Honor and Stigma in Mechanisms for Environmental Protection

Prasenjit Banerjee, Rupayan Pal, and Jason F. Shogren

Honor and Stigma in Mechanisms for Environmental Protection

Prasenjit Banerjee, Rupayan Pal, and Jason F. Shogren
Indira Gandhi Institute of Development Research (IGIDR)
General Arun Kumar Vaidya Marg
Goregaon (E), Mumbai- 400065, INDIA
Email(corresponding author): prasenjit.banerjee@manchester.ac.uk

Abstract

Honor and stigma play a role in environmental protection. Environmental honors are bestowed on people and firms who go out of their way to do right by the environment. Similarly, environmental stigma is put on people or firms who are publicly taken to task for their poor environmental record. We design a voluntary incentive mechanism by incorporating honor and stigma to induce heterogeneous firms to protect the environment at less cost. We encounter a motivational costs incurred by the green firm-it loses its leadership rents. Our result suggests (i) an additional social reward is needed for a green firm; and (ii) the brown firm may sacrifice information rent.

Keywords: Mechanism, reputation, environmental risk, honor, stigma, social norm

JEL Code: D03, Q52, Q58

Honor and Stigma in Mechanisms for Environmental Protection

Prasenjit Banerjee

University of Manchester, Economics
Arthur Lewis Bldg, Oxford Road, Manchester M13 9PL, United Kingdom
Email: prasenjit.banerjee@manchester.ac.uk

Rupayan Pal

Indira Gandhi Institute of Development Research (IGIDR)
Film City Road, Santosh Nagar, Goregaon (E), Mumbai 400065, India
Email: rupayan@igidr.ac.in

Jason F. Shogren

Department of Economics and Finance, University of Wyoming
Laramie, Laramie, WY 82071-3985, USA
Email: jramses@uwyo.edu

Abstract

Honor and stigma play a role in environmental protection. Environmental honors are bestowed on people and firms who go out of their way to do right by the environment. Similarly, environmental stigma is put on people or firms who are publicly taken to task for their poor environmental record. We design a voluntary incentive mechanism by incorporating honor and stigma to induce heterogeneous firms to protect the environment at less cost. We encounter a motivational costs incurred by the green firm—it loses its leadership rents. Our result suggests (i) an additional social reward is needed for a green firm; and (ii) the brown firm may sacrifice information rent.

Key Words: Mechanism, reputation, environmental risk, honor, stigma, social norm.

JEL Classification: D03, Q52, Q58.

Corresponding Author and Address: Prasenjit Banerjee, Arthur Lewis Building-2.009, Department of Economics, School of Social Sciences, The University of Manchester, Manchester M13 9PL, United Kingdom. Email: prasenjit.banerjee@manchester.ac.uk

1. Introduction

Social psychologists describe honor and stigma as two effective emotions that can motivate people to take pro-social actions voluntarily (Nathanson, 1987; Broucek, 1991; Lea and Webley, 1997). Honor is a feeling of self-respect and people seek honor; whereas people want to hide stigma as it is caused by a strong sense of guilt. Honor and stigma pertain to the gain or loss of "credit, reputation, or a good name" (French, 2002, p. 1). These emotions play a role in environmental protection. Environmental honors are bestowed on people and firms who go out of their way to do right by the environment. Examples abound. Honors bestowed range from local banquets and galas (Kern County (CA) Green Awards banquet; Clarksville-Montgomery County (TN) Green Certification Awards Banquet) to international prizes (Japan's Cosmos prize, National Audubon Society's Rachel Carson Award, European Union's European Business Awards for the Environment). In the United States, the Green Building Council established the LEED rating system—Leadership in Energy & Environmental Design—to honor the degree of sustainability achieved in buildings.

Similarly, environmental stigma is put on people or firms who are publicly taken to task for their poor environmental record. Social psychologists argue that people or firms have a tendency to comply with the norm, especially when actions are observable, to avoid stigma from non-compliance (Posner and Rasmusen 1999; Tadelis 2007). For example, voters' turn-out (in Switzerland) depends on surveillance effect which creates a stigma for non-participants (Kropf and Knack 1003; Funk 2006; Gerber et al 2008); Toxic Releases Inventory in the US reduced releases of toxic chemicals by 45% due to the effect of using public 'shame' tool (Fung and O'Rourke 2000); performance of NHS hospitals in England as measured by patients' waiting time has improved by using "naming and shaming" policy (Besley et al 2009; Bevan and Wilson

2013). This "naming and shaming" policy of global peer pressure also underlies the 2015 Paris agreement on climate change—each country's emission reduction plans are voluntary, with the requirement that they monitor, verify, and report progress. The expectation is that countries will not want to be viewed as global shirkers (see e.g., Tavoni et al., 2011). In addition, numerous bloggers, non-profits, and environmental groups frequently post their environmental scorecards or "hall of shame" to draw attention to people and firms for their environmental irresponsibility (e.g., League of Conservation Voters scorecard, the Public Eye Awards, Newsweek's Green Rankings on corporate environmental performance). Evidence suggests honor motivates compliance with the norms only amongst high-propensity norm-followers, while stigma mobilizes both high- and low-propensity norm-followers (e.g., see Gerber et al. 2008).

