

# Communicating Clean Technology: Green Premium, Competition, and Ecolabels

## **Abstract**

In markets where differences in environmental performance of competing firms arise due to differences in technology that cannot be altered in the short run and firms have private information about their own current technology, ecolabels can allow clean firms to directly and credibly communicate this private information to environmentally conscious consumers. Though direct communication through ecolabels can ameliorate the distortions that occur when firms need to signal their private information indirectly, I show that market competition creates a strategic disincentive for adopting ecolabels (even if the cost of adoption is negligible). Firms adopt ecolabels only if the green premium that buyers are willing to pay is large relative to the production cost advantage of dirty firms. In the latter case, ecolabels reduce market power, increase the market share of clean firms, and reduce expected environmental damage. I analyze the strategic (long run) incentive of firms to invest in the development of clean technology where the outcome of such investment is uncertain. Relative to a benchmark with no ecolabels, the availability of an ecolabel to directly communicate private information about the technology outcome (resulting from investment) actually reduces the ex ante strategic incentive to invest and eventually lowers industry investment in cleaner technology.

# 1 Introduction

Consumers' willingness to pay a *green premium* for the goods produced with lower environmental damage<sup>1</sup> can, in principle, play an important role in creating market incentives for firms to reduce the environmental damage caused by their production process and thus, foster voluntary compliance. An important barrier to the effectiveness of the green premium is that buyers are very unlikely to be able to directly observe attributes of a product that are relevant to its environmental impacts, such as the production technology and the nature of inputs used in the current production process. In fact, firms are more likely to have private information about these attributes. Ecolabels<sup>2</sup> (and other third party certification mechanisms) are instruments that can be used by the environmentally cleaner firms to credibly and voluntarily disclose their private information to the conscious buyers in order to shift and increase the demand for their own product. While the number of firms and industries where ecolabels have been adopted has increased rapidly in recent years, there are a large number of markets where firms do not appear to be adopting ecolabels (or other means of credible voluntary disclosure)<sup>3</sup>. This, however, does not mean that the environmentally conscious buyers necessarily make an uninformed purchase decisions in such markets. Market behavior of firms such as their pricing decisions may indirectly signal hidden information about current production technology and cost structure and this, in turn, can convey information

---

<sup>1</sup>The recent theoretical literature in environmental economics considers environmental friendliness as a vertical attribute of a product and shows that environmentally conscious (green) consumers pay a price premium for an environment-friendly product (see Cremer and Thisse (1989), Arora and Gangopadhyay (2003), Bansal and Gangopadhyay (2003)). Teisl et al. (2002) find that introduction of "dolphin-safe" labels increases the market share of canned tuna. Galarraga and Markandya (2004) show that consumers in the UK pay significant price premium for organic and fair trade coffee. Casadesus-Masanell et al. (2009) find that consumers are willing to pay more for sportswear made of organic cotton that involves lower use of pesticides and fertilizers.

<sup>2</sup>According to the Environmental Protection Agency in the US, "ecolabels aim to define and communicate "what is greener." They can be developed by private entities, by public agencies, or jointly by stakeholders and experts from the public and private sectors." Examples of ecolabels include Energy Star (US), EPA Lead Certification (US), Green-e Energy (US), Eco3Home (US), Agriculture Biologique (EU), Blue Angel (Germany), EU Ecolabel (EU), Swan (Nordic Countries) etc.

<sup>3</sup>According to ecolabelindex.com, a comprehensive online directory of ecolabels, to date there are 465 ecolabels across 25 different industries in 199 countries. There are 148, 74, 22, and 19 ecolabels in the food, electronics, carbon, and water industry respectively. Among the major economies, the US and EU have 203 and 237 ecolabels whereas China and India have 58 and 31 ecolabels respectively.

about the environmental performance of firms to buyers<sup>4</sup>. One should therefore view the adoption of ecolabels (and other means of direct disclosure) as an alternative to communicating information through market based signaling to buyers with green consciousness. The difference between the efficacy of these alternative channels of communication may not lie so much in the information conveyed to buyers as in the way they affect competition, prices, market shares, and profits of different types of firms.

This paper has two objectives. The first objective is to analyze the strategic incentive of a competing firm to adopt a credible ecolabel to disclose private information about how green its current production technology is when the firm can also signal the same information through pricing. This can help us to understand some of the factors behind why ecolabels are more likely to be prevalent in certain kinds of markets than others. Secondly, this paper attempts to understand how, by moving away from a pure signaling based equilibrium, adoption of an ecolabel can affect market allocation, the extent of environmental damage, and eventually the long run incentive to invest in cleaner technology.

The existing literature on ecolabels has highlighted the role of ecolabels in allowing cleaner firms to credibly disclose hidden information or actions (such as abatement) to buyers, and to focus on this role, the literature generally uses a framework where clean firms have no way to credibly convince consumers about their types or actions if they do not adopt ecolabels or if no ecolabels are available<sup>5</sup>. As mentioned above, this ignores the fact that buyers can make inferences about the cost structure and therefore, the underlying nature of production technology of firms from prices and other market variables chosen by firms. In principle, the market signaling mechanism can be as effective as direct disclosure in informing uninformed buyers. This paper builds on a small literature<sup>6</sup> that focuses on the role of signaling (for instance, through prices) in communication of information about the environmental impact of production technology to potential consumers to study the trade-off between ecolabels and price signaling in order to understand the strategic incentives for adoption of ecolabels by competing firms. In my framework, a firm may choose not

---

<sup>4</sup>Hwang et al. (2005) find that consumers use price as a signal of the quality of genetically modified food (corn, bread, and egg).

<sup>5</sup>In this context, it is important to understand that ecolabel resolves asymmetric information problems, namely moral hazard (hidden performance) as well as adverse selection (hidden type) (see Crampes and Ibanez (1996), Kuhn (1999), and Ibanez and Grolleau (2008)). Fischer and Lyon (2014, 2015) focus on the competition among multiple ecolabels that are essentially designed to reveal more information about the nature and degree of the abatement effort by the polluting firms.

<sup>6</sup>See, among others, Sengupta (2015).

to adopt an ecolabel and find it profitable to signal via prices even if there is a credible ecolabel available at a negligible cost. In order to focus on the role of ecolabels<sup>7</sup> and prices in conveying hidden information, I abstract from issues related to current actions of a firm being unobservable, by assuming that a firm cannot alter its production process in the short run in order to affect its environmental impact, and that its long run investment in making the production process cleaner is observable to all (though the final outcome is not).

In particular, I consider an imperfectly competitive industry where two firms compete in prices. All consumers are environmentally conscious and are willing to pay more for the product produced with a technology that causes lower environmental damage. The production technology of a firm can be of two potential types, *dirty* and *clean*. All firms are initially endowed with a dirty production technology and may invest in the development of a cleaner one where the outcome of the investment i.e., whether the realized production process is clean or remains dirty, is intrinsically uncertain. The latter may reflect uncertainty about the success of the project or the environmental impact of the new technology. Investment is observed publicly but not the realized technology<sup>8</sup>. In the next stage, firms choose prices and (clean) firms decide whether to adopt an ecolabel to disclose their actual environmental performance; here disclosure of the type through the ecolabel is assumed to be fully credible. Buyers update their beliefs about the true type of firms' production processes by observing the ecolabel and/or the prices charged before making their purchase decisions.

I find that *the clean firms adopt an ecolabel to credibly disclose their types only when the green premium (that the buyers are willing to pay) is significantly high*. If the green premium is low relative to the production cost advantage of the dirty technology, an ecolabel is not adopted by the clean firms (even if the cost of adopting ecolabels is negligible). This is because a clean type firm

---

<sup>7</sup>A strand of literature in environmental economics critically analyzes the effectiveness of ecolabel as a pollution control instrument (See Kuhn (1999), Matto and Singh (1994) and Dosi and Moretto (2001)). Amacher et. al. (2004) examine whether firms have incentive to invest in abatement technology to be able to qualify for and retain an ecolabel in the absence of any asymmetric information between competing firms and environmentally conscious consumers. In particular, they find that the nature of cost differentials determines whether firms invest in cleaner technology and adopt ecolabel; however, unlike the present paper an ecolabel is not implemented to resolve any incomplete information problem among firms and consumers.

<sup>8</sup>Coal energy plants roughly fit the basic characteristics of the model described in this paper. The final outcome of any R&D investment in the carbon capture and sequestration (CCS) technology (i.e., clean coal) is stochastic per se. Here, the success in (cleaner technology) CCS can be interpreted as "higher" or "better" than the existing level of carbon capture by CCS and failure means not being able to improve the percentage of carbon capture. Further, precise information about the carbon capture is likely to be observed only by the power generating plants.

with an ecolabel continues to be at a competitive disadvantage when the rival is of the dirty type and thus, focuses on Bertrand like aggressive competition with a rival of the clean type which, in turn, eliminates the rent necessary to cover the cost of adopting the ecolabel. On the other hand, when the green premium is significantly high (such that it exceeds the relative production cost advantage of the dirty technology), adopting an ecolabel enables the clean type to undercut a dirty rival, without being punished by the beliefs of buyers who may associate lower prices with lower production cost i.e., dirty technology. Armed with the ecolabel, a clean firm is able to exercise its full competitive advantage and in fact, to capture the entire market when the rival is dirty.

Interestingly, if no ecolabels are available, the market outcome for this case of high green premium is one where firms signal their types only through prices, with the clean firms charging high prices but ceding the market to the dirty rivals that charge relatively low prices. In fact, signaling through prices requires the dirty types to earn sufficient rent so as to not imitate the prices of the clean types. Consequently, the clean type has to charge a very high price to deter any imitation. This, in turn, allows a competing dirty firm to steal the entire market at relatively higher prices when the rival is of the clean type. The reason, this can be sustained as an equilibrium outcome, is primarily because the buyers cannot directly observe the types and thus, the beliefs (of the buyers) punish the clean firms that try to undercut and gain market share from the dirty rivals. As a result, the market power of firms tends to be high which creates multiple distortions; for example, the buyers buy from the dirty firms even when the clean product is available and the clean type creates higher social surplus. Buyers, however, learn about the true type of the firms from prices even though no ecolabels are available. Thus, the provision of a credible ecolabel (at small cost) is not necessary to enable the buyers to make more informed purchase decisions. The present analysis indicates that the main reason why provision of a credible ecolabel is desirable, is because it *removes the market power and the distortions associated with the price signaling outcome*; it shifts the market share from the dirty to the clean firms that can now use the ecolabel to exercise their competitive advantage. Consequently, the environmental outcome is improved and so is the social welfare.

Note that in a monopoly market, a clean seller would always adopt an ecolabel (with negligible adoption cost) to directly communicate its type as it allows him to remove the price distortion needed to convince environmentally conscious buyers about his true type through price signaling. Thus, our results indicate that competition plays an important role in deterring adoption of ecolabels in certain markets. With competition, buyers need to be willing to pay a fairly high green

premium in order to motivate firms to adopt ecolabels and to reap the benefits of lower prices and distortions,

Finally, I analyze a firm's long-run incentive to invest in the development of a clean technology where investment is publicly observable but the outcome (success or failure) is only known to the firm. Whether or not disclosure through ecolabels is a feasible option, a firm always has a positive unilateral incentive to invest in clean technology (when its rival does not invest); however, its reciprocal incentive to invest (when its rival invests) is strictly positive only if the green premium that buyers are willing to pay is significantly high. Relative to a benchmark where no ecolabels are available and firms signal only through prices, the ability to credibly disclose through an ecolabel improves a firm's unilateral incentive to invest especially when the consumers pay a high green premium. However, the presence of an ecolabel does not necessarily increase a firm's reciprocal incentive to invest relative to the price signaling benchmark. This is because, in the price signaling outcome, incomplete information (created by uncertainty about the realized technology of a firm that invests) can eventually increase market power to deter imitation of the clean type's price by the dirty type; this can create additional incentive to invest that is not present when types are disclosed directly through the ecolabel. I show that when both firms simultaneously make their investment decisions, then compared to a benchmark with no ecolabels (pure price signaling scenario), the availability of an ecolabel never increases aggregate investment (number of investing firms) and for a range of parameters, actually decreases aggregate investment. Interestingly, the effect of an increase in the green premium (that buyers are willing to pay for the clean product) on investment (i.e., a firm's strategic incentive to invest as well as the equilibrium investment behavior) is non-monotonic; increase in environmental consciousness may have a perverse effect on the environmental cleanness of the production technology.

The remainder of the paper is organized as follows. The next section describes the basic model. Section 3 discusses market outcomes when there is no ecolabel and firms signal their environmental performances through prices. In Section 4, I examine how the green premium affects firms' decisions to adopt an ecolabel to disclose their environmental performances as well as other market outcomes. Section 5 presents an augmented version of the basic model, discusses whether competing firms have strategic incentive to invest, studies their equilibrium investment behaviors, and also compares the equilibrium investment outcomes with and without ecolabel. In Section 6, I briefly discuss market outcomes (pricing as well as investment equilibria) under mandatory disclosure laws. The last section concludes.

## 2 The basic model

I consider a market where the production process of two firms ( $i = 1, 2$ ) that compete in prices may cause an environmental damage. The production technology of each firm can be of two potential types *dirty* ( $D$ ) and *clean* ( $C$ ). The products of the firms are not differentiated in any dimension other than the environmental impacts of their respective production technologies. Each firm produces at a constant unit cost. The unit production costs of a clean type and a dirty type are defined by  $m_C$  and  $m_D$  respectively. A firm can adopt an ecolabel (or signal through prices) to directly and credibly communicate the environmental impact (or nature) of its *current* production technology. I assume that only clean firms are eligible to adopt an ecolabel and a (clean) firm's decision to adopt an ecolabel is a *short run* decision (as it pertains to current production process).

There is a unit mass of risk-neutral consumers in the market. Consumers have unit demand i.e., each consumer buys at most one unit of the good. All consumers are environmentally conscious<sup>9</sup> in the sense that they are willing to pay a premium,  $\Delta > 0$ , for a unit of the clean type's product; I call this as the *green premium* in the rest of the paper. The *green premium* that the consumers are willing to pay for the product produced by the cleaner technology acts as the *measure of environmental consciousness* of consumers. The consumers have identical valuation  $V$  for a unit of the dirty product and  $(V + \Delta)$  for a unit of a clean product. I assume that  $V > m_C$ , all consumers are aware of the unit production cost of the clean type ( $m_C$ ) as well as of the dirty type ( $m_D$ ), the green premium does not exceed the maximum surplus that a dirty type can earn<sup>10</sup> i.e.,

$$\Delta \leq \bar{\Delta} = (V - m_D), \quad (\text{Assumption 1})$$

and the clean type produces at a higher unit cost<sup>11</sup> i.e.,

$$0 < m_D < m_C. \quad (\text{Assumption 2})$$

---

<sup>9</sup>An alternative and more general structure is where a fraction of consumers are environmentally conscious and a strictly positive fraction does not care i.e., this group of people are not willing to pay the green premium for the product produced by cleaner technology. If I introduce this heterogeneity among the consumers then the qualitative nature of the equilibria in this paper do not change; rather it just makes the characterization of the equilibrium technically messy and complicated.

<sup>10</sup>If the green premium exceeds this threshold i.e.,  $\Delta > V - m_D$ , in the absence of any ecolabel the clean type has to charge its own full information monopoly price  $(V + \Delta)$  to prevent imitation by the dirty type; this implies that some consumers may not buy at all when both firms are of clean type. To ensure that all consumers buy in the equilibrium even when the clean type does not have an ecolabel to adopt, I impose this upper bound  $(V - m_D)$  on the green premium.

