Fat-tail Climate Risks, Mechanism design, and Reputation*

Prasenjit Banerjee[a] and Jason F. Shogren[b]

4 June 2012

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2012 AAEA Annual Meeting, Seattle, Washington, August 12-14, 2012

[a]University of Manchester, Oxford Road, Manchester M13 9PL UK

[b]University of Wyoming, Laramie, WY 82071, USA

Abstract

This paper investigates the interaction between consumers and producers in designing incentive mechanism for climate protection. Firms have material interests in building a moral reputation for those consumers who prefer buying from socially responsible firms. We examine optimal monetary transfer by addressing crowding out effect due to reputation. We find green reputation leads to *overprotection* and brown firms buy reputation if consumers have strong preference on green products. When consumers care less about firms' reputation, firms do not have any incentive to buy reputation.

Key Words: Asymmetric Information, Climate change, Crowding out, Mechanism Design, Reputation.

JEL Classification: D02, D03, Q15, Q34, Q57.

1. Introduction

Climate change policy requires behavioral change, both in production and consumption. Firms see this change as "going green;" to be viewed as socially responsible by stakeholders and consumers by reducing CO2 emissions and investing in energy efficient technology (Arora and Gangopadhyay 1995). Producers must believe this behavior will pay them off by attracting green consumers and stakeholders or they may be truly concerned about climate change (Friedman 1970; Besley and Ghatak 2007). If firms undertake costly investments towards climate protection to buy a good reputation, a policy maker can exploit this behavior by designing incentive mechanisms that explicitly addresses the desire to buy a "green reputation" (see Banerjee and Shogren, 2011).

Herein we explore whether the decision maker can design mechanisms to help reduce climate change risk at less cost given she allows the firm to purchase a socially responsible "reputation". We investigate the interaction between consumers and producers in designing incentive mechanism for climate protection when firms have material interests in building a moral reputation for those consumers who prefer buying from socially responsible firms. Following Be´nabou and Tirole (2003) and Frey (1994), we examine mechanism design by focusing on both the extrinsic material interests of the firms (e.g., Baliga and Maskin 2003; Laffont, 1995) and the intrinsic motives (e.g., warm glow). For an intrinsically-motivated person or firm, monetary rewards create doubt about his true motives for doing a good deed, i.e., the crowding out effect (see Bowles, 2008).

As Be´nabou and Tirole (2006) note intrinsic motives are not the end of the story—people and firms also care about their personal reputation. For example, since mid-90s, firms started adopting new technology and other different policies to comply with climate protection either for

social reasons or to capture a good reputation or to prepare for expected regulation (Kolk and Pinkse 2004; Kolk 2000; Levy and Egan 2003). If we witness a firm engaging in a pro-climate action with little or no compensation it might be because it is "buying" a good reputation, not intrinsic motivation. If so, the open question is whether the regulator can redesign an incentive mechanism to save public funds by exploiting the firm's concern about its reputation.

Our model examines how private information about reputation affects the performance of a Laffont-style (1995) mechanism designed for an optimal regulation of a project with a risk of a climate catastrophe, e.g., a fat-tailed risk (see Weitzman 2010, 2009; Pindyck 2010). A firm cares about its reputation of being socially responsible because consumers get utility when they buy from a socially responsible firm. Consumers gain intrinsic satisfaction if they could buy from a climate-friendly firm (Klein 1990; Cairncross 1992; Arora and Gangopadhaya 1995). We define reputation based on the firms' intrinsic valuation for money – a high intrinsic valuation for money implies a brown reputation and vice-versa (Be´nabou and Tirole, 2006). Climate protection without compensation increases a firm's reputation, whereas protection only for the monetary reward decreases reputation. This is the crowding out effect—because the firm loses reputation if it is perceived within society as "money hungry" (Be´nabou and Tirole, 2006). Firms have social preferences for climate protection even though it is privately costly to them. They have social preferences for reduce climate risk without compensation from those who benefit. Actually paying them to protect nature might be counter-productive.

Other firms, however, do not have strong social preferences for reducing the impact of climate change. They are uninterested in paying a private cost to protect a public good—unless they get something out of it themselves. These firms might invest in climate protection to "buy" a green reputation. As consumers care about green reputation, firms can increase their profit by

gaining green reputation. The regulator's challenge is that she does not know which firm is which—social preferences or reputation buyer. She does not want to chase away the firm with social preferences by crowding out their incentives to do the right; she does not want to reward the reputation seeker by paying out extra money that could be spent elsewhere. The challenge for mechanism design is to identify by observing firms' behavior why they contribute to a social project—is this due to intrinsic motivation or sociality (Dana et al. 2003)—as firms care about reputation too.

In our model, firms do not have complete information about consumers' preference. Some consumers want to purchase from a green firm. Others might not have the same preference. Firms care reputation as it increases future profit. If consumers do not care, firms are unable to extract some surplus from gaining a green reputation. Firms face a challenge—they do not know consumers' true preference. The regulator faces a challenge too—she lacks information on which firm is socially responsible and which one is reputation buyer. Also, she does not know about consumers' preference. The open question is whether she can design a mechanism that specifies a menu of monetary transfer-to-effort that maximized efficiency and minimizes information rents given both types of firms.

Our results suggest that brown firms with a poor reputation sacrifice their information rent; and they can go so far as to pay the regulator to gain reputation for climate protection– provided consumers prefer to buy from a green firm. However, brown firms do not act like reputation buyer when consumers do not have strong preference on green product. In contrast, green firms exert more than optimal climate protection when consumers value green reputation. The social planner can save public funds if consumers like to buy from green firms as brown

firms want to buy reputation. The more a regulator publicizes private climate protection actions, the more all firms contribute, both those with high and low concern for reputation.