Both environmental honor and stigma are aimed achieving behavioral change, especially in production. Firms see this change as "going green"—they gain a green reputation. Consumers see the firms' costly actions as a real economic commitment to socially responsible behavior (e.g., protecting the environment by investing in energy efficient technology, see Akerlof 1980; Besley and Ghatak, 2007; Kitzmueller and Shimshack 2012). The firm has earned a reputation for having solid social preferences. Here a typical policy that provides pure monetary incentives to change a firm's behavior may well backfire in the eye of public perception (e.g., see Frey 1997). Consumers might view a firm as "money-hungry" when it accepts money for "doing the right thing" rather than it having genuine social preferences for the environment (see Bénabou and Tirole 2006; Bowles 2008).

But such social preferences are not always fixed or exogenous—rather they can depend on how honor and stigma within the prevailing social norm (e.g., average action or opinion) interact with preferences (see Bowles, 1998; B; Lindbeck et al 1999; Bowles and Polanía-Reyes,

2012). Social preferences now arise endogenously from the interplay of honor/stigma and preferences (see Bénabou and Tirole, 2006 , 2011). Given honor implies a firm is worthy of high esteem, while stigma implies the opposite, both elements are valuable to a firm if they draw in or avoid repelling potential consumers. For honor, if a firm with social preferences is among the few who do *the right thing,* it enjoys the honor of leading-by-example and being considered a "green firm" (Hermalin 1998; Vesterlund 2003; Potters et al., 2007). The outward perception that 'no one else does it' makes this behavior valuable to the green firm (e.g., a scarcity rent).

For stigma, a firm without social preferences (a brown firm or reputation seeker) might still increase contributions to avoid the stigma of being a "bad player" on the scene. This stigma-driven behavior arises when a *brown* firm's contribution toward environmental protection is less than the average contributions of other firms, and it decides to comply with the norm by doing more protection.[1] Policy makers can exploit stigma-driven behavior by publishing high average contributions (Kuhfuss et al 2015; Allcott 2011; Genius et al 2013). But as Bénabou and Tirole (2011) argue, this policy could backfire too: if a firm sees its leadership role is now less important such that the scarcity value of its reputation is lower, it may do less. The policy maker then faces the dilemma of asymmetric information—the regulator does not know a firm's true motivation with certainty. He does not want to pay extra money to a brown firm to participate if he can use stigma to motivate more environmental protection. He does not also want to undermine the green firm's honor value due to an increase in participation by others.

---

[1] Business groups once within the Global Climate Coalition (GCC), a group of multinationals resisting efforts to reduce GHG emissions, have moved in to a different tent – Center for Climate and Energy Solution (CCES) (previously known as Pew Center for Global Climate Change), the largest U.S.-based association of companies committed to climate protection. This behavioral change may be driven by stigma from abstention as the public view has become more favorable to anthropogenic climate change. A survey by McKinsey reveals nearly 70 percent of global executives believe climate change matters for their strategy for their corporate reputation and brand (http://www.mckinseyquarterly.com/How_companies_think_about_climate_change_A_McKinsey_Global_Survey_2099).

Herein we address how a regulator can re-design voluntary incentive mechanism to use honor and stigma to induce heterogeneous firms to protect the environment at less cost. We extend Banerjee and Shogren (2010) by incorporating endogenous norms (average actions of other firms) into a firm's reputational value. We encounter a motivational costs incurred by the green firm—the green firm loses its leadership rents. Our result suggests an additional social reward is needed to keep them in the program. In addition, we find the brown firm would exert optimal effort and sacrifice information rent only when the stigma effect dominates the honor value. This finding contradicts Banerjee and Shogren (2010) finding that reputation seeker always buys reputation (also see Fischer and Huddart 2008; Bowles and Hwang 2008; Arce 2012) as they did not allow the interaction between honor and stigma. In contrast to Qin and Shogren (2015), the green firm may not always over-comply since its leadership role is now in play.

2. Analytical Model

Using Laffont (1995) as the motivating model, consider a model on optimal regulation of an environmental project with the risk of environmental catastrophe with a probability $(1 - \pi)$, $0 \leq \pi \leq 1$, and E is the realized damage. The project has social value S and cost $C = \beta - e_1$, where β is an efficiency characteristic of the firm running the project and e_1 is an effort variable to reduce cost. The market is a competitive market and in mechanism design the regulator considers one firm holding the other firms' behavior constant. The regulator observes cost ex-ante. The project's expected value is $S - (1 - \pi)E$. The firm self-protects, e_2, to reduce the probability of a catastrophe, $\big(1 - \pi(e_2)\big)$, where $\pi'(e_2) > 0$. Assume $e_2 \in \{0, 1\}$, with $\pi(1) > \pi(0)$. Assume E is so large that the firm always finds it optimal to invest $e_2 = 1$. Let t be the monetary net transfer to the firm from the regulator.

Now we add honor and stigma of a firm to the model. A firm gains good reputation when others (e.g., consumers, stakeholders) think that the firm genuinely contributes to the environment by exerting effort, e_1. Taking monetary compensation, t, from the regulator for doing the right thing may affect the reputation adversely. This reputation depends on the firm's unobservable characteristics whether it truly cares about the environment or does everything for money, given its observed effort and t. The firm gains *honor* when its contribution to the environment is greater than others (i.e., average contribution) as it shows the firms is truly concerned about the environment. In contrast, when its contribution is lower than the average or received monetary compensation is greater than the average, the firm will be perceived as socially irresponsible—stigma attached to the firm.

