경제학석사 학위논문

An Agency Perspective on Stakeholderism

대리인 이론으로 바라본 이해관계자 중심주의

2022년 8월

서울대학교 대학원

경제학부 경제학 전공

김 장 원

An Agency Perspective on Stakeholderism

지도 교수 Jihong Lee

이 논문을 경제학석사 학위논문으로 제출함

2022년　7월

서울대학교 대학원

경제학부 경제학 전공

김 장 원

김장원의 경제학석사 학위논문을 인준함

2022년　7월

위 원 장 Dmitry Shapiro

부위원장 Jihong Lee

위　　원 Yves Gueron

Abstract

An Agency Perspective on Stakeholderism

Jang Won Kim
Department of Economics
The Graduate School
Seoul National University

Stakeholderism is the idea that firms should maximize the total welfare of stakeholders, instead of pursuing only profit. Using a multitask agency model, we examine whether stakeholderism can fully address the externalities generated by firm activity. We demonstrate that a greater emphasis on stakeholder interests not only reduces profits, but could also discourage stakeholder−friendly activities, if it leads to less informative signals. Moreover, if some outputs of stakeholder−friendly activities are unmeasurable, stakeholderism may provide no advantage over shareholder−value maximization. These arguments characterize shareholder−value maximization as a second−best solution, in determining the appropriate objective function of the firm.

Keyword : stakeholderism, agency theory, multitask, corporate governance, corporate finance
Student Number : 2020−25470

Table of Contents

Chapter 1. Introduction

What should be the objective function of the firm? Conventionally, firms are conceptualized as profit-maximizing entities. Of course, imperfections such as agency costs prevent actual firms from achieving this goal, yet profit maximization is still held as the benchmark. Furthermore, that firms *should* maximize profit is often invoked as a normative ideal. For example, Milton Friedman argued that firms should pursue the interests of shareholders, and that other welfare concerns, such as the need for charity, should not be pursued at the expense of profit.[①]

Yet, by focusing solely on profit, firms often impose significant externalities on third parties, such as the community or the environment. These parties, whose welfare is affected by firm activity, are referred to as "stakeholders" (which includes shareholders). For example, an oil company might be lured into drilling excessively, damaging the environment. Furthermore, conventional remedies against externalities, such as Pigouvian taxes or delineating property rights, may be unavailable if the regulator has insufficient information.[②] For example, an outside regulator may not be able to observe the externality or to identify the affected parties.

An alternative method for internalizing externalities is for the firm to adopt "stakeholderism." Stakeholderism is the idea that firms should maximize the total welfare of stakeholders, instead of simply pursuing shareholder value. It is distinct from more traditional methods of controlling externalities in that it attempts to modify the preferences of (or the objective function pursued by) the firm itself. The unorthodox objective function of maximizing total welfare may be implemented by, for example, a state-run pension fund exerting influence on the governance of a firm. In fact, stakeholderism as a

---

① For more detail on Milton Friedman's argument, see Hart and Zingales (2017).

② For an introduction to the externality problem, see Mas-Colell, Whinston, and Green (1995).

solution to addressing externalities is garnering more attention in recent days, along with the growing popularity of ESG investment.

A priori, stakeholderism seems promising – but does it hold up to its promise? In this paper, we ask whether stakeholderism achieves its goal of advancing the collective welfare of stakeholders. Using an agency model with multiple tasks, we demonstrate that if there are insufficient signals to motivate prosocial activities, embracing more stakeholderism may actually *negatively* affect the welfare of stakeholders. In this case, profit maximization may be second-best.

In order to study the impact of stakeholderism on firm activity and on welfare, we modify Holmström and Tirole (1997)'s model of credit rationing, so that the manager of a firm chooses his multidimensional effort from a continuum. In Chapter 3, we outline the model. There are two parameters of interest: α, which is the degree of stakeholderism adopted by the firm, and β, which indicates the informativeness of the signal available for the externality-generating activity. In Chapter 4, we characterize the optimal financing contract signed by the firm.

In Chapter 5, we examine what happens to profit and externality as the degree of stakeholderism α is increased. We find that surplus decreases with more stakeholderism. Also, if β depends on α with a sufficiently negative relation, then even externality may fall with more stakeholderism. These results suggest that when agency costs are a serious problem, stakeholderism may be suboptimal for stakeholder welfare. In Chapter 6, we consider a more general situation where prosocial effort may positively impact surplus. While some of the comparative static results carry over, we find in addition that the amount of credit rationing may also be affected.

In Chapter 7, we adopt the multitask framework of Holmström and Milgrom (1991). When there are many stakeholders, but not enough signals relative to the number of stakeholders, it may be optimal for a stakeholderist principal to provide muted incentives for prosocial activities in general. Thus, when unseen effort benefiting certain stakeholders may be crowded out by high-powered incentives on select prosocial activities, it may be second-best even for a stakeholderist principal to resort to conventional profit maximization. Chapter 8 concludes. Chapter 9, which we refer to as the appendix, contains proofs of propositions.

Chapter 2. Related Work

2.1. Stakeholderism

Tirole (2001) sets the tone for the debate on stakeholderism. He defines corporate governance as "the design of institutions that induce or force management to internalize the welfare of stakeholders."[③] He then explains the various obstacles society may encounter in implementing stakeholderism, including agency issues, which is the central theme of our paper. However, Tirole (2001) does not give a fully formal account of these problems. As such, our paper contributes in formalizing Tirole (2001)'s intuition, and in continuing the debate on stakeholderism.

Magill, Quinzii, and Rochet (2015) provide a set of conditions under which stakeholderism may be justified. They study a competitive equilibrium where externalities generated by large firms, and felt by stakeholders such as consumers and employees, lead to inefficiency. They show that if there is only one firm, a "stakeholder equilibrium," in which the firm maximizes the total welfare of stakeholders, is Pareto optimal. In order to implement stakeholderism, they propose a Coasian solution, in which employee and consumer rights are defined and utilized for incentive provision to the manager. In contrast, our paper concerns a different context where informational and incentive problems cannot be completely eliminated.

Like the stakeholderist camp, Hart and Zingales (2017) argue

[③] However, for a more traditional view of governance focusing on private parties instead of stakeholders, see Hermalin (2012).

that the appropriate objective function of the firm may not necessarily coincide with maximizing market value. However, their justification for this deviation is that shareholders are often prosocial. Therefore, their view still falls under the category of shareholder-value maximization, and is conceptually distinct from stakeholderism, which requires public companies to internalize externalities even if shareholders care only about profit.

Hart and Zingales (2017) then discuss how to implement the prosocial shareholder value, via governance. The point is to prevent "amoral drift," whereby a company's decision strays away from the prosocial one because shareholders are apathetic towards the social cause. However, once a policy is determined via governance, there is no moral hazard in implementing it. The difference in our paper is that we take as given the governance structure and focus on agency issues. As for the governance structure, we assume that it is given in a simple form: the firm's policy is decided by a stakeholderist owner.[④] There is no concern of amoral drift (i.e., susceptibility to profit-enhancing hostile takeovers) because the stakeholderist owner does not exhibit an asymmetric attitude towards externalities vis-a-vis profit.

2.2. Multitask Principal-Agent Models

In Chapter 7, we draw insights from the multitask principal-agent literature, initiated by Holmström and Milgrom (1991).[⑤] The first insight is that providing high-powered incentives for a particular task may crowd out effort in other tasks. Therefore, if strong incentives cannot be provided for a certain subset of tasks (due to lack of measurability), then incentives for the rest of the tasks should be low-powered as well.