<sup>11</sup>This happens to be a standard assumption in the existing literature.

I consider *three* different levels of green premium, namely

- (1) Low-green premium:  $0 < \Delta \leq \frac{(m_C - m_D)}{2}$ ,
- (2) Intermediate-green premium:  $\frac{(m_C - m_D)}{2} < \Delta \leq (m_C - m_D)$ , and
- (3) High-green premium:  $(m_C - m_D) < \Delta \leq V - m_D = \bar{\Delta}$ .

The *low-green premium* and the *intermediate-green premium* scenarios arise when the green premium that the environmentally conscious consumers are willing to pay is lower than the cost difference between the clean and the dirty production technologies. In both these cases,  $V + \Delta - m_C \leq V - m_D$  which implies that the surplus generated by the dirty type is greater than that of the clean type. When the green premium exceeds the cost difference ( $\Delta > m_C - m_D$ ), the clean type generates strictly higher surplus ( $V + \Delta - m_C > V - m_D$ ) and has competitive advantage over the dirty type.

Formally, I have a three-stage game. First, nature independently draws the type (environmental performance) of the production technology of each firm from a distribution that assigns probabilities  $\mu \in (0, 1)$  and  $(1 - \mu)$  to  $C$  and  $D$  respectively; it is common knowledge. However, the realization of the type of the production technology of a firm remains private knowledge to the firm and unknown to the rival firm as well as to the consumer. In the next stage, (clean) firms (having observed their own types) simultaneously decide whether to adopt the ecolabel and also choose prices to disclose the environmental performance to consumers<sup>12</sup>. Finally, consumers, after observing the firms' decisions to adopt the ecolabel and their prices, decide to buy. It is perhaps important to note that irrespective of a (clean) firm's decision to adopt an ecolabel, the equilibrium price always signals the firm's environmental performance. The pay-off of each firm is its *expected profit* whereas the pay-off of each consumer is her *expected net surplus*.

Further, I assume that there is only one *credible* binary ecolabel offered by a third-party certifier and the cost of adopting such an ecolabel is strictly positive. This cost of adoption is denoted by  $\epsilon$ . In order to focus on the strategic incentive to adopt an ecolabel, it is assumed that  $\epsilon$  is sufficiently small. In particular,

when  $\frac{m_C - m_D}{2} < \Delta \leq m_C - m_D$

$$\epsilon < \min \left\{ \mu(V + \Delta - m_C), \frac{\mu}{2} (m_D + 2\Delta - m_C) \right\} \quad (\text{Assumption 3})$$

---

<sup>12</sup>Observe that, in this paper, firms simultaneously choose price and decide whether to adopt an ecolabel in the same stage. If one assumes that the decision to adopt an ecolabel for the production process is a long run one, firms may choose prices after they decide to adopt an ecolabel. Moreover, this not only depicts an alternative economic structure compared to the present paper, the qualitative nature of the market outcomes is quite different as well.



and when  $\Delta > m_C - m_D$

$$\epsilon < \min \{(1 - \mu)(m_D + \Delta - m_C), \mu(m_D + \Delta - m_C)\}. \quad (\text{Assumption 4})$$

I investigate the effects of the availability of the ecolabel and the green premium on certain market variables. Though most of these variables are typical market outcomes, to avoid any confusion, I define them explicitly in the context of this specific model. The *market power* of a firm is defined by the firm's ability to charge a price higher than its own marginal cost and is measured by the mark-up i.e., difference between the equilibrium price and the respective marginal cost. The *ex ante expected profit* of a firm is

$$\mu\pi_C + (1 - \mu)\pi_D \quad (\text{Expected profit})$$

where  $\pi_C$  is the clean type's expected profit and  $\pi_D$  is the dirty type's expected profit<sup>13</sup>. Note that I do not explicitly define an environmental damage function, but I consider the probability that the dirty type sells in the equilibrium as a good proxy for the *expected environmental damage* created by the industry.

One of the main objectives of this paper is to compare firms' market behavior under three different information disclosure mechanisms that the firms can use to reveal their own environmental performance to rivals and to the consumers. In the rest of the paper, I denote *ex ante* expected profit of a firm (and later other market outcomes) when firms can disclose their environmental performances (1) with the ecolabel (as well as by signaling via prices), (2) by (pure) signaling via prices without any credible ecolabel, and (3) under mandatory disclosure laws with superscripts  $E$ ,  $S$ , and  $M$  respectively.

### 3 Market outcomes without any ecolabel

Consider a scenario, where there is no ecolabel available for the firms to reveal their respective environmental performances to the consumers. In the absence of any voluntary disclosure mechanism, the firms can signal their environmental quality via their respective prices when the consumers are aware of the marginal cost associated with both types of technology. In a somewhat different context, this particular case has been extensively discussed in Sengupta (2015).<sup>14</sup> In this section,

<sup>13</sup>The expected profit of any type is the product of the probability that this particular type sells (which depends on the type of its rival), its market share, and mark up (price minus marginal cost and the cost of adoption of the ecolabel (the latter cost is considered only if the firm is of clean type and it decides to adopt)).

<sup>14</sup>However unlike this paper, in Sengupta (2015) not only the environmental consciousness of consumers but also the expected future liability associated with the future environmental damage motivates the competing firms to

I highlight the relevant adopted results that will help the reader to understand importance of an ecolabel in communicating environmental performance (in the next section).

I consider the three-stage game as described in Sengupta (2015). In the first stage, nature independently draws the type (environmental performance) of the production technology of each firm from a distribution that assigns probabilities  $\mu \in (0, 1)$  and  $(1 - \mu)$  to  $C$  and  $D$  respectively. The realization of the type of the production technology of a firm remains private knowledge to the firm and unknown to the rival firm as well as to the consumer. In the next stage, the firms choose prices simultaneously to signal the environmental performance to the consumers. The solution concept used in the signaling game is that of the Perfect Bayesian Equilibrium (PBE) which is supported by the out-of-equilibrium beliefs<sup>15</sup> that satisfy Cho-Sobel (1990) D1 Criterion<sup>16</sup>. Finally, the consumers observe the prices charged by the firms, update their beliefs, decide whether to buy, and from which firm to buy. The following lemma summarizes the pricing equilibria (in the light of Sengupta (2015)).

**Lemma 1** *In the absence of any ecolabel, a clean type charges a deterministic price  $p_C$  which is higher than any price charged by a dirty type, and the dirty type follows a mixed strategy with support  $[\underline{P}_D, \overline{P}_D]$  and a continuous distribution function  $F_D(p)$ , where*

$$p_C = \begin{cases} m_C & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ m_D + 2\Delta & \text{if } \frac{m_C - m_D}{2} \leq \Delta \leq \overline{\Delta} \end{cases}, \quad (1)$$

---

strategically invest in the cleaner technology, and there is a fraction of consumers who are not willing to pay the green premium for the product of the clean firm.

<sup>15</sup>Consumers form beliefs about the quality of the product after observing the prices charged by the firms on the equilibrium path. Firms compute their expected profits on the basis of these consumer beliefs. However, in order for the firms to decide that the price they charge on the equilibrium path is actually optimal they need to figure out the profit they would earn if they charged a different price (an "out of equilibrium" price). In the latter event (a zero probability event in the conjectured equilibrium), the buyers' optimal buying behavior would depend on what they would believe about the quality of the product if they were to observe such an out of equilibrium price. Criteria such as the Intuitive Criteria and the D1 refinement impose strong restrictions on these out of equilibrium beliefs and therefore refine the set of Perfect Bayesian Equilibrium. In this paper, I have mostly used D1 refinement criterion which happens to be one of the stronger "reasonableness" restriction on specification of out of equilibrium beliefs.

<sup>16</sup>I consider the D1 equilibrium of the pricing game in the second stage. This strong refinement criterion is originally developed by Cho and Sobel (1990) in the context of pure signaling games with one sender. Janssen and Roy (2010) modify and adapt D1 criterion in their model with multiple senders (firms). An out-of-equilibrium belief satisfies D1 criterion if consumers believe that the off equilibrium price is charged by the type which has relatively higher incentive to deviate to that price (given the equilibrium strategy of the rival) compared to the other type.

$$\bar{P}_D = p_C - \Delta, \underline{P}_D = \mu [p_C - \Delta] + (1 - \mu) m_D, \quad (2)$$

$$\text{and } F_D(p) = 1 - \frac{\mu}{(1 - \mu)} \left( \frac{p_C - \Delta - m_D}{p - m_D} - 1 \right) \quad (3)$$

in the unique perfect Bayesian (separating) equilibrium; it is supported by out-of-equilibrium beliefs that satisfy D1 criterion.

The clean type sells only when the rival is of clean type too, otherwise the dirty type caters to all consumers.

**Proof.** See Appendix. ■

In a separating equilibrium without any ecolabel, a clean type must charge a price such that after observing the price consumers believe that it is a clean type with probability one i.e., consumers should be convinced that a dirty type will not charge such a price as it is not profitable for the dirty type to do so. From the above mentioned lemma, one can see that there is no separating equilibrium in pure strategies. In the separating equilibrium, the dirty type (with lower marginal cost) ought to earn sufficient positive rent otherwise it will imitate clean type's equilibrium price. If the rival is of clean type (with higher marginal cost), a dirty type can earn a strictly positive rent by charging a lower price and does not have any incentive to imitate the clean type's higher price. However, in a state where the rival is of dirty type, it has an incentive to undercut the dirty rival (with the same marginal cost). Therefore, the dirty type (with lower marginal cost) randomizes over a price interval. Lack of information about the actual environmental performances of firms does not only allow the clean type but also the dirty type to enjoy stochastic market power even when all consumers are environmentally conscious (willing to pay more for the products of the clean type).

When the green premium is *not high* enough ( $\Delta \leq m_C - m_D$ ) a dirty type generates higher surplus and has competitive advantage over the clean type. In particular, the low green premium ( $\Delta \leq \frac{m_C - m_D}{2}$ ) drives the clean type's equilibrium price to as low as its own marginal cost ( $m_C$ ); the dirty type randomizes over  $[\mu(m_C - \Delta) + (1 - \mu)m_D, m_C - \Delta]$  (where at the upper bound ( $m_C - \Delta$ ) a consumer is indifferent between buying from the dirty type at this price and from the clean type at  $m_C$ ), earns strictly positive profit, and sells to all consumers. However, as the incentive to imitate increases with increase in the green premium the clean type needs to charge a higher price than its own marginal cost which, in turn, helps to increase the market power and profitability of both types. If  $\Delta \geq \frac{m_C - m_D}{2}$  ( $\implies m_D + 2\Delta \geq m_C$ ), then the dirty type has an incentive to imitate the clean type's marginal cost pricing; thus, the clean type raises its price to

$m_D + 2\Delta$  as long as  $\Delta \leq V - m_D = \bar{\Delta}$ .

Observe that even though all consumers are environmentally conscious, the clean type can only sell when both firms are of the clean type and can make strictly positive profit only when the green-premium is not low ( $\Delta \geq \frac{m_C - m_D}{2}$ ). The equilibrium expected profits of the clean type and the dirty type are

$$\pi_C = \begin{cases} 0 & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ \frac{\mu}{2} (m_D + 2\Delta - m_C) & \text{if } \frac{m_C - m_D}{2} \leq \Delta \leq \bar{\Delta} \end{cases}$$

and

$$\pi_D = \begin{cases} \mu (m_C - \Delta - m_D) & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ \mu \Delta & \text{if } \frac{m_C - m_D}{2} \leq \Delta \leq \bar{\Delta} \end{cases}$$

respectively. Thus, the *ex ante* expected profit of a firm when there is no ecolabel is given by

$$\pi^S = \begin{cases} (1 - \mu) \mu (m_C - \Delta - m_D) & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ \frac{\mu^2}{2} (m_D + 2\Delta - m_C) + (1 - \mu) \mu \Delta & \text{if } \frac{m_C - m_D}{2} \leq \Delta \leq \bar{\Delta} \end{cases}. \quad (4)$$

## 4 Market outcomes with an ecolabel

In this section, I critically examine a firm's incentive to disclose its environmental performance to consumers via an ecolabel (as well as through market prices). I also investigate how the green premium affects short run market outcomes such as market power, expected profit, and expected environmental damage in the presence of the ecolabel. Finally, I compare the market outcomes with and that of without an ecolabel.

I find that a clean type does not always find it profitable to adopt the ecolabel to credibly disclose its environmental performance to consumers. In particular, this is true when the green premium is low ( $\Delta \leq \frac{m_C - m_D}{2}$ ). When the green premium is in the intermediate range ( $\frac{m_C - m_D}{2} < \Delta \leq m_C - m_D$ ), only one of the clean firms adopts an ecolabel. However, when the green premium is sufficiently high ( $m_C - m_D \leq \Delta \leq \bar{\Delta}$  such that it generates comparatively higher surplus for the clean type), all clean type firms adopt the ecolabel in the equilibrium. Thus, one can conclude that the level of the green premium (which measures the environmental consciousness) determines a firm's incentive to adopt the ecolabel.

Further, I show that the availability of an ecolabel yields less distortionary equilibrium market outcomes compared to a disclosure regime without any ecolabel (pure signaling). To be more specific, under the high-green premium, unlike in the case of pure signaling (without any ecolabel), the clean type does not have to charge an exceptionally higher price to prevent the dirty type from

imitating; rather adoption of an ecolabel in the equilibrium helps the clean type to charge a less distortionary price. The clean type always sells to the entire market when it discloses its type by adopting the ecolabel. This, in turn, implies that the expected environmental damage is lower with the ecolabel compared to that of pure signaling via price.

First, I solve the (two-sided) *incomplete information* simultaneous move game played by the firms in the second stage (described in Section 2), where the firms with private knowledge about their own environmental performance decide whether to adopt the ecolabel and compete in prices. The solution concept used is that of perfect Bayesian equilibrium.

**Proposition 1 (Low-green premium)** *Consider the low-green premium  $\left(\Delta \leq \frac{(m_C - m_D)}{2}\right)$ . A clean type does not adopt any ecolabel and charges a deterministic price which is equal to its marginal cost whereas a dirty type randomizes over an interval viz.,*

$$p_C = m_C, p_D \in [\mu(m_C - \Delta) + (1 - \mu)m_D, m_C - \Delta], \quad (5)$$

*in the unique perfect Bayesian (separating) equilibrium which is supported by out-of-equilibrium beliefs that satisfy D1 criterion.*

*The clean type sells only when the rival is of clean type too, otherwise all consumers buy from the dirty type.*

**Proof.** See Appendix. ■

The above proposition depicts two major sources of *distortion*. One stems from the fact that all consumers, though they are willing to pay more for the product produced by the cleaner technology, often end up buying from the dirty type except when both firms are of clean type in the equilibrium. Even though the equilibrium prices reveal the actual environmental impact of the production process of the dirty firms, the inability of a clean type to offer a competitive price drives the cleaner firms away from the market. This, in turn, leads to additional distortion in the equilibrium market outcomes. In particular, even if the clean type has an option to voluntarily adopt the ecolabel to disclose its own environmental performance to the consumer, individual clean firm chooses not to do so in the equilibrium.

