KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY
Department of Business Economics and Public Policy

Greenwash: Corporate Environmental Disclosure under Threat of Audit*

Thomas P. Lyon† and John W. Maxwell‡

May 23, 2007

Abstract

We present an economic model of greenwash, in which a firm strategically discloses environmental information and a non-governmental organization (NGO) may audit and penalize the firm for failing to fully disclose its environmental impacts. We show that disclosures increase when the likelihood of good environmental performance is lower. Firms with intermediate levels of environmental performance are more likely to engage in greenwash. Under certain conditions, NGO punishment of greenwash induces the firm to become less rather than more forthcoming about its environmental performance. We also show that complementarities with NGO auditing may justify public policies encouraging firms to adopt environmental management systems.

1 Introduction

The most notable environmental trend in recent years has been the shift away from traditional regulation and towards voluntary programs by government and industry. Thousands of firms participate in the Environmental Protection Agency's partnership programs, and many others participate in industry-led environmental programs such as those of the World Business Council for Sustainable Development, the Chicago Climate Exchange, and the American Chemistry Council's "Responsible Care" program.[1] However, there is growing scholarly concern that these programs fail to deliver meaningful environmental

*We would like to thank Mike Baye, Rick Harbaugh, Charlie Kolstad, John Morgan, Michael Rauh, and participants in seminars at the American Economic Association meetings, Canadian Resource and Environmental Economics meetings, Dartmouth, Indiana University, Northwestern, UC Berkeley, UC Santa Barbara and University of Florida for their helpful comments.

†Ross School of Business, School of Natural Resources and Environment, and Erb Institute, University of Michigan.

‡Kelley School of Business, Indiana University.

[1] Voluntary programs have proliferated in Europe, Japan, and developing countries, as well. For an introduction to voluntary environmental programs and corporate environmental strategy, see Lyon and Maxwell (2004b).

improvement.[2] Furthermore, non-governmental organizations (NGOs) increasingly complain that voluntary corporate efforts are mere "greenwash."

At the 2002 Earth Summit in Johannesburg, a group of NGOs held a Greenwash "Academy Awards" event to criticize companies that falsely promote themselves as environmentally responsible and to "recognize these companies for what they are: hypocrites." Winner for Best Greenwash was "BP for their Beyond Petroleum rebranding campaign," which highlights the company's investments in renewable energy without mentioning their major efforts in petroleum exploration. Among the other awards, South African electricity firm Eskom was Runner up for Best Picture "for being a key member of Business Action for Sustainable Development while generating electricity from coal and nukes." Monsanto was Runner Up for the Lifetime Achievement Award for its "tireless promotion of Roundup Ready GM [genetically modified] crops as a solution to world hunger."[3]

Ralph Nader reveals a similar skepticism regarding corporate social efforts:

> "One recent misstep is the U.N.'s 'Global Compact.' With the disappointing support of some international human rights and environmental organizations, the U.N. has asked multinational corporations to sign on to the compact's unenforceable and overly vague code of conduct. Companies are able to sign on to the compact and 'bluewash' themselves, as critics at the Transnational Research and Action Center in San Francisco have labeled the effort by image-impaired corporations to repair public perceptions by hooking up with the U.N..."[4]

To what extent can market and stakeholder pressures internalize corporate social and environmental impacts without resort to government regulation? There is a growing literature showing that if stakeholders wield strong enough threats they can force firms to internalize externalities. For example, Segerson and Miceli (1998) and Maxwell, Lyon and Hackett (2000) show that the threat of regulation can lead firms to make welfare-enhancing preemptive investments in environmental abatement.[5] Similarly, Baron and Diermeier (2005) and Innes

[2] In a recent review of government voluntary programs by the National Research Council, Wilbanks and Stern (2002, p. 337) concluded that "[T]he effects of voluntary measures appear so far to have been rather modest." King and Lenox (2000) actually find that chemical companies that participated in the industry's Responsible Care program achieved fewer reductions of toxic chemical emissions than did non-participants. In a recent review of voluntary programs, Morgenstern and Pizer (2007, p. 184) conclude: "[N]one of the case study authors found truly convincing evidence of dramatic environmental improvements. Therefore, we find it hard to argue for voluntary programs where there is a clear desire for major changes in behavior."

[3] For details, see "Greenwash Academy Awards Announced at Earth Summit," http://www.corpwatch.org/article.php?id=3648.

[4] Ralph Nader, "Corporations And The UN: Nike And Others "Bluewash" Their Images," San Francisco Bay Guardian, September 18, 2000. Available at http://www.commondreams.org/views/091900-103.htm

[5] Lyon and Maxwell (2004) present a thorough analysis of corporate environmental initiatives conducted against the backdrop of regulatory threats.

(2006) show that an NGO's threat of a corporate boycott can induce companies to undertake environmental improvements.

Previous analyses, however, have not focused on the ability of the firm to manipulate the flow of information to external stakeholders. Is the governance role played by stakeholder pressure undermined when firms control information flows?[6] Much of financial regulation is intended to prevent managers from exploiting information advantages to the detriment of investors, but these safeguards were not designed with broader environmental and social concerns in mind.[7] The literature linking stakeholder pressure and corporate environmental information flows is still quite small, much of it empirical work in accounting journals.[8] Economic analyses typically focus on NGO communications, as the firm is assumed to be unable to credibly communicate environmental information.[9] However, none of these papers study direct corporate communication regarding social and environmental issues, which is fundamental to understanding the phenomena of bluewash and greenwash. In the remainder of this paper, we will focus on greenwash in order to streamline the discussion, but it should be clear that the phenomenon encompasses social as well as environmental issues.

In this paper, we present what is to our knowledge the first economic analysis of greenwash. We construct a financial disclosure model in the spirit of Shin (2003) in which a company conducts multiple activities with environmental impacts that may turn out well or turn out poorly, and has the option whether or not to reveal its performance on any activities. To this basic structure we add an NGO that may audit corporate performance when the firm does not fully disclose its performance information, and punishes greenwash when it is discovered. We characterize fully how the possibility of NGO punishment influences the firm's disclosure decisions, and show how corporate disclosure varies with the firm's probability of success in its environmental activities, and the probability the firm is informed about the outcome of its activities at the time it makes a disclosure. We show that disclosures increase when the likelihood of good environmental performance is lower. In particular, we find that firms with intermediate levels of environmental performance are more likely to engage in greenwash. We find that NGO auditing strengthens incentives for some types of firms to disclose fully, but deters other types of firms from disclosing at all. We also consider complementarities between NGO auditing of greenwash and corporate adoption of an environmental management system (EMS), thereby providing a new rationale for encouraging firms to adopt EMSs.

Our focus is on the optimal firm response to NGO campaigns against greenwash, rather than on optimal NGO behavior in itself. The literature suggests

[6] There has been considerable work on government-mandated information disclosure programs. For an overview, see Tietenberg (1998). However, these analyses do not address the problems that arise when the firm controls the flow of information.

[7] See, for example, Dewatripont and Tirole (1994) or Goodhart et al. (1998).

[8] See, for example, Patten (1992), Blacconiere and Patten (1994), and Deegan and Rankin (1996). These papers lack a firm economic underpinning; they are either atheoretical or grounded in sociological notions such as legitimacy theory.

[9] For example, Feddersen and Gilligan (2001) study NGO endorsements of particular firms' products, and Lyon and Maxwell (2004a) study corporate funding of NGO communications.

a variety of motivations for NGOs, and there is no accepted model comparable to the model of maxizing profits for firms or maximizing re-election probability for politicians. Some economic models assume NGOs attempt to maximize environmental benefits (Innes 2006), and others assume they have an objective that deviates from social welfare in some linear fashion (Grossman and Helpman 2001). Baron (2003) argues that some NGOs may be "intransigent" types that do not behave as rational actors. Here we opt to model NGO behavior, rather than assuming any particular objective function for NGOs. As we argued above, the evidence shows that many NGOs attack firms they see as engaging in greenwash, and that is the behavior we incorporate into our model.

The remainder of the paper is organized as follows. Section 2 presents our definition of greenwash, and distinguishes it from other "disinformation" strategies. In section 3, we present the basic disclosure model without NGO auditing. Section 4 adds an NGO that audits firm performance whenever the firm does not fully disclose its performance information, and penalizes greenwash whenever it is discovered. We characterize the pure strategy equilibria of this game. Section 5 provides a full characterization of the model's equilibria as a function of changes in the expected penalty for greenwash. In section 6, we use our analysis to draw out testable hypotheses for empirical study, and to consider the link between greenwash and whether a company has implemented an environmental management system. Section 7 concludes.

2 What is Greenwash?

Formal analysis of greenwash requires a clear definition of the phenomenon. Unfortunately, popular usage of the term tends to be broad and vague; indeed, in their book on greenwash, Greer and Bruno (1996) never actually define the term.[10] *Webster's New Millenium Dictionary of English* defines greenwash as "The practice of promoting environmentally friendly programs to deflect attention from an organization's environmentally unfriendly or less savory activities." The *Concise Oxford English Dictionary (10th Edition)* defines it as: "Disinformation disseminated by an organization so as to present an environmentally responsible public image; a public image of environmental responsibility promulgated by or for an organization etc. but perceived as being unfounded or

[10] Even academic discussions can be surprisingly broad. Laufer (2003), for example, presents a set of elements of greenwashing that include "confusion," "fronting," and "posturing." Confusion (p. 257) is achieved through "careful document control and strict limits on the flow of information made available to regulators and prosecutors." Fronting (p. 257) "is realized by subordinate scapegoating or reverse whistle blowing," and may involve such strategies as "cast doubt on the severity of the problem" or "emphasize uncertainty associated with the problem." Posturing (p. 256) involves the use of "front groups" to influence legislation or suggest that particular policies enjoy widespread "grassroots" support. While we find these distinctions useful, in our view, these activities differ too much to be viewed as a single phenomenon; indeed, we have already modeled the use of "astroturf lobbying" through "front groups" in Lyon and Maxwell (2004a). Astroturf lobbying involves the provision of soft information targeted at a public decisionmaker to influence policy decisions. Greenwash involves public disclosure of hard information targeted to influence shareholder value.

intentionally misleading." Both of these definitions emphasize the idea that the public has limited information about corporate environmental performance, and that corporations therefore can manipulate the dissemination of information to mislead the public.

