\overline{\zeta} > 0$ such that the principal prefers an ex-ante commitment to an ex-post bargaining contract at the beginning of the first period for $\zeta \in [0, \bar{\zeta})$, and prefers an ex-post bargaining to an ex-ante commitment contract for $\zeta \in (\bar{\zeta}, \infty)$. Next, if condition (C1) is satisfied, the same result can be obtained for the limited liability constraint of type (i).

Corollary 3 *For the limited liability constraint of type (ii), there exists a $\overline{z} > 0$ such that*

the principal prefers an ex-ante commitment to an ex-post bargaining contract at the beginning of the first period for $z \in [0, \overline{z})$, and prefers an ex-post bargaining to an ex-ante commitment contract for $z \in (\overline{z}, \infty)$. Next, if condition (C1) is satisfied, the same result can be obtained for the limited liability constraint of type (i).

Proof of Proposition 4: I begin with the case of the limited liability constraint of type (ii). When $\zeta = 0$ so that $I_y^* = I_y^{**} = 0$, I can prove that the principal's utility is larger under the ex-ante commitment contract than under the ex-post bargaining contract. Indeed, set $w_2^L = \frac{1}{2}x^L > 0$, $w^H = \frac{1}{2}x^H > 0$, and

$$w_1^L = D_x(I_x^*) - \frac{1}{2}\sum_{j=H,L} P^j(I_x^*)x^j + 2\underline{u}.$$

I show that under the ex-ante commitment contract, the manager chooses I_x^* (see (A2a) in Appendix A), the manager's utility at the beginning of period 1 is equal to $2\overline{u}$, and the principal obtains the same utility as she does under the ex-post bargaining contract (see (A4), (A7), and (A8) in Appendix A). In fact, under the ex-ante commitment contract, the principal can make the wage difference, $w_2^H - w_2^L$, larger than $\frac{1}{2}(x^H - x^L)$ so that the level of I_x chosen by the manager becomes larger than I_x^*. The principal can also keep the manager's expected wage constant. Hence, her gain is larger under the ex-ante commitment contract than under the ex-post bargaining contract.

When $\zeta > 0$, Proposition 1.3 still holds with a small $\bar{\zeta}$. This is because if I consider the ex-post bargaining contract, the principal obtains the same gain as in the absence of the limited liability constraints. Alternatively, if I consider the ex-ante commitment contract, the principal's gain is smaller under the limited liability constraint.

For the limited liability constraint of type (i), if (C1) is satisfied, the same results are obtained using the same argument as that of the limited liability constraint of type (ii). ∎

Proof of Corollary 3: When $z > 0$, Corollary 1 still holds with a small $\overline{z}$. The same logic as in proof of Proposition 4 holds. ∎

D. Risk-averse agent

In this Appendix, I adopt the same model as in Section 3, except that the manager's utility

regarding his wage, w, is expressed as $U(w) = w^{1-\rho}$, where $0 \leq \rho < 1$, i.e., the case of constant relative risk aversion, and that the domain of w is the set of nonnegative real numbers, i.e., I impose the limited liability constraint of type (i).[26] I can show that the same results obtained in Section 3 hold for ρ close to zero, because all equilibrium values can be shown to be continuous functions of (ρ, ζ). Note that the risk-neutral case with the limited liability constraint of type (i) corresponds to the case of $\rho = 0$.

Proposition 5 *Suppose that*

$$D_x(I_x^*) + D_y(I_y^*) - V_2^m(I_x^*, I_y^*) + 2\underline{u} > 0,$$

where I_x^* and I_y^* are the optimal investment levels chosen under the ex-post bargaining contract when $\rho = 0$, and that $P^H(I_x) \in (0,1)$ for all I_x. Then, there exists a $\bar{\rho} \in (0,1)$ such that the following properties hold for all $\rho \in (0, \bar{\rho}]$:

1. Under the ex-ante commitment contract, w_2^H is larger than w_2^L, and under the ex-post bargaining contract, the fixed wage, i.e., $w_2^H = w_2^L$, is offered.