④ For a discussion of the motivations for pursuing corporate social responsibility, see Bénabou and Tirole (2010).
⑤ Dewatripont, Jewitt, and Tirole (2000) provide a general survey of the multitask literature.

One implication is that it may be desirable to separate tasks across agents. Equivalently, the literature, as well as this paper, warns against the dangers of combining tasks. However, Laux (2001) reaches a different conclusion. Like our paper, Laux (2001) examines a multitask principal–agent situation where the agent is risk–neutral and protected by limited liability. In his paper, it is shown that combining multiple tasks under a single manager saves costs, because doing so relaxes the limited liability constraint. Basically, when a manager is responsible for multiple projects, the principal can devise a contract that effectively punishes the manager for failure in certain projects (while respecting limited liability), by denying reward from successful projects. This contrasts with our setting, where an additional dimension of stakeholder externality does not relax limited liability, since the surplus to be divided is unaffected. This has to do with a feature of our model that incentives for the externality dimension can only be provided contingent on monetary success.

Finally, there is a parallel literature concerning implicit incentives (or career concerns) in a multitask setting (Dewatripont, Jewitt, & Tirole, 1999). Theirs is a setting where explicit contracts cannot be signed ex ante to motivate the manager. Instead, the market pays according to expected productivity, which motivates the manager to shape the market's expectation by exerting effort. As pointed out in Tirole (2001), when explicit incentives are muted (which in our setting is deliberate), implicit incentives become prominent. Dewatripont et al. (1999) show that under certain conditions, the larger the number of tasks, the lower is total effort. Although they reach the familiar conclusion that task exclusion may be desirable, they dispense with the assumption that tasks may vary in measurability; instead, a superfluous number of tasks may endogenously weaken the link between performance and the market's appraisal of talent.

Chapter 3. The Model

To study the impact of stakeholderism on firm activity and on welfare, we modify Holmström and Tirole (1997)' s model of credit rationing, so that the manager of a firm chooses his multi-dimensional effort from a continuum.

3.1. The Firm' s Project

A firm faces a project that requires I units of capital. The firm's initial capital endowment is $A < I$ so that it needs to raise at least $I - \mathrm{A}$ units of outside investment to undertake the project.

The nature of the project is multi-dimensional, involving profit and externality. Examples of externality include pollution and local employment. The manager of the firm chooses a pair (x, y), where $x \in [0,1]$ is the effort to produce profit and $y \in [0,1]$ is the effort to produce externality. The efforts are unobservable. The manager's private cost of efforts is given by a function $C(x, y)$. We sometimes refer to x as the "lucrative" activity and y as the "prosocial" activity.

In the profit dimension, there are two possible outcomes of the project: success (S_1) or failure (F_1). If the project succeeds, it generates revenue $R > 0$, which is observable. If the project fails, revenue is zero. The probability of S_1 is equal to x.

Similarly, there are two possible outcomes for the externality dimension: success (S_2) or failure (F_2). In case of success, an externality of T is generated, whereas zero externality is generated in the case of failure. The probability of S_2 is equal to y and independent of the profit distribution.

Note that if the firm's objective is to maximize profit, the manager will be instructed to put zero effort in y. Thus, the project will fail to generate any benefit other than monetary surplus. However, by encouraging y, the firm could have a positive impact on its stakeholders. Therefore, the firm's project, by denying stakeholders their well−being, is generating an externality.

The externality outcome is not observable. Instead, there is a signal $\sigma \in \{\sigma_G, \sigma_B\}$ which provides partial information about y. This signal is observable and verifiable. We assume that σ is correlated with success in the externality dimension such that

$$P(\sigma_G|S_2) = P(\sigma_B|F_2) = \frac{1}{2} + \frac{\beta}{2},$$

where $0 < \beta < 1$. Thus, β represents the informativeness of the signal. Also note that

$$q(y) \coloneqq P(\sigma_G|y) = \frac{1}{2} + \beta\left(y - \frac{1}{2}\right).$$

Remark 1 (Informativeness). A larger β represents a more informative signal σ, in Blackwell's sense (Blackwell and Girshick (1979)). That is, whenever $\beta_1 > \beta_2$, $\sigma(\beta_2)$ garbles $\sigma(\beta_1)$, where $\sigma(\beta)$ is understood as the information system associated with β. To see this, observe that for all $y \in [0,1]$,

$$\begin{bmatrix} q(y;\beta_2) \\ 1 - q(y;\beta_2) \end{bmatrix} = \begin{bmatrix} a & b \\ 1-a & 1-b \end{bmatrix} \begin{bmatrix} q(y;\beta_1) \\ 1 - q(y;\beta_1) \end{bmatrix},$$

where $a = (1 + \beta_2/\beta_1)/2$ and $b = (1 - \beta_2/\beta_1)/2$. Since $0 \leq \beta_2/\beta_1 \leq 1$, we have $0 \leq a, b \leq 1$, so $\begin{bmatrix} a & b \\ 1-a & 1-b \end{bmatrix}$ is indeed a stochastic matrix.[6]

Another approach to model informativeness can be found in the literature on repeated games with almost−perfect monitoring. Hörner and Olszewski (2006) define monitoring to be ϵ−perfect if for any player i and any action profile $a = (a_i, a_{-i})$, player i obtains signal $\sigma_i = a_{-i}$ with probability at least $1 - \epsilon$, where a player's set of signals is the set of action profiles of his opponents. However, we cannot take the same approach in our setup, since the size of the signal space (success or failure) is smaller than the action space $[0,1]$.

Remark 2 (Profit vs. Externality). We have assumed that the two

[6]Grossman and Hart (1983) also use Blackwell's criterion to compare informativeness in their principal-agent setup where action is chosen from a compact subset of Euclidean space with finitely many outcomes.

dimensions, profit and externality, of the project do not interact, in the sense that the two outcomes are statistically independent, and the probability of success in each dimension depends on effort in that dimension only. While they simplify the analysis, these modeling features are chosen above all to clearly separate our focus from the perspective that corporate activities promoting stakeholder interests actually have positive effects on profitability.

There are some potentially interesting generalizations, nonetheless. For example, contributing to local employment or more environmental friendliness may also benefit profit. We later consider these generalizations in Chapter 6.

3.2. Financing the Project

The firm must turn to an outside investor to fund its project. The capital market is competitive so that the investor has no bargaining power. The investor has an outside option which guarantees an interest rate of γ. External financing is facilitated by a contract.

We assume that all parties involved are protected by limited liability. In other words, under no circumstances can any party end up with a negative cash position. The inside capital not invested in the project, and the interest earnings resulting thereof, are private gains by the owners, and cannot be used as collateral.

There are four possible outcome combinations given the two dimensions. Given limited liability, if the project results in monetary failure, both parties receive 0, regardless of the outcome of the prosocial activity. If the project is monetarily successful, the contract stipulates the following returns, contingent on the outcome of the prosocial activity: if the externality signal is good (σ_G), profit is shared by $\left(R_f^S, R_u^S\right)$, and if the signal is bad (σ_B), profit is shared by $\left(R_f^F, R_u^F\right)$.

Then, given our assumption on bargaining power, the firm offers to the investor a contract of the following form

$$\{(x,y),(I_f,I_u),(R_f^S,R_u^S),(R_f^F,R_u^F)\},$$

where the subscript f indicates the "firm" and u indicates the "uninformed investor" and (I_f,I_u) is the pair of investments. We assume that the portion of inside capital which is not invested in the project, i.e. $A-I_f$, is immediately distributed back to the owners of the firm according to their shares, who then *privately* invest this amount in the outside option.