The term "disinformation" implies the provision of deliberately false or fraudulent messages. To us, however, corporate greenwashing does not seem to fit this definition. Instead, the typical concerns raised by NGOs are that companies present positive information out of context in a way that could be misleading to individuals who lack background information about the company's full portfolio of activities. Consider the following example, taken from *Don't Be Fooled: The Ten Worst Greenwashers of 2003*:[11]

> "Royal Caribbean points to its advanced wastewater treatment systems as a sign of environmental progressiveness, yet they are installed on just 3 of the company's 26 cruise ships. The advanced systems are only found on its Alaskan fleet, which due to Alaskan law are subject to the strictest environmental standards in the industry. Royal Caribbean deems them unnecessary on cruise ships that travel other routes."

This example, like those outlined in the Introduction, depicts a company making a statement that is true, yet not the whole truth. It supports our view, which is that greenwash can be defined as the *selective disclosure of positive information about a company's environmental or social performance, without full disclosure of negative information on these dimensions.*[12]

An interesting example of selective disclosure comes the Department of Energy's Voluntary Greenhouse Gas Reporting program, created by section 1605b of the Energy Policy Act of 1992. Kim and Lyon (2006) show that electric utility participants in the 1605(b) program reported reductions in their greenhouse gas emissions during the period 1995-2003, but their actual emissions rose. Furthermore, during the same period, non-participant utilities reduced their emissions. This misleading reporting behavior is not illegal, for the program allows participants great flexibility in how they choose to report emissions reductions. In particular, firms can choose to report at the "project level" or the "entity level." The former allows a firm to report only on the outcomes of successful projects, while remaining silent about its aggregate performance. This is precisely what we mean by the term greenwash.

3 The Basic Disclosure Game

Our model focuses on a single firm, whose stock is traded publicly, and a nongovernmental organization (NGO). The firm has N different activities that each

[11] See Johnson (2003).

[12] Empirical research in accounting suggests that this is a common practice for firms that choose to engage in corporate environmental disclosure; see, for example, Deegan and Rankin (1996).

have some potential environmental impact.[13] The magnitude of N is assumed to be common knowledge, e.g., available on the firm's web site or Annual Report; the non-environmental aspects of the firm's operations are assumed to be already incorporated into the firm's market value. However, the firm's environmental profile is not known at the outset of the model. We assume the market sets the firm's value at its actuarily fair level.[14]

There are 3 periods. Let V_t represent the expected value of the firm in period t. At period 0, there is common knowledge about the likelihood there will be a liability associated with any given activity. Each activity generates for the firm a "success" of value u (e.g., an outcome that improves the firm's public image) with probability $r \in (0,1)$, and a "failure" of value $d < u$ with probability $1-r$.[15] Thus, the expected number of environmental failures the firm faces is simply $(1-r)N$. Its market value in period 0 is

$$V_0 = N(ru + (1-r)d) + \widetilde{V}, \tag{1}$$

where $\widetilde{V}$ is the total value created by the firm aside from its environmental impacts. Throughout the remainder of the paper, we will simplify notation by normalizing $\widetilde{V}$ to 0. At period 2, all information about environmental impacts is revealed and becomes common knowledge, and is incorporated into stock prices. The important action in the model takes place in the interim period 1, during which the manager attempts to influence the firm's stock price through the information he discloses.[16]

We assume there is a probability $\theta \in (0,1)$ that the manager actually learns the environmental impact of the activity by period 1.[17] Thus, at the interim period, the expected number of activities for which the manager has information on environmental outcomes is θN. The expected number of activities known to have environmental liabilities at the interim period is $\theta(1-r)N$. The manager has the ability to disclose publicly the number of activities that are known to be successes. We assume that all such disclosures are verifiable by outside parties. Thus, the manager is free to selectively withhold information, but he cannot actually lie to outsiders. We assume the manager adopts a disclosure strategy that maximizes the value of the firm.

Let n be the actual number of activities whose liabilities are known at the

[13] We refer to environmental impacts for concreteness, but could just as easily refer to corporate social responsibility more generally.

[14] The model draws upon the work of Shin (2003), but departs from it by using an additive rather than a multiplicative structure for payoffs, and by incorporating monitoring and punishment of greenwash.

[15] Note that d may be greater than zero, since the project itself may add to corporate value even if its environmental impact is unfavorable. What is critical is that $u > d$, that is, projects that are viewed as harmful to the environment diminish firm value relative to environmentally successful projects.

[16] There are many reasons a manager wants to influence the stock price, e.g. compensation packages that are linked to stock price performance. For further details, see Milgrom and Roberts (1992).

[17] It is worth noting that we would expect θ to be greater for firms that have created an environmental management system. We return to this issue in section 6.

interim period, s be the number of successes and f the number of failures, so that $n = s + f$. Let the manager's disclosures of the number of successes and failures be given by $\widehat{s}$ and $\widehat{f}$. We assume $V_1 = E(V_2)$. If the market knows s and f, as would be the case if the manager fully disclosed his information in period 1, then

$$V_1 = E(V_2) = us + df + (N - s - f)(ru + (1 - r)d), \tag{2}$$

where u equals the additive impact of an environmental success on the firm's value and d equals the additive impact of an environmental failure on the firm's value.

If the manager discloses $\widehat{s} > 0$, and the total number of disclosures $\widehat{s} + \widehat{f}$ is less than N, the NGO may investigate the manager's report for the possibility of greenwash (i.e., that the manager has a bad outcome that he failed to disclose).[18] With probability α the NGO obtains hard (verifiable) information about the true values of s and f at the interim period and mounts a successful campaign against the firm that imposes a punishment of cost P on the firm; with probability $1-\alpha$ it learns nothing and takes no action against the firm. The punishment might come about because the NGO triggers a consumer boycott, because it creates an advertising campaign that damages the firm's value, or through some other channel that the firm finds costly.[19] We will use the notation $\eta = \alpha P/(u - d)$ to indicate the "cost/benefit ratio" for greenwash, where αP is the expected penalty for greenwash, and $u - d$ represents the maximum value the firm could possibly obtain from successful greenwash.

We are interested in Perfect Bayesian Equilibria (PBE), which involve specifying a disclosure strategy for the manager, a market valuation, and a set of NGO and market beliefs for each time t such that (a) the disclosure strategy $(\widehat{s}, \widehat{f})$ is a best response mapping for a firm with actual environmental profile (s, f), given the market's pricing policy and the beliefs of the market and the NGO, (b) $V_1 = E(V_2)$ given the market's beliefs at period 1 and the manager's disclosure strategy, and (c) at period 0 the market believes the expected number of environmental failures is rN, and at period 1 it computes the expected number of environmental failures using Bayes' rule, conditional on any environmental reports. We will focus on pure strategy equilibria in section 4, and provide a full characterization of the model's equilibria in section 5.

It is easy to see that if the market believed the manager always truthfully disclosed all successes and failures, then the manager would have incentives to report $f = 0$. The expected value of an activity whose social impact is unknown is greater than the value of a failure, that is, $ru + (1 - r)d > d$. As a result, the manager always prefers to minimize the number of failures reported, and report only the successes; full disclosure is not an equilibrium strategy.[20]

[18] To simplify the analysis, we will assume the NGO commits *ex ante* to audit with fixed probability whenever $\widehat{n} < N$.

[19] Baron and Diermeier (2005) present a model of strategic NGO activism in which firms are punished for bad social outcomes, rather than being punished for greenwashing.

[20] Shin (2003) refers to the strategy of not disclosing any failures as "sanitization," but

If the manager follows a strategy of partial disclosure in equilibrium, and the market knows this, then the firm's expected value at the interim stage is

$$V_{PD} = us + (N - s)(qu + (1 - q)d), \tag{3}$$

where

$$q = \frac{r - \theta r}{1 - \theta r} \tag{4}$$

is the probability of success of an activity conditional on the fact that the manager has not disclosed information about that activity. Note that this expression has the same structure as equation (2), except that r (the ex ante probability that an activity succeeds) in (2) is replaced by q (the conditional probability that an undisclosed activity succeeds) in (3). The partial disclosure equilibrium can be supported by a set of off-equilibrium beliefs on the part of the market that if the manager ever reports $f > 0$, then all undisclosed outcomes are failures.[21]

4 Pure-Strategy Equilibria in the Disclosure Game with NGO Auditing

In this section we assess how auditing by an NGO affects the manager's incentives to make environmental disclosures. We fully characterize the set of pure-strategy Perfect Bayesian Equilibria that can emerge in the model, and show how they are related to the underlying parameters of the model. This analysis prepares us for a detailed examination in section 5 of the full set of equilibria and how they change with expected penalties for greenwash.

It is natural to ask whether the NGO can effectively punish partial disclosure without auditing, e.g. by penalizing the firm retroactively based on the ultimate outcomes in period 2. It turns out this is not possible. As we noted in section 2 above, punishing partial disclosure is distinct from simply punishing the firm for bad social outcomes. Punishing partial disclosure involves punishing firms that were aware of, but failed to disclose, a failure. At period 2, however, all the NGO knows is the ultimate number of failures, *not* the number that were known at the interim period. Thus, it is impossible to punish partial disclosure *per se* by only observing period 2 outcomes. Instead, it is essential to have some sort of independent auditing structure in period 1. This is the issue to which we now turn.

4.1 The Disclosure Game with NGO Auditing

In order to keep the analysis tractable and focused, we present it in the context of a model with $N = 2$. This is the simplest setting in which partial disclosure

does not distinguish situations where the firm has positive as well as negative news to report, which are the sorts of situations in which greenwash may become a problem.

[21] While this is not the only set of off-equilibrium beliefs that support the partial disclosure strategy, it is the simplest.

can emerge as an equilibrium outcome.[22] We will use the notation $V_1(\widehat{s}, \widehat{f})$ to indicate the market's valuation of the firm at period 1 when the manager discloses $(\widehat{s}, \widehat{f})$. Note that when $\widehat{n} \equiv \widehat{s} + \widehat{f} = 2$ the market has no problem inferring the firm's true value, since information disclosures are verifiable. These values are easily seen to be $V_1(0,2) = 2d$, $V_1(2,0) = 2u$, and $V_1(1,1) = u + d$. It is only in states where $\widehat{n} \equiv \widehat{s} + \widehat{f} < 2$ that we must carefully analyze the market's inference problem. (It is also worth noting that if the firm faced no penalties it would always pursue the strategy of partial disclosure, because it raises the firm's value; this is precisely the case treated above in section 3.)