2. There exists a $\bar{\zeta} > 0$ such that the principal prefers an ex-ante commitment to an ex-post bargaining contract at the beginning of the first period for $\zeta \in [0, \bar{\zeta})$, and prefers an ex-post bargaining to an ex-ante commitment contract for $\zeta \in (\bar{\zeta}, \infty)$.

Corollary 4 *Under Proposition 5, there exists a $\bar{\rho} \in (0,1)$ such that the following properties hold for all $\rho \in (0, \bar{\rho}]$: There exists a $\bar{z} > 0$ such that the principal prefers an ex-ante commitment to an ex-post bargaining contract at the beginning of the first period for $z \in [0, \bar{z})$, and prefers an ex-post bargaining to an ex-ante commitment contract for $z \in (\bar{z}, \infty)$.*

Proof of Proposition 5:

An ex-post bargaining contract

The bargaining problem in period two is as follows: for a given (I_x, I_y),

$$\max_{w_2^H, w_2^L} \left\{ \sum_{j=H,L} P^j(I_x)(x^j - w_2^j) - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1 - \zeta y^L)z - \underline{u}\right] \right\}$$

[26] As I can derive similar results when the manager can save, I do not discuss savings for simplicity. Note that there is no need to consider savings for the risk-neutral manager if the saving interest rate is equal to zero.

$$\times \left\{ \sum_{j=H,L} P^j(I_x) U(w_2^j) - U(\underline{u}) \right\}.$$

Note that $w_2^H, w_2^L \geq 0$ is shown later. The first-order conditions with respect to w_2^H and w_2^L are then as follows:

$$\begin{aligned} & \sum_{j=H,L} P^j(I_x) U(w_2^j) - U(\underline{u}) \\ = \; & U'(w_2^i) \left\{ \sum_{j=H,L} P^j(I_x)(x^j - w_2^j) - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z - \underline{u} \right] \right\}, \end{aligned}$$

for $i = H, L$. This yields

$$w_2^H = w_2^L.$$

That is, a fixed wage is offered. On the other hand, it follows from the above first-order conditions with $U'(w) = (1-\rho)w^{-\rho}$ that

$$\begin{aligned} w_2^\circ &= w_2^H = w_2^L \\ &= \frac{1-\rho}{2-\rho} \left\{ \sum_{j=H,L} P^j(I_x) x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z \right] \right\} + \frac{\underline{u}\left[\underline{u}^{-\rho} \left(w_2^\circ \right)^\rho + 1 - \rho \right]}{2-\rho} \\ &\geq 0. \end{aligned} \tag{D1}$$

Thus, the utility for the manager in the second period, denoted by $V_2^m(I_x, I_y, \rho, \zeta)$, is equal to $(w_2^\circ)^{1-\rho}$. Note that $V_2^m(I_x, I_y, 0, \zeta)$ is equal to the utility for the risk-neutral manager in the second period obtained in Section 3. The utility for the principal is then obtained as follows:

$$\begin{aligned} & V_2^p(I_x, I_y, \rho, \zeta) \\ = \; & \frac{1}{2-\rho} \left\{ \sum_{j=H,L} P^j(I_x) x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - \left[x^L - (1-\zeta y^L)z \right] \right\} + \underline{u} \\ & + x^L - (1-\zeta y^L)z - \underline{u} - \frac{\underline{u}\left[\underline{u}^{-\rho} \left(w_2^\circ \right)^\rho + 1 - \rho \right]}{2-\rho}. \end{aligned}$$

In the first period, the agent chooses I_x and I_y so as to satisfy the incentive compatibility

constraint:

$$\max_{I_x,I_y}\ \left(w_1^L\right)^{1-\rho} - D_x(I_x) - D_y(I_y)$$
$$+\left\{\frac{1-\rho}{2-\rho}\left[\sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - (x^L - (1-\zeta y^L)z)\right]\right.$$
$$\left.+\ \frac{\underline{u}\left[\underline{u}^{-\rho}\left(w_2^\circ\right)^\rho + 1 - \rho\right]}{2-\rho}\right\}^{1-\rho}. \tag{D2}$$