The parameters of the contract must satisfy the following conditions:

(i) $0 \le I_f \le A,\ I_u \ge 0,$ and $I_f + I_u = I.$

(ii) $0 \le R_f^S \le R,\ 0 \le R_u^S \le R,$ and $R_f^S + R_u^S = R.$

(iii) $0 \le R_f^F \le R,\ 0 \le R_u^F \le R,$ and $R_f^F + R_u^F = R.$

3.3. Ownership Structure

There are two types of owners of the firm: the stakeholderist owner and the profit-seeking owner. A typical example we have in mind is when the firm is owned together by a pension fund or a sovereign wealth fund trying to advance a social agenda, along with standard profit-maximizing shareholders.

The two owners' shares are exogenously given by $\tau \in (0,1)$ for the stakeholderist owner and $1-\tau$ for the profit-seeking owner. We assume that ownership influences only the cash flow rights, and do not discuss the relationship between ownership and its contribution to initial capital A. Also, the payments to both owners are assumed to be linear in expected surplus. It follows that the stakeholderist (or profit-seeking) owner is entitled to a proportion τ (or $1-\tau$) of the expected surplus, where the expected surplus is defined by

$$S := x\big(q(y)R_f^S + (1-q(y))R_f^F\big) + \gamma(A-I_f).$$

We assume that the profit-seeking owner manages the firm, and as such, bears the private costs of exerting effort. The manager's expected utility is

$$(1-\tau)S - C(x,y)$$
$$= (1-\tau)\,[x\{q(y)R_f^S + (1-q(y))\,R_f^F\} + \gamma(A-I_f)] - C(x,y).$$

The stakeholderist owner does not directly determine efforts (x,y) and takes into account the non-financial impact of the project. His expected utility is therefore a weighted sum of its share of the expected surplus and the expected externality, denoted by $T(y) \coloneqq yT$.

$$\alpha T(y) + (1-\alpha)\tau S$$
$$= \alpha y\,T + (1-\alpha)\,\tau[x\{q(y)R_f^S + (1-q(y))\,R_f^F\} + \gamma(A-I_f)],$$

where $\alpha \in [0,1]$ is the weight that the stakeholderist owner places on externality relative to its monetary gain.

3.4. Modeling Stakeholderism

To model stakeholderism, we assume that the financial contract is drafted by the stakeholderist owner. This is plausible if the stakeholderist owner is the controlling shareholder of the firm. The firm's objective function then coincides with that of the stakeholderist owner, and we can write the problem as a principal-agent problem, where the principal is the stakeholderist owner and the agent is the profit-seeking owner-manager.

Specifically, re-scaling T by $1/\tau$ and $C(x,y)$ by $1/(1-\tau)$, the principal designs a contract to maximize:

$$\begin{aligned} W(\alpha) &\coloneqq \alpha T(y) + (1-\alpha)\tau S \\ &= \alpha y\,T + (1-\alpha)[x\{q(y)R_f^S + (1-q(y))\,R_f^F\} + \gamma(A-I_f)], \end{aligned} \tag{1}$$

subject to the following constraints:

(i) $0 \le I_f \le A$, $I_u \ge 0$, and $I_f + I_u = I$.

(ii) $0 \le R_f^S \le R$, and $0 \le R_f^F \le R$.

(iii) The incentive compatibility (IC) constraint for the manager, which is

$$(x,y) \in \underset{\tilde{x},\tilde{y}\in[0,1]}{\operatorname{argmax}}\, \tilde{x}\,[q(\tilde{y})R_f^S + (1-q(\tilde{y}))\,R_f^F] - C(\tilde{x},\tilde{y}). \tag{2}$$

The FOCs are

$$q(\tilde{y})R_f^S + (1 - q(\tilde{y}))R_f^F = C_x(x, y), \tag{3}$$

and

$$x\left[\left(R_f^S - R_f^F\right)q'(y)\right] = C_y(x, y). \tag{4}$$

(iv) The individual rationality (IR) constraint for the manager, which is

$$x(q(y)R_f^S + (1 - q(y))R_f^F) + \gamma(A - I_f) - C(x, y) \geq 0. \tag{5}$$

(v) The breakeven constraint for the investor, which is

$$x\left[R - (q(y)R_f^S + (1 - q(y))\, R_f^F)\right] \geq \gamma I_u. \tag{6}$$

Note that the left-hand side is the pledgeable expected income. Also, note that (6) must hold with equality at the optimum:

$$x\left[R - (q(y)R_f^S + (1 - q(y))\, R_f^F)\right] = \gamma I_u. \tag{7}$$

Otherwise, I_u can be made larger, which leads to higher surplus.

Note that $\alpha \in [0,1]$ can be thought of as the degree of stakeholderism. When $\alpha = 0$, the objective function is τS, so the firm maximizes profit. When $\alpha = 1$, the objective function is T, so the firm is solely focused on the externality.

Remark 3 (Validity). The first-order approach is valid as long as

$$x[q(y)R_f^S + (1 - q(y))\, R_f^F] - C(x, y)$$

is concave in (x, y). Thus, we require the Hessian

$$H = \begin{bmatrix} -C_{xx}(x,y) & \beta\left(R_f^S - R_f^F\right) - C_{xy}(x,y) \\ \beta\left(R_f^S - R_f^F\right) - C_{xy}(x,y) & -C_{yy}(x,y) \end{bmatrix}$$

to be negative semi-definite, which is equivalent to

$$C_{xx} \geq 0, \text{ and } C_{xx}C_{yy} \geq \left(\beta\left(R_f^S - R_f^F\right) - C_{xy}\right)^2.$$

Since $0 \leq \beta\left(R_f^S - R_f^F\right) \leq R$,[⑦] the above is satisfied under the following

⑦ It is without loss of generality to assume that $R_f^S \geq R_f^F$. Otherwise, $y = 0$ will be induced, but in that case, the principal can do just as good by adjusting R_f^S and R_f^F to be the same. See the argument in Section 9.3, case (ii).

assumption, which we shall maintain throughout the rest of the paper.[8]

Assumption 1. For all $(x,y) \in [0,1]^2$,

$C_{xx} > 0,\ C_{yy} > 0,\ C_{xy} > 0,$ and $C_{xx}C_{yy} \geq max\{R^2, (R - C_{xy})^2\}.$

Note that $C_{xy} > 0$ means that the lucrative activity x and the prosocial activity y are indeed linked, via the agent's cost function. Specifically, a higher level of effort in one of the activities makes it more costly to exert effort in the other.

Chapter 4. Optimal Contract

To characterize the optimal contract, let us assume that marginal costs are sufficiently high.

Assumption 2. $C_x(0,1) > R$ and $C_x(1,0) > R - \gamma(I - A).$

Our first result characterizes the optimal financing contract as one that maximizes a simple objective function, a weighted sum of the two activities x and y, over a specific region derived from the various constraints of the problem. The region can be interpreted as a Pareto frontier in x and y. In other words, the stakeholderist owner' s problem is one of choosing a point on the two−dimensional space with Pareto weights α and $1 - \alpha$. This will enable us to analyze the effects of a change in α.

[8] We only need weak inequalities on the second−derivatives here. The strict inequalities in Assumption 1 help us characterize optimal contract.