We focus on the case in which the true state is $(1,1)$, as this is the only possible case—for $N = 2$—when partial disclosure can occur. Specifically, partial disclosure consists of reporting $(1,0)$ in state $(1,1)$. This is the type of behavior that we label greenwash. The firm receives no punishment for any report except when the state is $(1,1)$ and the manager reports $(1,0)$. Hence our focus is on what the manager will report when $(s,f) = (1,1)$. There are four reporting possibilities: $(\widehat{s}, \widehat{f}) \in \{(0,0),(1,0),(0,1),(1,1)\}$. Given the arguments we have made above, however, it is clear that the manager will never report $(\widehat{s}, \widehat{f}) = (0,1)$.

In order to understand the manager's reporting incentives, we must know how the market will interpret each of the three possible reports. Consider them in turn. The probability that the state is actually $(1,1)$ can then be computed via Bayes' Rule. Table 2 below presents the prior probability of each state at the interim period, along with the value the market attaches to that state. It is easy to see that reporting $(1,0)$ earns the firm a better value than does reporting $(1,1)$.

Type	Probability	$V_1(s,f)$
$(2,0)$	$r^2\theta^2$	$2u$
$(1,0)$	$2r\theta(1-\theta)$	$u + (ru + (1-r)d)$
$(1,1)$	$2r(1-r)\theta^2$	$u + d$
$(0,0)$	$(1-\theta)^2$	$2(ru + (1-r)d)$
$(0,1)$	$2(1-r)\theta(1-\theta)$	$d + (ru + (1-r)d)$
$(0,2)$	$(1-r)^2\theta^2$	$2d$

Table 2: Interim Period States, Probabilities, and Values

We will use the notation $\mu(\widehat{s}, \widehat{f}; s, f)$ to indicate the probability the market assigns to the manager playing reporting strategy $(\widehat{s}, \widehat{f})$ when the state is (s,f).[23] Thus market beliefs μ constitute a set of values $\mu(\widehat{s}, \widehat{f}; s, f)$ for all $(s,f) \in \{(0,0),(0,1),(0,2),(1,0),(2,0),(1,1)\}$. We will denote market beliefs corresponding to a particular pure-strategy equilibrium with the notation μ_i for $i \in \{F, N, P\}$ defining the full-disclosure, non-disclosure and partial-disclosure

[22] Conducting the analysis for general N would significantly complicate the notation, but is unlikely to yield qualitatively new insights. Even with $N = 2$, some derivations of formulae are complicated enough that we relegate them to the Appendix.

[23] In equilibrium, of course, we must have $\mu(\widehat{s}, \widehat{f}; s, f)$ equal to the firm's true probability of playing a given strategy.

equilibria, respectively. To avoid unnecessary notational clutter, we will use this notation sparingly, reserving it for expressions representing the firm's expected values under different sets of beliefs. That is, we will denote the firm's expected value from a particular disclosure strategy by the notation $E[\widehat{s},\widehat{f}|s,f,\mu]$, where μ identifies the beliefs of the market and NGO regarding the firm's behavior.

We define $\Psi(\widehat{s},\widehat{f})$ as the probability the market assigns to observing a report $(\widehat{s},\widehat{f})$; this is the sum of the probabilities of each interim state multiplied by the probability that the firm reports $(\widehat{s},\widehat{f})$ in that state. For example,

$$\begin{aligned}\Psi(0,0) &= (1-\theta)^2\mu(0,0|0,0)+2(1-r)\theta(1-\theta)\mu(0,0|0,1)\\ &\quad +(1-r)^2\theta^2\mu(0,0|0,2)+2r(1-r)\theta^2\mu(0,0|1,1).\end{aligned}$$

We turn now to the expected value the firm obtains in state $(1,1)$ from alternative possible disclosure strategies.

If the firm reports $(1,1)$, the market knows the state with certainty, and the firm has market value

$$E[1,1|1,1,\mu]=u+d. \tag{5}$$

If the firm in state $(1,1)$ reports $(1,0)$, then the market believes the state is either $(1,0)$ and the firm is revealing truthfully; $(2,0)$ and the firm is failing to report a success; or $(1,1)$ and the firm is engaging in greenwash. Thus, $\Psi(1,0)=2r\theta(1-\theta)\mu(1,0|1,0)+r^2\theta^2\mu(1,0|2,0)+2r(1-r)\theta^2\mu(1,0|1,1)$. If the NGO audits, and finds that the state is really $(1,1)$ but the firm engaged in greenwash, then the NGO launches a campaign against the firm that imposes a penalty P. The firm's expected value from reporting $(1,0)$ is

$$\begin{aligned}E[1,0|1,1,\mu] &= [u+(ru+(1-r)d)]\frac{2r\theta(1-\theta)\mu(1,0|1,0)}{\Psi(1,0)}+2u\frac{r^2\theta^2\mu(1,0|2,0)}{\Psi(1,0)}\\ &\quad +[u+d]\frac{2r(1-r)\theta^2\mu(1,0|1,1)}{\Psi(1,0)}-\alpha P.\end{aligned} \tag{6}$$

If the firm in state $(1,1)$ reports $(0,0)$, the market recognizes that the state may be $(0,0)$, $(0,1)$, $(0,2)$ or $(1,1)$.[24] Note that there is no possibility of a punishment in this case, since a report of $(0,0)$ does not constitute greenwash, as it does not claim any positive outcomes. The firm's expected value is

$$\begin{aligned}E[0,0|1,1,\mu] &= [ru+(1-r)d]\frac{(1-\theta)^2 2\mu(0,0|0,0)}{\Psi(0,0)}\\ &\quad +[d+(ru+(1-r)d)]\frac{2(1-r)\theta(1-\theta)\mu(0,0|0,1)}{\Psi(0,0)}\\ &\quad +2d\frac{(1-r)^2\theta^2\mu(0,0|0,2)}{\Psi(0,0)}+[u+d]\frac{2r(1-r)\theta^2\mu(0,0|1,1)}{\Psi(0,0)}\end{aligned} \tag{7}$$

[24] A firm in state $(1,0)$ or $(2,0)$ has no incentive to report $(0,0)$.

Expressions (6) and (7) appear complicated, but are actually quite simple in equilibrium. For example, the manager never has incentives to hide a success, so a firm in state $(2,0)$ will never report $(1,0)$. Thus $\mu(1,0|2,0) = 0$. Since the NGO is assumed to only punish what it views as greenwash, there is no punishment for reporting $(0,0)$; thus, firms in states $(0,1)$ or $(0,2)$ always have incentives to report $(0,0)$, and $\mu(0,0|0,0) = \mu(0,0|0,1) = \mu(0,0|0,2) = 1$. Furthermore, when we solve for the truthful disclosure equilibrium, it must be the case that in equilibrium the manager truthfully reports the firm's state when it is a $(1,1)$, that is, $\mu(1,1;1,1) = 1$ and $\mu(0,0;1,1) = 0$, and the manager does not report falsely, that is, $\mu(1,0;1,1) = 0$. Substituting in these values of $\mu(\cdot)$ greatly simplifies equations (6) and (7).

There are three types of pure-strategy equilibria that can emerge in this model in state $(1,1)$: a) the firm fully discloses the state, b) the firm engages in partial disclosure, or c) the firm does not disclose at all. We now examine each of these three equilibria in turn. Details of calculations are in the Appendix.

4.2 Full Disclosure Equilibrium

The incentive compatibility constraints necessary for a firm in state $(1,1)$ to disclose fully are $E[1,1|1,1,\mu_F] > E[0,0|1,1,\mu_F]$ and $E[1,1|1,1,\mu_F] > E[1,0|1,1,\mu_F]$. In addition, if market participants believe the full disclosure equilibrium is being played, their beliefs must reflect the nature of this equilibrium, that is, they believe that with probability one a firm in state $(1,1)$ discloses fully rather than engaging in partial disclosure or not disclosing at all. Formally, this means that $\mu(0,0|1,1) = \mu(1,0|1,1) = 0$, and $\mu(1,1|1,1) = 1$.

Recall from equation 5 that

$$E[1,1|1,1,\mu_F] = u + d.$$

Understanding the payoff for non-disclosure is more complex. By definition, in the full disclosure equilibrium the market believes that a firm in state $(1,1)$ will fully disclose. Hence, when the market observes non-disclosure, it concludes the state is $(0,0), (0,1)$, or $(0,2)$. The market then assigns the firm an expected value that reflects the payoff of each of these three states, weighted by the probability of each one occurring, conditional on the observation that the firm disclosed nothing. Some algebraic manipulation reveals that

$$E[0,0|1,1,\mu_F] = \frac{2\left(d(1-r) + ru(1-\theta)\right)}{(1-r\theta)}.$$

The expected value of partial disclosure is

$$E[1,0|1,1,\mu_F] = u + (ru + (1-r)d) - \alpha P.$$

The intuition for this value is simple: market participants believe the full disclosure equilibrium is being played, so the only time a firm would report $(1,0)$ is when the state is $(1,0)$. One can see immediately that if the expected penalty were $\alpha P = 0$, then the firm would always prefer to disclose $(1,0)$ rather

than $(1,1)$, since by so doing the firm creates an impression of being "greener" than it is in fact. The only thing that will prevent the firm in state $(1,1)$ from making such a disclosure is the threat of a punishment if it is found guilty of greenwash.

The condition that $E[1,1|1,1,\mu_F] > E[0,0|1,1,\mu_F]$ is equivalent to

$$u + d > \frac{2\left(d(1-r) + ru(1-\theta)\right)}{(1-r\theta)},$$

which simplifies to

$$r < r_{FN} \equiv \frac{1}{2-\theta}. \tag{8}$$

The notation r_{FN} refers to the fact that this value of r is defined by the firm's comparison of full disclosure vs. non-disclosure as alternative strategies. In general we will use the notation r_{ij} to indicate the value of r at which the firm is just indifferent between strategy i and strategy j where $i, j \in \{F, N, P\}$ and market beliefs are that the firm is playing strategy i. The expression r_{FN} turns out to be a very important determinant of firm behavior in the model, as it determines the firm's disclosure strategy when the punishment for greenwash is so great as to eliminate partial disclosure as a viable strategy. In this case, a firm in state $(1,1)$ must choose between full disclosure or non-disclosure.