2. Analytical Model

We begin by defining our basic analytical model of mechanism design. We introduce the idea of social preferences and reputation under asymmetric information, assuming homogenous technology.

2.1. Basic model

Following Laffont (1995), consider a model on mechanism design for a project with a risk of a climate catastrophe, e.g., a fat-tailed risk. A project of climate protection has social value S and cost $C = \beta - e$, where β is the efficiency of the agent and e is effort to reduce cost. The regulator observes the cost ex-post. The project's expected value is S [1]. Let t be the monetary net transfer to firm from the regulator. The regulator is going to offer up this transfer *t* in exchange for cost control on the project. This is standard mechanism design, in which the regulator offers the firm a "menu" of transfers and effort. The greater the transfer, the greater the effort expected—this is the transfer-menu mechanism. The firm chooses from this mechanism the transfer-effort combination that maximizes its private payoffs.

The regulator's goal is to prevent the climate change catastrophe. To protect the climate, the regulator wants the firms to reduce carbon emission, or to change production practices to be more eco-friendly. A firm suffers a monetary loss when it introduces eco-friendly technology. They would voluntarily participate in the program of climate protection if they get monetary transfer from the regulator to compensate their loss. The regulator designs a contract to

[1] This expected value captures a benefit from the project and expected costs from a catastrophe. The firm self-protects to reduce the probability of a fat-tail climate catastrophe. Following Laffont (1995), assume the damage from catastrophe is so large that the firm always invests optimally in self-protection.

maximize social welfare from climate protection subject to the firms' participation constraints. Social welfare is the utility of private firms and consumers, including benefit from climate protection, minus monetary compensation scaled by social value of public funds. Under complete information about the cost structure of the firms, the regulator knows their loss and offers a contract specifying a monetary compensation for abatement. Under asymmetric information, the regulator maximizes social welfare by choosing an optimal contract designed to extract private information cost-effectively and to protect the climate efficiently.

2.2. Benchmark case with different costs and no reputation

We begin by showing the outcome of voluntary incentive mechanism with asymmetric information about costs structure. The regulator designs a voluntary incentive mechanism under private information about cost efficiency. By constructing the mechanism, the regulator offers an optimal contract $\{e^i, t^i\}$ so a firm is no worse off when voluntarily choosing the contract.

A firm's ex-post utility on successful completion of project is expressed as:

$$U = t - \psi(e) \tag{1}$$

where $\psi(e)$ is the cost function of effort with $\psi'(.) > 0 \; and \; \psi''(.) > 0$. Consumer's expected value of the project is,

$$V = S - (1 + \lambda)(C + t) \tag{2}$$

where λ is the social value of public funds used by the regulator to compensate the firm.The objective function of the regulator is,

$$W = V + U = S - (1 + \lambda)\big(\beta - e + \psi(e)\big) - \lambda U \tag{3}$$

Under full information, the regulator maximizes W by choosing e and U subject to the constraint $U \geq 0$. Assuming interior solution, optimal regulation implies: (i) $\frac{\partial \psi(e)}{\partial e} = 1$ (MB = MC); and (ii) $U = 0$ (zero rents for firms). The compensation ensures the firms are exactly compensated for their loss in rents.

Consider now the case of asymmetric information about cost type. Suppose, the regulator cannot observe β, e ex-ante but can observe the cost and if a catastrophe takes place ex-post. Assume two types of firms exist, low or high cost $\left(i.e., \beta \in \left(\underline{\beta}, \overline{\beta}\right)\right)$, and ω is the probability the firm is low cost. A high-cost firm should earn higher rents than the low-cost firm. Under asymmetric information, however, the regulator only knows that the two types exist, not who is who. According to Revelation Principle, the regulator offers a contract to a firm. If the firm accepts it, it exerts effort and receives compensation as specified in the contract. The mechanism provides the incentive so each firm reveals its private information. The regulator considers six incentive compatibility constraints (see, Laffont (1995)). The constraints say each type should not deviate both in truthful announcements of its characteristic and in safety care.

The regulator's challenge is to choose a contract to maximize social welfare given the firms' voluntarily participation. Social welfare is the weighted average of the utility of firms, including the benefit from climate protection, net of cost to fund the project. The regulator maximizes the following objective function by choosing $\left(\underline{e}, \overline{e}, U\left(\underline{\beta}\right), U(\overline{\beta})\right)$ subject to the binding participation and incentive compatibility constraints,

$$W = \omega\left[S - (1+\lambda)\left(\underline{\beta} - \underline{e} + \psi(\underline{e})\right) - \lambda U(\underline{\beta})\right] + (1-\omega)\left[S - (1+\lambda)\left(\overline{\beta} - \overline{e} + \psi(\overline{e})\right) - \lambda U(\overline{\beta})\right] \quad (4)$$

The binding constraints are: (i) incentive constraint of the low type firm (the efficient firm), and (ii) participation or individual rationality constraint of the high type firm (the inefficient firm). The other constraints can be satisfied at no additional social cost. Solving the problem, we have,

$$\psi'(\underline{e}) = 1 \tag{5}$$

$$\psi'(\overline{e}) = 1 - \frac{\lambda}{1+\lambda}\frac{\omega}{1-\omega}\,\Phi'(\overline{e}), \quad \text{with } \Phi(x) = \psi(x) - \psi(x - \Delta\beta) \tag{6}$$

We also have,

$$U(\overline{\beta}, \overline{\beta}, 1) = 0 \tag{7}$$

$$U\left(\underline{\beta}, \underline{\beta}, 1\right) = \Phi(\overline{e}) \tag{8}$$

Compared to the full information case, the expressions (7) and (8) imply the low cost firm invests the optimal level of effort $\underline{e} = e^*$ and captures some information rents, $U\left(\underline{\beta}, \underline{\beta}, 1\right) = \Phi(\overline{e})$. The high cost firm under-invests and captures no information rent. The low-cost firm earns information rents because it mimics the high-cost firm. The high-cost firm does not want to mimic low-cost firm because it would incur a monetary loss. It forces the regulator to give up a positive rent to the efficient firm, as she wants the inefficient firm to be active. Also, the regulator accepts high-cost firm's contribution less than the optimal level to minimize the information rents paid out.