Following Bénabou and Tirole (2006), denote a firm's intrinsic valuation to contribute to an environmental project by ϑ_{e_1} and intrinsic valuation for money by ϑ_t. ϑ_{e_1} improves the firm's reputation and ϑ_t has adverse effect on reputation—i.e., contributing to the environmental project only for money is perceived as irresponsible to society. Defining reputational value as observers' posterior expectation of the firm's type, the *reputational payoff* by choosing e_1 given t is

$$R = x\left[\gamma_{e_1}E\left(\vartheta_{e_1}\middle|e_1,t\right) - \gamma_t E(\vartheta_t|e_1,t)\right] ; \qquad x > 0, \ \gamma_{e_1} \geq 0, \ \gamma_t \geq 0, (1)$$

If the firm wants to be perceived as public-spirited, the signs of γ_{e_1} and γ_t are strictly positive. Let x capture the visibility of firm's environmental effort, i.e., the probability others observe their actions. Visibility and the weight a firm assigns to its reputation define the firm's concern about reputation, μ, $\mu = x\gamma$. Assume μ is identical across firms with fixed γ and x.

Consider a continuum of firms, each of them decides whether to exert positive effort, i.e., $e_1 \in (0, 1]$. ϑ_{e_1} is normalized to one. Assume the distribution function of ϑ_t is $F(\vartheta_t)$ with finite support $V \equiv [\vartheta_{t_{min}}, \vartheta_{t_{max}}]$, and density $f(\vartheta_t)$, where $f(.)$ is continuously differentiable, with mean $\overline{\vartheta_t}$. Define X^+ and X, as the means in the upper and lower tails, for any candidate cut-off ϑ_t,

$$X^+ = E(\widetilde{\vartheta_t} | \widetilde{\vartheta_t} < \vartheta_t) \qquad (2)$$

$$X^- = E(\widetilde{\vartheta_t} | \widetilde{\vartheta_t} > \vartheta_t) . \qquad (3)$$

The expression (2) and (3) govern 'honor' and 'stigma' conferred by participation and abstention. Since the intrinsic valuation for money is below the average level expression (2) corresponds to virtue. The difference between the conditional moments defines net reputational gain, Ω (Benabou and Tirole, 2011):

$$\Omega = \mu\big(X^+(\vartheta_t) - X^-(\vartheta_t)\big), \ \forall \vartheta_t \in V \text{ and } \mu \text{ is fixed.} \qquad (4)$$

The utility function of the firm is

$$U = e_1 + \vartheta_t t - \psi(e_1 + e_2) + \Omega \qquad (5)$$

where $\psi(e_1 + e_2)$ is the cost of effort with $\psi'(.) > 0$ and $\psi''(.) > 0$.

Given the green norm and monetary transfer t, a firm exerts positive effort (i.e, $e_1 > 0$) if $\left.\frac{\partial U}{\partial e_1}\right|_{e_1=0} > 0$. Define a threshold level of ϑ_t, ϑ^*, such that

$$\left.\frac{\partial U}{\partial e_1}\right|_{e_1=0} = 1 - \psi'(e_1 + e_2) + \Omega'(\vartheta^*) = 0 . \qquad (6)$$

Assuming an interior solution, the net reputational incentive at the cut-off level is

$$\Omega(\vartheta^*) = \mu_t\big(X^+(\vartheta^*) - X^-(\vartheta^*)\big)$$

When more firms start doing 'the right thing', ϑ^* decreases, honor from scarcity value decreases, but stigma from abstention worsens. The effect on net reputational incentive depends on the relative strength of honor and stigma. If honor decreases, $\Omega'(\vartheta^*) > 0$, the decisions become substitutes—i.e., a firm that was among the few heroic early birds, withdraws participation when others join the program. An honor effect dominates when few firms participate in the program (i.e., "do the right thing"). A complementarity effect is observed when the stigma effect dominates (i.e., $\Omega'(\vartheta^*) < 0$)—when only a few deviants fail to comply with the norm.

We now explore the effect of the extrinsic incentive on the contribution level. A comparative static result shows,

$$\text{Max}_{e_1} U = U = e_1 + \vartheta_t t - \psi(e_1 + e_2) + \Omega$$

$$\frac{\partial e_1}{\partial t} = -\frac{\Omega_{e_1 t}}{-\psi'' + \Omega''} \tag{7}$$

where $\Omega_{e_1 t}$ is the cross partial derivative of net reputational payoff. The marginal effect of receiving money for doing the *right thing* reduces the reputational value (i.e., $\Omega_{e_1 t} < 0$). Monetary incentive crowds out intrinsic motive as a firm's social esteem gets adversely affected (by SOC, $-\psi'' + \Omega'' < 0$, as cost is convex and net reputational gain increases at a decreasing rate)—i.e., material incentives are effective neither for honor-driven behavior nor for stigma-driven behavior. This holds under the following conditions: (i) net cost of efforts is low; and (ii) efforts are easily observable (i.e., high μ).