The condition $E[1,1|1,1,\mu_F] > E[1,0|1,1,\mu_F]$ simplifies to

$$r < r_{FP} \equiv \frac{\alpha P}{u-d}. \tag{9}$$

Proposition 1 summarizes the foregoing analysis of the existence of a full-disclosure equilibrium

Proposition 1 *A full disclosure equilibrium exists for all* $r \leq \min\{r_{FP}, r_{FN}\}$.

The basic intuition regarding full disclosure is that when the probability of success is low, there is little advantage to the firm in hiding a failure, since undisclosed activities will essentially be branded as failures by the market anyway.

4.3 Non-Disclosure Equilibrium

The incentive compatibility requirements for a non-disclosure equilibrium are $E[0,0|1,1,\mu_N] > E[1,1|1,1,\mu_N]$ and $E[0,0|1,1,\mu_N] > E[1,0|1,1,\mu_N]$. The beliefs consistent with the equilibrium are $\mu(1,1|1,1) = \mu(1,0|1,1) = 0$, and $\mu(0,0|1,1) = 1$.

Once again, the payoff to full disclosure does not depend upon beliefs because disclosures are fully verifiable. As in the previous section, the payoff to full disclosure is

$$E[1,1|1,1,\mu_N] = u + d.$$

The payoff to partial disclosure is also unchanged from the previous section. Here, the beliefs associated with the equilibrium are that a firm in state $(1,1)$ chooses not to disclose *any* information. If the market sees a firm disclose $(1,0)$ then, it believes the firm is in state $(1,0)$. Thus, a firm in state $(1,1)$ that engages in greenwash obtains payoff

$$E[1,0|1,1,\mu_N] = u + (ru + (1-r)d) - \alpha P.$$

The payoff to non-disclosure is different than it was in the full disclosure equilibrium. Specifically, the market now believes there are four states in which a firm chooses to not disclose: $(0,0),(0,1),(0,2)$, and $(1,1)$. The total probability a firm chooses to not disclose is

$$\Psi(0,0) = 1 - \theta r\left(2 - (2-r)\theta\right).$$

To the firm that does not disclose, the market assigns an expected value of

$$\begin{aligned} E[0,0|1,1,\mu_N] &= \frac{(1-\theta)^2 2(ru+(1-r)d) + 2(1-r)\theta(1-\theta)(d+(ru+(1-r)d))}{1-\theta r\left(2-(2-r)\theta\right)} \\ &\quad + \frac{(1-r)^2\theta^2 2d + 2r(1-r)\theta^2(u+d)}{1-\theta r\left(2-(2-r)\theta\right)}. \end{aligned} \tag{10}$$

A non-disclosure equilibrium requires $E[0,0|1,1,\mu_N] > E[1,1|1,1,\mu_N]$ and $E[0,0|1,1,\mu_N] > E[1,0|1,1,\mu_N]$. The first of these simplifies to

$$r > r_{NF} \equiv r_{FN} \equiv \frac{1}{2-\theta}. \tag{11}$$

The second requirement, $E[0,0|1,1,\mu_N] > E[1,0|1,1,\mu_N]$ can be shown to be equivalent to

$$\frac{(1-r)\left(r^2\theta^2+1\right)}{\left(1-\theta r\left(2-(2-r)\theta\right)\right)} < \frac{\alpha P}{(u-d)}. \tag{12}$$

We denote the value of r for which the foregoing condition holds with equality as r_{NP}, so that $(1-r_{NP})\left(r_{NP}^2\theta^2+1\right)/(1-\theta r_{NP}\left(2-(2-r_{NP})\theta\right)) = \alpha P/\left(u-d\right)$. Then the following proposition summarizes the above analysis regarding the non-disclosure equilibrium.

Proposition 2 *A non-disclosure equilibrium exists for* $r > \max\{r_{NF,}r_{NP}\}$.

Intuitively, the non-disclosure equilibrium exists when the probability of a success is high, in which case a firm with a failure gains significantly from hiding it. Peloza (2005) finds that "Many managers worry that by overtly promoting their participation stakeholders might view the activity as self-serving. In fact, many respondents reported minimal or no attempts of self-promotion." For example, one of his survey respondents commented that "We're pretty sensitive. We don't want to go out thumping our chests saying 'oh, aren't we wonderful

and here's all the great things we do!' We want people to see for themselves and they can draw their own conclusions."[25] This is consistent with the results of Proposition 2, which suggests that the firms most likely to remain silent regarding their environmental performance are those with high values of r.

4.4 Partial-Disclosure Equilibrium

The incentive compatibility conditions for this type of equilibrium are $E[1,0|1,1,\mu_P] > E[1,1|1,1,\mu_P]$ and $E[1,0|1,1,\mu_P] > E[0,0|1,1,\mu_P]$. The beliefs consistent with a partial disclosure equilibrium are $\mu(1,1|1,1) = \mu(0,0|1,1) = 0$, and $\mu(1,0|1,1) = 1$.

Since punishment for greenwash is critical to the partial-disclosure equilibrium, in this subsection we introduce the notation $\eta = \alpha P/(u-d)$ for the "cost/benefit ratio" for greenwash, where αP is the expected penalty for greenwash and $u-d$ represents the maximum value the firm could possibly obtain from successful greenwash. When the expected penalty for greenwash is zero, then $\eta = 0$, and we expect the firm to engage in greenwash. When the expected penalty for greenwash rises to $u-d$, then $\eta = 1$, and any potential benefits of greenwash are outweighed by the expected penalty, and we expect the firm to avoid greenwashing.

As in the previous sections, the payoff to full disclosure does not depend upon beliefs, and

$$E[1,1|1,1,\mu_P] = u + d.$$

In the partial disclosure equilibrium, the market believes the firm in state $(1,1)$ will disclose $(1,0)$. Hence, the payoff to making this disclosure is different than it was in the two previous types of equilibrium. Now, there are two situations when firms disclose $(1,0)$—when the state is $(1,0)$ and when the state is $(1,1)$. Thus, the total probability that a firm discloses $(1,0)$ is

$$\begin{aligned}\Psi(1,0) &= 2r\theta(1-\theta)\mu(1,0|1,0) + 2r(1-r)\theta^2\mu(1,0|1,1) \\ &= 2r\theta(1-r\theta).\end{aligned}$$

Using this information, we can compute the expected payoff to partial disclosure as

$$E[1,0|1,1,\mu_P] = \frac{u(1+r(1-2\theta)) + d(1-r)}{1-r\theta} - \alpha P.$$

The non-disclosure payoff is now the same as it was in the full disclosure equilibrium, since the market believes there are three types of firms that opt not to disclose: $(0,0), (0,1)$, and $(0,2)$. Thus, the total probability of non-disclosure in this equilibrium is

$$\begin{aligned}\Psi(0,0) &= (1-\theta)^2\mu(0,0|0,0) + 2(1-r)\theta(1-\theta)\mu(0,0|0,1) + (1-r)^2\theta^2\mu(0,0|0,2) \\ &= (1-r\theta)^2.\end{aligned}$$

[25] See Peloza (2005), p. 16.

The expected payoff to non-disclosure is

$$E[0,0|1,1,\mu_P] = \frac{2\left(d(1-r)+ru(1-\theta)\right)}{(1-r\theta)}.$$

Some algebraic manipulation shows that $E[1,0|1,1,\mu_P] > E[0,0|1,1,\mu_P]$ if

$$r < r_{PN} \equiv \frac{1-\eta}{1-\theta\eta}, \tag{13}$$

where the notation r_{PN} indicates the boundary between partial disclosure and non-disclosure.

Similarly, $E[1,0|1,1,\mu_P] > E[1,1|1,1,\mu_P]$ reduces to

$$r > r_{PF} \equiv \frac{\eta}{\theta\eta+(1-\theta)}. \tag{14}$$

where the notation r_{PF} indicates the boundary between partial disclosure and full disclosure. A partial disclosure equilibrium exists for $r \in [r_{PF}, r_{PN}]$. One can see from inspection of (13) and (14) that as η goes to zero, r_{PF} goes to zero and r_{PN} goes to one. Thus, as the expected penalty becomes negligible, partial disclosure is the unique pure-strategy equilibrium for all values of r and θ. It is also immediate that there is no partial disclosure equilibrium in pure strategies if $r_{PF} > r_{PN}$. The next proposition characterizes conditions for the existence of a partial disclosure equilibrium.

Proposition 3 *If $\eta = 0$, partial disclosure is the unique equilibrium for all values of r and θ. For $\eta \in (0,1/2)$ a partial disclosure equilibrium exists for $r \in [r_{PF}, r_{PN}]$. The difference between r_{PF} and r_{PN} is decreasing in θ.*

Proof. *As shown in the text above, the incentive compatibility conditions for a partial disclosure equilibrium imply that such an equilibrium exists for $r \in [r_{PF}, r_{PN}]$. If $\alpha P = 0$, then $r_{PF} = 0$ and $r_{PN} = 1$, so partial disclosure is the unique equilibrium strategy for all r and θ. If $\eta = 1/2$ then $r_{PF} = r_{PN}$ and there is no region of partial disclosure. For $\eta \in (0,1/2)$ the partial disclosure equilibrium exists. Finally, let $R_P = r_{PN} - r_{PF}$ Some calculation shows that*

$$R_P = \frac{(1-\theta)\left(u-d-2P\alpha\right)(u-d)}{\left((u-d)(1-\theta)+P\theta\alpha\right)\left(u-d-P\theta\alpha\right)}.$$

Differentiating with respect to θ yields

$$\frac{dR_P}{d\theta} = \frac{-\alpha\theta(u-d)P\left(u-d-2P\alpha\right)\left(u-d-P\alpha\right)(2-\theta)}{\left(u-d-P\theta\alpha\right)^2\left((u-d)(\theta-1)-P\theta\alpha\right)^2}.$$

The denominator is positive. Assuming $\eta < 1/2$, which is the condition for the existence of a partial disclosure equilibrium, we must have $(u-d-2P\alpha) > 0$ and $(u-d-P\alpha) > 0$, so $dR_P/d\theta < 0$. ■

Intuitively, a partial-disclosure equilibrium can only exist when the expected penalty is not too high. As we will show in more detail in the following section, if the penalty is made large enough, it will deter any type of firm from engaging in partial disclosure. Furthermore, it is worth noting that the types of firms most likely to engage in partial disclosure are not those with particularly high or low values of r, but rather those with an intermediate likelihood of positive outcomes. The intuition for this observation is straightforward. Firms with low values of r fully disclose: they gain a lot from trumpeting a success, and lose little by withholding information about a failure (since they are already expected to fail); thus, there is little value in risking public backlash by refusing to disclose. At the other extreme, firms with high values of r do not disclose anything: they gain little by disclosing information about successes (since they are already expected to succeed), and lose a lot by disclosing a failure; thus, there is little value in risking public backlash by disclosing a success. For firms with moderate values of r partial disclosure is attractive: disclosing a success can produce a significant improvement in public perception, and withholding information about a failure can prevent a significant negative public perception; thus, they are willing to risk public backlash by disclosing only partially.