2.3. Reputation and homogenous cost

We now introduce a firm's *reputation* into the model to explore how it affects firms' behavior in climate protection. A firm can utilize its reputation to increase private selfish gain.

As consumers in our model explicitly care about firm's behavior towards climate protection, it could potentially increase firm's profit. Following Benabou and Tirole (2006), let ϑ_e represent a firm's intrinsic valuation and $\vartheta_e\ e$ be its benefit. The firm also values money as a medium of exchange. Denote ϑ_t as firm's intrinsic valuation for money; the firm's benefit is ϑ_t given t transfer from the regulator. Protecting the environment for money lowers a firm's reputation. Defining reputational value as observers' posterior expectation of the firm's type, the *reputational payoff* by choosing e given t is

$$R = x[\gamma_e E(\vartheta_e|e,t) - \gamma_t E(\vartheta_t|e,t)]\ ; \qquad x > 0,\ \ \gamma_e \geq 0,\ \ \gamma_t \geq 0, \quad (9)$$

where, γ_e and γ_t are the weights a firm puts on how it wants to be perceived – *socially responsible* (γ_e) or *selfish* (γ_t). For example, if a firm wants to be perceived as socially responsible, the sign of γ_e is strictly positive. Assume γ_e and γ_t are exogenously given.

Let x capture the *visibility* of a firm's contribution to climate protection, i.e., the probability others observe the firm's actions. Visibility and the weight a firm assigns to reputation define his concern about reputation, $x\gamma_i, i = e, t$. For simplicity, assume reputational concern is identical across firms with fixed γ_i and *x*. Denote

$$R_e = x\left[\gamma_e \frac{\partial E(\vartheta_e|e,t)}{\partial e} - \gamma_t \frac{\partial E(\vartheta_t|e,t)}{\partial t}\right] = x[\cdot] \qquad (10)$$

The marginal impact of expending effort on reputation is positive, $R_e > 0$. A firm who exerts effort for climate protection sends a positive signal about its social preferences, i.e., $\frac{\partial E(\vartheta_e|e,t)}{\partial e} > 0$. Also, other people think the firm's decision might not be driven by money, i.e., $\frac{\partial E(\vartheta_t|e,t)}{\partial t} < 0$ (see the proof of proposition 2 in Bénabou and Tirole, 2006).

A firm can use its green reputation for selfish motive—raise its profit through increase in sale as consumers like to buy products from green firms. Consider consumer's utility as,

$$V = S - (1 + \lambda)(C + t) + \vartheta^C R \quad (11)$$

where, ϑ_e^C is intrinsic satisfaction of consumer if he buys from a firm that has green reputation.

The utility function of a reputation-concerned firm is

$$U = t - \psi(e) + \vartheta^C \pi(R) \quad (12)$$

The firm enjoys gaining green reputation and gets benefit with higher profit as consumers cares about firm's green reputation.

The first order condition is

$$-\psi'(e) + \vartheta^C \pi_e(R) = 0 \quad (13)$$

The second order condition requires $\vartheta^C \pi_{ee} - \psi'' < 0$. Expression (13) implies marginal cost of effort to climate protection equates marginal benefits from monetary compensation and profit due to reputation. The firm gains good reputation as people see its contribution as 'noble' work. This behavior helps firm to increase its profit as consumers love to buy from a firm who cares about protecting the climate.

The regulator encourages social behavior through public displays and advertising (for simplicity, assume visibility of firm's contribution is costless), as

$$\frac{\partial e}{\partial x} = \frac{\vartheta^C \pi_{ex}}{\psi'' - \vartheta^C \pi_{ee}} > 0 \quad (14)$$

By the second-order condition ($\vartheta^C \pi_{ee} - \psi'' < 0$), the denominator of (14) is positive. Since greater publicity of a firm's contribution leads to greater marginal reputational value due to participation in climate protection (since $R_e > 0$, then $\pi_e > 0$ and therefore $\vartheta^C \pi_{ex} > 0$), firms

contribute more to the social project ($\frac{\partial e}{\partial x} > 0$). For simplicity, assume x is exogenously set by the regulator.

Monetary gains increase a firm's utility; but the money could also reduce contributions toward the social project, i.e., the *crowding out effect*, in which extrinsic incentives reduce the incentives of reputation driven firms. The intuition behind crowding out effect due to reputation is firms care more about reputation than money, as taking money for a social work reduces reputation in society and thereby they lose profit (as consumers care about firms' green reputation).