Proposition 1. Under Assumptions 1 and 2, the optimal contract solves

$$\max_{(x,y)\in\tilde{X}} \alpha y T + (1-\alpha) x R$$

where $\tilde{X} = K \cap L(\beta) \cap P$ such that

$$K \coloneqq \{(x,y) \in [0,1]^2 : x\left(R - C_x(x,y)\right) = \gamma(I - A)\},$$

$$L(\beta) \coloneqq \{ (x,y) \in [0,1]^2 :$$

$$0 \le C_x(x,y) - \frac{C_y(x,y)}{x}\left(\left(y - \frac{1}{2}\right) + \frac{1}{2\beta}\right) \le R,$$

$$0 \le C_x(x,y) - \frac{C_y(x,y)}{x}\left(\left(y - \frac{1}{2}\right) - \frac{1}{2\beta}\right) \le R\}.$$

$$P \coloneqq \{(x,y) \in [0,1]^2 : x C_x(x,y) - C(x,y) \ge 0\}.$$

Proof. See the appendix. ■

Chapter 5. The Effects of Stakeholderism

We are interested in the effects of stakeholderism, which can be studied by exploring the response of the optimal contract and corresponding values to a change in the parameter α.

We consider in turn the effects on profit and externality.

5.1. Effects on Profit

Let us begin by examining how α influences surplus. First, there is credit rationing, as in Holmström and Tirole (1997), but this is independent of stakeholderism in our model.

Proposition 2. α does not affect the level of credit rationing.

Proof. For financing to take place, it is necessary that the manager's IC constraint and the investor's breakeven constraint hold. By viewing (x, y) as a function of $\left(R_f^S, R_f^F\right)$ (implicitly defined by (2)), the two constraints reduce to $A \geq \tilde{A}(\gamma)$, where $\tilde{A}(\gamma)$ is such that

$$\max_{R_f^S, R_f^F \in [0,R]} x\left[R - \left(q(y)R_f^S + \left(1 - q(y)\right)R_f^F\right)\right] = \gamma\left(I - \tilde{A}(\gamma)\right). \quad (8)$$

The left-hand side of (8) is the *maximum* pledgeable expected income. Notice indeed that this does not depend on α, which follows from the independence of the manager's IC constraint from α. ■

Even though stakeholderism does not alter the chance of the project taking place, it may still affect the resulting surplus. Let

$$S(\alpha) \coloneqq x^*\left(q(y^*)R_f^{S*} + \left(1 - q(y^*)\right)R_f^{F*}\right) + \gamma\left(A - I_f^*\right)$$

be the firm's expected surplus under α, where $\left(x^*, y^*, I_f^*, R_f^{S*}, R_f^{F*}\right)$ denote the optimal values. Clearly, $S(0) \geq S(\alpha)$ for all $\alpha \in [0,1]$. This means, unsurprisingly, that stakeholderism cannot be associated with a higher surplus than standard profit maximization.

Under certain conditions, we can claim a stronger proposition, that $S'(\alpha) < 0$.

Assumption 3. (i) $C_{xxx} = C_{xxy} = C_{xyy} = 0$ for all (x, y).⑨
(ii) For a positive measure of α in $(0,1)$, the optimal choice of (x, y) is an interior point of $L(\beta) \cap P$.

Proposition 3. Under Assumptions 1–3, $S'(\alpha) < 0$ for a positive measure of α in $(0,1)$.

Proof. Note that

$$\left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}\right) = \left((1 - \alpha)R, \alpha T\right),$$

so the effort level chosen will be a point on the graph of K, where

⑨ Consider, for instance, $C(x, y) = (c_1 x + c_2 y)^2$, where $c_1, c_2 > 0$.

either (i) the slope equals $-(1-\alpha)R/\alpha T$ or (ii) one or more of the inequalities in $L(\beta)$ and P hold with equality. We focus on the first case, which occurs for a positive measure of α in $(0,1)$ by the second part of Assumption 3.

Since $-(1-\alpha)R/\alpha T$ is increasing in α, $x'(\alpha) < 0$ as long as the function $\phi(x)$ is concave in x, where $y = \phi(x)$ is implicitly defined by

$$x[R - C_x(x,y)] = \gamma(I - A).$$

We show that the assumptions guarantee the concavity of ϕ. Differentiating ϕ by x yields

$$C_{xx} + C_{xy}\phi' = \frac{\gamma(I - A)}{x^2}.$$

Differentiating again by x, and using the first part of Assumption 3, we obtain

$$\phi''(x) = \frac{-2\gamma(I - A)}{C_{xy}x^3}.$$

Since $I - A > 0$ and, by Assumption 1, $C_{xy} > 0$, we have $\phi''(x) < 0$. Therefore, $x'(\alpha) < 0$ and $S'(\alpha) < 0$ for a positive measure of α in $(0,1)$. ■

In Proposition 3, profit falls as the stakeholderist owner cares more about externality. The result is intuitively obvious; as stakeholderism intensifies, the optimal contract will induce higher effort towards the prosocial activity while reducing incentives for the lucrative activity. To maintain x, the manager must be compensated with a higher return, but this is unsustainable because it reduces the investor's share of the surplus. Given this conflict of interest between the manager and the outside investor, a stakeholderist owner who cares more about stakeholders opts for a smaller x, even if it harms overall profit.

Though unsurprising, this result gives us a reason to suspect the feasibility of stakeholderism in the real world. By reducing profitability, it will likely face opposition from profit-driven shareholders and may also squeeze the firm's future financing opportunities.

We now turn to the issue of whether stakeholderism will improve externality if not profit.

5.2. Effects on Externality

What happens to externality as α increases? Is it possible that a greater emphasis on stakeholders actually hurt their interests? One possible source of the paradoxical outcome can be found in the informativeness of signals.

Specifically, let us assume that β falls with α. That is, more extensive stakeholderism forces the signal on externality to become less informative. There are several reasons to think why this might be the case in reality.

First, it may be that, with higher α, the scope of the stakeholder interests encompassed in activity y widens and this makes less informative the aggregate signal which simply indicates an "overall" success or failure. Second, the manager may choose from multiple activities that unanimously generate the same outcome, but with different informativeness of their signals. As advocated by Bebchuk and Tallarita (2020), stakeholderism can increase managerial insulation, and less informative activities may better suit the manager's private interests.[10]

We next identify conditions under which more intense stakeholderism worsens externality due to this informativeness channel. Specifically, we derive a set of sufficient conditions on the cost structure such that, keeping α fixed, y, and hence expected externality $T(y)$, fall as the corresponding signal becomes less informative. Since an increase in α leads to higher externality for a given value of β, if the tradeoff between α and β is large enough then the net effect of higher α is going to be lower y.

---

[10] One possible example is "greenwashing", whereby the firm fosters a false perception that its products are more environment-friendly than they actually are.

Proposition 4. Assume that when $\alpha = 0$, inducing $y > 0$ is optimal. Also, in addition to Assumptions 1 and 2, suppose $C(x, y) = (c_1 x + c_2 y)^2$ and that $4c_1 + c_2 \leq 2\sqrt{R}$. Then, keeping α fixed, y falls as β falls. Furthermore, if β depends on α, and if $d\beta/d\alpha$ is sufficiently negative, y is decreasing in α.

Proof. See the appendix. ■

Remark 4 (When $\alpha = 0$). One concern that may be raised is that when $\alpha = 0$, $y = 0$ is induced, so that there is no room for y to decrease further as α increases. For this, we must restrict attention to the case where $\alpha = 0$ is followed by $y > 0$. This is possible because of the constraint $L_2(\beta)$.