Given (D1), the first-order conditions yield

$$\frac{dD_x(I_x)}{dI_x} = \left[\frac{(1-\rho)^2}{2-\rho}\frac{dP^H(I_x)}{dI_x}(x^H - x^L) + \frac{1-\rho}{2-\rho}\underline{u}^{1-\rho}\left(w_2^\circ\right)^{\rho-1}\rho\frac{\partial w_2^\circ}{\partial I_x}\right]$$
$$\times\left\{\frac{1-\rho}{2-\rho}\left[\sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - (x^L - (1-\zeta y^L)z)\right]\right.$$
$$\left.+\frac{\underline{u}\left[\underline{u}^{-\rho}\left(w_2^\circ\right)^\rho + 1 - \rho\right]}{2-\rho}\right\}^{-\rho}, \tag{D3a}$$

and

$$\frac{dD_y(I_y)}{dI_y} = \left[\frac{\zeta(1-\rho)^2}{2-\rho}\frac{dQ^H(I_y)}{dI_y}(y^H - y^L)z + \frac{1-\rho}{2-\rho}\underline{u}^{1-\rho}\left(w_2^\circ\right)^{\rho-1}\rho\frac{\partial w_2^\circ}{\partial I_y}\right]$$
$$\times\left\{\frac{1-\rho}{2-\rho}\left[\sum_{j=H,L} P^j(I_x)x^j - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z - (x^L - (1-\zeta y^L)z)\right]\right.$$
$$\left.+\frac{\underline{u}\left[\underline{u}^{-\rho}\left(w_2^\circ\right)^\rho + 1 - \rho\right]}{2-\rho}\right\}^{-\rho}, \tag{D3b}$$

where

$$\frac{\partial w_2^\circ}{\partial I_x} = \frac{1-\rho}{2-\rho}\frac{\frac{dP^H(I_x)}{dI_x}(x^H - x^L)}{1 - \frac{1}{2-\rho}\underline{u}^{1-\rho}\left(w_2^\circ\right)^{\rho-1}\rho}, \tag{D4a}$$

$$\frac{\partial w_2^\circ}{\partial I_y} = \frac{1-\rho}{2-\rho}\frac{\frac{dQ^H(I_y)}{dI_y}\zeta(y^H - y^L)z}{1 - \frac{1}{2-\rho}\underline{u}^{1-\rho}\left(w_2^\circ\right)^{\rho-1}\rho}. \tag{D4b}$$

Note that by Assumption 1, the solutions of the above equations, denoted by $I_x^*(\rho,\zeta)$ and $I_y^*(\rho,\zeta)$, are continuous functions of (ρ,ζ) if ρ is sufficiently close to zero. On the other hand, by the individual rationality constraint, the principal must set

$$\left(w_1^L\right)^{1-\rho} = D_x(I_x^*(\rho,\zeta)) + D_y(I_y^*(\rho,\zeta)) - V_2^m(I_x^*(\rho,\zeta), I_y^*(\rho,\zeta)) + 2\left(\underline{u}\right)^{1-\rho}. \tag{D5}$$

Then, the principal's value,

$$x^L - w_1^L - (1-\zeta y^L)z + V_2^p(I_x^*(\rho,\zeta), I_y^*(\rho,\zeta)), \tag{D6}$$

is a continuous function of (ρ,ζ), because I_x^* and I_y^* are continuous functions of (ρ,ζ). Note that w_1^L is positive for ρ sufficiently close to 0, because $I_x^*(0,\zeta)$ and $I_y^*(0,\zeta)$ are the investments for a risk-neutral manager and wages are then positive. That is, the limited liability constraint is satisfied.