Proposition (3) shows that the band of r values that constitute a partial disclosure equilibrium becomes smaller as θ increases. The reason is that a manager of a low-θ firm is less likely to be informed about the performance of his activities, and hence the market does not draw strongly negative inferences if the firm fails to report two outcomes. For high-θ firms, however, the market is likely to weight heavily the likelihood that projects are environmental failures when faced with managerial silence. Note also that as θ increases, the pure-strategy greenwash region shifts upward toward higher values of r. As θ rises, the market becomes increasingly sure that an undisclosed outcome is a failure. Only for firms with very high levels of r will the market grant a non-trivial probability that an undisclosed outcome is really a success.[26]

5 A Full Characterization of Equilibria in the Disclosure Game

The analysis in section 4 established conditions for the existence of different types of pure-strategy disclosure equilibria. These equilibria depend upon different sets of beliefs on the part of participants in the disclosure game, and depend upon the parameters r, θ and η. We are now ready to fully characterize the set of equilibria of the disclosure game as a function of the costs and benefits of greenwash. The expression $r_{FN} = 1/(2-\theta)$ plays a key role in parts (c) and (d) of Proposition 4; if greenwash were prohibitively costly, then for $r > r_{FN}$ the firm prefers non-disclosure and for $r < r_{FN}$ the firm prefers full disclosure.

[26] Technically, as θ approaches one, the numerator of (4) becomes very small, and the market highly discounts any failure to disclose. Only for firms with very high levels of r will the denominator of (4) be small enough that the market grants any credibility to a firm's profession that it has not yet observed an outcome.

Proposition 4 *(a) If $\eta = 0$, partial disclosure is the unique pure strategy equilibrium for all (r, θ). (b) If $\eta \in (0, 1/2)$, then for any $\theta \in (0, 1)$, there exists a series of non-negative values $r_{FP} < r_{PF} < r_{PN} < r_{NP} < 1$ such that for $r \in (0, r_{FP})$ the unique equilibrium is full disclosure, for $r \in [r_{FP}, r_{PF})$ the equilibrium is a mixed strategy mixing between full disclosure and partial disclosure, for $r \in [r_{PF}, r_{PN})$ the unique equilibrium is partial disclosure, for $r \in [r_{PN}, r_{NP})$ the equilibrium is a mixed strategy mixing between partial disclosure and non-disclosure, and for $r \in [r_{NP}, 1)$ the unique equilibrium is non-disclosure. (c) If $\eta \in (1/2, 1)$, then for $r < \min(r_{FN}, r_{FP})$ the unique equilibrium is full disclosure, for $r \in (\min\{r_{FN}, r_{FP}\}, r_{FN})$ the equilibrium is a mixed strategy mixing between full disclosure and partial disclosure, for $r \in [r_{FN}, \max(r_{FN}, r_{NP}))$ the equilibrium is a mixed strategy mixing between partial disclosure and non-disclosure, and for $r \in [\max(r_{NP}, r_{FN}), 1)$ the unique equilibrium is non-disclosure. (d) If $\eta \geq 1$, then for $r < r_{FN}$ full disclosure is the unique pure-strategy equilibrium, and for $r > r_{FN}$ non-disclosure is the unique pure-strategy equilibrium.*

Proof. See the Appendix. ■

When $\eta = 0$, as was shown in Proposition 3, partial disclosure is the only equilibrium strategy for a firm in state $(1, 1)$ Disclosing $(1, 0)$ produces a positive effect on external beliefs about the firm, and carries with it no penalty. Thus, partial disclosure dominates either full disclosure or no disclosure.

For $\eta \in (0, 1/2)$, each of the three types of pure-strategy equilibria exists for at least some value of (r, θ). However, it is also true that pure-strategy equilibria do not exist for all (r, θ) pairs. This is illustrated in Figure 1. There are three regions with pure-strategy equilibria: the non-disclosure region lies above the upper solid curve, the partial-disclosure region lies between the two dotted curves, and the full-disclosure region lies below the lower solid line. (The two dotted curves converge toward the dashed curve lying between them as η approaches 1/2, at which point greenwash is eliminated as a pure strategy equilibrium.) There are also two regions in which there are no pure-strategy equilibria: between the upper solid curve and the upper dotted curve, the region labeled "Mix$_{NP}$," firms employ a mixed strategy that involves mixing between non-disclosure and partial disclosure;.between the lower dotted curve and the lower solid line, the region labeled "Mix$_{FP}$," firms mix between full disclosure and partial disclosure.

[Figure 1 about here]

At the point where $\eta = 1/2$, partial disclosure is just eliminated as a pure strategy equilibrium. Two types of pure-strategy equilibria continue to exist, but there are again regions in which pure-strategy equilibria do not exist. This is illustrated in Figure 2. Again, the non-disclosure region lies above the upper solid curve, and the full-disclosure region lies below the lower solid line. Now there is no pure-strategy partial disclosure region, because the penalty is large enough to eliminate it as an equilibrium. From a graphical perspective, the

two former dotted curves bounding the partial disclosure region have collapsed together into the dashed curve in the middle of the graph. Once again, there are two regions in which there are no pure-strategy equilibria: the region labeled "Mix$_{NP}$,"in which firms employ a mixed strategy that involves mixing between non-disclosure and partial disclosure, and.the second, labeled "Mix$_{FP}$," in which firms mix between full disclosure and partial disclosure. Thus, even though partial disclosure is not a pure strategy equilibrium for any (r, θ) pairs, it is still part of the mixed strategies in the aforementioned regions. Note that at $\eta = 1/2$, the left intercept is at $r = 1/2$ for all three curves defining the full disclosure, non-disclosure and mixing regions.

[Figure 2 about here]

For $\eta \in (1/2, 1)$, the mixed strategy regions shrink as η increases, as can be seen by comparing Figure 3 to Figure 2. For $\eta > 1/2$, there is a critical value $\theta^* = 2 - 1/\eta$ where $r_{FN} = r_{NP} = r_{FP} = \eta$. For $\theta < \theta^*$, the unique equilibrium is full disclosure for $r < r_{FN}$ and non-disclosure for $r > r_{FN}$. For $\theta > \theta^*$, we know that $r_{FP} < r_{FN}$, and the unique equilibrium is full disclosure for $r < r_{FP}$. For $r \in (r_{FP}, r_{FN})$ the equilibrium involves mixing between full disclosure and partial disclosure. For $r \in (r_{FN}, r_{NP})$ the equilibrium involves mixing between non-disclosure and partial disclosure. Finally, for $r > r_{NP}$, the unique equilibrium is non-disclosure. In the limit, as η goes to 1, θ^* goes to 1 as well; then non-disclosure is the unique pure-strategy equilibrium for all $r > r_{FN}$ and full disclosure is the unique pure-strategy equilibrium for $r < r_{FN}$.

[Figure 3 about here]

The fact that the greenwash regions shrink as the expected penalty grows is intuitive. The maximum benefit the firm can possibly obtain from greenwash is $u - d$. This occurs if the firm has a very high value of r, so the market grants the firm expected value of u for undisclosed outcomes, whereas it would have gotten a d if it revealed the failure. If the penalty is large enough to outweigh this maximum possible benefit to partial disclosure, then it will deter firms from using this strategy. Thus, if $\alpha P \geq (u - d)$, the firm in state $(1, 1)$ simply chooses between full disclosure or non-disclosure. As shown in section 4, this decision turns upon whether or not $r < 1/(2 - \theta)$, with full disclosure the equilibrium if the inequality holds, and non-disclosure the equilibrium if it does not.

6 Implications for Firm Strategy and Empirical Analysis

To this point, our analysis has been strictly theoretical. In this section we use it to shed light on disclosure behavior in a variety of practical examples. Our emphasis is on the two key underlying parameters of the model that describe

firms, namely r, the probability that a given activity has a positive environmental outcome, and θ, the probability the manager is informed about an activity's outcome at the time he makes a report.

Proposition 4 showed that when NGO penalties for greenwash are low, then greenwash is the communications strategy of choice. Empirical research in accounting suggests that this is indeed a common practice for firms that choose to engage in corporate environmental disclosure.[27]

More generally, Figures 1-3 illustrate the full set of possible equilibria and how they vary with these two parameters. The following subsections discuss changes in r and θ, respectively, and how they relate to disclosure behavior.

6.1 Changing Perceptions of Environmental Performance

One of the key dimensions on which firms differ in our model is their propensity for good environmental performance, as captured by the variable r. Our analysis predicts that changes in r may change a firm's disclosure behavior. For example, if a firm with a high likelihood of good performance suffers a drop in r, then it may shift from a policy of non-disclosure to one of partial disclosure. A careful reading of Proposition 4 shows that decreases in r lead (weakly) to increases in disclosure. Loosely speaking, large enough decreases in r cause non-disclosing firms to begin greenwashing, and cause greenwashers to shift toward full disclosure. Modest reductions in r may not induce any change in the firm's strategy, but whenever the decrease in r moves the firm from one region to another in Figures 1-3, this corresponds to a move toward greater disclosure. For example, in Figure 1, a firm at the lower end of the non-disclosure range transitions to a mixed strategy that involves both non-disclosure and partial disclosure. Further reductions in r eventually cause the firm to transition to a pure strategy involving greenwash. Continued reductions in r cause transitions to a mix of greenwash and full disclosure, and finally to full disclosure.[28] We record this observation in the following Corollary to Proposition 4.