Consumer's expected value of the project is

$$V = S - \big(1 - \pi(e_2)\big)E - (1 + \lambda)(C + t), \tag{8}$$

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

$$W = V + U$$

$$= S - \left(1 - \pi(1)\right)E - (1 + \lambda)(\beta - e_1) - (1 + \lambda)t + e_1 + \vartheta_t t - \psi(e_1 + 1) + \Omega$$

$$= S - \left(1 - \pi(1)\right)E - (1 + \lambda)\beta + (2 + \lambda)e_1 - (1 + \lambda - \vartheta_t)t - \psi(e_1 + 1) + \Omega \quad (9)$$

Consider now an aggregate exogenous shift in the distribution of firms' preferences such that average opinion becomes *green*, i.e., more and more firms' now think exerting effort to protect the environment is the right thing to do and abstain from such action is 'just not done'. The original distribution of intrinsic valuation shifts by θ, $F(\vartheta_t - \theta)$ with density $f(\vartheta_t - \theta)$ and with support: $v_\theta = [\vartheta_{t_{min}} + \theta, \vartheta_{t_{max}} + \theta]$. Net reputational return becomes: $\Omega_\theta(\vartheta_t) \equiv \Omega(\vartheta_t - \theta)$. We normalize ϑ_t such that $\Omega(\vartheta_t)$ is minimum at $\vartheta_t = 0$ and $\Omega_\theta(\vartheta_t)$ is minimum at $\vartheta_t = \theta$. Now, we can define type as follows. A firm can be one of two types—green and brown—based on ϑ_t— high ϑ_t leads to low reputational value (i.e., *brown* firm). Following the exogenous shift of the distribution of preferences, a firm's type remains the same. That is, a *green* type firm with low intrinsic valuation for money remains green and a *brown* type firm with high intrinsic valuation for money remains brown.[2]

Under complete information, the regulator maximizes the following objective function such that the firm i's participation constraint is satisfied, $i = G$ *(Green)*, B *(Brown)*,

---

[2] Since contribution to environmental protection is a public good, firms have incentive to free ride. But, in this paper, we are not modeling a game of strategic interaction between firms per se. We are considering a game between the regulator and one firm (holding all other firms' behavior constant, which we recognize as a strict restriction—one that future research can relax to examine the thorny issue of mechanism design, reputation, and free riding) (e.g., see Shogren 1987).

$$\max W = S - \left(1 - \pi(1)\right)E - (1+\lambda)\beta + (2+\lambda)e_1^i - \left(1+\lambda-\vartheta^i\right)t^i - \psi\left(e_1^i+1\right) + \Omega_\theta^{iG} \quad (10)$$

subject to $U^{iG} \geq \underline{U}^i$, where, $i = G, B$ and $\underline{U}^i$ is reservation utility of firm i. From the the binding participation constraint, we have,

$$t^i\left(e_1^i\right) = \frac{1}{\vartheta^i}\left[\underline{U}^i - e_1^i - \Omega_\theta^{iG} + \psi\left(e_1^i+1\right)\right], \text{ and } \frac{\partial t^i}{\partial e_1^i} = \frac{1}{\vartheta^i}\left[-1 - \Omega_\theta^{iG\prime} + \psi'\left(e_1^i+1\right)\right],$$

where, $\Omega_\theta^{iG\prime} = \frac{d\Omega_\theta^{iG}}{de_1^i}$.

Optimality requires,

$$(2+\lambda) - \left(1+\lambda-\vartheta^i\right)\frac{\partial t^i}{\partial e_1^i} + \Omega_\theta^{iG\prime} = \psi'\left(e_1^i+1\right)$$

$$\Rightarrow 1 + \vartheta^i + \Omega_\theta^{iG\prime} = \psi'\left(e_1^i+1\right) \quad (11)$$

This implies that the sum of marginal intrinsic satisfaction and marginal net reputational gain from environmental project should be equal to the marginal cost. Solving this we obtain optimal effort,e_1^{i*}; substituting e_1^{i*} into the participation constraint gives the optimal monetary transfer, t^{*i}, for firm i.

$$t^{*i} = \frac{1}{\vartheta^i}\left[U^{iB} - e_1^{i*} - \Omega_\theta^{iG} + \psi\left(e_1^{i*}+1\right)\right]. \quad\quad (12)$$

When the regulator knows the type of the firm, she can offer monetary transfer according to the firm's type and induce the firm to exert effort. No firm gets any information rent.

Under incomplete information, the regulator only knows that a firm can be one of two types – green (with probability *p*) and brown. A *green firm* does not want to hide her private information when the average opinion is green. When more firms join the 'green club' due to the

change in the community standard, she loses the scarcity value. A *brown* firm wants to exert effort, when others doing the same, with less or no money to avoid the shame form abstention. The regulator does not want to pay a brown firm when he can use this stigma to induce the firm and save some public fund. He also does not want to lose the green firm by reducing their honor value. The challenge for him then is to design the contract such a way that both types join in the social project. The contract induces both types if following binding constraints are satisfied.

Honor. The green firm faces an incentive problem. She wants to keep her leadership role even after others become green. The regulator wants to keep green firms in the project as they tend to do substantially more in investing in new green-innovation and practice compared to a brown firm who merely 'tick the box'. Their behavior help encourage future behavior that adapts *green* standard (see Barrett 1995; Tangney et al 2007)[3]. This type of firm will participate if she can maintain its honor value, even after a change in average behavior or opinion. That is, the firm's participation constraint, which will ensure that it will not be worse off by choosing the contract even when more firms become green, is satisfied,

$$U^{GG} \geq U^{GB} \Rightarrow \vartheta^G t^G = \Omega_\theta^{GB} - \Omega_\theta^{GG} \equiv \Delta \Rightarrow t^G(e_1^G) = \frac{\Delta}{\vartheta^G} \qquad (13)$$

where, Ω_θ^{GG} and Ω_θ^{GB} represent the net reputational gain of a green firm under the green and brown community standard. The RHS of the inequality (i.e., U^{GB}) shows the status quo level of utility of the green firm when the average norm is brown. She receives no transfer (i.e., $t = 0$) and exerts e_1^G at the status quo level.