Remark 5 (Governance). As already mentioned, we mainly focus on agency issues in this paper; nevertheless, the issue of governance may also be important. In reality, implementing stakeholderism requires that those who care about stakeholder welfare (whether it be the stakeholders themselves, sympathetic shareholders, or some government official) be in charge. For this, we require that a stakeholderist party, such as a pension fund, has made ex ante investments to become the owner. However, since pension funds only have finite capital, it may not be a sustainable way of implementing stakeholderism at a nation-wide level.

A more sustainable method of implementing stakeholder society would be to modify the institution of the public company itself. One example is the German example of codetermination, which mandates labor representation on corporate boards. However, even then, the governance structure and managerial moral hazard may interact. This is an important question and is related to why β may decrease with higher α. Aghion and Tirole (1997)'s model of formal and real authority may be useful. One possibility is that more dispersed formal authority (i.e., board includes various stakeholders) leads to lower externalities, because it becomes optimal to delegate more real authority to the agent, aggravating moral hazard.

Chapter 6. Profitable Prosocial Activity

So far, we have maintained the assumption that more effort towards the prosocial activity (i.e., higher y) does not contribute to profit. An unsatisfactory aspect of this restriction is that the level of credit rationing is determined independently of the degree of stakeholderism α or the informativeness of the signal β. Indeed, the next proposition shows that under the assumption of orthogonality between profit and externality, β does not affect the severity of credit rationing.

Proposition 5. Suppose that $\max_{x \in [0,1]} x[R - C_x(x,0)]$ is achieved uniquely at $x^* \in (0,1)$. Also assume that $C_x(x^*,0) \in [0,R]$, and that $x^* C_x(x^*,0) - C(x^*,0) \geq 0$. Then, β does not affect the level of credit rationing.

Proof. See the appendix. ■

However, credit rationing may depend on β, if we generalize the manner in which effort affects the probability of success in each dimension. To this end, for the rest of this chapter, we make the following generalized assumption: When effort (x,y) is exerted, the probability of success in the externality dimension is still y. However, the probability of success in the profit dimension is now given by

$$p(x,y;\theta) \equiv (1-\theta)x + \theta y. \tag{9}$$

The parameter $\theta \in [0,1]$ indicates how much effort y affects profit *in addition to* externalities.

Proposition 6. There exists a $\theta_0 \in (0,1)$ such that whenever $\theta \geq \theta_0$, the maximum pledgeable income increases with β (equivalently, credit rationing decreases with β).

Proof. See the appendix. ■

However, note that when θ is small so that the FOC for the IC constraint holds, then the pledgeable income can be expressed as $\{(1-\theta)x+\theta y\}[R-C_x(x,y)/(1-\theta)]$, and the dependence of credit rationing on β is lost. Thus β affects credit rationing insofar as maximizing pledgeable income does not involve a positive x.

Now we look at the principal's problem, who puts weight α on stakeholderism, and whose project has informativeness β and type θ. We fix a $\theta \in [0,1]$, and examine how surplus and externality varies with α and β.

The principal's problem is to find $R_f^S, R_f^F \in [0, R]$ and (x, y) in order to solve the maximization program (after rearranging):

$$max\, \alpha\, yT + (1-\alpha)[p(x,y)R - \gamma(I-A)]$$
$$= [\alpha T + (1-\alpha)\theta R]y + (1-\alpha)(1-\theta)Rx - constant$$

subject to the IC constraint

$$(x,y) \in \underset{\tilde{x},\tilde{y}\in[0,1]}{\operatorname{argmax}}\; p(\tilde{x},\tilde{y})[q(\tilde{y})R_f^S + (1-q(\tilde{y}))R_f^F] - C(\tilde{x},\tilde{y}),$$

the IR constraint

$$p(x,y)\left[q(y)R_f^S + (1-q(y))R_f^F\right] - C(x,y) \geq 0,$$

and the capital constraint

$$p(x,y)\left[R - q(y)R_f^S - (1-q(y))R_f^F\right] \geq \gamma(I-A).$$

The resulting expected surplus is then

$$p(x,y)R = [(1-\theta)x + \theta y]R,$$

and the resulting expected externality is yT.

The next proposition summarizes the effect of varying α and β, under our generalized production technology involving θ.

Proposition 7. Assume that the y induced is greater than $1/2$, and that the IR condition holds. Also, assume that $C^*(p,y) \equiv C(p(x,y),y)$ satisfies (in place of $C(x,y)$) Assumptions 1 – 3, as well as the assumption made in Proposition 4.

When $\theta < 1$, an increase in α reduces surplus while increasing externality, and an increase in β increases externality. When $\theta = 1$, α does not affect either surplus or externality, while an increase in β increases both surplus and externality.

Proof. See the appendix. ■

Chapter 7. Multitask Analysis

There is a more fundamental issue surrounding stakeholderism. Typically, a firm has many stakeholders, and this means that the firm may have to pursue separate goals to promote the different interests of different stakeholders. But then, the following question arises. If the manager's diverse activities are all unobservable, and if there are many goals to meet, can we ever have enough performance signals with which to discipline the multidimensional moral hazard problem? If not, the scope of explicit incentive provision and corporate stakeholderism may be limited.

To address this question, we now borrow the approach of Holmström and Milgrom (1991) and extend the model by adding another externality dimension. Specifically, there are now two externalities, T_1 and T_2, that the manager has to contend with. The two corresponding efforts, both unobservable, are denoted by y_1 and y_2, respectively. For $i = 1,2$, the probability of success is y_i, which yields externality T_i when successful. There is a signal for the first prosocial activity, y_1, with the same conditional distribution as before, but we assume that there is no signal for the second activity, y_2.

Let $C(x, y_1, y_2)$ be the private cost to the manager with the following specification.

Assumption 4. $C(x, y_1, y_2) = c(x) + h(y_1, y_2)$ for some $c(\cdot)$ and $h(\cdot,\cdot)$ such that

$h(y_1, y_2) = 0$ if $y_1 = 0$ and $0 \leq y_2 \leq \bar{y}_2$ for some $\bar{y}_2 \in [0,1]$;
$h(y_1, y_2) > 0$ otherwise.

This assumption implies that, even in the absence of explicit incentives, task y_2 will be carried out to some extent as long as $y_1 = 0$.

Given the outside investor's breakeven constraint (6), the stakeholderist owner's objective becomes

$$W(\alpha) = \alpha(y_1 T_1 + y_2 T_2) + (1 - \alpha)\big(xR + \gamma(A - I)\big). \tag{10}$$

We can immediately apply the insight of Holmström and Milgrom (1991) to obtain the following result.[11]

Proposition 8. Under Assumption 4, if $\bar{y}_2 T_2 > T_1$, then it is optimal to induce $y_1 = 0$.

Proof. First, consider the case in which the solution to (10) involves $y_1 > 0$. Then, since $h_2(y_1, y_2) > 0$ if $y_1 > 0$, it must be that $y_2 = 0$, and the value of the optimization problem becomes

$$\max_{x, y_1} \alpha\, y_1 T_1 + (1 - \alpha)\big(xR + \gamma(A - I)\big),$$

subject to the relevant constraints. Let (x^*, y_1^*) be the optimal level of efforts.