The objective function of the regulator is

$$W = V + U$$

$$= S - (1+\lambda)(C+t) + \vartheta^C R + \vartheta^C \pi(R) + t - \psi(e) \qquad (15)$$

$$= S - (1+\lambda)\big(\beta - e + \psi(e)\big) - \lambda U + (1+\lambda)\vartheta^C \pi(R) + \vartheta^C R$$

We show the crowding out effect by solving firm's problem described in (12). Using the FOC in (13), comparative static result gives

$$\frac{\partial e}{\partial t} = \frac{\vartheta^C \pi_{et}}{\psi'' - \vartheta^C \pi_{ee}} \quad . \qquad (16)$$

Since monetary rewards for social projects suggest the firm is socially irresponsible, and it reduces profit, the cross partial derivative of reward for additional profit π_{et} is negative. By the second order condition of utility maximization ($\vartheta^C \pi_{ee} - \psi'' < 0$), the denominator is positive. Increase in monetary reward reduces effort on climate protection project for reputation-concerned firm ($\frac{\partial e}{\partial t} < 0$). Since monetary rewards adversely affect firm's reputation – firm looses

profit as reputation of being socially irresponsible reduces consumers' satisfaction of buying from the firm – firm reduces its effort.

The regulator designs the incentive mechanism in such a way that it would overcome the crowding out effect. We consider the case where the intrinsic valuation for contributing to climate protection project is same for all the firms (assume ϑ_e is normalized) and allow intrinsic valuation for money (ϑ_t) to vary across firms. Firms' reputation depends on intrinsic valuation for money – high intrinsic valuation for money implies brown reputation and vice-versa. We derive an optimal monetary contract (e^i, t^i) for ith firm to induce the firm to participate in the project. This contract is obtained by optimizing firm's utility in that the firm is no worse off by voluntarily selecting the contract. Under full information, the regulator knows the intrinsic valuation for money for different firms. The regulator maximizes welfare (expression (15)) by choosing e and U subject to $U \geq 0$. Assuming interior solution, optimal regulation implies: (i) information rents are zero, $U = 0$; and (ii) marginal cost of effort equates marginal benefit from monetary reward, gain in profit due to socially responsible reputation, and satisfaction from gaining good reputation as consumers value it

$$\frac{\partial \psi(e)}{\partial e} = 1 + \vartheta^C \pi'(R) + \frac{\vartheta^C}{(1+\lambda)} R'(e,t) \qquad (17)$$

and optimal monetary transfer,

$$t^* = U - \vartheta^C \pi(R(e^*, t)) + \psi(e^*) \qquad (18)$$

The expression (18) is the optimal money transfer to the firm that captures: (i) direct material cost of effort; (ii) gain in profit due to reputation partially offset the material costs of contributing.

Now consider *incomplete information*. Assume two types of reputation: firms with high or low intrinsic valuation for money: "brown" or "green" firms $\left(i.e., \vartheta_t \in (\overline{\vartheta}_t, \underline{\vartheta}_t)\right)$. Let q be the probability of green firm. The green firm participates in climate protection projects by exerting high effort; and has no incentive to hide its private information. The brown firm, however, loses reputation if people indentify its true intrinsic valuation for climate protection. Under asymmetric information, the brown firm gains a good reputation by pretending it has low intrinsic valuation (low ϑ_t). Denote the reputational value for the green firm as $\overline{R}$ $\left(= \gamma_t x\left[E\left(\underline{\vartheta}_t \middle| e_1, t\right)\right]\right)$ and for the brown firm as $\underline{R}$. Define $r\left(\overline{\vartheta_T}, \underline{\vartheta_T}\right)$ as the reputation of the brown firm pretending to be a green firm. The mechanism is incentive compatible if reporting the true information is the dominant strategy of each firm. The incentive compatibility constraints for green and brown firms are:

$U(\underline{\vartheta}_t, \underline{\vartheta}_t) \geq U(\underline{\vartheta}_t, \overline{\vartheta}_t)$ *(L1a)*

$U(\underline{\vartheta}_t, \underline{\vartheta}_t) \geq U(\underline{\vartheta}_t, \underline{\vartheta}_t)$ *(L2a)*

$U(\overline{\vartheta}_t, \overline{\vartheta}_t) \geq U(\overline{\vartheta}_t, \underline{\vartheta}_t)$ *(H1a)*

$U(\overline{\vartheta}_t, \overline{\vartheta}_t) \geq U(\overline{\vartheta}_t, \overline{\vartheta}_t)$ *(H2a)*

and participation constraints:

$U(\underline{\vartheta}_t) \equiv U(\underline{\vartheta}_t, \underline{\vartheta}_t) \geq 0$ *(L4a)*

$U(\overline{\vartheta}_t) \equiv U(\overline{\vartheta}_t, \overline{\vartheta}_t) \geq 0$ *(H4a)*

As the green firm has no incentive to reveal its private information, rather it participates if it gains positive utility, the participation constraint is (L4a) binding. Brown firm can raise benefit by pretending to be a green firm, the incentive compatibility constraint (H1a) is binding. From the binding incentive compatibility and participation constraints, we have

$$\underline{U} = \Phi = \vartheta^C(\pi(r\left(\overline{\vartheta_T}, \underline{\vartheta_T}\right)) - \pi(\overline{R})) \qquad \textit{(L1a)}$$

$$U\left(\overline{\vartheta}_t\right) = 0 \qquad \textit{(H4a)}$$

where $r\left(\overline{\vartheta_T}, \underline{\vartheta_T}\right) = x\gamma_t E\left(\overline{\vartheta}_t \middle| e_1^*, \underline{t}\right)$. Since the intrinsic valuation for money of the brown firm exceeds the green firm (i.e., $r\left(\underline{t}\right) < \underline{R}_t$ and $\pi(r\left(\underline{t}\right)) < \pi(\overline{R})$ as $\underline{\vartheta}_t < \overline{\vartheta}_t$), Φ is negative. Regulator optimizes welfare with respect to $\left(\underline{e},\ \overline{e}, U\left(\underline{\vartheta}_t\right), U(\overline{\vartheta}_t)\right)$ subject to the binding constraints