Corollary 5 *Decreases in the probability of environmental success r lead (weakly) to increases in disclosure.*

A natural experiment for testing this hypothesis was provided when the Exxon Valdez struck a reef in Prince William Sound, Alaska, on March 24, 1989. The 11 million gallon spill caused oil company stakeholders, including citizens and shareholders, to re-evaluate the environmental riskiness of oil company operations. As one might expect, the incident also had a strong negative impact on the company's finances: within a year of the accident, Exxon had already spent over $2 billion to clean up the spill. Patten (1992) found that on average major oil companies more than doubled their environmental disclosures in the wake of the Valdez accident. This increase in disclosures was virtually

[27] See, for example, Deegan and Gordon (1996) and Guthrie and Parker (1990).

[28] A similar set of transitions can be observed in Figures 2 and 3 as well.

required for Exxon, which had to describe the event to shareholders. The interesting finding was that other major firms in the industry also increased their disclosure behavior. In our framework, this is consistent with a decrease in r across the entire industry, which moved firms toward greater disclosure.

Deegan and Rankin (1996) examine the practices of 20 Australian firms that were subject to successful prosecution by the Environmental Protection Agency. They find that these firms increased their environmental disclosures in the year in which they were prosecuted. Nevertheless, most appeared to employ a greenwash strategy, with the overwhelming majority of disclosures being of a positive nature. In our model, this could be viewed as a move from a region in which the firms employed a strategy involving a mix of non-disclosure and greenwash into a region in which greenwash is a pure strategy.[29]

Our perspective also helps interpret events in the news, such as Walmart's recent conversion to a more promotional stance regarding its social contributions.[30] Historically, Walmart kept a low profile on social issues, but this has begun to change as the company has come under attack for its low pay and lack of benefits. As the company has expanded beyond its rural origins and into suburban America, Wal-Mart is increasingly seen as offering a Faustian bargain involving "everyday low prices" in exchange for negative impacts on social issues such as living wages, health care and global outsourcing. A McKinsey and Company study leaked to the press by walmartwatch.com found that up to 8% of shoppers had stopped patronizing the chain because of its negative reputation. In response, the company has begun to promote its social and environmental contributions more prominently. In the past, says CEO Lee Scott, "We would put up the sandbags and get out the machine guns."[31] Today it takes a much more pro-active stance in communicating its positive contributions. In addition, according to the company's web site, on January 8, 2007, "Wal-Mart Stores, Inc. today announced that it will begin airing two national television advertisements as part of its continued effort to inform the public about the company's positive impact on communities, including its core values, affordable health care, customer savings, and charitable contributions."[32] The company has also been widely touting its environmental sustainability. In terms of our model, this newfound interest in communication represents a shift from non-disclosure to partial disclosure. Such a shift is consistent with the notion that Walmart has experienced a reduction in r, that is, in the probability that its actions produce socially desirable outcomes.

[29] Deegan, Rankin and Voght (2000) examine five major environmental incidents and the resulting changes in disclosure behavior. They find a common pattern in all five cases, with firms in the industry increasing disclosure after an environmental disaster. This was true for firms listed on U.S. stockmarkets but also for firms on other markets too.

[30] A speech on the topic by Walmart's CEO can be found at http://www.walmartstores.com/Files/21st%20Century%20Leadership.pdf

[31] See Gunther (2006).

[32] See http://www.walmartfacts.com/articles/4675.aspx

6.2 Environmental Management Systems and NGO Auditing

The threat of NGO audits does not lead all firms to increase their disclosures. In particular, NGO auditing is likely to lead to reduced disclosures from poorly informed firms in clean industries, that is, when r is large and/or θ is small. On the other hand, NGO auditing may lead to an increase in disclosures from well-informed firms in dirty industries, that is, where r is small and/or θ is large. This can be understood by thinking about how increases in expected penalties affect equilibria. When $\eta = 0$, all firms engage in greenwash. However, when $\eta \geq 1$, firms no longer engage in greenwash at all. At this point, firms separate themselves into two groups. One group, positioned below the dashed line in Figure 3, elects to fully disclose. The other group, positioned above the dashed line in Figure 3, elects not to disclose at all. Thus, the increasing threat of NGO audits leads to diametrically opposite shifts in behavior for firms above and below the dashed line.

It is worth noting that as θ increases, more and more firms fall below the dashed line in Figure 3. This means that NGO auditing is more likely to increase disclosures for firms that have a high value of θ. In our model, θ measures the likelihood that a firm knows its environmental impacts at period 1. In practice, an increase in θ is likely to correspond to a firm's adoption of an environmental management system (EMS), such as ISO 14001. An EMS is a set of management processes and procedures that allows an organization to integrate environmental issues into day-to-day decisions. Of course, a necessary component of an EMS is a reliable system for measuring a firm's environmental impacts.[33] Thus, whether NGO auditing is likely to increase disclosures depends on the presence of environmental management systems (EMS) within the audited firms. In our model, however, firms have no incentive to adopt an EMS, since the firm's market value in the interim period is lower when it adopts an EMS, as is shown in the following proposition.

Proposition 6 *In the interim period, the firm's value in the partial disclosure equilibrium is decreasing in θ. Its value in the full disclosure or non-disclosure equilibria is unaffected by θ.*

Proof. Differentiating equation (3) with respect to θ yields $dV_{PD}/d\theta = (u - d)(N - s)(dq/d\theta)$. All terms in this expression are positive with the possible exception of $dq/d\theta$. Recalling that $q = (r - \theta r)/(1 - \theta r)$, and differentiating this

expression yields $dq/d\theta = -r(1 - r)/(1 - \theta r)^2 < 0$. Thus, $dV_{PD}/d\theta < 0$. It is straightforward to see that $V_{FD} = u + d$ and $V_{ND} = N(ru + (1 - r)d)$, neither of which is a function of θ. ■

The intuition for the proposition is as follows. In the partial disclosure equilibrium, the manager withholds unfavorable information to increase its market

[33] For more details on EMSs, see Coglianese and Nash (2001) and Delmas (2000).

value. This strategy works because for each withheld piece of information, the market valuation of the firm reflects only the possibility, not the certainty, of a failure. However, as the likelihood increases that the manager knows the environmental outcomes of the firm's activities, the market increasingly interprets non-disclosure as withheld negative information rather than as true uncertainty on the part of the manager. Adopting an EMS improves the manager's internal information, and thus makes the market increasingly skeptical when the manager does not fully disclose all possible environmental information.

Of course, our model does not incorporate the benefits of an EMS in terms of improved internal control and ability to comply with environmental regulations. Nevertheless, our analysis does identify a countervailing incentive that may deter firms from adopting EMSs. Furthermore, our story is broadly consistent with the empirical results of Delmas (2000), who finds that many U.S. firms elect not to adopt ISO 14001 (a particular form of EMS) because they wish to limit public access to internal information about their environmental performance.

Our results suggest that public policy pressures may be required to induce a broad cross-section of firms to adopt EMSs. Interestingly, Coglianese and Nash (2001, p. 15) find that there has been "an explosion of programs in the United States that offer financial and regulatory incentives to firms that implement EMSs." These programs are being implemented at both the federal and state levels. Whether these programs are likely to achieve their objectives is unclear. Coglianese and Nash (2001, p. 16) point out that "[a]ll of these policy initiatives are premised on the assumption that EMSs make a difference in environmental performance. Yet this question merits research and evidence rather than untested optimism."

Our analysis points to a new rationale for encouraging firms to adopt EMSs, one that does not appear to have been recognized in prior literature, either by academics or practitioners. We do not presume that an EMS makes any difference in environmental performance, but instead simply assume an EMS improves the manager's internal information about the firm's environmental performance. In this capacity, an EMS operates as a complement to NGO auditing of environmental disclosure and greenwash. With an EMS in place, when a manager discloses nothing about the firm's environmental performance, the market infers that the manager is failing to disclose some negative information, and thus downgrades its rating of the firm's value. In turn, this means that an NGO's threat to punish greenwash is more likely to drive the manager to disclose fully rather than to not disclose at all. In effect, the presence of the EMS brings the market closer to a state of common knowledge, thereby increasing market efficiency. With an EMS in place, the manager is more likely to be well informed about his firm's own environmental impact, and *the market knows that the manager is more likely to be well informed*. As a result the manager is unable to hide behind the veil of ignorance when he fails to fully disclose the impacts of his firm's actions, and is thereby pressured to fully disclose.

7 Conclusions

This paper has presented what is to our knowledge the first economic analysis of greenwash. We defined greenwash as the selective disclosure of positive information about a company's environmental performance, without full disclosure of negative information on these dimensions. We then modeled the phenomenon using tools from the literature on financial disclosure. In our model, a non-governmental organization (NGO) can audit corporate environmental reports, and penalize firms caught engaging in greenwash. Our model is relatively simple, yet produces some interesting positive implications. We show that the types of firms most likely to engage in partial disclosure are those with an intermediate probability of producing positive environmental and social outcomes. For such firms, disclosing a success can produce a significant improvement in public perception, and withholding information about a failure can prevent a significant negative public perception; thus, they are willing to risk public backlash by disclosing only partially. We also show that when there is a decline in a firm's likelihood of producing environmental successes, it tends to increase its disclosures, a result consistent with empirical findings in the accounting literature. In addition, we find that NGO auditing of corporate disclosure behavior is more likely to induce a firm to become more open and transparent if the firm operates in an industry that is likely to have socially or environmentally damaging impacts, and if the firm is relatively well informed about its environmental or social impacts. This description fits quite well with the broad types of firms typically singled out for scrutiny and outrage by activists.

The model also has interesting normative implications. We show that there is a real possibility that the threat of public backlash for greenwash will cause firms to "clam up" rather than become more open and transparent. In particular, such a response is likely from socially responsible firms with a high probability of successful projects, yet who are not fully informed about the social impacts of their actions. In an environmental context one might characterize such firms as "poorly informed firms in clean industries." For firms such as this, activist pressures designed to increase disclosure may backfire and produce exactly the opposite of the intended results.

The likelihood that a firm responds to the threat of NGO auditing by opting for non-disclosure is reduced if the firm has adopted an environmental management system (EMS), and the complementarity between EMSs and NGO auditing of greenwash points to a benefit from public policies that mandate the adoption of EMSs. Indeed, our analysis points to a new rationale for encouraging firms to adopt EMSs An EMS brings the market closer to a state of common knowledge, thereby increasing market efficiency. With an EMS in place, the manager is better informed about his firm's environmental impact, and the market knows that the manager is better informed. As a result the manager is unable to hide behind the veil of ignorance when he fails to fully disclose the impacts of his firm's actions, and is thereby pressured to fully disclose.