An ex-ante commitment contract

The principal's problem is as follows:

$$\max_{w_1^L\geq 0, I_x, I_y, w_2^H\geq 0, w_2^L\geq 0} x^L - w_1^L - (1-\zeta y^L)z + \sum_{j=H,L} P^j(I_x)(x^j - w_2^j) - z + \sum_{i=H,L} Q^i(I_y)\zeta y^i z, \tag{D7}$$

$$\text{s.t.} \quad (w_1^L)^{1-\rho} - D_x(I_x) - D_y(I_y) + \sum_{j=H,L} P^j(I_x)(w_2^j)^{1-\rho} \geq 2\,(\underline{u})^{1-\rho}\,, \tag{D8}$$

$$\begin{aligned}&(w_1^L)^{1-\rho} - D_x(I_x) - D_y(I_y) + \sum_{j=H,L} P^j(I_x)(w_2^j)^{1-\rho}\\ &\geq (w_1^L)^{1-\rho} - D_x(I_x') - D_y(I_y') + \sum_{j=H,L} P^j(I_x')(w_2^j)^{1-\rho}, {}^\forall I_x', I_y'.\end{aligned} \tag{D9}$$

As the principal has no incentive scheme for I_y, it has to be zero, that is, $I_y = 0$.

Below, I follow Berge's Maximum Theorem (see Hildenbrand, 1974) and Adachi-Sato and Kamiya (2013) to show that the value of the above problem is a continuous function of ρ. Let $B = D_x(I_x^{**}+1) + 2\,(\underline{u})^{1-\rho} + 1$, where I_x^{**} is the investment level that maximizes the expected total net payoff generated by the firm without considering social costs in the case of a risk-neutral manager. Then, I can restrict the domain of investments and wages in the compact set $\Omega = \{(I_x, w_1^L, w_2^H, w_2^L) \mid 0 \leq I_x \leq I_x^{**}+1,\ 0 \leq (w_1^L)^{1-\rho}, (w_2^H)^{1-\rho}, (w_2^L)^{1-\rho} \leq B\}$.

I now prove that the feasible set in the above problem is a continuous correspondence of ρ. Let $\Gamma(\rho)$ be the feasible set of the principal's problem, i.e., the set of $(I_x, w_1^L, w_2^H, w_2^L)$ satisfying (D8) and (D9). Let $\Pi(\rho) = \Gamma(\rho) \cap \Omega$. The upper hemicontinuity of $\Pi(\rho)$ follows from the continuity of the functions in the constraints. The lower hemicontinuity of $\Pi(\rho)$ can be obtained as follows. Note that by the strict concavity of P^H and the strict convexity of D_x, the optimal I_x in (D9) is a continuous function of (ρ, w_2^H, w_2^L), denoted by $I_x(\rho, w_2^H, w_2^L)$. For $\hat{\rho} \in [0,1)$,

let $(\widehat{I}_x, \hat{w}_1^L, \hat{w}_2^H, \hat{w}_2^L) \in \Pi(\hat{\rho})$ and let $\rho^k \in [0,1), k = 1, 2 \ldots$, be a sequence converging to $\hat{\rho}$. If $(\hat{w}_1^L)^{1-\rho}, \delta(\hat{w}_2^H)^{1-\rho}$, and $\delta(\hat{w}_2^L)^{1-\rho}$ are larger than 0 and smaller than B, it is easy to find a subsequence $(w_1^{Lk}, w_2^{Hk}, w_2^{Lk}), k = 1, 2 \ldots$, satisfying (D8) with $\rho = \rho^k$ and $I_x = I_x(\rho^k, w_2^{Hk}, w_2^{Lk})$, and converging to $(\hat{w}_1^L, \hat{w}_2^H, \hat{w}_2^L)$. Suppose that some of $(\hat{w}_1^L)^{1-\rho}$,$(\hat{w}_2^H)^{1-\rho}$, and $(\hat{w}_2^L)^{1-\rho}$ are equal to 0 or to B. If all such wages are equal to zero, then (D8) is not satisfied because $\underline{u} > 0$. Thus, some of these wages must be positive. If at least one such positive one is less than B, it is easy to find a subsequence $(w_1^{Lk}, w_2^{Hk}, w_2^{Lk}), k = 1, 2 \ldots$, satisfying (D8) with $\rho = \rho^k$ and $I_x = I_x(\rho^k, w_2^{Hk}, w_2^{Lk})$, and converging to $(\hat{w}_1^L, \hat{w}_2^H, \hat{w}_2^L)$. If all are equal to B, then (D8) is satisfied with strict inequality, and thus it is possible to find $(w_1^{Lk}, w_2^{Hk}, w_2^{Lk})$ satisfying (D8) with $\rho = \rho^k$ and $I_x = I_x(\rho^k, w_2^{Hk}, w_2^{Lk})$, and converging to $(\hat{w}_1^L, \hat{w}_2^H, \hat{w}_2^L)$. It is clear that $I_x(\rho^k, w_2^{Hk}, w_2^{Lk})$ converges to $\widehat{I}_x$. Then, Π is a lower hemicontinuous correspondence. Consequently, Π is a continuous correspondence of ρ.