Stigma. A *brown* firm's decision of taking part into the project is driven by stigma. When more firms join the project, it feels socially outskirt if it does not participate, which leads to a

---

[3] Andreoni and Petrie (2004) argue that presence of leaders (who consistently contribute more than others to a public good) in a group is important as, according to their experimental study, leaders' presence in a group appear to increase others' contribution to a public good.

stigma from abstention. Under asymmetric information, the brown firm may want to avoid this cost of stigma and be seen as green—it wants to mimic a green firm by hiding her private information when the average opinion is green. A brown type's incentive is to avoid social stigma by gaining reputation as good as the green type. Thus, the incentive compatibility constraint of the brown firm, which ensures that it will not be better off by mimicking green type's behavior, is as follows.

$$U^{BG} \geq U^{GG}$$

$$\Rightarrow \vartheta^B t^B = (e^G - e^B) - \left(\psi(e_1^G + 1) - \psi(e_1^B + 1)\right) + \left(\Phi_\theta^{GG} - \Omega_\theta^{BG}\right) + \frac{\vartheta^B}{\vartheta^G}\Delta$$

$$\Rightarrow t^B(e^B, e^G) = \frac{1}{\vartheta^B}\left[(e^G - e^B) - \left(\psi(e_1^G + 1) - \psi(e_1^B + 1)\right) + \left(\Phi_\theta^{GG} - \frac{\vartheta^B}{\vartheta^G}\Omega_\theta^{GG}\right) + \right.$$

$$\left.\left(\frac{\vartheta^B}{\vartheta^G}\Omega_\theta^{GB} - \Omega_\theta^{BG}\right)\right] \tag{14}$$

The term Φ_θ^{GG} represents a brown firm's net reputational gain when she pretends to be a green firm under green community standard. This term captures the reputation of the brown firm when it mimics the green firm's behavior by choosing the same level of effort e^G as that of the green firm.

The regulator wants to design a voluntary contract that would (i) induce both type of firms by satisfying their honor and stigma; and (ii) maximize social welfare which is a weighted average of two firm-types' and consumer's utilities,

$$Max\ W = p\left[S - (1 - \pi(e_2))E - (1 + \lambda)\beta + (2 + \lambda)e^G - (1 + \lambda - \vartheta^G)t^G(e^G) - \right.$$

$$\psi(e_1^G + 1) + \Omega_\theta^{GG}\right] + (1 - p)\left[S - (1 - \pi(e_2))E - (1 + \lambda)\beta + (2 + \lambda)e^B - (1 + \lambda - \right.$$

$$\left.\vartheta^B)t^B(e^B, e^G) - \psi(e_1^B + 1) + \Omega_\theta^{BG}\right].$$

Substituting binding constraints into the objective function, we have,

$$Max\ W = p\left[S - (1 - \pi(e_2))E - (1+\lambda)\beta + (2+\lambda)e^G - \left(\frac{1+\lambda-\vartheta^G}{\vartheta^G}\right)\Delta - \psi(e_1^G + 1) + \Omega_\theta^{GG}\right] +$$

$$(1-p)\left[S - (1 - \pi(e_2))E - (1+\lambda)\beta + (2+\lambda)e^B - \left(\frac{1+\lambda-\vartheta^B}{\vartheta^B}\right)\left\{(e^G - e^B) - \left(\psi(e_1^G + 1) - \right.\right.\right.$$

$$\psi(e_1^B + 1)\right) - \left(\Phi_\theta^{GG} - \Omega_\theta^{GG}\right) + \frac{\vartheta^B}{\vartheta^G}\Delta\right\} - \psi(e_1^B + 1) + \Omega_\theta^{BG}\right].$$

The necessary first order conditions imply,

$$e_1^G: (2+\lambda) - (1+\lambda - \vartheta^G)\frac{\partial t^G(e_1^G)}{\partial e_1^G} + \Omega_\theta^{GG'} - \frac{1-p}{p}(1+\lambda - \vartheta^B)\frac{\partial t^B(e_1^G, e_1^B)}{\partial e_1^G} = \psi'(e_1^G + 1)$$

$$\Rightarrow 1 + \vartheta^G + \Omega_\theta^{GG'} + \frac{H}{1-H}\left[\vartheta^G + \Omega_\theta^{GG'}\right] - \frac{1}{1-H}\frac{\vartheta^G}{1+\lambda}\left[\left(\frac{1+\lambda-\vartheta^G}{\vartheta^G}\right)\Omega_\theta^{GB'} + \left(\frac{1-p}{p}\right)\left(\frac{1+\lambda-\vartheta^B}{\vartheta^G}\right)\Delta'\right] =$$

$$\psi'(e_1^G + 1) \tag{15}$$

$$e_1^B: 1 + \vartheta^B + \Omega_\theta^{BG'} = \psi'(e_1^B + 1) \tag{16}$$

where, $H = \frac{\vartheta^G}{(1+\lambda)}\left[\frac{1+\lambda-\vartheta^G}{\vartheta^G} + \left(\frac{1-p}{p}\right)\left(\frac{1+\lambda-\vartheta^B}{\vartheta^B}\right)\right]$, and $(1 - H) = \frac{\vartheta^G}{\vartheta^B}\frac{[p(1+\lambda)-(1+\lambda-\vartheta^B)]}{p(1+\lambda)}$.