In addition, note that the pure signaling equilibrium (discussed in the previous section in Lemma 1) is sustained even in the presence of an ecolabel under the low-green premium  $\left(\Delta \leq \frac{m_C - m_D}{2}\right)$ . Since the clean type does not have any competitive advantage and thus, fails to earn a strictly positive profit, it does not adopt an ecolabel (even if the cost of adoption is zero) in the equilibrium.

In Lemma 1, I consider the symmetric Perfect Bayesian signaling equilibrium that satisfies D1 refinement; one can show that all possible symmetric signaling Perfect Bayesian equilibria (PBE)<sup>17</sup> collapse with an ecolabel once the green premium exceeds the lowest threshold i.e.,  $\Delta > \frac{m_C - m_D}{2}$ .

**Lemma 2** *In the presence of an ecolabel, no symmetric Perfect Bayesian signaling (separating) equilibrium can be sustained when consumers are willing to pay more than the low-green premium (i.e.,  $\Delta > \frac{m_C - m_D}{2}$ ).*

**Proof.** See Appendix. ■

If there is no ecolabel available and the consumers are willing to pay more than the low-green premium, then in a symmetric (separating) PBE the clean types charge a deterministic price which is greater than its own marginal cost ( $m_C$ ) (such that it earns a strictly positive expected profit) as well as any price that a dirty type charges (so that the latter cannot imitate the clean type's price) and sell only in the state where the rival is of clean type too. However, if an ecolabel is available (only for the clean type to adopt) at a strictly positive but sufficiently small cost (depicted by Assumption 3 and Assumption 4) then the clean types can adopt the ecolabel (and decide on prices in the same stage) and charge a lower price than the signaling equilibrium price to credibly convince the consumers of their environmental performances. Note that the out-of-equilibrium beliefs used to justify a signaling equilibrium (in the absence of any ecolabel) become redundant when a clean type adopts an ecolabel. In particular, the restrictions on the cost of an ecolabel guarantees a gainful deviation (from pure signaling to prices with ecolabel) by a clean type as *ex ante* expected profit of the clean type without any ecolabel is greater than the cost of ecolabel.

**Proposition 2 (Intermediate-green premium)** *Consider the intermediate-green premium case where*

$$\frac{m_C - m_D}{2} < \Delta \leq m_C - m_D.$$

*Suppose that the cost of adopting an ecolabel is small enough*

$$0 < \epsilon < \min \left\{ \mu(V + \Delta - m_c), \frac{\mu}{2} (m_D + 2\Delta - m_C) \right\}$$

*Then:*

*(i) A symmetric separating perfect Bayesian equilibrium does not exist.*

---

<sup>17</sup>The complete set of symmetric PBE can be adopted from Proposition 2 of Janssen and Roy (2010).

(ii) There exist (two) asymmetric separating perfect Bayesian equilibria where only one of the two firms adopts an ecolabel (when it is of clean type) and firms follow identical pricing strategies. In particular, clean types charge a deterministic price

$$p_C = m_C + \frac{\epsilon}{\mu} \quad (6)$$

and the dirty types randomize over an interval  $[\underline{p}_D, \bar{p}_D]$  where

$$\underline{p}_D = \mu \left( m_C + \frac{\epsilon}{\mu} - \Delta \right) + (1 - \mu)m_D, \quad \bar{p}_D = p_C - \Delta, \quad (7)$$

and the distribution function is given by

$$F_D(p) = 1 - \frac{\mu}{(1 - \mu)} \left( \frac{m_C + \frac{\epsilon}{\mu} - \Delta - m_D}{p - m_D} - 1 \right), p \in [\underline{p}_D, \bar{p}_D]$$

Buyers buy from dirty firms except in the state where both firms are of clean type, and in the latter case, they buy only from the firm adopting the ecolabel. The clean type of the firm that does not adopt the ecolabel never sells. The clean type of the firm adopting the ecolabel breaks even. These equilibria are supported by out-of-equilibrium beliefs that satisfy the Cho-Kreps (1987) Intuitive Criterion.

**Proof.** See Appendix. ■

Recall the two distortions observed in case of the low-green premium. A clean type is not able to offer a competitive price even under the intermediate-green premium and thus loses the entire market to its rival of dirty type. However, the second distortion goes away as one of the clean types does find it profitable to adopt an ecolabel in the equilibrium. Next, I show that both distortions vanish from the equilibrium outcomes when the green premium is high enough to generate competitive advantage for the clean type.

**Proposition 3 (High-green premium)** Consider the high-green premium ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ). The clean type adopts an ecolabel and randomizes over an interval whereas the dirty type charges its own marginal cost viz.,

$$p_C \in [\mu m_C + (1 - \mu)(m_D + \Delta), m_D + \Delta] \text{ and } p_D = m_D \quad (8)$$

in the unique perfect Bayesian (separating) equilibrium that satisfies D1 refinement.

The clean type caters to the entire market. The dirty type sells only when the rival is of dirty type as well.

**Proof.** See Appendix. ■

From the above proposition, one can conclude that the least distortionary equilibrium outcomes can be obtained under the high-green premium. In the equilibrium, the clean type not only adopts the ecolabel to disclose its environmental performance to the consumers but also manages to capture the entire market and earns strictly positive expected profit. Incomplete information about the type of the rival softens the price competition and creates positive expected profit for the clean type even when the rival is of the same type. The *ex ante* expected profit of a firm is

$$\pi^E = \begin{cases} \mu(1 - \mu)(m_C - \Delta - m_D) & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ (1 - \mu)\mu \left( m_C + \frac{\epsilon}{\mu} - \Delta - m_D \right) & \text{if } \frac{m_C - m_D}{2} < \Delta \leq m_C - m_D \\ \mu[(1 - \mu)(m_D + \Delta - m_C) - \epsilon] & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases} . \quad (9)$$

The following corollary illustrates the effect of the environmental consciousness of the consumers on the market power and expected profit of a firm. Rise in the level of environmental consciousness among consumers is measured by the increase in the green premium that the consumers are willing to pay for the cleaner product.

**Corollary 1** *Consider the high-green premium ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ) scenario when the ecolabel is available. An increase in the environmental consciousness ( $\Delta$ ) among consumers increases the market power as well as the expected profit of a firm. The expected environmental damage is lower compared to that under the low-green premium ( $\Delta \leq \frac{m_C - m_D}{2}$ ).*

The next Corollary summarizes the interesting features of the comparative analyses across these two disclosure regimes.

**Corollary 2** *Consider the high-green premium ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ).*

(i) *Availability of an ecolabel yields less distortionary equilibrium outcomes compared to signaling via prices. The clean type always sells under the high-green premium equilibrium with the ecolabel whereas it may not sell when firms can only signal through price.*

(ii) *The expected environmental damage is lower with the ecolabel compared to that of the signaling via price.*

The underlying reason behind less distortionary and environmentally superior market outcomes with the ecolabel compared to that of pure signaling (via price) is as follows. In the pure signaling equilibrium, the dirty type needs to have market power and earn strictly positive rent, so that it



does not have the incentive to imitate the clean type's price. The incentive to imitate increases with increase in the green premium; this implies that to prevent the dirty type, the clean type has to increase its price as well. Therefore, even when environmental consciousness among all consumers is pretty high ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ), in the pure signaling equilibrium the clean type cannot sell if there is a dirty type in the market (see Lemma 1); however, under the high-green premium a clean type always sells with an ecolabel (see Proposition 3). To be precise, the probability that the dirty type sells in the pure signaling equilibrium is higher than that of with the ecolabel (i.e.,  $(1 - \mu^2) > (1 - \mu)^2$ ). Note that the expected profit of a firm is higher in case of signaling compared to the regime with the ecolabel, also because of the positive rent earned by the dirty type (compare (4) and (9)). However, the market power of the dirty type under signaling always decreases with increase in the green-premium.

## 5 Investment in clean technology

In this section, I consider an augmented version of the basic model (discussed in Section 2). In particular, I assume that firms are initially endowed with a dirty production technology, which means that each firm incurs a unit cost of  $m_D$ . In the first stage, firms simultaneously decide whether or not to invest in the development of the clean technology. Suppose investment in clean technology requires a fixed cost  $F > 0$ . This is a long run decision that cannot be modified in the short run. The actions chosen by each firm at this stage i.e., whether or not it has invested is observed by both firms and consumers. If it does not invest, a firm remains dirty with probability one, and this is known to all. If it invests then the realized production technology is clean with probability  $\mu \in (0, 1)$ <sup>18</sup> and dirty with probability  $(1 - \mu)$ ; this is common knowledge. However, the realized production technology is pure private information - unknown to the rival firm as well as to consumers. The realizations of production technology after investments are independent across firms. If a firm attains the clean technology as a result of investment then the firm incurs a marginal cost of  $m_C$ . In the next stage, firms simultaneously decide whether to adopt the ecolabel as well as choose prices to disclose the environmental performance to consumers. Finally, consumers decide to buy.

Note that there are three possible information structures in the second stage of the augmented

---

<sup>18</sup>It can be alternatively interpreted as the probability that the newly developed technology eliminates accidents that can damage the environment or prevents currently unknown externalities on the environment in the future.

game, following any profile of investment (long run) decision made in the first stage. In the first case, both firms decide not to invest, both remain dirty for sure, and the pricing game in the second stage degenerates to a standard (1) *full information* symmetric Bertrand price competition game. For any value of green premium, both firms charge a common price equal to the marginal cost of production of the dirty type ( $m_D$ ), and both earn zero profit. A more interesting case arises under the second situation viz., when only one firm invests. Here the pricing game is a (2) *one-sided incomplete information* game; the firm that invests becomes clean with probability  $\mu$  and remains dirty with probability  $(1 - \mu)$ , while the firm that does not invest stays dirty for sure. Lastly, I consider the scenario where both firms invest in the first stage. In this case, the final outcomes of the investment undertaken by both firms are private information i.e., firms do not know each other's type when they strategically compete in terms of prices and decide whether to adopt ecolabel. I have already discussed the relevant pricing equilibria under (3) *two-sided incomplete information* framework with as well as without an ecolabel in the previous sections.

I denote the *ex ante* expected profit of an investing firm by  $\pi_{II}$  and  $\pi_{INI}$  if the rival invests and does not invest respectively, whereas the *ex ante* expected profits of a non-investing firm given that the rival invests and does not invest are denoted by  $\pi_{NII}$  and  $\pi_{NINI}$  respectively. In this paper, the *strategic incentive* of a firm to invest in cleaner technology is defined by the difference between the *ex ante* expected profit of the firm if it invests and the expected profit when it does not invest. The strategic incentive to invest differs between situations where the rival firm does not invest and the rival invests. In particular, the *unilateral incentive (UI)* to invest in cleaner technology is defined as the difference between *ex ante* expected profit of an investing firm when the rival does not invest and the expected profit of a firm when both firms do not invest

$$UI = \pi_{INI} - \pi_{NINI} \quad (\text{Unilateral incentive to invest})$$

whereas the *reciprocal incentive (RI)* to invest is the *ex ante* expected profit of an investing firm when both firms invest minus the *ex ante* expected profit of a non-investing firm when the rival invests

$$RI = \pi_{II} - \pi_{NII}. \quad (\text{Reciprocal incentive to invest})$$

If  $UI \geq 0$  then a firm has an incentive to invest in cleaner technology even if the rival does not invest; moreover if  $RI \geq 0$  then a firm has reciprocal incentive to invest. In equilibrium, at least one firm invests if the unilateral incentive to invest is at least as high as the fixed cost ( $UI \geq F$ ), and both firms invest when the reciprocal incentive to invest exceeds the fixed cost of investment

$(RI \geq F)$ .

First, I discuss the market outcomes if only one firm invests (i.e., one-sided incomplete information case) when there is an ecolabel available to adopt as well as in the absence of any ecolabel. In the second and third subsection, I discuss the effect of the green premium (i.e., environmental consciousness of consumers) on a firm's strategic (unilateral as well as reciprocal) incentive to invest and their equilibrium investment behavior with and without an ecolabel respectively. Next, I compare the strategic incentives and equilibrium investment behavior of firms with and without the option of credibly disclosing environmental performance via the ecolabel.

I find that when the green premium is high enough (to generate relatively higher surplus for the clean type compared to the dirty i.e.,  $m_C - m_D < \Delta \leq \bar{\Delta}$ ) the unilateral incentive to invest is strictly positive with ecolabel whereas it is zero in the absence of any ecolabel. However, presence of an ecolabel does not improve a firm's reciprocal incentive to invest. In particular, when the green premium is not low (i.e.,  $\frac{m_C - m_D}{2} < \Delta \leq \bar{\Delta}$ ) a firm with the option of adopting an ecolabel has a strictly higher reciprocal incentive to invest relative to the pure price signaling case. Further, I find that the *total investment* (which is defined by the number of firms that invest in the equilibrium) is higher when there is no ecolabel available.

## 5.1 Only one firm invests

Suppose that in the first stage of the augmented investment game, only one firm invests in cleaner technology. This leads to a one-sided incomplete information structure in the following pricing game. First, I discuss the case where if the investing firm becomes clean, it tries to convince the consumers that it is of clean type by adopting an ecolabel (whenever it is profitable to do so). I argue that there exists a unique perfect Bayesian equilibrium of the one-sided incomplete information pricing game where the investing firm charges a higher price when it is of clean type than when it is of dirty. The clean type has more incentive to charge higher price because of its relatively higher marginal cost.

**Lemma 3** *Suppose there is an ecolabel available, only one firm invests, and the green premium is low enough to generate higher surplus for the dirty type ( $\Delta \leq m_C - m_D$ ). No (clean) firm adopts any ecolabel, the clean type charges its own marginal cost ( $m_C$ ) and sells zero whereas the dirty type randomizes over a price interval  $[\mu(m_C - \Delta) + (1 - \mu)m_D, m_C - \Delta]$  in the unique perfect Bayesian (separating) equilibrium which is supported by out of equilibrium beliefs that satisfy D1 criterion.*

All consumers buy from the dirty type.

**Proof.** See Appendix. ■

The above lemma depicts that if only one firm invests the clean type can never sell in the equilibrium with  $\Delta \leq m_C - m_D$ . The green premium is not significantly high enough to beat the cost differential of the two types; this leads to higher surplus and creates competitive advantage for the dirty type of the non-investing firm (and the investing firm as well). In other words, the non-investing dirty type always undercuts the investing clean type; this, in turn, attributes to the failure of the clean type of the investing firm to capture any market segment. Therefore, it is not at all profitable for the clean type to adopt an ecolabel to disclose its environmental performance. However, the clean type does adopt an ecolabel when the opposite holds true, i.e., when the green premium is high enough to exceed the cost differential which generates competitive advantage for the clean type. Following lemma illustrates the pricing equilibrium under the high-green premium.