There are a number of areas in which further research would be valuable.

One need is for empirical study of greenwash, its effects on firm valuation, and its interaction with NGO information campaigns. Ramus and Montiel (2005) represents one needed step in this direction, as does Kim and Lyon (2006). Yet more work is needed before we have a robust empirical understanding of the phenomenon. A second need is to explore more fully the motivations of activist groups that monitor and punish corporate hypocrites. Articulating their objective functions—maximizing membership, maximizing financial contributions, affecting change in the industry, or some mix of the above—would allow for a strategic analysis of activist behavior, and the equilibrium of such a model would produce further insights into corporate non-market strategy. Third, it would also be interesting to extend the model so that the firm's activities are heterogeneous in nature, varying in cost, likelihood of success, and social or environmental impact. This would allow for an analysis of firms' incentives to invest in projects known to have a high probability of success but low social or environmental value, an accusation leveled against some firms. In this case, partial disclosure may divert scarce funds from valuable risky projects to relatively certain but low-value projects.

Appendix

In this appendix we present some derivations of formulae that appear in simplified form in the text.

Full Disclosure Equilibrium

$$\begin{aligned}\Psi(0,0) &= (1-\theta)^2 + 2(1-r)\theta(1-\theta) + (1-r)^2\theta^2 \\ &= (1-r\theta)^2\end{aligned}$$

$$\begin{aligned}E[0,0|1,1,\mu_F] &= \frac{(1-\theta)^2 2(ru+(1-r)d)\mu(0,0|0,0)}{\Psi(0,0)} \\ &\quad + \frac{2(1-r)\theta(1-\theta)[d+(ru+(1-r)d)]\mu(0,0|0,1)}{\Psi(0,0)} \\ &\quad + \frac{(1-r)^2\theta^2 2d\mu(0,0|0,2) + 2r(1-r)\theta^2(u+d)\mu(0,0|1,1)}{\Psi(0,0)} \\ &= \frac{(1-\theta)^2 2(ru+(1-r)d) + 2(1-r)\theta(1-\theta)[d+(ru+(1-r)d)]}{\Psi(0,0)} \\ &\quad + \frac{(1-r)^2\theta^2 2d}{\Psi(0,0)} \\ &= \frac{2\left(d(1-r)+ru(1-\theta)\right)}{(1-r\theta)}\end{aligned}$$

$$\begin{aligned}\Psi(1,0) &= 2r\theta(1-\theta)\mu(1,0|1,0) + r^2\theta^2\mu(1,0|2,0) + 2r(1-r)\theta^2\mu(1,0|1,1) \\ &= 2r\theta(1-\theta)\end{aligned}$$

$$\begin{aligned}E[1,0|1,1,\mu_F] &= (u+(ru+(1-r)d))\frac{2r\theta(1-\theta)\mu(1,0|1,0)}{\Psi(1,0)} + 2u\frac{r^2\theta^2\mu(1,0|2,0)}{\Psi(1,0)} \\ &\quad + (u+d)\frac{2r(1-r)\theta^2\mu(1,0|1,1)}{\Psi(1,0)} - \alpha P \\ &= (u+(ru+(1-r)d))\frac{2r\theta(1-\theta)}{\Psi(1,0)} - \alpha P \\ &= u+(ru+(1-r)d) - \alpha P\end{aligned}$$

Non-Disclosure Equilibrium

$$\begin{aligned}\Psi(0,0) &= 1 - \Pr(1,0) - \Pr(2,0) \\ &= 1 - 2\theta r(1-\theta) - r^2\theta^2 \\ &= 1 - \theta r\left(2-(2-r)\theta\right)\end{aligned}$$

$$E[0,0|1,1,\mu_N] = \frac{(1-\theta)^2 2(ru+(1-r)d)+2(1-r)\theta(1-\theta)(d+(ru+(1-r)d))}{1-\theta r\left(2-(2-r)\theta\right)} + \frac{(1-r)^2\theta^2 2d+2r(1-r)\theta^2(u+d)}{1-\theta r\left(2-(2-r)\theta\right)}.$$

Partial Disclosure Equilibrium

$$\begin{aligned}\Psi(1,0) &= 2r\theta(1-\theta)\mu(1,0|1,0)+2r(1-r)\theta^2\mu(1,0|1,1)\\ &= 2r\theta(1-r\theta)\end{aligned}$$

$$\begin{aligned}E[1,0|1,1,\mu_P] &= (u+(ru+(1-r)d))\frac{2r\theta(1-\theta)\mu(1,0|1,0)}{\Psi(1,0)}+2u\frac{r^2\theta^2\mu(1,0|2,0)}{\Psi(1,0)}\\ &\quad +(u+d)\frac{2r(1-r)\theta^2\mu(1,0|1,1)}{\Psi(1,0)}-\alpha P\\ &= (u+(ru+(1-r)d))\frac{2r\theta(1-\theta)}{\Psi(1,0)}+(u+d)\frac{2r(1-r)\theta^2}{\Psi(1,0)}-\alpha P\\ &= (u+(ru+(1-r)d))\frac{2r\theta(1-\theta)}{2r\theta(1-r\theta)}+(u+d)\frac{2r(1-r)\theta^2}{2r\theta(1-r\theta)}-\alpha P\\ &= \frac{u(1+r(1-2\theta))+d(1-r)}{1-r\theta}-\alpha P.\end{aligned}$$

$$\begin{aligned}\Psi(0,0) &= (1-\theta)^2\mu(0,0|0,0)+2(1-r)\theta(1-\theta)\mu(0,0|0,1)\\ &\quad +(1-r)^2\theta^2\mu(0,0|0,2)+2r(1-r)\theta^2\mu(0,0|1,1)\\ &= (1-r\theta)^2\end{aligned}$$

$$\begin{aligned}E[0,0|1,1,\mu_P] &= \frac{(1-\theta)^2 2(ru+(1-r)d)\mu(0,0|0,0)+2(1-r)\theta(1-\theta)(d+(ru+(1-r)d))\mu(0,0|0,1)}{\Psi(0,0)}\\ &\quad +\frac{(1-r)^2\theta^2 2d\mu(0,0|0,2)+2r(1-r)\theta^2(u+d)\mu(0,0|1,1)}{\Psi(0,0)}\\ &= \frac{(1-\theta)^2 2(ru+(1-r)d)+2(1-r)\theta(1-\theta)(d+(ru+(1-r)d))+(1-r)^2\theta^2 2d}{\Psi(0,0)}\\ &= \frac{2\left(d(1-r)+ru(1-\theta)\right)}{(1-r\theta)}.\end{aligned}$$

Lemma 7 *For* $r > \frac{1-\sqrt{2(1-\theta)}}{\theta}$, $N(r,\theta) \equiv \frac{(1-r)\left(r^2\theta^2+1\right)}{(1-\theta r(2-(2-r)\theta))} > G(r,\theta) \equiv \frac{1-r}{1-r\theta}$.

Proof. Define

$$\begin{aligned}\Delta(r,\theta) &= N(r,\theta)-G(r,\theta) = \frac{(1-r)\left(r^2\theta^2+1\right)}{\left(1-\theta r\left(2-(2-r)\theta\right)\right)}-\frac{1-r}{1-r\theta}\\ &= \frac{\left(2\theta-2r\theta+r^2\theta^2-1\right)}{\left(2r\theta-2r\theta^2+r^2\theta^2-1\right)}\frac{(1-r)\,r\theta}{(1-r\theta)},\end{aligned}$$

the second term of which is always non-negative. Hence the sign of $\Delta(r,\theta)$ is the same as the sign of

$$\frac{\left(2\theta - 2r\theta + r^2\theta^2 - 1\right)}{\left(2r\theta - 2r\theta^2 + r^2\theta^2 - 1\right)}.$$

Differentiating the denominator shows that it is increasing in both r and θ for $r \in (0,1)$ and $\theta \in (0,1)$. Evaluating the denominator, we find it is negative for all for $r \in (0,1)$ and $\theta \in (0,1)$. Hence the denominator is negative. Thus we focus on signing the numerator. Solving for $2\theta(1-r) + (r\theta+1)(r\theta-1) = 0$ we find there are two roots. The one which is defined for $r \in (0,1)$ and $\theta \in (0,1)$ is $r = \frac{1-\sqrt{2(1-\theta)}}{\theta}$. Hence the numerator is negative for $r > \frac{1-\sqrt{2(1-\theta)}}{\theta}$. Thus, we have shown that $\Delta(r,\theta) > 0$ for $r > \frac{1-\sqrt{2(1-\theta)}}{\theta}$. ■