In expression (15), we consider $\Phi_\theta^{GG'} = 0$. This is because when the brown firm successfully mimics the green firm's behavior (which is captured by Φ_θ^{GG}), observers (i.e., citizens, members of the society) cannot update their prior beliefs regarding firms' types – green or brown. In such a situation, additional observed effort choice cannot have any effect on expectation formation regarding the true value of v and, thus, the green firm's marginal reputational value will be unaffected.

The second order sufficient conditions for expression (15) and (16) require,

$$\frac{\partial^2 W}{\partial e_1^{G^2}} = \frac{1}{\vartheta^G}\Omega_\theta^{GB''}[\{p\vartheta^G + (1-p)\vartheta^B\} - (1+\lambda)] - \frac{1}{\vartheta^G}\Omega_\theta^{GG''}[(1-p)\vartheta^B - (1+\lambda)] -$$

$$\frac{1}{\vartheta^B}\psi''(e_1^G + 1)[\vartheta^B - (1-p)(1+\lambda)] < 0 \tag{17}$$

$$\frac{\partial^2 W}{\partial e_1^{B^2}} = \frac{(1-p)(1+\lambda)}{\vartheta^B}\left[\Omega_\theta^{BG''} - \psi''(e_1^B + 1)\right] < 0 \tag{18}$$

To satisfy the SOC for the green firm's optimization problem (i.e., expression (17)), it would be sufficient to have the following: (i) $\vartheta^B > (1+\lambda)$; and (ii) $\vartheta^G > (1+\lambda)$, since $\psi''(e_1^G + 1) > 0$, $\Omega_\theta^{GG''} < 0$, $\Omega_\theta^{GB''} < 0$. Following these two conditions, in expression (15), $H < 0$ and $(1-H) > 0$. In case of the brown firm, the SOC is satisfied as $\Omega_\theta^{BG''} < 0$, and $\psi''(e_1^B + 1) > 0$.

Expression (15) shows the green firm's optimal effort under asymmetric information case. The first three terms are same as complete information case. Compared to the complete information, she incurs a marginal motivational costs due to a loss of her leadership role (captured by the first square bracket term, as $\frac{H}{(1-H)} < 0$) and enjoys a marginal benefit in reputational value if she could retain some of her leadership rents (captured by the last square bracket term). Her optimal effort then depends on her relative net marginal reputational gain— she would over-comply if her net marginal reputational gain dominates over her net reputational loss as she loses her leadership role (i.e., if the last square bracket term dominates the first square bracket term).

This finding contradicts Qin and Shogren (2015) who argue that the green firms always over-comply. Previous studies on voluntary compliance by firms document over-compliance by 'green' firms—conceptually and empirically (e.g., McClelland and Horowitz 1999; Bansal and Gangopadhaya 2008; Gunningham, Kagan, and Thorton 2002; Haan 2015). This contradiction may be because previous studies do not consider the possibility of relative reputational gain/loss

by incorporating the effect of social norm on social preferences. As we add this, we encounter this motivational costs incurred by the green firm. Standard empirical methodology may not be able to capture this effect, as the behavior of small number of true 'green leader' firms may not affect the overall empirical finding significantly. Nevertheless, it is important to understand and address this trade-off by the green firms to keep them in the pool to encourage and maintain participation by others in a social project. Andreoni and Petrie (2004) supports such conjecture that leaders' presence in a group appear to increase others' contribution to a public good.

In addition, as the binding participation constraint (13) reveals, the green firm needs extra rent to compensate the relative reputation loss due to a change in average behavior (i.e., scarcity rents). This extra rent may not be a monetary compensation as that would crowd out the private incentive of the green firm. We argue this needs to be a 'social reward' that recognizes and celebrates this firm's 'early bird' role in environmental protection when the average behavior was still brown. This finding is in line with Besley and Ghatak (2008) who argue that motivated agents need a 'status incentive' to participate in a pro-social programme.

Compared to the full information case, the brown firm would exert optimal effort (expression (16)), which is similar to Banerjee and Shogren (2010). Using the 'stigma' tool, given that the firm's effort level is observable, the regulator can ensure optimal contribution from the brown firm which may not be possible otherwise. This finding supports the view that stigma can enhance social welfare by increasing the social costs of non-compliance (see Rasmusen 1996).[4] Examples abound— toxic chemicals reduction by Toxic Releases Inventory in the US, NHS hospital performance improvement in England. The brown firm's binding incentive compatibility constraint (expression (14)) can be expressed as,

---

[4] As sociologists have emphasized, however, stigmatization can backfire and reduce social welfare—welfare falls if a person who failed one-time to comply with social norms is then forced or induced to remain deviant due to the stigma (e.g., welfare-trap, creation of career criminals) (see Goffman 1963; Furuya 2002).

$$U^{BG} = U^{GB} - \frac{(\vartheta^G - \vartheta^B)}{\vartheta^G} \Omega_\theta^{GB} + \left(\Phi_\theta^{GG} - \frac{\vartheta^B}{\vartheta^G} \Omega_\theta^{GG} \right).$$