**Lemma 4** *Suppose there exists an ecolabel for adoption. If only one firm invests and the green premium is high enough to generate relatively higher surplus for the clean type compared to the dirty type ( $m_C - m_D \leq \Delta \leq \bar{\Delta}$ ), then in the D1 perfect Bayesian separating equilibrium, the clean type adopts the ecolabel, charges  $(m_D + \Delta)$ , and caters to the entire market whereas the dirty type charges its own marginal cost ( $m_D$ ) and equally shares the market only if the investing rival remains dirty.*

**Proof.** See Appendix. ■

The equilibrium pricing behavior of the firms under the high-green premium is somewhat desirable compared to the outcomes under the low and the intermediate green premiums, in the sense that the clean type not only adopts an ecolabel but also sells to the entire market. To summarize,

$$\pi_{INI}^E = \begin{cases} \mu(1 - \mu)(m_C - \Delta - m_D) & \text{if } \Delta \leq m_C - m_D \\ \mu(m_D + \Delta - m_C) - \epsilon & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases} \quad \text{and} \quad (10)$$

$$\pi_{NII}^E = \begin{cases} \mu(m_C - \Delta - m_D) & \text{if } \Delta \leq m_C - m_D \\ 0 & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases}. \quad (11)$$

It is indeed interesting to note that when  $\Delta \leq m_C - m_D$ , the non-investing firm enjoys a kind of positive externality (i.e.,  $\pi_{NII}^E > \pi_{INI}^E$ ) due to its rival's decision to invest in the cleaner technology.

Finally, I briefly discuss the case when only one firm invests and there is no ecolabel (i.e., pure signaling through prices) in the light of Sengupta (2015). Observe that when only one firm invests,

there is no ecolabel, and  $\Delta \leq m_C - m_D$ , the pricing equilibria are qualitatively similar to the pricing equilibria with ecolabel (discussed in Lemma 3). The one-sided incomplete information pricing equilibrium under the high-green premium without any ecolabel is described below.

**Lemma 5** *Suppose there is no ecolabel available. When only one firm invests and the green premium is high ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ), the dirty type charges its marginal cost ( $m_D$ ) and all consumers buy from the dirty type with probability one whereas the clean type charges a higher price ( $m_D + \Delta$ ) and sells zero in the unique D1 perfect Bayesian separating equilibrium.*

Interestingly, even though the clean type yields higher surplus than the dirty type (as  $\Delta \geq m_C - m_D$ ), the clean type can never sell in the equilibrium. In the separating equilibrium the non-investing dirty type sells with probability one in the state where the rival investing firm is of clean type otherwise it equally shares the market with the rival. Note that if the clean type happens to sell with a strictly positive probability then the dirty type of the investing firm will always have an incentive to imitate the clean type. Thus, in this pure strategy unique separating equilibrium both types earn zero profit.

The above unique separating equilibrium can be supported by the following out-of-equilibrium beliefs of consumers: if a firm charges any off equilibrium price  $p < m_D + \Delta$  or  $p > m_D + \Delta$  then consumers believe that the firm is of dirty or clean type respectively with probability one. Note that for any level of quantity if it is profitable for a clean type to deviate to any price  $p < m_D + \Delta$  then the dirty type also finds it profitable to deviate, whereas for any level of quantity if it is profitable for the dirty type to deviate to a price  $p > m_D + \Delta$  then the clean type finds it strictly profitable to deviate as well; thus, the out-of-equilibrium beliefs satisfy the D1 Criterion.

Observe that if *one firm invests* the clean type earns zero expected profit ( $\pi_C = 0$ ) for any level of green-premium whereas the dirty type earns strictly positive expected profit ( $\pi_D = \mu(m_C - \Delta - m_D)$ ) for  $\Delta \leq m_C - m_D$ . Thus, I find

$$\pi_{INI}^S = \begin{cases} \mu(1 - \mu)(m_C - \Delta - m_D) & \text{if } \Delta \leq m_C - m_D \\ 0 & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases}, \text{ and} \quad (12)$$

$$\pi_{NII}^S = \begin{cases} \mu(m_C - \Delta - m_D) & \text{if } \Delta \leq m_C - m_D \\ 0 & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases} \quad (13)$$

respectively.

## 5.2 Investment outcomes with ecolabel

In the previous sections, the analysis of market outcomes with availability of ecolabel is carried out under the assumption that  $\epsilon$ , the cost of adopting an ecolabel, is strictly positive but sufficiently small. For ease of exposition and in order to abstract from the effect of cost of adopting ecolabel on the incentive to invest, in this section the expected profit of a firm (when ecolabel is adopted) is taken to be the limiting value as  $\epsilon$  is reduced to zero. I calculate the unilateral incentive to invest (when the ecolabel is available)

$$UI^E = \pi_{INI}^E - \pi_{NINI}^E = \begin{cases} \mu(1 - \mu)(m_C - \Delta - m_D) & \text{if } 0 \leq \Delta \leq m_C - m_D \\ \mu(m_D + \Delta - m_C) & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases}; \quad (14)$$

it is strictly positive for any level of green premium. Observe that, surprisingly, even when the green premium is as low as zero a firm has a strictly positive incentive to invest given that the rival does not invest. Recall from Lemma 3 that since the dirty type has competitive advantage over the clean type ( $\Delta \leq m_C - m_D$ ), clean type of the investing firm charges its own marginal cost, does not adopt any ecolabel, sells zero, and earns zero profit ( $\pi_C = 0$ ), whereas the dirty type of the investing (as well as the non-investing) firm randomizes over a price interval which is strictly above its own marginal cost, always sells strictly positive quantity, and earns strictly positive profit ( $\pi_D = \mu(m_C - \Delta - m_D)$  from (42)). The uncertainty about the actual environmental performance of the investing firm (when only one firm invests) reduces competition and generates market power for the dirty type. The *ex ante* expected profit of an investing firm is essentially equal to profit of the investing firm if it remains dirty with probability  $(1 - \mu)$  i.e.,  $\pi_{INI}^E = (1 - \mu)\pi_D = (1 - \mu)\mu(m_C - \Delta - m_D)$ . The reciprocal incentive to invest is given by

$$RI^E = \pi_{II}^E - \pi_{NII}^E = \begin{cases} -\mu^2(m_C - \Delta - m_D) & \text{if } 0 \leq \Delta \leq m_C - m_D \\ \mu(1 - \mu)(m_D + \Delta - m_C) & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases}. \quad (15)$$

The values of  $\pi_{II}^E$  are given by (9), and (11) depicts  $\pi_{NII}^E$ . Note that for any  $\Delta \leq m_C - m_D$ , a firm does not have any reciprocal incentive to invest, whereas it is strictly positive when the green premium is high. The intuition is as follows. Recall that when the green premium is not high, the clean type has zero market power, whereas the non-investing dirty type enjoys market power with probability one. Thus, the non-investing dirty type earns higher strictly positive expected profit compared to the investing firm. This implies that a firm does not have any incentive to invest if the rival invests; in other words, the firm has negative reciprocal incentive to invest when  $\Delta \leq m_C - m_D$ . However, the positive externality (negative of reciprocal incentive) enjoyed

by the non-investing dirty type decreases as the green premium increases. Finally, it disappears when the green premium is high enough to generate competitive advantage for the clean type ( $m_C - m_D < \Delta$ ). In this case, the clean type has market power whereas the dirty type does not. This, in turn, generates higher expected profit for the investing firm compared to the non-investing one and creates strictly positive reciprocal incentive to invest.

Figure 1 depicts how a firm's strategic, unilateral ( $UI^E$ ) as well as reciprocal ( $RI^E$ ), incentives to invest change with increase in the green premium ( $\Delta$ ). The dashed line and solid line represent unilateral and reciprocal incentives to invest respectively. Note that  $UI^E > 0$  even at  $\Delta = 0$ ;  $UI^E$  decreases as  $\Delta$  increases in the range  $0 \leq \Delta \leq (m_C - m_D)$ , becomes equal to zero at  $\Delta = (m_C - m_D)$ , and starts increasing again as the green premium rises over the range  $m_C - m_D < \Delta \leq \bar{\Delta}$ . For any strictly positive fixed cost of investment ( $F$ ) depicted by the horizontal line),  $\Delta_1$  and  $\Delta_2$  are critical levels of green premium at which  $UI^E = F$ . It can be shown that

$$\Delta_1 = (m_C - m_D) - \frac{F}{\mu(1 - \mu)} \text{ and} \quad (16)$$

$$\Delta_2 = (m_C - m_D) + \frac{F}{\mu}. \quad (17)$$

Unlike unilateral incentive,  $RI^E \leq 0$  for any green premium  $0 \leq \Delta \leq (m_C - m_D)$ ; though it increases with increase in  $\Delta$  and becomes strictly positive over the range  $m_C - m_D < \Delta \leq \bar{\Delta}$ . For any strictly positive fixed cost,  $\Delta_3$  is the critical level of green premium at which  $RI^E = F$ ; this implies that

$$\Delta_3 = (m_C - m_D) + \frac{F}{\mu(1 - \mu)}. \quad (18)$$

Finally, I investigate the equilibrium investment behavior of firms. In particular, I consider the first-stage investment game where firms simultaneously decide whether to invest in cleaner technology where the outcome is uncertain and find the equilibrium by comparing the ex ante expected profit of the investing as well as the non-investing firms. I summarize the investment equilibria in the following proposition and in Figure 2(a).

**Proposition 4 (With ecolabel)** *Consider the case where the ecolabel is available. Let  $\Delta_1$ ,  $\Delta_2$ , and  $\Delta_3$  be defined by (16), (17), and (18). In the investment equilibrium, as the green premium ( $\Delta$ ) increases*

(i) *when  $0 \leq \Delta \leq \Delta_1$  only one firm invests,*

(ii) *when  $\Delta_1 \leq \Delta \leq \Delta_2$  no firm invests,*

- (iii) when  $\Delta_2 \leq \Delta \leq \Delta_3$  only one firm invests, and  
(iv) when  $\Delta_3 \leq \Delta \leq \bar{\Delta}$  both firms invest in cleaner technology.

**Proof.** See Appendix. ■

It might be useful to explicitly summarize the sub game perfect Nash equilibrium in the presence of an ecolabel. For  $0 \leq \Delta \leq \Delta_1 < (m_C - m_D)$  only one firm invests and  $\Delta_1 \leq \Delta \leq \Delta_2$  (where  $(m_C - m_D) < \Delta_2$ ) no firm invests, no firm adopts the ecolabel, and the dirty type is the only one that caters to the consumers (even though they are willing to pay a higher price for the product of a firm with cleaner technology). When  $(m_C - m_D) < \Delta_2 \leq \Delta \leq \Delta_3$ , only one firm invests and if it becomes clean then it adopts an ecolabel and sells to the entire market. Finally for  $\Delta_3 \leq \Delta \leq \bar{\Delta}$ , both firms invest, the clean type adopts the ecolabel, charges a price higher than its own marginal cost, and always sells in the market. In other words, higher green premium creates competitive advantage for the clean type. This, in turn, generates incentive for the firms to invest in the cleaner technology. Incomplete information about the final investment outcome softens price competition and enables clean type to earn strictly positive expected profit. Therefore, the clean type decides to disclose its environmental performance to consumers by adopting the ecolabel.

### 5.3 Investment outcomes without any ecolabel

Next, I consider an augmented version of the three-stage signaling game (without any ecolabel) as described in Section 3. In the first stage, the firms simultaneously decide whether to invest in a cleaner technology where the final outcome is uncertain. The final outcome of the investment remains private knowledge. In the next stage, the firms choose prices simultaneously to signal the environmental performance to the consumers. Finally, the consumers observe the prices charged by the firms, update their beliefs, decide whether to buy, and from which firm to buy. Similar to the case with an ecolabel, we have three different information structures. The case where only one firm invests is discussed in Subsection 5.1, and Section 3 depicts the pricing equilibria when both firms invest.

The *ex ante* expected profit of an investing as well as a non-investing firm given the rival invests and does not invest are given by

$$\pi_{II}^S = \begin{cases} (1 - \mu) \mu (m_C - \Delta - m_D) & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ \frac{\mu^2}{2} (m_D + 2\Delta - m_C) + (1 - \mu) \mu \Delta & \text{if } \frac{m_C - m_D}{2} \leq \Delta \leq \bar{\Delta} \end{cases}, \quad (19)$$



(12), (13), and  $\pi_{NINI}^S = 0$ . I calculate the unilateral and the reciprocal incentive to invest as follows

$$UI^S = \pi_{INI}^S - \pi_{NINI}^S = \begin{cases} \mu(1-\mu)(m_C - \Delta - m_D) & \text{if } \Delta \leq m_C - m_D \\ 0 & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases} \quad \text{and} \quad (20)$$

$$RI^S = \pi_{II}^S - \pi_{NII}^S = \begin{cases} -\mu^2(m_C - \Delta - m_D) & \text{if } \Delta \leq \frac{m_C - m_D}{2} \\ \mu[(m_D + 2\Delta - m_C) - \frac{\mu}{2}(m_C - m_D)] & \text{if } \frac{m_C - m_D}{2} < \Delta \leq m_C - m_D \\ \frac{\mu^2}{2}(m_D + 2\Delta - m_C) + (1-\mu)\mu\Delta & \text{if } m_C - m_D < \Delta \leq \bar{\Delta} \end{cases} \quad (21)$$

respectively. In Figure 3, I plot  $UI^S$  and  $RI^S$  on the vertical axis and  $\Delta$  on the horizontal axis. Note that a firm has strictly positive unilateral incentive only when  $\Delta < (m_C - m_D)$ . The intuition is not too different from the case where the firms can voluntarily choose an ecolabel to disclose their environmental performances (discussed in the previous subsection). In particular, when the clean type does not have any competitive advantage over the dirty type ( $\Delta \leq m_C - m_D$ ), if only one firm invests in cleaner technology, the uncertainty about the type of the investing rival creates market power and positive externality for the non-investing (as well as the investing) dirty type; thus,  $UI^S \geq 0$  when  $\Delta \leq m_C - m_D$ . Recall that the positive externality enjoyed by the non-investing dirty type vanishes once the green premium is high enough to generate higher surplus for the clean type (i.e.,  $m_C - m_D < \Delta \leq \bar{\Delta}$ ); the clean type fails to sell any positive quantity in the separating equilibrium. This explains why  $UI^S = 0$  when  $m_C - m_D < \Delta \leq \bar{\Delta}$ . For any strictly positive fixed cost of investment ( $F$ ) depicted by the horizontal line),  $\Delta_1$  (defined by (16)) is the critical level of green premium where  $UI^S = F$ .

Observe that there are two opposing forces that affect the reciprocal incentive to invest when firms signal their type through prices, namely the positive externality enjoyed by the non-investing firm and the market power enjoyed by the clean type of the investing firm. Under the low-green premium ( $\Delta \leq \frac{m_C - m_D}{2}$ ), the clean type does not have any market power and thus, the positive externality dominates and  $RI^S < 0$ . For  $\frac{m_C - m_D}{2} < \Delta \leq \bar{\Delta}$ , the clean type enjoys higher market power when both firms invest ( $p_C = m_D + 2\Delta$ ) compared to the case where only one firm invests ( $p_C = m_D + \Delta$ ). Moreover, recall that the positive externality becomes zero for any  $m_C - m_D < \Delta \leq \bar{\Delta}$ . Combining these two opposing forces, I find that  $RI^S \leq 0$  if  $\Delta \leq \frac{3(m_C - m_D)}{4}$  and  $RI^S \geq 0$  if  $\frac{3(m_C - m_D)}{4} \leq \Delta \leq \bar{\Delta}$ . For any  $F > 0$ ,  $\Delta_4$  is the critical level of green premium at which  $RI^S = F$ ,

and it can be shown that

$$\Delta_4 = \begin{cases} \frac{(1-\frac{\mu}{2})}{2} (m_C - m_D) + \frac{F}{2\mu} & \text{if } F \leq \mu \left(1 - \frac{\mu}{2}\right) (m_C - m_D) \\ \frac{\mu}{2} (m_C - m_D) + \frac{F}{\mu} & \text{if } F \geq \mu \left(1 - \frac{\mu}{2}\right) (m_C - m_D) \end{cases}. \quad (22)$$

The investment equilibria without ecolabel are described in the next proposition and in Figure 2(b).