$$W = q\left[S - (1+\lambda)\left(\beta - \underline{e} + \psi\left(\underline{e}\right)\right) - \lambda U\left(\underline{\vartheta}_t\right) - \vartheta^C\overline{R}\left(\underline{\vartheta}_t \middle| e_1^*, t\right) - (1+\lambda)\vartheta^C\overline{\pi}\left(\overline{R}\right)\right] +$$

$$(1-q)\left[S - (1+\lambda)\left(\beta - \overline{e} + \psi(\overline{e})\right) - \lambda U\left(\overline{\vartheta}_t\right) - \vartheta^C\underline{R}\left(\overline{\vartheta}_t \middle| e_1^*, t\right) - (1+\lambda)\vartheta^C\underline{\pi}\left(\underline{R}\right)\right] \quad (19)$$

Solving regulator's problem we have

$$\frac{\partial\psi(\underline{e})}{\partial\underline{e}} = 1 + \vartheta^C\overline{\pi}'(R) + \frac{\vartheta^C}{(1+\lambda)}\overline{R}'\left(\underline{e}, t\right) - \frac{(1-q)}{q}\frac{\lambda}{(1+\lambda)}\frac{\partial\Phi}{\partial\underline{e}} \qquad (20)$$

$$\frac{\partial\psi(\overline{e})}{\partial\overline{e}} = 1 + \vartheta^C\underline{\pi}'\left(\underline{R}\right) + \frac{\vartheta^C}{(1+\lambda)}\underline{R}'(\overline{e}, t) \qquad (21)$$

As increase in effort level raises reputation (i.e., $\frac{\partial R}{\partial e} > 0$), $\frac{\partial\Phi}{\partial\underline{e}}$ is negative. Equations (20) and (21) state the marginal cost of effort equates marginal benefit of effort when the firms are intrinsically motivated and reputation-concerned. Compared to full information, the green firm exerts more

effort than optimal. Low intrinsic valuation for money, high reputational value, and the feelings of superiority to other firms (i.e., the difference between its own reputation and the brown firm's reputation) motivate the green firm to put higher effort. In contrast, the brown firm exerts the optimal level of effort.

From binding incentive compatible and individual rationality constraints, we find information rents are negative for the brown firm and zero for the green firm:

$$U(\overline{\vartheta}_t) \equiv U(\beta, \overline{\vartheta}_t, \overline{\vartheta}_t) = \Phi \text{ and } U(\underline{\vartheta}_t) \equiv U(\beta, \underline{\vartheta}_t, \underline{\vartheta}_t) = 0 \ . \qquad (22)$$

The brown firm pays the regulator to increase its reputation ($\Phi < 0$). The green firm does not—it has no incentive to increase reputation via the regulator's advertisement of its responsibility to the environment. These findings are similar to our previous work (see, Banerjee and Shogren 2010) where we observe green firms exerts more than optimal effort and brown firms buy reputation.

3. Two Sources of Asymmetric Information: Reputational value and Consumers' preference

We now explore how to design the mechanism when firms do not have complete information about consumers' preference – whether they love to buy from green firms or not. We still have two types of firms – high and low reputation, (green and brown firm), corresponding to low- and high-intrinsic valuation for money ($i.e., \vartheta_T \in (\underline{\vartheta}_T, \overline{\vartheta}_T)$). We now assume two levels consumers' preference on green products: high and low $\left(i.e., \vartheta^C \in (\overline{\vartheta^C}, \underline{\vartheta^C})\right)$, where $\overline{\vartheta^C} > \underline{\vartheta^C}$. We have four possible cases to consider they are described by θ, where $\theta \in (GH, GL, BH, BL)$ and, for example, *GH* implies green (i.e., low ϑ_T) firm and consumers' have 'high' (strong)

preference to buy from green firms (see Table 1). In the full information benchmark case, the regulator faces the same problem as in (14). The optimal conditions also are similar to (6).

Now consider two levels of asymmetric information. Assume f is the probability a firm's type is θ, where, $f \in (p, q, r, s)$ (e.g., the probability a firm is *green* with *high preference on green product* is *p*). We define four different types of firms: (i) the *green firm and 'high' preference on green product* (*type-GH*) who no intention to hide private information; (ii) the *green firm and 'low' preference on green product* (*type-GL*) who has incentive to say it has high intrinsic valuation for money to offset the profit compared to the situation where consumers were 'high' green; (iii) the *brown firm and 'high' preference on green product* (*type-BH*) who pretends it is as a green type; and (iv) the *brown firm and 'low' preference on green product* (*type-BL*) who has no incentive to hide its private information.

The regulator faces a trade-off in designing an efficient mechanism: firm with low intrinsic valuation for money could reduce supply due to the crowding out effect of extrinsic motivation. In contrast, a firm with a high intrinsic valuation for money could refuse to participate with less monetary reward. Here we define an efficient mechanism to induce the two types of firms in four different situations. Under asymmetric information, efficient mechanism should satisfy the incentive compatibility constraints and individual rationality constraints.

As defined in Appendix 2, only (GL1), (BH2), (BL3), and (GH3) are binding. We now explain this in some detail. No incentive compatibility constraints in Appendix 2 for *type-GH* firm are binding. Its dominant strategy is to tell the truth as utility of revealing the true private information about reputation is no less than reporting that it has high intrinsic valuation for money (it would lose reputation and thereby profit as consumers have high preference on green products). It participates if its utility from participation is at least equal to the utility from not

participating (i.e., its participation constraint or individual rationality constraint is satisfied). Having a low intrinsic valuation for money, a green firm should not have any incentive to hide its private information about its valuation for money; otherwise it would lose his reputation. However, in *type-GL* case, as consumers are not much interested in buying from green firms, firms' green reputation would not be helpful in raising profit.