Since the true green reputational value dominates the *pretended* green reputation ($\Phi_\theta^{GG} <$ Ω_θ^{GG}), $\left(\Phi_\theta^{GG} - \frac{\vartheta^B}{\vartheta^G} \Omega_\theta^{GG} \right) < 0$. The brown firm may want to sacrifice some of its information rents if $\left| \Phi_\theta^{GG} - \frac{\vartheta^B}{\vartheta^G} \Omega_\theta^{GG} \right| > \left| \frac{(\vartheta^G - \vartheta^B)}{\vartheta^G} \Omega_\theta^{GB} \right|$, i.e., if the stigma of being brown dominates the honor of the green firm. If honor value is too high, not having the same honor may not be viewed as irresponsible. But, if achieving that honor level is not too difficult (i.e., everybody does it), then not reaching to that level would lead to a stigma value. The brown firm may then sacrifice rents to overcome that stigma effect. Otherwise, the brown firm may be better off by not giving up information rents to buy a *green* reputation (when the green firm's baseline reputation is too high before the norm-shift). This finding contradicts previous studies (e.g., Banerjee and Shogren 2012; Bowles and Hwang 2008) which demonstrate that reputation seeker will always buy reputation. These studies did not explore the possibility of interaction between honor and stigma and preferences. Once we add such a possibility, a firm compares her own reputation before and after the norm (i.e., average opinion) shift and she compares her reputational value with other type of firm and the trade-off therein. Our work shows that the brown firm's information rents depend on the relative strength of honor and stigma.

3. Concluding remarks

Honor and stigma are two moral emotions, given a social norm, that can affect a firm's effort toward improving environmental protection. Herein we design an incentive mechanism that incorporates social preferences and norms by allowing the interplay between honor and stigma generated from the complementarity and substitutability effects between individual firm's behavior and average behavior in a community. We identify a potential motivational cost for an

intrinsically motivated firm once we add the effect of endogenous norms on social preferences. Our results show that (a) an additional social reward can induce the 'early birds' who used to be green even before other firms undertake voluntary environmental protection actions; (b) that a green firm's over-compliance would depend on her loss/gain of her leadership role; and that (c) a decision maker can protect the environment at a lower cost by allowing firms who are merely interested in social reputation to purchase a 'social responsibility reward'.

Our specific assumption about the social norm is stronger than concluding that the proposed mechanism incorporating honor and stigma is welfare enhancing. This will depend on the relationship between the norm and policy—in particular, whether the policy enforcement and the existing norms are conflicting (e.g., see Acemoglu and Jackson 2013). For example, historically some of British Laws in colonies were unsuccessful because those policies conflicted with existing norms. In contrast, the civil-right legislation in the US had a pervading impact on society as the law and norm-change went hand-in-hand (see Parsons (2013) and Wright (2013) as cited in Acemoglu and Jackson, 2013). In our case, the successful implementation of our mechanism is conditional on well-accepted existing social norm. Otherwise, non-compliance will not influence the social costs of stigma and participation will not ensure honor value. Moreover, our results of honor-value loss and stigma-costs depend crucially on the exogenous shift of the average behavior (or, norm).

Our claim about cost-effectiveness can compete against the costs of influencing such a shift by the policy maker. Public displays, education, mass-awareness, campaign can be useful means of establishing and influencing a norm. For example, recently the newly elected Prime Minister of India launched a huge campaign called 'Clean India Campaign' ('Swach Bharat Avijan') through which he induced participation of schools, colleges, Govt. offices, media,

popular film-actors, and crickters. The government of India introduced a fixed levy for cleaner India without much controversy or protest. Such a campaign might have created a favorable social norm for introducing this monetary tool. It is unclear though whether the costs and benefits of the campaign balanced out. Future research could explore such possibilities and their welfare impact in more details.

References

Acemoglu, D. and Jackson, M.O., 2014. Social norms and the enforcement of laws (No. w20369). National Bureau of Economic Research.

Akerlof, G. 1980. A theory of social custom, of which unemployment may be one consequence. Quarterly Journal of Economics 94(4): 749-775.

Allcott, H., 2011. Social norms and energy conservation. Journal of Public Economics 95(9-10): 1082-1095.

Arce, D. G., 2012. Principals' preferences for agents with social preferences. Journal of Economic Behavior & Organization, 90: 154-163.

Arora, S., and Gangopadhyay, S., 1995. Toward a theoretical model of voluntary overcompliance. Journal of Economic Behavior and Organization, Vol. 28, pp. 289–309.

Banerjee, P., and Shogren, J.F., 2010. Regulation, Reputation, and Environmental Risk. Economics Letters 106: 45-47.

_____ , 2012. Material Interests, Moral Reputation, and Crowding out Species Protection on Private Land. Journal of Environmental Economics and Management 63(1): 137-149.

Batson, D., 1967. Altruism and Prosocial Behavior, in: D. Gilbert, S. Fiske, G. Lindzey (Eds.), Handbook of Social Psychology, Vol. 2, McGraw Hill, New York, pp. 282–316.

Bénabou, R., and Tirole, J., 2006. Incentives and prosocial behavior. American Economic Review 96 (5): 1652–1678.

_____, 2011. Laws and social norms. NBER Working Paper.

Besley, T., and Ghatak, M., 2007. Retailing public goods: The economics of corporate social responsibility. Journal of Public Economics, 91: 1645-1663.

_____ , 2008. Status Incentives. American Economic Review 98(2): 206-11.

Besley, T., Konrad, B. B., and Bevan, G., 2009. Naming and shaming: the impacts of different regimes on hospital waiting times in England and Wales. Discussion paper, 7306. Centre for Economic Policy Research, London, UK.