Proof of Proposition 4: (a) See Proposition 3. (b) Since $r_{FP} \equiv \frac{\alpha P}{u-d} = \eta$, it is clear that $r_{FP} > 0$ for $\eta > 0$. By the definition of $r_{PF} \equiv \frac{\eta}{\theta\eta+(1-\theta)}$, we see that $r_{PF} - r_{FP} = \frac{\theta\eta(1-\eta)}{\theta\eta+(1-\theta)} > 0$, so $r_{PF} > r_{FP}$. Proposition 3 shows that $r_{PN} > r_{PF}$. To show that $r_{NP} > r_{PN}$, we construct a proof by contradiction. Define $G(r,\theta) = \frac{1-r}{1-r\theta}$ and $N(r,\theta) = \frac{(1-r)\left(r^2\theta^2+1\right)}{(1-\theta r(2-(2-r)\theta))}$. Recall that r_{NP} is defined implicitly by the relation $N(r_{NP},\theta) = \eta$ and r_{PN} is defined implicitly by $G(r_{PN},\theta) = \eta$. Now suppose that for a given η it is the case that $r_{NP} < r_{PN}$. It is straightforward to show that $dG/dr = \frac{\theta-1}{(r\theta-1)^2} \leq 0$, so $G(r_{NP},\theta) > G(r_{PN},\theta) = \eta$. Since $N(r_{NP},\theta) = \eta$ it follows directly that $G(r_{NP},\theta) > N(r_{NP},\theta)$. The Lemma shows that this can only be true for $r < \frac{1-\sqrt{2(1-\theta)}}{\theta}$. Solving for the minimimum of $N(r,\theta)$ subject to the constraints $r < \frac{1-\sqrt{2(1-\theta)}}{\theta}$, $\theta > 0$, and $\theta < 1$, we find the minimum occurs at $\theta^* = 2(\sqrt{2}-1)$ and $r^* = (1-\sqrt{6-4\sqrt{2}})/\theta^*$; substituting in r^* and θ^* yields a value of $N(r^*,\theta^*) \simeq .85355$. Since we are only concerned with $\eta < .5$, we know that $N(r_{NP},\theta) > \eta$ on the range $r < \frac{1-\sqrt{2(1-\theta)}}{\theta}$ and there cannot exist a solution $r_{NP} < \frac{1-\sqrt{2(1-\theta)}}{\theta}$. Thus it must not be true that $r_{NP} < r_{PN}$. Finally, to show that $r_{NP} < 1$, note that $N(1,\theta) = 0$, so for any $\eta > 0$ it is impossible to have $N(r_{NP},\theta) = \eta$ with $r_{NP} = 1$. (c) Because $\eta > 1/2$, partial disclosure is not a pure-strategy equilibrium, as shown in Proposition 3. Proposition 1 establishes that there is a full disclosure equilibrium for $r < \min(r_{FN}, r_{FP})$. For $r_{FP} < r < r_{FN}$, there is no pure-strategy equilibrium, but both full disclosure and partial disclosure are undominated strategies. Hence Dasgupta and Maskin (1986a) applies and there is a mixed strategy equilibrium involving a mix of full and partial disclosure. For $r_{FN} < r < r_{NP}$, there is no pure-strategy equilibrium, but both non-disclosure and partial disclosure are undominated strategies. Hence Dasgupta and Maskin (1986a) applies and there is a mixed strategy equilibrium involving a mix of non-disclosure and partial disclosure. Proposition 2 establishes that there is a non-disclosure equilibrium for $r > \max\{r_{FN}, r_{NP}\}$. (d) For $\eta \geq 1$, the maximum gain from greenwashing is less than $u - d$ and

the firm never has incentives to engage in greenwash. Thus it simply compares full disclosure and non-disclosure, and Proposition 2 implies that the unique equilibrium is full disclosure for $r < r_{FN}$ and non-disclosure for $r > r_{FN}$.

References

[1] Baron, David P. 2003. "Private Politics," *Journal of Economics and Management Strategy*, 12: 31-66.

[2] Baron, David P and Daniel Diermeier. 2005. "Strategic Activism and Nonmarket Strategy," Research Paper No. 1909, Stanford University.

[3] Coglianese, Cary and Jennifer Nash. *Regulating from the Inside: Can Environmental Management Systems Achieve Policy Goals?*, Resources for the Future Press, 2001.

[4] Dasgupta, Partha and Eric Maskin. 1986a. "The Existence of Equilibrium in Discontinuous Economic Games 1: Theory," *Review of Economic Studies*, 53: 1-26.

[5] Dasgupta, Partha and Eric Maskin. 1986b. "The Existence of Equilibrium in Discontinuous Economic Games 2: Applications," *Review of Economic Studies*, 53: 27-42.

[6] Deegan, Craig and B. Gordon. 1996. "A Study of the Environmental Disclosure Policies of Australian Corporations," *Accounting and Business Research,* 26: 187-199.

[7] Deegan, Craig and Michaela Rankin. 1996. "Do Australian Companies Report Environmental News Objectively? An Analysis of Environmental Disclosures by Firms Prosecuted Successfully by the Environmental Protection Authority," *Accounting, Auditing and Accountability Journal*, 9: 50-67.

[8] Deegan, Craig, Michaela Rankin, Peter Voght. 2000. "Firms' Disclosure Reactions to Major Social Incidents: Australian Evidence," *Accounting Forum* 24: 101–130.

[9] Delmas, Magali A. 2000. "Barriers and Incentives to the Adoption of ISO 14001 by Firms in the United States," *Duke Environmental Law and Policy Forum*, 11: 1-38.

[10] Dewatripont, Mathias and Jean Tirole. 1994. *The Prudential Regulation of Banks*. Cambridge, MA: MIT Press.

[11] Feddersen, Timothy J. and Thomas W. Gilligan. 2001. "Saints and Markets: Activists and the Supply of Credence Goods," *Journal of Economics and Management Strategy*, 10: 149-171.

[12] Goodhart, Charles, Philipp Hartmann, David Llewellyn, Lilian Rojas-Suarez, and Steven Weisbrod. 1998. *Financial Regulation: Why, How, Where Now?* London: Routledge Press.

[13] Greer, Jed and Kenny Bruno. 1996. *Greenwash: The Reality Behind Corporate Environmentalism.* Penang, Malaysia: Third World Network and New York: the Apex Press.

[14] Grossman, Gene M. and Elhanan Helpman. 2001. *Special Interest Politics.* Cambridge, MA: The MIT Press.

[15] Gunther, Mark. 2006. "The Green Machine," *Fortune*, v. 154, no. 3, August 7, 2006.

[16] Guthrie, J. and L. Parker. 1990. "Corporate Social Disclosure Practice: A Comparative International Analysis," *Advances in Public Interest Accounting*, 3: 159-176.

[17] Innes, Robert. 2006. "A Theory of Consumer Boycotts under Symmetric Information and Imperfect Competition," *Economic Journal*, 116: 355–381.

[18] Johnson, Geoff. 2003. *Don't Be Fooled: The Ten Worst Greenwashers of 2003*, Boston, MA: Earthday Resources for Living Green. Available at http://www.thegreenlife.org/dontbefooled.html.

[19] Kim, Eun Hee and Thomas P. Lyon, 2006, "Greenhouse Gas Reductions or Greenwash?: The DOE's 1605b Program," Working Paper, Ross School of Business, University of Michigan.

[20] Laufer, William S. 2003. "Social Accountability and Corporate Greenwashing," *Journal of Business Ethics*, 43: 253-261.

[21] Lyon, Thomas P. and John W. Maxwell. 2004a. "Astroturf: Interest Group Lobbying and Corporate Strategy," *Journal of Economics and Management Strategy*, 13: 561-597.

[22] Lyon, Thomas P. and John W. Maxwell. 2004b. *Corporate Environmentalism and Public Policy.* Cambridge: Cambridge University Press.

[23] Margolis, Joshua D. and James P. Walsh. 2003. "Misery Loves Companies: Rethinking Social Initiatives by Business," *Administrative Science Quarterly*, 48: 268-305.

[24] Milgrom, Paul and John Roberts. 1992. *Economics, Organization and Management.* Englewood Cliffs, NJ: Prentice Hall.

[25] Morgenstern, Richard D. and William A. Pizer, editors. 2007. *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe and Japan*, RFF Press.

[26] Patten, Dennis M. 1992. "Intra-Industry Environmental Disclosures in Response to the Alaskan Oil Spill: A Note on Legitimacy Theory," *Accounting, Organizations and Society*, 17: 471-475.

[27] Ramus, Catherine and Ivan Montiel. 2005. "When Are Corporate Environmental Policies a Form of Greenwashing?," Business and Society, 44: 377-414

[28] Shin, Hyun Song. 2003. "Disclosures and Asset Returns," *Econometrica*, 71: 105-133.

[29] Tietenberg, Thomas. 1998. "Disclosure Strategies for Pollution Control," *Environmental and Resource Economics*, 11: 587-602.

[30] Wilbanks, Thomas J. and Paul C. Stern. 2002. "New Tools for Environmental Protection: What We Know and What We Need to Know," in *New Tools for Environmental Protection: Education, Information, and Voluntary Measures*, Washington, DC: National Research Council.

Figure 1: Disclosure Equilibria when Greenwash Penalties are Low (η < .5)

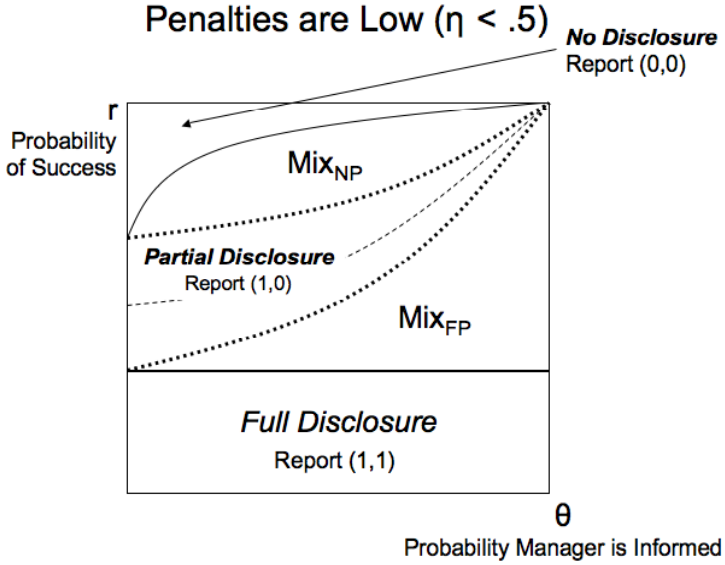

Figure 2: Disclosure Equilibria when Greenwash Penalties are Moderate (η = .5)

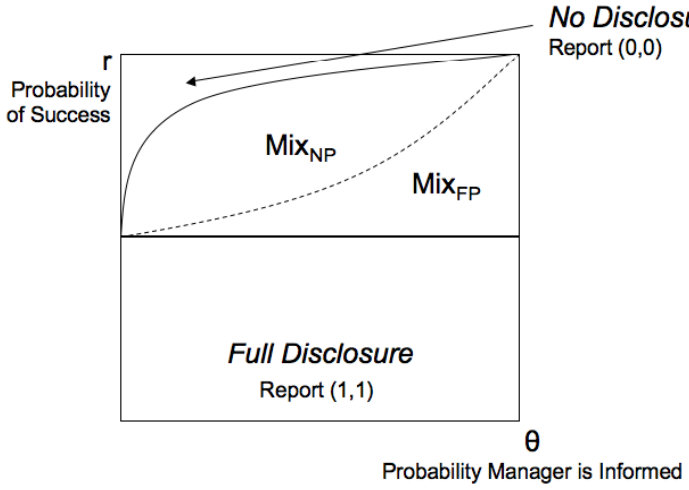

Figure 3: Disclosure Equilibria when Greenwash Penalties are High (.5 < η < 1)