Brown firms want to gain a good reputation by pretending they have low intrinsic valuation for money when they know consumers have strong inclination to buy from green firms. In this case, brown firms want to report they have low intrinsic valuation for money. This would help the brown firms to raise their profit. In contrast, a brown firm would reveal its true type when consumers are less interested about green reputation of firms.

The regulator designs the mechanism to maximize social welfare: the sum of the utilities of different types of firms with their probability of occurrence. She optimizes social welfare by selecting $(e^{GH}, e^{GL}, e^{BH}, e^{BL}, t^{GH}, t^{GL}, t^{BH}, t^{BL})$ subject to (GH1) to (GH5), (GL1) to (GL5), (BH1) to (BH5), and (BL1) to (BL5). Substituting the binding constraint for θth case into the utility of firm under θth case, where $\theta \in (GH, GL, BH, BL)$, the regulator's problem is,

$$W = p\left[S - (1+\lambda)\left(\beta - \underline{e} + \psi(\underline{e})\right) - \lambda U\left(\overline{\vartheta}^{C}, \underline{\vartheta}_t\right) + \overline{\vartheta}^{C}\overline{R}(\underline{\vartheta}_t|e,t) + (1+\lambda)\overline{\vartheta}^{C}\overline{\pi}(\overline{R})\right] + q\left[S - (1+\lambda)\left(\beta - \underline{e} + \psi(\underline{e})\right) - \lambda U\left(\overline{\vartheta}^{C}, \underline{\vartheta}_t\right) + \underline{\vartheta}^{C}\overline{R}(\underline{\vartheta}_t|e,t) + (1+\lambda)\underline{\vartheta}^{C}\overline{\pi}(\overline{R})\right] + r\left[S - (1+\lambda)(\beta - \overline{e} + \psi(\overline{e})) - \lambda U\left(\overline{\vartheta^{C}}, \overline{\vartheta}_t\right) + \overline{\vartheta}^{C}\underline{R}(\overline{\vartheta}_t|e,t) + (1+\lambda)\overline{\vartheta}^{C}\underline{\pi}(\underline{R})\right] + s[S - (1+\lambda)(\beta - \overline{e} + \psi(\overline{e})) - \lambda U(\underline{\vartheta}^{C}, \overline{\vartheta}_t) - \underline{\vartheta}^{C}\underline{R}(\overline{\vartheta}_t|e,t) + (1+\lambda)\underline{\vartheta}^{C}\underline{\pi}(\underline{R})] \quad (23)$$

The necessary conditions imply

$$e^{GH}: \frac{\partial \psi(e^{GH})}{\partial e^{GH}} = 1 + \overline{\vartheta}^{C}\overline{\pi}'\left(\overline{R}\right) + \frac{\overline{\vartheta}^{C}}{(1+\lambda)}\overline{R}'(e^{GH}, t) - \frac{S}{p}\frac{\lambda}{(1+\lambda)}\frac{\partial \Phi}{\partial e^{GH}} \quad (24)$$

$$e^{GL}: \frac{\partial \psi(e^{GL})}{\partial e^{GL}} = 1 + \underline{\vartheta}^{C}\underline{\pi}'\left(\overline{R}\right) + \frac{\underline{\vartheta}^{C}}{(1+\lambda)}\overline{R}'(e^{GL}, t) - \frac{S}{q}\frac{\lambda}{(1+\lambda)}\frac{\partial \Phi}{\partial e^{GL}} \quad (25)$$

$$e^{BH}: \frac{\partial \psi(e^{BH})}{\partial e^{BH}} = 1 + \overline{\vartheta}^{C}\underline{\pi}'\left(\underline{R}\right) + \frac{\overline{\vartheta}^{C}}{(1+\lambda)}\underline{R}'(e^{BH}, t) \quad (26)$$

$$e^{BL}: \frac{\partial \psi(e^{BL})}{\partial e^{BL}} = 1 + \underline{\vartheta}^{C}\underline{\pi}'\left(\underline{R}\right) + \frac{\underline{\vartheta}^{C}}{(1+\lambda)}\underline{R}'(e^{BL}, t) \quad (27)$$

The first order conditions show the marginal benefit and cost of effort exertion for four different cases. Green firms contribute more than optimal because of low intrinsic valuation for money (i.e., high reputational value leads to high profit as consumers love buying from green firms) and the feelings of superiority to other firms (expression (24)). However, a green firm faces a trade-off whether to exert optimal effort or not when consumers are less interested about firms' green reputation (expression (25)). A green firm's feeling of superiority to others motivates the firm to exert greater effort whereas this behavior does not help the firm raising profit. This behavior is not observed in Banerjee and Shogren (2011) model where consumers' preference on environment does not affect firms' utility.

A brown firm exerts optimal level of effort when consumers have high preference on making their purchase from a socially responsible firm. The firm wants to influence consumers by giving optimal effort so that it can gain higher profit (expression (26)). In contrast, brown firms put less than the optimal when consumers care less about firms' reputation (expression (27)).

We now emphasize our two key findings of the model. The first key result is that the green firm exerts *more* than the optimal effort and the brown firm retires *at least* the optimal

provided consumers are green buyer. This overmeeting of climate protection standard by green firms is not new among firms as they do not want to gain a bad or brown reputation (Smart 1992 in Arora and Gangopadhyay 1995). This result is similar to Banerjee and Shogren (2011) that green people overprotect the environment to maintain their good reputation (also, see Milinski et al). In contrast, Laffont (1995) found firms with low cost retire *at least* the optimal and the firms with high costs retire *less than* the optimal.