Bevan, G., and Wilson, D., 2013. Does 'naming and shaming' work for schools and hospitals? Lessons from natural experiments following devolution in England and Wales. Public Money & Management, 33:4, 245-252.

Bowles, S., and Hwang, S., 2008. Social preferences and public economics: mechanism design when social preferences depend on incentives. Journal of Public Economics 92 (8–9): 1811–1820.

Bowles, S., and Polanía-Reyes, S., 2012. Economic incentives and social preferences: substitutes or complements? Journal of Economic Literature 50(2): 368-425.

Bowles, S., 1998. Association Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. Journal of Economic Literature, 36 (1): 75-111.

_____ , 2008. Policies designed for self-interested citizens may undermine the moral sentiments: evidence from economic experiments. Science 320: 1605–1609.

Brennan, G. and Pettit, P., 2004. The Economy of Esteem, Oxford: Oxford University Press.

Bryan, J. H., and Test, M. A., 1967. Models and Helping: Naturalistic Studies in Aiding Behavior. Journal of Personality and Social Psychology, 6 (4): 400-4007.

Fischer, P., and Huddart, S., 2008. Optimal Contracting with Endogenous Social Norms. American Economic Review 98(4): 1459-1475.

French, P., 2002. Honor, Shame, and Identity. Public Affairs Quarterly, Vol. 16, No. 1 (Jan.),. 1-15.

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

Fung, A. and ORourke, D., 2000. Reinventing Environmental Regulation from the Grassroots Up: Explaining and Expanding the Success of the Toxics Release Inventory. Environmental Management, 25 (2), 115-127.

Funk, P., 2006. Modern voting tools, social incentives and voter turnout: Theory and evidence. Mimeo, Universitat Pompeu Fabra.

Furuya, K., 2002. A socio-economic model of stigma and related social problems. Journal of Economic Behavior & Organization, 48: 281-290.

Genius, M., Koundouri, P., Nauges, C., and Tzouvelekas, V., 2013. Information Transmission in Irrigation Technology Adoption and Diffusion: Social Learning, Extension Services, and Spatial Effects. American Journal of Agricultural Economics 96(1): 328-344.

Gerber, A. S., Green, D. P., and Larimer, C. W., 2008. Social pressure and voter turnout: Evidence from a large scale field experiment. American Political Science Review, 102, 33–48.

Goffman, E., 1963. Stigma: Notes on the Management of Spoiled Identity. Simon and Schuster, New York.

Gunningham, N., Kagan, R. A., and Thorton, D., 2002. Social License and Environmental Protection: Why Business Go Beyond Compliance. Law & Social Inquiry.

Hermalin, B.E,. 1998. Toward an economic theory of leadership: leading by example. American Economic Review 88(5): 1188-1206.

Kropf, M., and Knack, S., 2003. Viewers like you: Community norms and contributions to public broadcasting. Political Research Quarterly, 56(2), 187–197.

Kuhfuss, L., Preget, R., Thoyer, S., Hanley, N., Le Coent, P., and Desole, M. 2015. Nudges, Social Norms and Permanence in Agri-Environmental Schemes, http://purl.umn.edu/204233

Laffont, J.J., 1995. Regulation, moral hazard and insurance of environmental risks. Journal of Public Economics 58: 319–336.

Lea, S. E. G., and Webbley, P., 1997. Pride in economic psychology. Journal of Economic Psychology, 18: 323-340.

Lindbeck, A., Nyberg, S. and Weibull, J.W., 1999. Social norms and economic incentives in the welfare state. Quarterly Journal of Economics 114 (1): 1-35.

McClelland, J. D., and Horowitz, J. K., 1999. The Costs of Water Pollution Regulation in the US. Pulp and Paper Industry. Land Economics 75(2):220-32.

Parsons, T., 2010. The Rule of Empires: Those Who Built Them, Those Who Endured Them, and Why They Always Fall. Oxford University Press.

Posner, R., and Rasmusen, E., 1999. Creating and enforcing norms, with special reference to sanctions. International Review of Law and Economics, 19, 66–84.

Potters, J., Sefton, M., and Vesterlund, L., 2007. Leading-by-example and signaling in voluntary contribution games: an experimental study. Economic Theory 33: 169-182.

Qin, B., and Shogren, J. F., 2015. Social norms, regulation, and environmental risk. Economics Letters, 129: 22-24.

Rasmusen, E., 1996. Stigma and self-fulfilling expectations of criminality. Journal of Law and Economics 39,519–543.

Shogren, J. F. 1987. Negative Conjectures and Increased Public Good Provision. Economics Letters 23, 181-184.

Tadelis, S., 2007. The power of shame and the rationality of trust. Available at SSRN 1006169.

Tangney, J. P., Stuewig, J. and Mashek, D. J., 2007. Moral Emotions and Moral Behavior. Annual Review of Psychology, 58: 345-372.

Tavoni, A., Dannenberg, A., Kallis, G., and Löschel, A., 2011. Inequality, communication, and the avoidance of disastrous climate change in a public goods game. Proceedings of the National Academy of Sciences 108: 11825-11829.

Vesterlund, L., 2003. Informational value of sequential fundraising. Journal of Public Economics 87: 627–657

Wright, G., 2013. Sharing the Prize: The Economics of the Civil Rights Revolution in the American South. Harvard University Press, Cambridge, MA.