Second, consider the case in which the solution to (10) involves $y_1 = 0$. Then, by Assumption 4, $y_2 = \bar{y}_2$ can be induced. The value in this case is

$$\max_{x} \alpha\, \bar{y}_2 T_2 + (1 - \alpha)\big(xR + \gamma(A - I)\big),$$

subject to the relevant constraints. Observe that if $(x, y_1) = (x^*, y_1^*)$ is feasible in the first case, then $(x, y_1) = (x^*, 0)$ is feasible in the second case. Therefore, the claim holds if $\bar{y}_2 T_2 > T_1 \geq y_1^* T_1$. ■

If the principal is guaranteed a certain level of prosocial activities at no cost, represented by $\bar{y}_2$, and if the resulting benefit is larger than the benefit under explicit control, the principal chooses to overlook the latter completely. Thus, adopting stakeholderism is

[11] See Proposition 1 of Holmström and Milgrom (1991).

inconsequential. This “crowding out” of effort is driven by the fact that the overall informativeness of (two) signals is not enough relative to the overall number of (three) unobservable activities. Note, however, that Proposition 8 differs from Proposition 4 in that here, adopting stakeholderism, or greater α, does not reduce $T \coloneqq T_1 + T_2$.[12]

How can we interpret $\bar{y}_2$? It could simply be some unmeasurable activity that is “very important,” as in the own words of Holmström and Milgrom (1991). Another potential source in our context may be found in government intervention. Some of the objectives of the stakeholderist owner, such as pollution and governance, may be addressed by the government as well, and regulation could ensure a minimum level of corresponding activities. The measures that the government employs to regulate those activities may not be available for the firm. We can then think of y_1 as prosocial activities that lie outside the reach of government intervention. In this interpretation, Proposition 8 suggests that, if the government is already dealing with the majority of relevant stakeholder interests, explicit corporate stakeholderism may not make any additional difference.

Chapter 8. Conclusion

This paper was an attempt to weigh in on the recent debate on corporate stakeholderism. In particular, we sought to formally address the claims that explicit consideration of stakeholders in addition to profits might paradoxically end up hurting their own

[12] Since the lucrative activity x is important for obtaining outside investment, we cannot obtain this crowding-out result with just one prosocial activity y.

interests (e.g. (Bebchuk & Tallarita, 2020)).

We developed an agency model that extends in several directions the credit-rationing model of Holmström and Tirole (1997). Our multitask moral hazard analysis suggests that a greater emphasis on stakeholder interests not only reduces profits but could also end up discouraging activities promoting stakeholder interests if it leads to less informative signals.

Another lesson from the analysis came from applying the insight of Holmström and Milgrom (1991). If the output of some stakeholder-friendly activities is unmeasurable, and if those activities are important enough that the agent expends a certain level of effort even without the principal' s incentives, then there may actually be no gains to be made from pursuing other measurable stakeholder objectives.

These arguments provide a possible (albeit partial) justification for viewing shareholder-value maximization as a second-best solution, in determining the appropriate objective function of the firm.

Chapter 9. Appendix: Proof of Propositions

9.1. Proof of Proposition 1

Proof. From the outside investor's breakeven constraint (6), we can re-write the surplus as

$$\begin{aligned} & x\big(q(y)R_f^S + \big(1-q(y)\big)R_f^F\big) + \gamma\big(A - I_f\big) \\ & = x\big(q(y)R_f^S + \big(1-q(y)\big)R_f^F\big) + \gamma I_u - \gamma(I-A) \\ & = xR - \gamma(I-A). \end{aligned} \tag{11}$$

That is, the firm extracts as surplus the entire net present value of the project, due to the assumption that the capital market is competitive.

Then, the objective function of the stakeholderist owner (1) becomes

$$W(\alpha) = \alpha yT + (1-\alpha)\big(xR + \gamma(A-I)\big),$$

where the choice of efforts (x,y) is constrained by (i) $0 \le R_f^S \le R$ and $0 \le R_f^F \le R$, (ii) the FOCs for the IC constraint, (3) and (4), (iii) the firm's capital constraint, i.e., $I_u \ge I - A$, and (iv) the IR constraint (5). We now show that these four conditions are equivalent to $(x,y) \in \tilde{X}$ as in the claim.

First, note that the FOCs determine $\left(R_f^S, R_f^F\right)$ as a function of the effort level we want to induce. (Conversely, any (x,y) can be induced by the optimal level of $\left(R_f^S, R_f^F\right)$, as long as $0 \le R_f^S \le R$ and $0 \le R_f^F \le R$.) In order for (i) to hold, by using (3) and (4) together with the specification for $q(y)$, it is straightforward to see that (x,y) must belong to

$$L(\beta) := \{\, (x,y) \in [0,1]^2:$$
$$0 \le C_x(x,y) - \frac{C_y(x,y)}{x}\left(\left(y - \frac{1}{2}\right) + \frac{1}{2\beta}\right) \le R,$$
$$0 \le C_x(x,y) - \frac{C_y(x,y)}{x}\left(\left(y - \frac{1}{2}\right) - \frac{1}{2\beta}\right) \le R\}.$$

Second, using (3) and (7), (iii) becomes

$$x[R - C_x(x,y)] \ge \gamma(I - A). \tag{12}$$

If Assumption 2 holds, the region defined by (12) is contained within the square $[0,1] \times [0,1]$ since

$$R - C_x(1,y) < \gamma(I - A), \qquad \forall y,$$
$$0(R - C_x(0,1)) \ < \ \gamma(I - A), and$$
$$\frac{\partial}{\partial x}\left[x\big(R - C_x(x,1)\big)\right] = R - \ C_x(x,1) - xC_{xx}(x,1) < 0.$$

Then, at the optimum, (12) holds with equality, which corresponds to the set K.

Finally, to check the IR condition (5), observe that, if (12) holds with equality, the surplus (11) becomes $xC_x(x,y)$. Therefore, the IR constraint corresponds to the set P. ■

9.2. Proof of Proposition 4

Proof. Note from Proposition 1 that β affects y through $L(\beta)$ and P. Assume that $C(x,y) = (c_1 x + c_2 y)^2$. Then, the inequalities in $L(\beta)$ become

$$0 \leq 2c_1 x - 2c_2 y + c_2\left(1 - \frac{1}{\beta}\right) \tag{13}$$

and

$$2c_1 x - 2c_2 y + c_2\left(1 + \frac{1}{\beta}\right) \leq \frac{Rx}{c_1 x + c_2 y}. \tag{14}$$

The first condition is equivalent to $R_f^F \geq 0$, while the second condition is equivalent to $R_f^S \leq R$. Let us call the first region $L_1(\beta)$ and the second region $L_2(\beta)$. Also, recall from Proposition 1 that P is the region such that $xC_x(x,y) - C(x,y) \geq 0$, which under our specification for C becomes $c_1 x \geq c_2 y$. Since $0 < \beta < 1$, we have $L_1(\beta) \subset P$, so it suffices to consider the more restrictive set $L_1(\beta)$.

We claim that $L_1(\beta) \cap L_2(\beta) \neq \emptyset$, and $L_1(\beta)$ is increasing in β while $L_2(\beta)$ is decreasing in β, where "increasing" and "decreasing" are understood in terms of inclusion (i.e., $L(\beta)$ is a "strip" that "moves to the right" as β decreases).

First, note that the boundaries of the two regions do not cross. If $R_f^F = 0$, then

$$2c_1 x - 2c_2 y + c_2\left(1 - \frac{1}{\beta}\right) = 0.$$

However,

$$R_f^S - R_f^F = \frac{2c_2(c_1 x + c_2 y)}{x\beta} \leq R$$

if

$$y \leq \frac{\beta R - 2c_1 c_2}{2c_2^2} x.$$

This is true if (comparing to the case $x = 1$)

$$2\frac{c_1}{c_2} + \frac{1}{2} \leq \frac{\beta R}{2c_2^2} + \frac{1}{2\beta}$$

for all $\beta \in [0,1]$, which holds if

$$4c_1 + c_2 \leq 2\sqrt{R}$$

by using the arithmetic−geometric inequality. Under this condition, the region is a "strip". Next, note that (13) holds for higher values of β and (14) holds for lower values of β. Hence, $L(\beta)$ "moves to the right" as β decreases.