**Proposition 5 (Without any ecolabel)** *Suppose there is no ecolabel and the firms signal their respective environmental performances only via prices. Let  $\Delta_1$  and  $\Delta_4$  be defined by (16) and (22).*

*In the investment equilibria, as the green premium increases*

- (i) *when  $0 \leq \Delta \leq \Delta_1$  only one firm invests,*
- (ii) *when  $\Delta_1 \leq \Delta \leq \Delta_4$  no firm invests, and*
- (iii) *when  $\Delta_4 \leq \Delta \leq \bar{\Delta}$  both firms invest.*

**Proof.** See Appendix. ■

The equilibrium paths of the augmented investment game (in the absence of any ecolabel) discussed here look as follows. For  $0 \leq \Delta \leq \Delta_1$  only one firm invests in the first stage, the dirty type enjoys market power and sells to all consumers. When  $\Delta_1 \leq \Delta \leq \Delta_4$ , no firm invests, both firms remain dirty and share the market by charging its own marginal cost. Finally, if  $\Delta_4 \leq \Delta \leq \bar{\Delta}$ , both firms invest, the clean type sells only when both firms are of the clean type; otherwise the dirty type captures the entire market. Both the clean and the dirty type earn strictly positive rent.

#### 5.4 Investment outcomes: Ecolabel vs. Signaling

To the best of my knowledge, this is the first attempt to analyze and compare the equilibrium investment behavior of strategically competing firms and other market outcomes with and without an ecolabel. The comparisons between firms' strategic (unilateral as well as reciprocal) incentives to invest (see Figure 4 and Figure 5) and investment equilibria (see Figure 2(a) and Figure 2(b)) with and without any ecolabel are summarized in the following corollary.

**Corollary 3 (Incentive to invest)** *Consider the scenario where the green premium that buyers are willing to pay (for the cleaner product) is high; in particular ( $m_C - m_D \leq \Delta \leq \bar{\Delta}$ ). Then, (relative to the situation when no ecolabel is available for adoption) availability of an ecolabel increases a firm's (unilateral) incentive to invest in cleaner technology when its rival does not invest, but reduces*

the (reciprocal) incentive to invest when its rival firm invests.

(Equilibrium investment) Total investment by firms in cleaner technology when there is no ecolabel (available for adoption) is always at least as high as when an ecolabel is available and it is strictly higher for a particular range of green premium ( $\Delta_4 \leq \Delta \leq \Delta_3$ ).

The strictly positive unilateral incentive to invest under lower (such that  $\Delta \leq m_C - m_D$ ) green premium arises only because the dirty type can take the advantage of incomplete information about its rival's type ( $UI^E = UI^S > 0$ ). As the premium increases (such that the clean type generates higher surplus) the unilateral incentive becomes zero without any ecolabel ( $UI^S = 0$ ), since the clean type cannot sell to any consumer when the equilibrium price acts as a signal for the environmental performance of the firms. However, the competitive advantage over a dirty type drives a clean type to adopt an ecolabel (when it is available) in the equilibrium; consequently, the clean type earns strictly positive profit. Thus, in this case ( $m_C - m_D \leq \Delta \leq \bar{\Delta}$ ), a firm has a strictly positive unilateral incentive to invest with an ecolabel ( $UI^E > 0$ ).

Figure 5 shows that for  $\Delta > \frac{m_C - m_D}{2}$  a firm has a higher reciprocal incentive to invest when there is no ecolabel compared to the case where the firm can adopt an ecolabel. The intuition is as follows. If both firms invest then we have a situation of two-sided incomplete information i.e., the firms are not aware of each other's type. Note that *without any ecolabel*, a clean type has to charge a significantly higher price to credibly signal its environmental performance (so that the dirty type has no incentive to imitate the clean type's price). Thus, the incomplete information generates market power and strictly positive rent for the clean as well as the dirty type which, in turn, creates strictly positive reciprocal incentive to invest. Moreover, as the green premium increases, it fuels the dirty type's incentive to imitate clean type's price and consequently, increases the *ex ante* expected profit of an investing firm and eventually the reciprocal incentive to invest (evident from Figure 5). However, under the high-green premium ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ) when there is an ecolabel available, the clean type finds it profitable to adopt it and does not need to charge a very high price to prevent imitation by the dirty type. Therefore, presence of an ecolabel reduces a firm's market power, profitability, and reciprocal incentive to invest (compared to the pure signaling case).

Observe that (see Figure 2(a) and Figure 2(b)),  $\Delta_1 < \Delta_4 < \Delta_2 < \Delta_3$ . When  $0 \leq \Delta \leq \Delta_1$  one firm invests in the equilibrium with as well as without an ecolabel. For any green premium  $\Delta_4 \leq \Delta \leq \Delta_2$ , no firm invests in the presence of an ecolabel whereas both firms invest without any ecolabel. Similarly, for  $\Delta_4 < \Delta_2 \leq \Delta \leq \Delta_3$  only one firm invests with an ecolabel but both firms invest in case of pure signaling. This implies that for any level of green premium, the total

investment (number of investing firms in the equilibrium) without any ecolabel is never lower than that of with an ecolabel. More specifically, one can conclude that the presence of an ecolabel does not necessarily augments the aggregate investment in cleaner technology relative to the case where firms signal their environmental performances through prices.

## 6 Discussion: Mandatory Disclosure

In this section, I assume that regulatory authorities mandate firms to publicly disclose the final outcome of their investments in cleaner technology. Alternatively, the regulatory authorities can also on their own acquire information about actual environmental performance of firms and disseminate the information among public. As a result, the consumers as well as the rival firms become completely aware of the actual environmental performance of a firm. Thus, under mandatory disclosure laws, there is no need for adoption of ecolabels (or signaling via prices) by the firms. Toxic Release Inventory (USA), Environmental Reporting Decree (the Netherlands), Green Accounts (Denmark), and Pollutant Release and Transfer Register (UK) are few examples of mandatory policies that make firms disclose environmental performances of their production processes. The standard belief suggests that since the consumers are willing to pay a green premium for the product produced by relatively cleaner technology, the firms should always have significant incentive to become cleaner when the consumers are indeed aware of the actual environmental performance of the firms. I critically examine the need for mandatory disclosure laws to motivate firms to invest in clean technology.

Under mandatory disclosure laws, the multi-stage game described in the previous sections remains almost the same except the firms are forced to publicly disclose their realized production technologies before engaging in strategic price competition. First, I describe the full information equilibrium of the pricing game. If no firm invests then both firms remain dirty for sure; as a result they involve in aggressive price competition and charge a price equal to the dirty type's marginal cost earning zero profit. The pricing equilibria when at least one firm invests (either only one firm or both firms) are summarized below (adopted from Sengupta (2015)).

**Lemma 6** *Under mandatory disclosure, when at least one firm invests,*

*(i) if the firms are of the same type then they charge their respective marginal costs and equally share the market.*

*(ii) if the firms are of different types (one clean and the other dirty)*

(a) when  $0 \leq \Delta \leq m_C - m_D$ , the clean type charges its marginal cost  $m_C$  and the dirty type charges  $m_C - \Delta$ .

(b) when  $m_C - m_D \leq \Delta \leq \bar{\Delta}$ , the clean type caters to the entire market at  $m_D + \Delta$  whereas the rival dirty type charges  $m_D$ .

If both firms are of the same type, they aggressively compete and bring down the price to their respective marginal costs sharing the market equally. Recall that  $\Delta \leq m_C - m_D$  implies that the dirty type generates higher surplus than the clean type; thus, the dirty type has competitive advantage over the clean type. The lowest price that the clean type can charge to compete is its own marginal cost ( $m_C$ ). In the pricing equilibrium, the dirty type caters to the entire market at a price ( $m_C - \Delta$ ) where a consumer is indifferent between buying from the dirty type and the rival clean type at  $m_C$ . In this case,  $\pi_{II}^M = (1 - \mu)\mu(m_C - \Delta - m_D)$ ,  $\pi_{INI}^M = 0$ ,  $\pi_{NII}^M = \mu(m_C - \Delta - m_D)$ , and  $\pi_{NINI}^M = 0$ . The unilateral and the reciprocal incentive of a firm to invest in the cleaner technology are  $UI^M = \pi_{INI}^M - \pi_{NINI}^M = 0$  and  $RI^M = \pi_{II}^M - \pi_{NII}^M = -\mu^2(m_C - \Delta - m_D) < 0$  respectively. Thus, a firm does not have any incentive to invest in the cleaner technology when the green premium is lower than the cost differences (see Figure 6).

However, at a reasonably higher green premium ( $m_C - m_D \leq \Delta \leq \bar{\Delta}$ ), the clean type has competitive advantage over the dirty rival and sells to all consumers at a price ( $m_D + \Delta \geq m_C$ ); this is the price at which a consumer is indifferent between buying from the clean type and the dirty type at its own marginal cost ( $m_D$ ). This implies that a clean type rules the market in the presence of a dirty low cost rival when the consumers are willing to pay reasonably higher green premium. Note that in this case,  $\pi_{II}^M = (1 - \mu)\mu(m_D + \Delta - m_C)$ ,  $\pi_{INI}^M = \mu(m_D + \Delta - m_C)$ ,  $\pi_{NII}^M = 0$ , and  $\pi_{NINI}^M = 0$ . The unilateral and the reciprocal incentives of a firm are  $UI^M = \pi_{INI}^M - \pi_{NINI}^M = \mu(m_D + \Delta - m_C) > 0$  and  $RI^M = \pi_{II}^M - \pi_{NII}^M = \mu(1 - \mu)(m_D + \Delta - m_C) > 0$  respectively; in other words, a firm has a strictly positive incentive to become cleaner i.e., invest in the cleaner technology. For a strictly positive fixed cost of investment ( $F > 0$ ),  $UI^M \geq F$  when  $\Delta \geq \Delta_2$  (defined by (17)) and  $RI^M \geq F$  for  $\Delta \geq \Delta_3$  (defined by (18)) (see Figure 6).

The following proposition and Figure 2(c) illustrate the investment equilibria under the mandatory disclosure.

**Proposition 6 (Mandatory disclosure)** *Let  $\Delta_2$  and  $\Delta_3$  be defined by (17) and (18). Under the mandatory disclosure, as the green premium increases the investment equilibria are as follows.*

(i) when  $0 \leq \Delta \leq \Delta_2$  no firm invests ,

- (iii) when  $\Delta_2 \leq \Delta \leq \Delta_3$  only one firm invests, and  
 (iv) when  $\Delta_3 \leq \Delta \leq \bar{\Delta}$  both firms invest in cleaner technology.

**Proof.** See Appendix. ■

A firm does not have any strategic incentive to invest when  $\Delta \leq m_C - m_D$ , but both the unilateral as well as the reciprocal incentives are strictly positive when the green premium is high enough to generate higher surplus for the clean type ( $m_C - m_D \leq \Delta \leq \bar{\Delta}$ ). The investment equilibrium under a lower premium contradicts the standard belief, as no firm invests to become cleaner even though the consumers are aware of the actual environmental performance of the firms. The consumers need to pay a significantly higher green premium to encourage firms to invest in the cleaner technology under the mandatory disclosure laws. In other words, contrary to the standard beliefs, the mandatory disclosure laws are not enough to generate incentive for the firms to invest. Comparison of the investment equilibria under different disclosure regimes (see Figure 2) yields the following important observation.

**Corollary 4** *For a lower range of green premium viz.,  $0 \leq \Delta \leq \Delta_1$ , no firm invests in the equilibrium under mandatory disclosure whereas at least one firm invests in the presence of an ecolabel or price as a disclosure device.*

## 7 Conclusion

Market incentives created by consumer's willingness to pay a premium for environmentally cleaner product can, in principle, increase the market share of firms with environmentally cleaner production processes and motivate firms to voluntarily invest in clean technology even in the absence of any form of emission taxes, pollution permits, liability laws, quality standards, etc. Recent political changes that favor deregulation in many countries have increased the potential role of such market incentive-based voluntary compliance mechanisms. The lack of information among consumers about firms' actual production technology is, however, a crucial stumbling block that can distort and even prevent the transmission of buyers' preferences through market incentives. Third party certification mechanisms such as ecolabels that allow firms to directly and credibly disclose information can play an important role in overcoming this informational asymmetry between buyers and firms. However, for this to work firms must actually adopt ecolabels.

This paper shows that strategic competition can deter firms from adopting ecolabels to directly disclose their private information about their own production unless the green premium that buyers

are willing to pay is significantly high relative to the production cost disadvantage of the clean firms (vis-à-vis the dirty firms). Ecolabels are not likely to be observed in markets where environmental consciousness is only moderate and/or dirty firms have a sizeable cost advantage even when there is no imperfection or friction in the certification process. These conclusions appear to match casual observations about the present state of the world but clearly require rigorous empirical investigation.

My analysis shows that markets, where no ecolabels are adopted, may still have perfect transmission of private information through prices though the signaling distortion leads to higher market power and consumption distortions with a loss of social surplus. A significant increase in buyers' consciousness (i.e., the premium they are willing to pay for the clean product) not only leads to the adoption of ecolabels but by removing the need for price signaling, it reduces prices and increases social surplus (even though the actual information transmitted may remain the same). However, from a long-run perspective, when the market fundamentals support adoption of ecolabels, the *ex ante* individual firm's incentive to invest in cleaner technology and the aggregate investment level of the industry may be lower than the case where no ecolabels were available for the firms to adopt. Therefore, the eventual effects of ecolabeling or certification mechanism on environmental outcomes and welfare are likely to be mixed even if the information transmission mechanism itself is perfect.

## Appendix

**Proof of Lemma 1.** In the symmetric separating perfect Bayesian equilibrium the dirty type follows a common probability distribution  $F_D(p)$  whose support is an interval  $[\underline{P}_D, \bar{P}_D]$ , and the clean type charges a common deterministic price  $p_C \in [m_C, V + \Delta]$  which is always higher than the price charged by the dirty type. At the upper bound of the support ( $\bar{P}_D$ ), a consumer is indifferent between buying from a clean type at  $p_C$  and from a dirty type at  $\bar{P}_D$  i.e.,  $\bar{P}_D = p_C - \Delta$ . The dirty type charges a price less than  $\bar{P}_D$  almost surely since otherwise the rival dirty type can undercut to earn higher rent. This, in turn, implies that a clean type can only sell in the state when the rival is of clean type. The equilibrium expected profit of the dirty type for charging any price  $p \in [\underline{P}_D, \bar{P}_D]$  is given by

$$\pi_D = [\mu + (1 - \mu)(1 - F_D(p))] (p - m_D). \quad (23)$$

In a state where its rival is a clean type, a dirty type can charge  $\bar{P}_D$ , sell to all consumers, and earns a strictly positive profit equal to

$$(\bar{P}_D - m_D) \mu = (p_C - \Delta - m_D) \mu \quad (24)$$

which is identical to the equilibrium expected profit of the dirty type  $\pi_D$ . The lower bound of the support ( $\underline{P}_D$ ) is the lowest price that the dirty type wants to undercut, given that it is going to capture entire market irrespective of the type of its rival; it earns strictly positive expected profit which is equal to  $\pi_D$ . This implies

$$\underline{P}_D - m_D = \pi_D = (p_C - \Delta - m_D) \mu.$$

Therefore, the lower bound of the support is

$$\underline{P}_D = \mu [p_C - \Delta] + (1 - \mu) m_D. \quad (25)$$

Note that the equilibrium price distribution and the expected profit  $\pi_D$  of the dirty type depend on the deterministic price charged by the clean type. At every price  $p \in [\underline{P}_D, \bar{P}_D]$ , the dirty type can sell to all consumers as long as the rival of dirty type does not undercut, and its expected profit at  $p$  is equal to  $[\mu + (1 - \mu)(1 - F_D(p))](p - m_D)$ ; this is equal to  $\pi_D$  for every price  $p \in [\underline{P}_D, \bar{P}_D]$  i.e.,

$$[\mu + (1 - \mu)(1 - F_D(p))](p - m_D) = (p_C - \Delta - m_D) \mu$$

(from (23) and (24)) which implies that

$$F_D(p) = 1 - \frac{\mu}{(1 - \mu)} \left( \frac{p_C - \Delta - m_D}{p - m_D} - 1 \right)$$

where  $F_D(p)$  is continuous on  $[\underline{P}_D, \bar{P}_D]$ ,  $F_D(\underline{P}_D) = 0$ , and  $F_D(\bar{P}_D) = 1$ . In the perfect Bayesian separating equilibrium, a clean type can sell only in the state where its rival is clean too, and they equally divide the market among themselves as consumers are indifferent between firms; in this case, all consumers buy from the clean type with probability one as long as  $p_C \leq V + \Delta$ .