The binding participation constraint of the green firm when consumers have strong preference to buy from green firm (expression (GH5)) shows that the firm earns zero information rents—it has no incentive to hide private information. These firms behave similarly even when consumers care less about firms' reputation. In contrast, the brown firms will buy reputation and they earn *negative* information rents. As gaining green reputation increases profit, firms want to buy reputation. But brown firms do not have enough incentive to buy reputation when consumers do not have strong preference for green products. In that case, brown firms prefer to reveal their private information and gain zero information rent. Our result contrasts traditional mechanism design, which predicts a person with an information advantage will exploit it. Here once we add reputation into the model, the possibility of *negative* information rent arises– firms with low reputation might buy reputation. This result supports Tullock's (1985) conjecture that someone with a poor reputation "would have to offer some side-payments until he had established something in the way of a reputation". This result also supports Banerjee and Shogren (2011) that agents with low reputation would buy reputation. However, this is true only if consumers care about climate protection. This is possibility was not emerged from Banerjee and Shogren 2011 as they did not consider consumers' preference dictate firms' reputation seeking behavior. This is more realistic to consider firms' reputation seeking behavior when consumers value it.

4. Concluding remarks

We examine interaction between consumers and producers in designing effective incentive mechanism for climate protection held in private hands when both material interests and moral sentiment motivate a firm (see Smith 1759). Our result suggests private firms who already have a green reputation will over-protect the climate to shelter their own good reputation. Secondly, accounting for moral sentiments in the form of *reputation* could save public resources as compared to a traditional mechanism which focuses strictly on material interests. Brown firms will sacrifice their information rent to "buy" a good reputation for protecting the climate provided that consumers care about green products.

Together these two results suggest one policy implication—more wide-spread publicity for climate protection decisions. A firm's reputational value increases as the lay-public learns more about his contribution towards climate protection, which translates into more climate protection.[2] This policy implication has a caveat, however—we assume firms know their reputational value and make choices accordingly. In reality, a person may not know his precise reputation within society. Future research could assume firms do not know their reputation a priori; rather they learn it after they take part in the project. The desire to learn owns reputation and to compete to establish a good reputation could lead to a socially beneficial outcome (see, e.g., Suurmond et al. 2004; Milinski et al. 2002).

[2] In 2002 in Ireland, for example, the regulators imposed a two pronged strategy to reduce the use of plastic bags: (i) 33-cent per unit tax on the use of plastic bag and (ii) an advertising campaign to stigmatize the use of the bags. Carrying plastic bags quickly became socially unacceptable and their usage dropped by 94 percent within weeks (Rosenthal 2008).

References

Andreoni, J. (1989), "Giving with Impure Altruism: Applications to Charity and Recardian Equivalence", *Journal of Political Economy*, Vol. 97 (6), 1447-1458.

_____, (1995), "Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments", *Quarterly Journal of Economics*, Vol. 110 (1), 1-21.

Ariely, D., A. Bracha, and S. Meier (2009), "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially", *American Economic Review*, Vol. 99 (1), 544-555.

Baliga, S. and E. Maskin (2003), "Mechanism Design for the Environment", *Handbook of Environmental Economics*, K. G Maler and J. R. Vincent (Ed.), Vol. 1, 305-324 (Ch. 7).

Batson, D. (1998), "Altruism and Prosocial Behavior", *Handbook of Social Psychology*, D. Gilbert, S. Fiske, and G. Lindzey (eds.), Vol. 2, 282-316, New York: McGraw Hill.

Banerjee, P. and J. F. Shogren (2010), "Regulation, Reputation, and Environmental Risk", *Economics Letters*.

_____(2011), "Material Interests, Moral Reputation, and Crowding out Species Protection on Private Land", *Journal of Environmental Economics and Management*, Forthcoming.

Bénabou, R. and J. Tirole, (2006), Incentives and Prosocial Behavior, *American Economic Review*, Vol. 96 (5), 1652-1678.

Beatty, R. P. and J. R. Ritter (1986), "Investment Banking, Reputation, and Underpricing of Initial Public Offerings", *Journal of Financial Economics*, Vol. 15, 213-232.

Bowles, S. and S. Hwang, (2008), "Social Preferences and Public Economics: Mechanism Design When Social Preferences Depend on Incentives", *Journal of Public Economics*, Vol. 92 (8-9), 1811-1820.

Bowles, S. (2008), "Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments", *Science*, Vol. 320, 1605-1609.

Dana, J. D., R. A. Weber, and J. X. Kuang, (2003), "Exploiting Moral Wriggle Room: Behavior Inconsistent with a Preference for Fair Outcomes", Cambridge Mellon University Behavioral Decision Research, Working Paper 349.

Deci, E. L., and R. M. Ryan (1985), "Intrinsic Motivation and Self-determination in Human Behavior", Plenum Press, New York.

Drumwright, E. M., (1994), "Socially Responsible organizational Buying: Environmental Concern as a Noneconomic Buying Criterion", *Journal of Marketing*, Vol. 58 (3), 1-19.

Frey, S.B., (1997), "Not Just for the Money: An Economic Theory of Personal Motivation", (Edward Elgar, Cheltenham).

Frey, S. B., and R. Jegen, (2001), "Motivation Crowding Theory", *Journal of Economic Surveys*, 15 (5), 589-611.

Klein, B. and K. B. Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance", *Journal of Political Economy*, Vol. 89 (4), 615-641.