We have already seen in Proposition 3 that assuming $C(x,y) = (c_1 x + c_2 y)^2$ produces a concave boundary for the set K. Thus, we conclude that y (weakly) decreases as β decreases. In particular, for large values of β, $R_f^S = R$ and for small values of β, $R_f^F = 0$, while, for some intermediate values, $0 < R_f^S < R$ and $0 < R_f^F < R$. The condition for y to be decreasing in α is

$$\frac{d\beta}{d\alpha} < -\frac{\partial y^*/\,\partial\alpha}{\partial y^*/\,\partial\beta}.$$

Note that $\partial y^*/\partial\alpha > 0$, and $\partial y^*/\partial\beta < 0$, so $d\beta/d\alpha$ must be sufficiently negative. ■

9.3. Proof of Proposition 5

Proof. Credit rationing is determined by the level of maximum pledgeable income, so it is enough to show that the maximum pledgeable income is independent of β. The maximum pledgeable income is

$$\max_{R_f^S, R_f^F \in [0,R]} x\left[R - \left(q(y) R_f^S + \left(1 - q(y)\right) R_f^F\right)\right]$$

subject to (2) and (5). We claim that the maximum is achieved at $R_f^S = R_f^F = r$, where r solves

$$\max_{r \in [0,R]} x\,[R - r],$$

subject to $x = argmax_{\tilde{x} \in [0,1]} r\tilde{x} - C(\tilde{x}, 0)$. Under our assumptions, $r = C_x(x^*, 0)$. The maximum pledgeable income is then obtained as $x^*[R - C_x(x^*, 0)]$, which is independent of β.

First, compare $(R_f^S, R_f^F) = (r, r)$ with choices of (R_f^S, R_f^F) which imply an interior solution in problem (2). In this case, (R_f^S, R_f^F) and (x, y) must satisfy the FOC's, (3) and (4). In particular, because of (3), the pledgeable income becomes

$$x[R - C_x(x, y)],$$

which cannot be larger than $x^*[R - C_x(x^*, 0)]$.

Second, compare $(R_f^S, R_f^F) = (r, r)$ with alternative choices of (R_f^S, R_f^F) which imply a corner solution in problem (2). There are four cases: (i) $x = 0$, (ii) $y = 0$, (iii) $x = 1$, and (iv) $y = 1$.

(i) When $x = 0$, pledgeable income is 0, which is smaller than $x^*[R - C_x(x^*, 0)]$.

(ii) When $y = 0$, it must be that $R_f^S \leq R_f^F$. However, in this case, we can guarantee the same level of pledgeable income by setting new compensation levels $(R_f^{S\prime}, R_f^{F\prime})$ such that $R_f^{S\prime} - R_f^{F\prime} = q(0)R_f^S + (1 - q(0))R_f^F$. Therefore, we can do no better than the proposed maximum.

(iii) When $x = 1$, it must be that $[q(y)R_f^S + (1 - q(y))R_f^F] \geq C_x(1, y)$. However, in this case, we can guarantee a weakly larger pledgeable income by setting new compensation levels $(R_f^{S\prime}, R_f^{F\prime})$ such that $R_f^{S\prime} - R_f^{F\prime} = R_f^S - R_f^F$ and $q(1)R_f^{S'} + (1 - q(1))R_f^{F'} = C_x(1, y)$. Therefore, we can do no better than the "interior solution" case, *a fortiori* as well as the proposed maximum.

(iv) When $y = 1$, it must be that $R_f^S - R_f^F \geq \frac{C_y(x,1)}{\beta x}$. However, in this case, we can guarantee the same level of pledgeable income by setting new compensation levels $(R_f^{S\prime\prime}, R_f^{F\prime\prime})$ such that $R_f^{S\prime} - R_f^{F\prime\prime} = \frac{C_y(x,1)}{\beta x}$ and $q(1)R_f^{S\prime\prime} + (1 - q(1))R_f^{F\prime\prime} = q(1)R_f^S + (1 - q(1))R_f^F$. Therefore, we can do no better than the "interior solution" case, *a fortiori* as well as the proposed maximum. ■

9.4. Proof of Proposition 6

Proof. First, focus on the case when $\theta = 1$. The maximum pledgeable income solves

$$\max_{R_f^S, R_f^F \in [0,R]} y[R - (q(y)R_f^S + (1 - q(y))R_f^F)]$$

subject to

$$(x, y) \in \underset{\tilde{x},\tilde{y}\in[0,1]}{\operatorname{argmax}}\ \tilde{y}\left[q(\tilde{y})R_f^S + \left(1 - q(\tilde{y})\right)R_f^F\right] - C(\tilde{x}, \tilde{y}).$$

We can see immediately that $x = 0$, and that

$$\mathrm{y} \in \underset{\tilde{y}\in[0,1]}{\operatorname{argmax}}\ \tilde{y}\left[q(\tilde{y})R_f^S + \left(1 - q(\tilde{y})\right)R_f^F\right] - C(0, \tilde{y}).$$

Suppose $0 \leq \beta_1 < \beta_2 \leq 1$, and let $q_i(y) = 1/2 + \beta_i(y - 1/2)$, for $i = 1,2$. Also, let the values which maximize the maximum pledgeable income under β_1 be $\left(R_1^S, R_1^F, y_1\right)$. Then, when β is equal to β_2, there exists $\left(R_2^S, R_2^F\right)$ such that incentive compatibility still induces effort y_1, i.e.,

$$y_1 \in \underset{\tilde{y}\in[0,1]}{\operatorname{argmax}}\ \tilde{y}\left[q_2(\tilde{y})\right)R_2^S + \left(1 - q_2(\tilde{y})\right)R_2^F\right] - C(0, \tilde{y}),$$

while requiring less expected cost, i.e.,

$$q_2(y)R_2^S + \left(1 - q_2(y)\right)R_2^F < q_1(y)R_1^S + \left(1 - q_1(y)\right)R_1^F.$$

To show this, we use the FOC for y, which is

$$y\beta\left(R_f^S - R_f^F\right) + \left[q(y)R_f^S + \left(1 - q(y)\right)R_f^F\right] = C_y(0, y).$$

Rearranging, we get

$$q(y)R_f^S + \left(1 - q(y)\right)R_f^F = C_y(0, y) - y\beta\left(R_f^S - R_f^F\right).$$

Therefore, it is enough to show that $\beta_1\left(R_1^S - R_1^F\right) < \beta_2\left(R_2^S - R_2^F\right)$. It is sufficient to show that $\left(R_2^S, R_2^F\right)$ can be found that satisfy $R_1^S - R_1^F = R_2^S - R_2^F$. Rearranging the FOC once more, we get

$$[\beta y + q(y)]\left(R_f^S - R_f^F\right) + R_f^F = C_y(0, y),$$

so the desired compensation scheme $\left(R_2^S, R_2^F\right)$ can be found by adjusting R_f^F from R_1^F, while keeping the difference, $R_f^S - R_f^F$, unchanged. (If $y_1 \geq 1/2$, then R_f^F is lowered, and if $y_1 < 1/2$, the direction of change in R_f^F is ambiguous.)