The Bayesian equilibrium can be supported by the following out-of-equilibrium beliefs of consumers: if the price  $p$  charged by a firm is such that  $p \neq p_C$  and  $p \notin [\underline{P}_D, \bar{P}_D]$ , then consumers believe that the firm is of dirty type with probability one. Given these out-of-equilibrium beliefs, no firm has an incentive to unilaterally deviate to any out-of-equilibrium price. It can be argued that these out-of-equilibrium beliefs satisfy the D1 refinement.<sup>19</sup> Consider any out-of-equilibrium

<sup>19</sup>For a formal proof see Janssen and Roy (2010).



price; observe that for any level of quantity, if it is profitable for a clean type to deviate to the out-of-equilibrium price then the dirty type also finds it strictly profitable to deviate to such a price.

The incentive compatibility constraint of the dirty type and the clean type are

$$\frac{\mu}{2}(p_C - m_D) \leq (p_C - \Delta - m_D)\mu \text{ and } \frac{\mu}{2}(p_C - m_C) \geq (p_C - \Delta - m_C)\mu$$

respectively which imply  $2\Delta + m_D \leq p_C \leq 2\Delta + m_C$ . To be more precise,

$$\max\{2\Delta + m_D, m_C\} \leq p_C \leq \min\{2\Delta + m_C, V + \Delta\}.$$

The strategies and the out-of-equilibrium beliefs described above constitute a perfect Bayesian equilibrium which satisfies the incentive compatibility constraints of the clean and the dirty type iff  $2\Delta + m_D \leq V + \Delta$  i.e.,  $\Delta \leq \bar{\Delta} = V - m_D$ . If  $2\Delta + m_D \leq m_C \implies \Delta \leq \frac{m_C - m_D}{2}$  then the clean type charges its effective marginal cost  $m_C$  such that the firm loses its market power whereas if  $\frac{m_C - m_D}{2} \leq \Delta \leq \bar{\Delta}$  then clean type charges  $m_D + 2\Delta$ . ■

**Proof of Proposition 1.** Under the low-green premium, the dirty type generates higher surplus and thus, has competitive advantage over the clean type. The dirty type can undercut and capture the entire market in the state when the rival firm is of clean type but faces a fear of being undercut when the rival is of dirty type as well. To strike a balance between these two opposing forces, the dirty type randomizes over a price interval  $[\underline{p}_D, \bar{p}_D]$ . The upper bound of the interval ( $\bar{p}_D$ ) is a price at which a consumer is indifferent between buying from the dirty type and from the clean type i.e.,  $p_C = \bar{p}_D + \Delta$ . The dirty type charges a price less than  $\bar{p}_D$  almost surely, since otherwise the rival dirty type can undercut to earn higher profit. This, in turn, implies that a clean type can only sell in the state when the rival is of clean type. Therefore, in the equilibrium, the clean type ends up charging a price as low as its marginal cost ( $m_C$ ). The existence of this equilibrium is guaranteed, since the upper bound of the price support of the dirty type ( $\bar{p}_D = m_C - \Delta$ ) is greater than its marginal cost ( $m_D \leq m_C - \Delta$ ). The equilibrium expected profit of the dirty type for charging any price  $p \in [\underline{p}_D, \bar{p}_D]$  is given by

$$\pi_D = [\mu + (1 - \mu)(1 - F_D(p))](p - m_D). \quad (26)$$

In a state where its rival is a clean type, a dirty type can charge  $\bar{p}_D$ , sells to all consumers, and earns a strictly positive profit equal to

$$(\bar{p}_D - m_D)\mu = (m_C - \Delta - m_D)\mu \quad (27)$$

which is identical to the equilibrium expected profit of the dirty type  $\pi_D$  given by (26). The lower bound of the support  $(\underline{p}_D)$  is the lowest price that the dirty type wants to undercut, given that it is going to capture entire market irrespective of the type of its rival; it earns strictly positive expected profit which is equal to  $\pi_D$ . This implies  $\underline{p}_D - m_D = \pi_D = (m_C - \Delta - m_D)\mu$ . Therefore, the lower bound of the support is

$$\underline{p}_D = \mu [m_C - \Delta] + (1 - \mu) m_D. \quad (28)$$

At every price  $p \in [\underline{p}_D, \bar{p}_D]$ , the dirty type can sell to all consumers as long as the rival of dirty type does not undercut, and its expected profit at  $p$  is equal to  $[\mu + (1 - \mu)(1 - F_D(p))](p - m_D)$ ; this is equal to  $\pi_D$  for every price  $p \in [\underline{p}_D, \bar{p}_D]$  i.e.,

$$\begin{aligned} \pi_D &= [\mu + (1 - \mu)(1 - F_D(p))](p - m_D) = (m_C - \Delta - m_D)\mu \\ \implies F_D(p) &= 1 - \frac{\mu}{(1 - \mu)} \left( \frac{m_C - \Delta - m_D}{p - m_D} - 1 \right) \end{aligned} \quad (29)$$

where  $F_D(p)$  is continuous on  $[\underline{p}_D, \bar{p}_D]$ ,  $F_D(\underline{p}_D) = 0$ , and  $F_D(\bar{p}_D) = 1$ . In the Bayesian equilibrium, a clean type can sell only in the state where its rival is clean too, and they equally divide the market among themselves as consumers are indifferent between firms. The incentive compatibility constraint of the dirty type is

$$\frac{\mu}{2}(m_C - m_D) \leq (m_C - \Delta - m_D)\mu \implies \Delta \leq \frac{(m_C - m_D)}{2}. \quad (30)$$

The clean type does not have any incentive to imitate the price charged by the dirty type (which is below  $m_C$ ). Therefore, separating pricing equilibrium exists only when (30) holds true. The expected equilibrium profit of the clean type is zero irrespective of its rival's type. This implies that no firm has an incentive to adopt an ecolabel in the equilibrium.

The incomplete information Bayesian equilibrium described above can be supported by the following out-of-equilibrium beliefs of consumers: if a firm charges any (off equilibrium) price other than the effective marginal cost of the clean type (viz.,  $p > m_C$  or  $p < m_C$ ) then consumers believe that the firm is of clean or dirty type respectively with probability one. These out-of-equilibrium beliefs satisfy D1 refinement. The argument is as follows<sup>20</sup>; the set of quantities for which it is profitable for a clean type to deviate to any price  $p > m_C$  is larger than that of the dirty type, and since a clean type will never deviate to any price below its own effective marginal cost D1 refinement is trivially satisfied in this case. Given these out-of-equilibrium beliefs, no firm has an incentive to

---

<sup>20</sup>For a more formal proof see Lemma 3 Janssen and Roy (2010).

unilaterally deviate to any off equilibrium price.

Next, I check whether any firm has any incentive to deviate from the above mentioned equilibrium strategies. There is no reason for the clean type to adopt the ecolabel at marginal cost pricing as the price does signal the environmental performance of the firm. A more pertinent question is whether the clean type wants to deviate to a higher price with the help of the ecolabel, and the answer is no; because in this case, all consumers attain higher surplus if they buy from the dirty type that can easily undercut the clean type. Note that the dirty type does not want to charge a price above its upper bound either, as it will lose all its market to the clean type. ■

**Proof of Lemma 2.** The lowest possible price that a clean type charges without any ecolabel is the D1 equilibrium price  $p_C = 2\Delta + m_D$  and the corresponding expected profit of the clean type is  $\frac{\mu}{2}(m_D + 2\Delta - m_C)$  when  $\Delta > \frac{m_C - m_D}{2}$ . Since

$$\epsilon < \frac{\mu}{2}(m_D + 2\Delta - m_C) \text{ when } \Delta > \frac{m_C - m_D}{2}$$

(from Assumption 3 and Assumption 4) a clean type that adopts an ecolabel at  $\epsilon$  has a strictly positive incentive to deviate from its signaling PBE price and capture the entire market from its rival of clean type. ■

**Proof of Proposition 2.** (i) Consider the intermediate-green premium case where  $\frac{m_C - m_D}{2} < \Delta \leq m_C - m_D$ . As the dirty type generates higher surplus and thus has competitive advantage over the clean type, it can undercut and capture the entire market in the state when the rival firm is of clean type. Therefore, there is no equilibrium where a clean firm gets strictly positive market share when it faces a dirty rival. As the clean types sell only in the state where both firms are clean, there is no equilibrium where the clean types randomize over prices.

Observe that there cannot be a symmetric separating PBE where neither firm adopts ecolabels. This is because the strategy profile in such an equilibrium would also have to be a symmetric separating PBE in the version of the model where there is no possibility of adopting an ecolabel which implies (see Lemma 1), that the clean types must charge a common price that is at least as large as  $(2\Delta + m_D)$  and sell only in the state where both firms are of clean type (sharing the market equally); but in this model where a firm may acquire an ecolabel (simultaneously with pricing), as

$$\epsilon < \frac{\mu}{2}(m_D + 2\Delta - m_C) \tag{31}$$

it is easy to check that a clean type can strictly gain by acquiring an ecolabel and undercutting the rival so as to capture the *entire* market in the state where the rival is of clean type (acquiring

the ecolabel "liberates" the clean firm from any unfavorable out-of-equilibrium belief of buyers associated with the lower price).

Finally, observe that there cannot be a symmetric separating PBE where the clean types of both firms adopt ecolabels. In any such equilibrium, in order to break even, clean type of each firm must charge a deterministic price  $p_C \geq m_C + \frac{\epsilon}{\mu}$  and using (31) a clean type firm can strictly gain by undercutting this price and increasing its market share from  $\frac{1}{2}$  to 1. Thus, a symmetric separating PBE does not exist.

(ii) Consider the asymmetric equilibrium where the clean type of firm 1 adopts the ecolabel and that of firm 2 does not (there is another equilibrium where the roles are interchanged between the firms).

Let the pricing strategies be as specified in the proposition (given by (6) and (7)). Impose the following restriction on out-of-equilibrium beliefs: any firm that does not adopt an ecolabel and charges a price  $p \in (\bar{p}_D, p_C)$  is of dirty type with probability one. At price  $\bar{p}_D$ , a dirty type sells only in the state where the rival is of clean type (it is undercut with probability one when rival is dirty type) and so the equilibrium expected profit of the dirty type is given by

$$\pi_D = (\bar{p}_D - m_D)\mu = \left(m_C + \frac{\epsilon}{\mu} - \Delta - m_D\right)\mu > 0. \quad (32)$$

The last inequality follows from the fact that  $\Delta < m_C - m_D$ . Clean types of both firms earn zero net profit.  $\underline{p}_D$ , the lower bound of the support is the lowest price that the dirty type wants to undercut, given that it is going to capture entire market irrespective of the type of its rival; it earns strictly positive expected profit which is equal to  $\pi_D$ . This implies that

$$\underline{p}_D = \mu \left[ m_C + \frac{\epsilon}{\mu} - \Delta \right] + (1 - \mu) m_D.$$

At every price  $p \in [\underline{p}_D, \bar{p}_D]$ , the dirty type can sell to all consumers as long as the rival of dirty type does not undercut, and its expected profit at  $p$  is equal to  $[\mu + (1 - \mu)(1 - F_D(p))](p - m_D)$  i.e.,

$$\pi_D = [\mu + (1 - \mu)(1 - F_D(p))](p - m_D) = \left(m_C + \frac{\epsilon}{\mu} - \Delta - m_D\right)\mu \quad (33)$$

which implies that

$$F_D(p) = 1 - \frac{\mu}{(1 - \mu)} \left( \frac{m_C + \frac{\epsilon}{\mu} - \Delta - m_D}{p - m_D} - 1 \right) \quad (34)$$

Note that  $F_D(\underline{p}_D) = 0$ , and  $F_D(\bar{p}_D) = 1$ .

Given the out-of-equilibrium beliefs of buyers, a firm not adopting an ecolabel and charging price in  $(m_C, p_C)$  cannot sell. Thus, there is no incentive for the clean type of either firm to deviate to a lower price without adopting an ecolabel. I, now, show that given the rival's strategy, the clean type of neither firm has any incentive to adopt an ecolabel and lower its price to grab some market when the rival is of the dirty type. If a clean type deviates to a price  $p < p_C = \left(m_C + \frac{\epsilon}{\mu}\right)$  while adopting an ecolabel, then it earns net expected profit

$$\begin{aligned}
& [\mu + (1 - \mu)(1 - F_D(p - \Delta))](p - m_C) - \epsilon \\
= & [\mu + (1 - \mu)(1 - F_D(p - \Delta))](p - \Delta - m_D) \frac{(p - m_C)}{(p - \Delta - m_D)} - \epsilon \\
= & \pi_D \frac{(p - m_C)}{(p - \Delta - m_D)} - \epsilon, \text{ from (32)}. \tag{35}
\end{aligned}$$

If  $\Delta = m_C - m_D$ , then the latter profit is exactly equal to

$$\pi_D - \epsilon = (m_C - \Delta - m_D)\mu < 0$$

so that the deviation is not strictly gainful. If  $\Delta < m_C - m_D$  then

$$\frac{(p - m_C)}{(p - \Delta - m_D)}$$

is strictly increasing in  $p$  (as  $m_C > \Delta + m_D$ ) and so the deviation profit in (35) is strictly decreasing in  $p$  at  $p_C = \left(m_C + \frac{\epsilon}{\mu}\right)$  i.e., the deviation profit is bounded above by

$$\mu(p - m_C) - \epsilon < 0$$

so that deviation is not strictly gainful. The dirty type of firm 1 cannot imitate its clean type because the latter acquires an ecolabel. The dirty type of firm 2 has no incentive to imitate its clean type as the latter does not sell at all. Given the out-of-equilibrium beliefs, a dirty type cannot gain by charging a price above  $\bar{p}_D$  (as it cannot acquire an ecolabel). A dirty type cannot imitate the action of a clean type with an ecolabel, and it does not have any incentive to imitate the price of a clean type without any ecolabel as the latter sells zero. Finally, observe that the out-of-equilibrium beliefs satisfy the intuitive criterion as no price in  $(\bar{p}_D, p_C)$  is equilibrium dominated for the dirty type<sup>21</sup>. ■

---

<sup>21</sup>Dirty type of a firm would earn strictly higher than equilibrium profit by deviating to such a price if all buyers would buy. The latter is an undominated action for buyers given the equilibrium strategy of the other firm (it is optimal for buyers to behave this way if they think the deviating firm is clean for sure).