Together with the continuity of the objective function, the continuity of the maximum value of problem (D7) in ρ follows from Berge's Maximum Theorem. Moreover, because $I_y = 0$ always holds, the maximum value for the principal under the ex-ante commitment contract is a continuous function of (ρ, ζ).

A comparison of two types of contract

As shown above, the values for the principal under the ex-post bargaining contract and the ex-ante commitment contract are continuous functions of (ρ, ζ), and coincide with those in the case of a risk-neutral manager at $\rho = 0$. Therefore, for $^{\exists}\bar{\rho} \in (0,1)$, I verify that the same property as in the case of a risk-neutral manager holds for $^{\forall}\rho \in (0, \bar{\rho}]$. ∎

Proof of Corollary 4:

Applying the procedure used in "A comparison of two types of contract" analyzed above, the statement of this corollary is verified. ∎

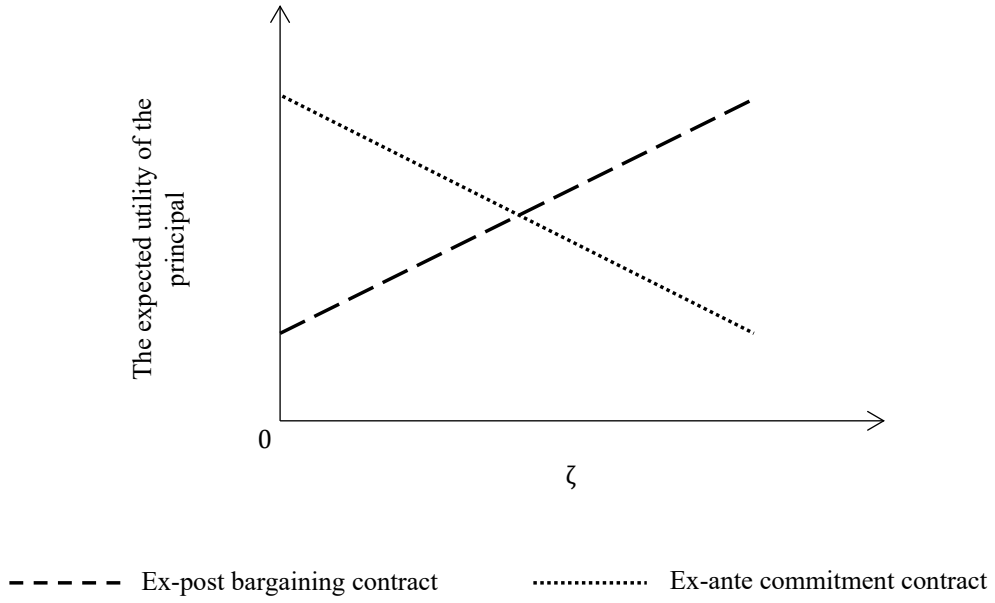

Fig.1. The choice of the contract **Notes.** The graph shows that the intersection of ex-post bargaining contract and ex-ante commitment contract is the threat point in which either contract is indifferent to the principal. To the right of the threat point, the principal is likely to choose the ex-post bargaining contract. To the left of the threat point, the principal is more likely to choose the ex-ante commitment contract.