Furthermore, we can see that there exists a $\theta_0 \in (0,1)$ such that whenever $\theta \geq \theta_0$, the maximum pledgeable income increases with β. For this, examine the x−derivative of the Agent's utility, which is

$$(1 - \theta)\left[q(y)R_f^S + \left(1 - q(y)\right)R_f^F\right] - C_x(x, y).$$

This value is negative when θ is close to 1, in which case $x = 0$. Then, basically the same argument given above carries over. ■

9.5. Proof of Proposition 7

Proof. First, let us examine the special case of $\theta = 1$. It is immediate that α does not influence either surplus or externality.

First, we look at the effect of β on y. Suppose that β increases from β_1 to β_2, that q_i is associated with β_i, and that the optimal values for β_1 are (y_1, R_1^S, R_2^S). Then, as in the proof of Proposition 5, we can find (R_2^S, R_2^F) which induce the same y_1, but increases the income pledged. In particular, this was achieved by only adjusting R_2^F, while keeping $R_2^S - R_2^F$ the same. Assume that $y_1 > 1/2$, in which case R_2^F is lowered from R_1^F. Also assume that the IR constraint holds at these new values also. This already shows that y weakly increase with β.

Next, we show that y *strictly* increases in β. First adjust (R_2^S, R_2^F) to (R_3^S, R_3^F), in such a way that $R_3^S - R_3^F$ is larger than $R_2^S - R_2^F$, while $R_3^F = R_2^F$. Exactly by how much will be specified later. Denote the new incentive compatible y by y_3. Let us find the derivatives with respect to y, at $y = y_1$, of the agent's utility function:

$$\begin{aligned}&[\beta_2 y_1 + q_2(y_1)](R_3^S - R_3^F) + R_3^F - C_y(0, y_1)\\&> [\beta_2 y_1 + q_2(y_1)](R_2^S - R_2^F) + R_2^F - C_y(0, y_1)\\&= 0.\end{aligned}$$

Therefore, $y_3 > y_1$. Also the IR condition is satisfied at (y_3, R_3^S, R_3^F), since

$$\begin{aligned}&y_1[q_2(y_1)R_3^S - (1 - q_2(y_1))R_3^F] - C(0, y_1)\\&> y_1[q_2(y_1)R_2^S - (1 - q_2(y_1))R_2^F] - C(0, y_1)\\&\geq 0.\end{aligned}$$

Next note that, for the income pledged,

$$y_1[R - q_2(y_1)R_2^S - (1 - q_2(y_1))R_2^F] > 0.$$

At this stage, we specify $R_3^S - R_3^F$ to be (larger than $R_2^S - R_2^F$ but) small enough so that

$$(R_3^S - R_3^F) - (R_2^S - R_2^F) < \epsilon,$$

and

$$\triangle y = y_3 - y_1 < \epsilon.$$

This can be done since the incentive compatible y varies continuously in $R_f^S - R_f^F$. Finally, note that the income pledged under

$\left(y_1, R_3^S, R_3^F\right)$ is

$$y_1\left[R - q_2(y_1)R_3^S - \left(1 - q_2(y_1)\right)R_3^F\right]$$

whose difference from the income pledged under $\left(y_1, R_2^S, R_2^F\right)$ is bounded by $q_2(y_1)\epsilon$. The $y-$derivative of the above expression is

$$\left[R - q_2(y_1)R_3^S - \left(1 - q_2(y_1)\right)R_3^F\right] - \left(R_3^S - R_3^F\right)\beta_2 y,$$

whose absolute value is bounded by

$$R + (1 + \beta_2)\left(R_3^S - R_3^F\right) + R_3^F.$$

Therefore, by the mean value theorem, if we allow ϵ to be small enough, the capital constraint will be satisfied for $\left(y_3, R_3^S, R_3^F\right)$. Therefore, we have shown that y strictly increases in β.

Next, let us examine the cases when $\theta < 1$. In this case, the marginal rate of substitution between x and y for the Principal is well-defined, and equals

$$\frac{(1 - \alpha)(1 - \theta)R}{\alpha T + (1 - \alpha)\theta R}.$$

To ease the exposition, let us make a change of variables and define

$$p \equiv p(x, y) = (1 - \theta)x + \theta y.$$

Then, the Principal's problem is actually the same as when $\theta = 0$, except that now the cost function is

$$C^*(p, y) \equiv C(p(x, y), y)$$

instead. If we assume that C^* exhibits the same properties as C, then Proposition 3 and 4 carry over. ∎

References

Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy, 105*(1), 1–29.

Bebchuk, L. A., & Tallarita, R. (2020). The illusory promise of stakeholder governance. *Cornell L. Rev., 106*, 91.

Bénabou, R., & Tirole, J. (2010). Individual and corporate social responsibility. *Economica, 77*(305), 1–19.

Blackwell, D. A., & Girshick, M. A. (1979). *Theory of games and statistical decisions*. Courier Corporation.

Dewatripont, M., Jewitt, I., & Tirole, J. (1999). The economics of career concerns, part ii: Application to missions and accountability of government agencies. *The Review of Economic Studies, 66*(1), 199–217.

Dewatripont, M., Jewitt, I., & Tirole, J. (2000). Multitask agency problems: Focus and task clustering. *European Economic Review, 44*(4-6), 869–877.

Grossman, S. J., & Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica, 51*(1), 7–45.

Hart, O., & Zingales, L. (2017). Companies should maximize shareholder welfare not market value. *ECGI-Finance Working Paper*(521).

Hermalin, B. E. (2012). Corporate Governance: A Critical Assessment. In Gibbons, R., & Roberts, J. (Eds.), *The handbook of organizational economics* (pp. 732–763). Princeton University Press.

Holmström, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization, 7*, 24–52.

Holmström, B., & Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics, 112*(3), 663–691.

Hörner, J., & Olszewski, W. (2006). The folk theorem for games with private almost-perfect monitoring. *Econometrica, 74*(6), 1499–1544.

Laux, C. (2001). Limited-liability and incentive contracting with multiple projects. *The RAND Journal of Economics, 32*(3), 514–526.

Magill, M., Quinzii, M., & Rochet, J.-C. (2015). A theory of the stakeholder corporation. *Econometrica, 83*(5), 1685–1725.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford University Press.

Tirole, J. (2001). Corporate governance. *Econometrica, 69*(1), 1–35.

국문초록

이해관계자 중심주의는 기업이 이윤만 추구할 것이 아니라, 기업활동으로 인해 영향을 받는 이해관계자의 후생도 함께 고려해야 한다는 주장이다. 본 논문은 다차원적 대리인 모형(multitask agency model)을 도입하여, 이해관계자 중심주의가 기업활동의 결과로 발생하는 외부효과의 문제를 효과적으로 해결할 수 있는지 알아본다. 먼저, 이해관계자의 이익에 대한 고려가 심화되면 이윤이 감소할 수 있다. 또한, 이해관계자 중심주의로 인해 기업활동에 대한 척도의 정보성이 악화된다면, 이해관계자에게 유리한 기업활동이 오히려 감소할 수 있다. 마지막으로, 이해관계자에게 유리한 기업활동 중 일부에 대한 측정이 불가능하다면, 이해관계자 중심주의를 도입하는 것이 이윤 극대화를 도입하는 것에 비해 이해관계자들의 이익을 더 잘 보장하지 못할 수 있다. 이러한 결과들은 이윤 극대화의 원칙이 기업의 알맞은 목적함수를 구하는 문제의 차선책이 될 수 있음을 시사한다.

주요어 : 이해관계자 중심주의, 대리인 이론, 다차원적 대리인 모형, 기업지배, 기업재무
학번 : 2020-25470