**Proof of Proposition 3.** Under the high-green premium, the clean type generates higher surplus and thus, has competitive advantage over the dirty type. The clean type can capture the entire market in the state when the rival firm is of dirty type but faces a fear of potential aggressive price competition when the rival is of clean type as well. Consequently, the clean type randomizes over a price interval  $[\underline{p}_C, \bar{p}_C]$ . The upper bound of the interval ( $\bar{p}_C$ ) is a price at which a consumer is indifferent between buying from the clean type and from the dirty type i.e.,  $p_D = \bar{p}_C - \Delta$ . The clean type charges a price less than  $\bar{p}_C$  almost surely, since otherwise the rival clean type can undercut to earn higher profit. This, in turn, implies that a dirty type can sell only in the state when the rival is of dirty type as well. Therefore, in the equilibrium, the dirty type ends up charging a price as low as its marginal cost ( $m_D$ ). The existence of this equilibrium is guaranteed, as the upper bound of the price support of the clean type ( $\bar{p}_C = m_D + \Delta$ ) is greater than its marginal cost ( $m_C < m_D + \Delta$ ). The equilibrium expected profit of the clean type for charging any price  $p \in [\underline{p}_C, \bar{p}_C]$  is given by

$$\pi_C = [(1 - \mu) + \mu(1 - F_C(p))] (p - m_C). \quad (36)$$

In a state where its rival is a dirty type, a clean type can charge  $\bar{p}_C$ , sell to all consumers, and earns a strictly positive profit equal to

$$(\bar{p}_C - m_C) (1 - \mu) = (m_D + \Delta - m_C) (1 - \mu) \quad (37)$$

which is identical to the equilibrium expected profit of the clean type  $\pi_C$  given by (36). The lower bound of the support ( $\underline{p}_C$ ) is the lowest price that the clean type wants to undercut, given that it is going to capture the entire market irrespective of the type of its rival. At this price, it earns strictly positive expected profit which is equal to  $\pi_C$ . This implies  $\underline{p}_C - m_C = \pi_C = (m_D + \Delta - m_C) (1 - \mu)$ . Therefore, the lower bound of the support is

$$\underline{p}_C = \mu m_C + (1 - \mu) (m_D + \Delta). \quad (38)$$

At every price  $p \in [\underline{p}_C, \bar{p}_C]$ , the clean type can sell to all consumers as long as the rival of clean type does not undercut, and its expected profit at  $p$  is equal to  $[(1 - \mu) + \mu(1 - F_C(p))] (p - m_C)$ ; this is equal to  $\pi_C$  for every price  $p \in [\underline{p}_C, \bar{p}_C]$  i.e.,

$$\pi_C = [(1 - \mu) + \mu(1 - F_C(p))] (p - m_C) = (m_D + \Delta - m_C) (1 - \mu) \quad (39)$$

which implies that

$$F_C(p) = \frac{1}{\mu} \left[ 1 - (1 - \mu) \left( \frac{m_D + \Delta - m_C}{p - m_C} \right) \right] \quad (40)$$

where  $F_C(p)$  is continuous on  $[\underline{p}_C, \bar{p}_C]$ ,  $F_C(\underline{p}_C) = 0$ , and  $F_C(\bar{p}_C) = 1$ .

In the equilibrium, a dirty type can sell only in the state where its rival is dirty too, and they equally divide the market among themselves as consumers are indifferent between firms. The expected equilibrium profit of the dirty type is zero irrespective of its rival's type. Since the clean type earns strictly positive expected profit and from *Assumption 4* we know that  $\epsilon < [(1 - \mu)(m_D + \Delta - m_C)]$ , it adopts the ecolabel in the equilibrium.

The incomplete information Bayesian equilibrium described above can be supported by the following out-of-equilibrium beliefs of consumers: if a firm does not adopt an ecolabel (off equilibrium) then consumers believe that the firm is of dirty type with probability one. This out-of-equilibrium belief satisfies the D1 refinement. The argument is as follows: the set of quantities for which it is profitable for a clean type to deviate to any price  $p > m_C$  *without any ecolabel* is lower than that of the dirty type; further, since a clean type will never deviate to any price below its own effective marginal cost D1 refinement is trivially satisfied in this case. Given these out-of-equilibrium beliefs, no firm has an incentive to unilaterally deviate to any off equilibrium ecolabel adoption decision and price.

Does the clean type have any incentive to deviate from the above mentioned equilibrium strategy? Observe that it is trivial to show that the clean type does not want to deviate to any out-of-equilibrium price *with the ecolabel*; for any price above  $\bar{p}_C$  the clean type loses all consumers whereas below the  $\underline{p}_C$  it does not gain any additional market share. I consider two other possible deviations; the clean type does not adopt an ecolabel and either charges the same price or a different price. In case the clean type does not adopt the ecolabel then the price charged by the clean type should act as a signal for the environmental performance. The out-of-equilibrium belief says that if consumers observe a price  $p \in [\mu m_C + (1 - \mu)(m_D + \Delta), m_D + \Delta]$  *without any ecolabel* they believe that the firm is of dirty type. Therefore, a clean type has no incentive to not adopt the ecolabel at this range of prices. The clean type does not have any incentive to charge a price lower or higher than its equilibrium lower or upper bound respectively as it will be regarded as a dirty type. ■

**Proof of Lemma 3.** When  $\Delta \leq m_C - m_D$ , the dirty type generates higher surplus and has a competitive advantage over the clean type. The non-investing dirty type can undercut the investing rival in case it is of the clean type but has a fear of being undercut by the rival dirty type. Thus, the non-investing dirty type randomizes over a price interval  $[\underline{p}_D, \bar{p}_D]$ , where the upper bound is set at the price where a consumer is indifferent between buying from the dirty type and the clean

type i.e.,  $p_C = \bar{p}_D + \Delta$ . The dirty type charges a price less than  $\bar{p}_D$  almost surely, since otherwise the rival dirty type can undercut to earn higher profit. This, in turn, implies that a clean type sells zero for sure and earns zero profit in the equilibrium. Therefore, in the equilibrium the clean type ends up charging a price as low as its marginal cost ( $m_C$ ). The existence of this equilibrium is guaranteed, because the upper bound of the price support of the dirty type ( $\bar{p}_D = m_C - \Delta$ ) is greater than its marginal cost ( $m_D \leq m_C - \Delta$ ). Since at price  $\bar{p}_D$ , the dirty type of the investing firm undercuts the non-investing firm with probability one, at price  $\bar{p}_D$  the non-investing firm sells only in the state where the rival investing firm is of the clean type; the equilibrium expected profit of the non-investing firm is given by

$$\pi_{NII}^E = \mu[\bar{p}_D - m_D] = \mu(m_C - \Delta - m_D) \quad (41)$$

for charging any price  $p \in [\underline{p}_D, \bar{p}_D]$ , and the dirty type of investing firm earns the same expected profit i.e.,

$$\pi_D = (m_C - \Delta - m_D) \mu \quad (42)$$

The lower bound of the support ( $\underline{p}_D$ ) is the lowest price that the dirty type wants to undercut, given that it is going to capture the entire market irrespective of the type of its rival; it earns strictly positive expected profit which is equal to  $\pi_D$ . This implies  $\underline{p}_D - m_D = \pi_D = (m_C - \Delta - m_D) \mu$ . Therefore, the lower bound of the support is

$$\underline{p}_D = \mu [p_C - \Delta] + (1 - \mu) m_D. \quad (43)$$

The non-investing firm assigns a probability mass of  $\mu$  to the upper bound of its price support ( $\bar{p}_D$ ); this is because the non-investing firm knows that it can sell at the upper bound only when the rival investing firm becomes clean (with probability  $\mu$ ). At every price  $p \in [\underline{p}_D, \bar{p}_D]$ , the non-investing firm can sell to all consumers as long as it is not undercut by the dirty type of the rival investing firm, and its expected profit at  $p$  is equal to  $\pi_{NII}^E$  viz.  $[\mu + (1 - \mu)(1 - F_I(p))](p - m_D) = (\bar{p}_D - m_D) \mu$ . This yields the probability distribution function of the dirty type of the investing firm

$$F_I(p) = 1 - \frac{\mu}{1 - \mu} \left[ \frac{\bar{p}_D - m_D}{p - m_D} - 1 \right], \quad p \in [\underline{p}_D, \bar{p}_D] \quad (44)$$

where  $F_I(p)$  is a continuous distribution function with no probability mass at any point,  $F_I(\underline{p}_D) = 0$ , and  $F_I(\bar{p}_D) = 1$ . Similarly, at every price  $p \in [\underline{p}_D, \bar{p}_D]$  the dirty type of the investing firm can sell to all consumers as long as it is not undercut by the rival non-investing firm, and its expected



profit at  $p$  is equal to  $\pi_{NII}^E$  viz.  $(p - m_D)(1 - F_{NI}(p)) = (\bar{p}_D - m_D)\mu$ ; this yields the probability distribution function of the non-investing firm

$$F_{NI}(p) = 1 - \mu \frac{\bar{p}_D - m_D}{p - m_D}, p \in [\underline{p}_D, \bar{p}_D] \quad (45)$$

where  $F_{NI}(\bar{p}_D) = 1 - \mu$  and  $F_{NI}(\underline{p}_D) = 0$ . The dirty type of the investing firm follows the same pricing strategy. There is no incentive to adopt the ecolabel for the investing firm, as the dirty type generates higher surplus and the clean type can never sell.

Note that this separating equilibrium is more or less similar (except the probability distribution of the non-investing firm) to the pricing equilibrium under pure signaling with low-green premium (described in Lemma 1). Thus, the above mentioned equilibrium is supported by the same set of out-of-equilibrium beliefs (that satisfies D1 criterion) described in the proof of Lemma 1. Further, neither the clean type nor the dirty type has any incentive to unilaterally deviate from the above mentioned equilibrium strategies (for details see the proof Lemma 1). ■

**Proof of Lemma 4.** The clean type enjoys competitive advantage over the dirty type, because consumers are willing to pay a green premium which is higher than the marginal cost differences between the clean and the dirty type ( $m_C - m_D \leq \Delta$ ). The clean type charges a price (viz.,  $m_D + \Delta$ ) at which a consumer is indifferent between buying from the clean type and from the dirty type at its marginal cost ( $m_D$ ), and this equilibrium price charged by the clean type is higher than its own marginal cost ( $m_D + \Delta \geq m_C$ ). Note that the dirty type still charges a lower price than that of the clean type and thus, have an incentive to imitate the clean type's price. To prevent this, the clean type adopts the ecolabel to credibly disclose its environmental performance and caters to the entire market. In other words, the dirty type cannot sell anything in the state where the investing rival firm is of the clean type. The best price that the non-investing dirty type can charge in the equilibrium is its own marginal cost irrespective of the rival's type. In the state where the investing firm fails to adopt the cleaner technology and thus remains dirty, it charges the same price as its non-investing rival; however, in this case the two dirty types equally share the market and earn zero profit.

The one-sided incomplete information Bayesian equilibrium can be supported by the following out of equilibrium beliefs of consumers: if a firm does not adopt an ecolabel (off equilibrium) then consumers believe that the firm is of dirty type with probability one. These out-of-equilibrium beliefs satisfy the D1 refinement; the argument is as follows. The set of quantities for which it is profitable for a clean type to deviate to any price  $p > m_C$  without any ecolabel is lower than that of the dirty

type, and since a clean type will never deviate to any price below its own effective marginal cost D1 refinement is trivially satisfied in this case. Given these out-of-equilibrium beliefs, no firm has an incentive to unilaterally deviate to any off equilibrium ecolabel adoption decision and price.

Does the clean type have any incentive to deviate from the above mentioned equilibrium strategy? Observe that it is trivial to show that there is no gainful deviation for the clean type to any out-of-equilibrium price *with the ecolabel*. In other words, if the clean type charges a price  $p > m_D + \Delta$ , then it loses all consumers to the dirty type and deviation to a lower price ( $p < m_D + \Delta$ ) is not a gainful deviation as the clean type does not earn any additional market share by lowering its price. I consider two other possible deviations; the clean type *does not adopt an ecolabel* and either charges the same price or a different price. In case the clean type does not adopt the ecolabel then the price charged by the clean type should act as a signal for the environmental performance. The out-of-equilibrium belief says that if consumers observe any price *without any ecolabel* they believe that the firm is of dirty type. Therefore, a clean type has no incentive to not adopt the ecolabel at any price. ■

**Proof of Proposition 4.**

	Invests	Does not Invest
Invests	$\pi_{II}^E - F, \pi_{II}^E - F$	$\pi_{INI}^E - F, \pi_{NII}^E$
Does not Invest	$\pi_{NII}^E, \pi_{INI}^E - F$	$\pi_{NINI}^E, \pi_{NINI}^E$

$\pi_{II}^E$  and  $\pi_{INI}^E$  are the *ex ante* expected profits of an investing firm when the rival invests and does not invest respectively, and  $\pi_{NII}^E$  and  $\pi_{NINI}^E$  are the *ex ante* expected profits of a non-investing firm when the rival invests and does not invest respectively in the presence of the ecolabel.

Consider  $0 \leq \Delta \leq (m_C - m_D)$ . Since  $\pi_{II}^E = \mu(1 - \mu)(m_C - \Delta - m_D) < \pi_{NII}^E = \mu(m_C - \Delta - m_D)$  an investing firm deviates to not invest given that its rival has invested. If  $\pi_{INI}^E - F = \mu(1 - \mu)(m_C - \Delta - m_D) - F > \pi_{NINI}^E = 0$  then an investing firm has no incentive to deviate from its investment decision given that its rival does not invest. Therefore, in the equilibrium, only one firm invests if  $\pi_{INI}^E = \mu(1 - \mu)(m_C - \Delta - m_D) \geq F \implies \Delta \leq \Delta_1 = (m_C - m_D) - \frac{F}{\mu(1-\mu)}$ , and no firm invests if  $\Delta_1 \leq \Delta \leq (m_C - m_D)$ .

If the green premium is high ( $m_C - m_D < \Delta \leq \bar{\Delta}$ ) then both firms invest if  $RI^E - F \geq 0 \implies \mu[(1 - \mu)(m_D + \Delta - m_C)] - F \geq 0 \implies \Delta_3 = (m_C - m_D) + \frac{F}{\mu(1-\mu)} \leq \Delta \leq \bar{\Delta}$ ; only one firm invests if  $\Delta \leq \Delta_3$  and  $\pi_{INI}^E = \mu(m_D + \Delta - m_C) - F \geq \pi_{NINI}^E = 0 \implies (m_C - m_D) + \frac{F}{\mu} = \Delta_2 \leq \Delta$ .