Klein, E., (1990), "The Selling of Green", *D&B Reports*, Vol. 38 (September/October), 30-31.

Laffont, J. J., (1995), "Regulation, moral hazard and insurance of environmental risks", *Journal of Public Economics*, Vol. 58, 319-336.

Mantymaa, E., A. Juutinen, M. Monkkonen, and R. Svento (2009), "Participation and Compensation Claims in Voluntary Forest Conservation: A Case of Privately Owned Forests in Finland", *Forest Policy and Economics*, forthcoming.

McKibbin, W. J., and P. J. Wilcoxen (2002), "The Role of Economics in Climate Change Policies", *Journal of Economic Perspectives*, Vol. 16 (2), pp. 107-129.
Milgrom, P. and J. Roberts (1982), "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory*, Vol. 27, 280-312.
_____ (1986a), "Relying on the Information of Interested Parties", *Rand Journal of Economics*, Vol. 17, 18-32.
_____, (1986b), "Price and Advertising Signals of Product Quality", *Journal of Political Economy*, Vol. 94, 796-821.
Milinski, M., D. Semmann, and H. Krambeck (2002), "Reputation Helps Solve the "Tragedy of the Commons", *Nature*, Vol. 415.
Nowak, M. A., and K. Sigmund (1998), "Evolution of Indirect Reciprocity by Image Scoring", *Nature*, Vol. 393, 573-577.
Rosenthal, E. (2008), "Motivated by a Tax, Irish Spurn Plastic Bags", *New York Times*, February 2, 2008, available at: http://www.nytimes.com/2008/02/02/world/europe/02bags.html
Sen, A. (1994), "The Formulation of Rational Choice", *American Economic Review*, Vol. 84 (2), 385-390.
Shapiro, C. (1983), "Premiums for High Quality Products as Returns to Reputations", *Quarterly Journal of Economics*, Vol. 98 (4), 659-680.
Smith, A., (1759), "The Theory of Moral Sentiment", New York: Prometheus Books.
Spence, A. M. (1974), "Market Signaling: Informational Transfer in Hiring and Related screening Process", Cambridge, Mass.: Harvard University Press.
Stavins, R. N. (1998), "What Can We Learn from Grand Policy Experiment? Lessons from SO_2 Allowance Trading", *Journal of Economic Perspectives*, Vol. 12 (3), pp. 69-88.
Titmuss, R., (1970), "The Gift Relationship", London: Allen and Unwin.
Tullock, G. (1985), "Adam Smith and the Prinsoners' Dilemma", *Quarterly Journal of Economics*, Vol. 100, 1073-1081.
Weigelt, K. and C. Camerer (1988), "Reputation and Corporate Strategy: A Review of Recent Theory and Applications", *Strategic Management of Journal*, Vol. 9, 443-454.
Weitzman, Martin L., "On Modeling and Interpreting the Economics of Climate Change," *Review of Economics and Statistics*, February 2009a, **91**, 1—19.
Weitzman, Martin L., "GHG Targets as Insurance Against Catastrophic Climate Damages," NBER Working Paper, May 2010a.
Weitzman, Martin L., "Fat-Tailed Uncertainty and the Economics of Climate Change," *Review of Environmental Economics and Policy*, forthcoming, 2010b.

Appendix 1

Table 1. Firm types under asymmetric information on consumers' preference and reputation

		Consumers' preference: Green	Consumers' preference: Brown
Firms' Reputation	Green	*(G, H)*	*(G, L)*
	Brown	*(B, H)*	*(B, L)*

Appendix 2.

Green firm with high consumers' green preference (*type-GH*):

$$U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \overline{\vartheta^C}\right) \geq U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \overline{\vartheta^C}\right) \qquad (GH1)$$

$$U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \overline{\vartheta^C}\right) \geq U\left(\underline{\vartheta}_t, \overline{\vartheta_t}, \overline{\vartheta^C}\right) \qquad (GH2)$$

Green firm with low consumers' green preference (*type-GL*):

$$U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \underline{\vartheta}^C\right) \geq U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \underline{\vartheta}^C\right) \qquad (GL1)$$

$$U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \underline{\vartheta}^C\right) \geq U\left(\underline{\vartheta}_t, \overline{\vartheta_t}, \underline{\vartheta}^C\right) \qquad (GL2)$$

Brown firm with high consumers' green preference (*type-BH*):

$$U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \overline{\vartheta^C}\right) \geq U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \overline{\vartheta^C}\right) \qquad (BH1)$$

$$U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \overline{\vartheta^C}\right) \geq U\left(\overline{\vartheta}_t, \underline{\vartheta}_t, \overline{\vartheta^C}\right) \qquad (BH2)$$

Brown firm with low consumers' green preference (*type-BL*):

$$U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \underline{\vartheta}^C\right) \geq U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \underline{\vartheta}^C\right) \qquad (BL1)$$

$$U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \underline{\vartheta}^C\right) \geq U\left(\overline{\vartheta}_t, \underline{\vartheta}_t, \underline{\vartheta}^C\right) \qquad (BL2)$$

and individual rationality constraints,

$$U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \overline{\vartheta^C}\right) \geq 0 \qquad (GH5)$$

$$U\left(\underline{\vartheta}_t, \underline{\vartheta}_t, \underline{\vartheta}^C\right) \geq 0 \qquad (GL5)$$

$$U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \overline{\vartheta^C}\right) \geq 0 \qquad (BH5)$$

$$U\left(\overline{\vartheta}_t, \overline{\vartheta}_t, \underline{\vartheta}^C\right) \geq 0 \qquad (BL5)$$