■

**Proof of Proposition 5.**

	Invests	Does not Invest
Invests	$\pi_{II}^S - F, \pi_{II}^S - F$	$\pi_{INI}^S - F, \pi_{NII}^S$
Does not Invest	$\pi_{NII}^S, \pi_{INI}^S - F$	$\pi_{NINI}^S, \pi_{NINI}^S$

$\pi_{II}^S$  and  $\pi_{INI}^S$  are the *ex ante* expected profits of an investing firm when the rival invests and does not invest respectively, and  $\pi_{NII}^S$  and  $\pi_{NINI}^S$  are the *ex ante* expected profits of a non-investing firm when the rival invests and does not invest respectively when the firms signal their environmental quality through prices.

When  $\Delta \leq \frac{m_C - m_D}{2}$ , observe that  $\pi_{II}^S = \mu(1 - \mu)(m_C - \Delta - m_D) < \pi_{NII}^S = \mu(m_C - \Delta - m_D)$  and  $\pi_{INI}^S - F = \mu(1 - \mu)(m_C - \Delta - m_D) - F \geq \pi_{NINI}^S = 0 \implies \Delta \leq \Delta_1 = (m_C - m_D) - \frac{F}{\mu(1 - \mu)}$  only one firm invests.

Consider the intermediate-green premium  $\frac{(m_C - m_D)}{2} < \Delta \leq (m_C - m_D)$ . Note that  $\pi_{II}^S - F = \frac{\mu^2}{2}(m_D + 2\Delta - m_C) + (1 - \mu)\mu\Delta - F \geq \pi_{NII}^S = \mu(m_C - \Delta - m_D) \implies RI^S \geq F$  when

$$\frac{(1 - \frac{\mu}{2})}{2}(m_C - m_D) + \frac{F}{2\mu} \leq \Delta \leq (m_C - m_D); \quad (46)$$

in this case, both firms invest in the equilibrium. Now suppose that  $\Delta \leq \frac{(1 - \frac{\mu}{2})}{2}(m_C - m_D) + \frac{F}{2\mu}$ ; an investing firm deviates given that its rival invests. If  $\pi_{INI}^S - F = (1 - \mu)\mu(m_C - \Delta - m_D) - F \geq \pi_{NINI}^S = 0 \implies \Delta \leq \Delta_1 = (m_C - m_D) - \frac{F}{\mu(1 - \mu)}$  then only one firm invests.

Suppose  $m_C - m_D < \Delta \leq \bar{\Delta}$ . Note that  $\pi_{II}^S - F = \frac{\mu^2}{2}(m_D + 2\Delta - m_C) + (1 - \mu)\mu\Delta - F \geq \pi_{NII}^S = 0 \implies$  if

$$\frac{\mu}{2}(m_C - m_D) + \frac{F}{\mu} \leq \Delta \leq \bar{\Delta} \quad (47)$$

then both firms invest, since  $\pi_{INI}^S = 0$  if  $\Delta \leq \frac{\mu}{2}(m_C - m_D) + \frac{F}{\mu}$  then only one firm invests in the equilibrium.

At  $\Delta = (m_C - m_D)$ ,  $RI^S = \mu[(m_D + 2\Delta - m_C) - \frac{\mu}{2}(m_C - m_D)] = \frac{\mu^2}{2}(m_D + 2\Delta - m_C) + (1 - \mu)\mu\Delta = \mu(1 - \frac{\mu}{2})(m_C - m_D)$ . I define the threshold green premiums derived in the above investment equilibrium analysis (given by (46) and (47))

$$\Delta_4 = \begin{cases} \frac{(1 - \frac{\mu}{2})}{2}(m_C - m_D) + \frac{F}{2\mu} & \text{if } F \leq \mu(1 - \frac{\mu}{2})(m_C - m_D) \\ \frac{\mu}{2}(m_C - m_D) + \frac{F}{\mu} & \text{if } F \geq \mu(1 - \frac{\mu}{2})(m_C - m_D) \end{cases}$$

Thus, one can conclude that no firm invests if  $\Delta_1 \leq \Delta \leq \Delta_4$  and both firms invest when  $\Delta_4 \leq \Delta \leq \bar{\Delta}$ . ■

**Proof of Proposition 6.**

	Invests	Does not Invest
Invests	$\pi_{II}^M - F, \pi_{II}^M - F$	$\pi_{INI}^M - F, \pi_{NII}^M$
Does not Invest	$\pi_{NII}^M, \pi_{INI}^M - F$	$\pi_{NINI}^M, \pi_{NINI}^M$

$\pi_{II}^M$  and  $\pi_{INI}^M$  are the *ex ante* expected profits of an investing firm when the rival invests and does not invest respectively, and  $\pi_{NII}^M$  and  $\pi_{NINI}^M$  are the *ex ante* expected profits of a non-investing firm when the rival invests and does not invest respectively under the mandatory disclosure.

Observe that when  $\Delta \leq m_C - m_D$ ,  $\pi_{II}^M = (1 - \mu) \mu (m_C - \Delta - m_D)$ ,  $\pi_{INI}^M = 0$ ,  $\pi_{NII}^M = \mu (m_C - \Delta - m_D)$ , and  $\pi_{NINI}^M = 0$ .  $\pi_{II}^M$  is the profit of an investing firm when it remains dirty (with probability  $(1 - \mu)$ ) whereas the rival has become clean (with probability  $\mu$ ) and the dirty type charges a price  $(m_C - \Delta)$  which is higher than the marginal cost of the dirty type. Similarly, I can explain other profit expressions. The unilateral and the reciprocal incentive of a firm to invest in the cleaner technology are

$$\begin{aligned}
 UI^M &= \pi_{INI}^M - \pi_{NINI}^M = 0 \text{ and} \\
 RI^M &= \pi_{II}^M - \pi_{NII}^M = -\mu^2 (m_C - \Delta - m_D) < 0
 \end{aligned}$$

respectively. Thus, the equilibrium of this investment game is as follows; no firm invests for any strictly positive fixed cost ( $F$ ). For zero fixed cost of investment there is an asymmetric invest equilibrium where only one firm invests.

When  $m_C - m_D \leq \Delta \leq \bar{\Delta}$ ,  $\pi_{II}^M = (1 - \mu) \mu (m_D + \Delta - m_C)$ ,  $\pi_{INI}^M = \mu (m_D + \Delta - m_C)$ ,  $\pi_{NII}^M = 0$ , and  $\pi_{NINI}^M = 0$ . Note that in this case, the unilateral and the reciprocal incentives of a firm are

$$\begin{aligned}
 UI^M &= \pi_{INI}^M - \pi_{NINI}^M = \mu (m_D + \Delta - m_C) > 0 \text{ and} \\
 RI^M &= \pi_{II}^M - \pi_{NII}^M = \mu(1 - \mu) (m_D + \Delta - m_C) > 0
 \end{aligned}$$

respectively. Both firms invest if  $RI^M - F \geq 0 \implies \mu [(1 - \mu) (m_D + \Delta - m_C)] - F \geq 0 \implies \Delta_3 = (m_C - m_D) + \frac{F}{\mu(1-\mu)} \leq \Delta \leq \bar{\Delta}$ ; only one firm invests if  $\Delta \leq \Delta_3$  and  $\pi_{INI}^M = \mu (m_D + \Delta - m_C) - F \geq \pi_{NINI}^E = 0 \implies (m_C - m_D) + \frac{F}{\mu} = \Delta_2 \leq \Delta$ . ■

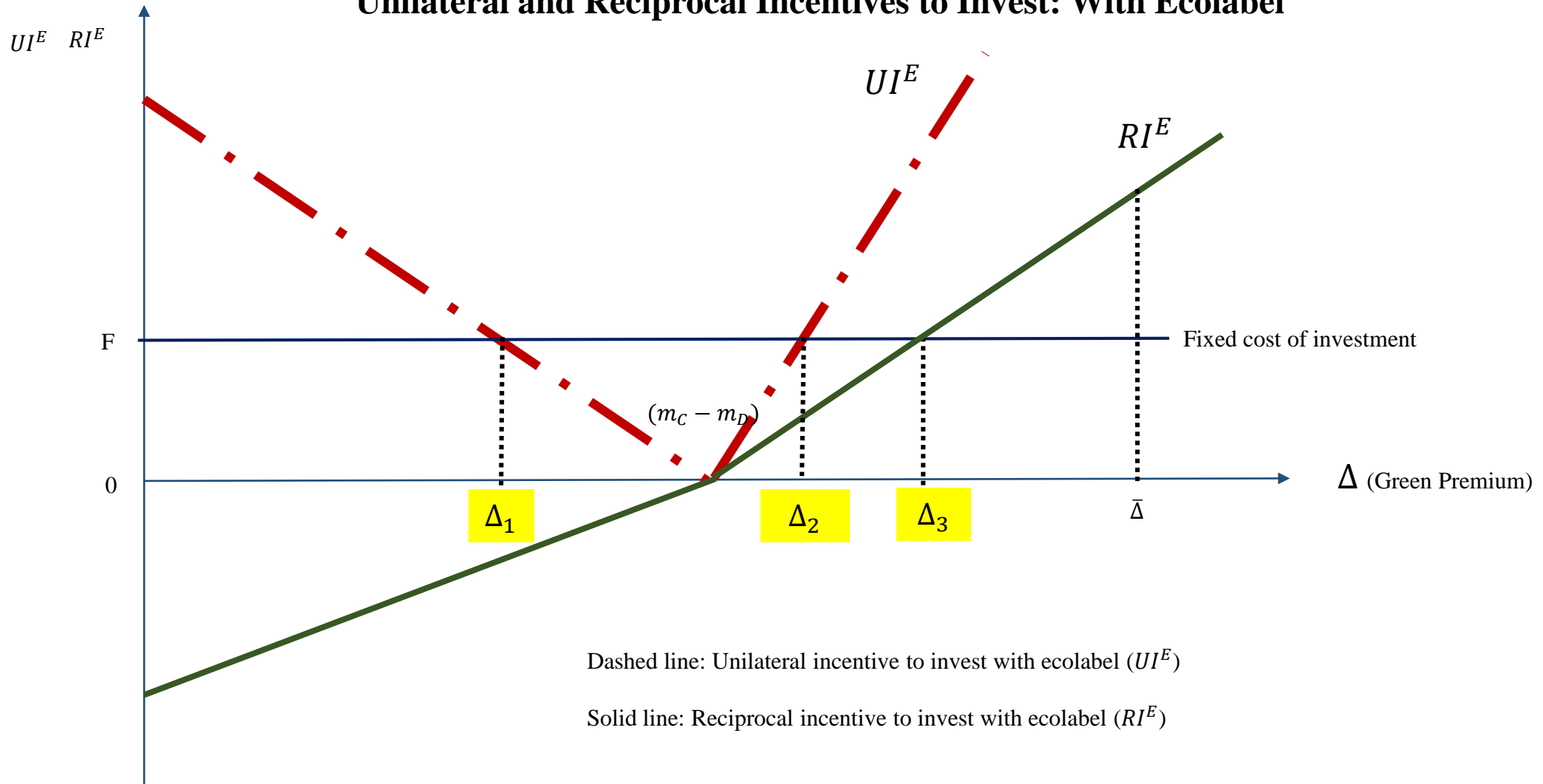
## References

- [1] Amacher, G. E Koskela, and M., Ollikainen. 2004. Environmental quality competition and ecolabeling. *Journal of Environmental Economics and Management*. 47, 284-306.

- [2] Arora, S and S. Gangopadhyay. 2003. Toward a theoretical model of voluntary Overcompliance. *Journal of Economic Behavior and Organization*. 28, 289-309.
- [3] Bansal, S. and S. Gangopadhyay. 2003. Tax/subsidy policies in the presence of environmentally aware consumers. *Journal of Environmental Economics and Management*. 45, 333–355.
- [4] Casadesus-Masanell, R., M. Crooke, F. Reinhardt, and V. Vasishth. 2009. Households’ willingness to pay for “green” goods: evidence from Patagonia’s introduction of organic cotton sportswear. *Journal of Economics and Management Strategy*. 18, 203–233.
- [5] Cho, I. K. and D. Kreps. 1987. Signaling games and stable equilibria. *Quarterly Journal of Economics*. 102, 179–222.
- [6] Cho, I. K. and J. Sobel. 1990. Strategic Stability and Uniqueness in Signalling Games. *Journal of Economic Theory*. 50, 381-413.
- [7] Cremer, H. and J. Thisse. 1999. On the taxation of polluting products in a differentiated industry. *European Economic Review*. 43, 575–594.
- [8] Dosi, C. and M. Moretto. 2001. Is ecolabeling a reliable environmental policy measure?. *Environmental and Resource Economics*. 18, 113-127.
- [9] Fischer, C. and T. Lyon. 2014. Competing Environmental Labels. *Journal of Economics and Management Strategy*. 23, 692-716.
- [10] Fischer, C. and T. Lyon. 2015. A Theory of Multi-Tier Ecolabels. Working paper.
- [11] Galarraga, I and A. Markandya. 2004. Economic techniques to estimate the demand for sustainable products: a case study for fair-trade and organic coffee for the United Kingdom. *Economía Agraria Recursos Naturales*. 4, 109–134.
- [12] Grolleau, G. and L. Ibanez. 2008. Can ecolabeling schemes preserve the environment?. *Environmental and Resource Economics*. 40, 233-249.
- [13] Hwang, Y., B. Roe, and M. Teisl. 2006. Does Price Signal Quality? Strategic Implications of Prices as a Signal of Quality for the Case of Genetically Modified Food. *International Food and Agribusiness Management Review*. 9, 93-115.

- [14] Janssen, M. and S. Roy. 2010. Signaling Quality Through Prices in an Oligopoly. *Games and Economic Behavior*. 68, 192-207.
- [15] Kuhn, M. 1999. Environmental Labels and Entry into an Environmentally Differentiated Market under Asymmetric Information. University of Rostock working paper no. 20.
- [16] Mattoo, A. and A. Singh. 1994. Eco-labelling: Policy considerations. *Kyklos*. 47., 53-65.
- [17] Sengupta, A. 2015. Competitive Investment in Clean Technology and Uninformed Green Consumers. *Journal of Environmental Economics and Management*. 71, 125-141.
- [18] Teisl, M., B. Roe, and R. Hicks. 2002. Can eco-labels tune a market? Evidence from Dolphin-Safe labeling. *Journal of Environmental Economics and Management*. 43, 339-359.

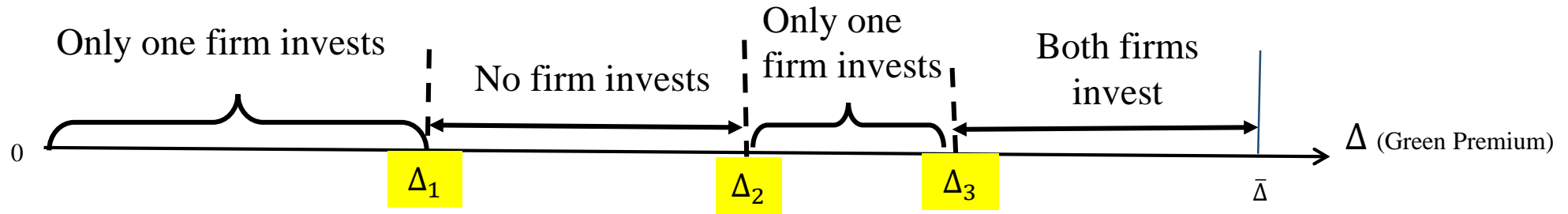
# Unilateral and Reciprocal Incentives to Invest: With Ecolabel



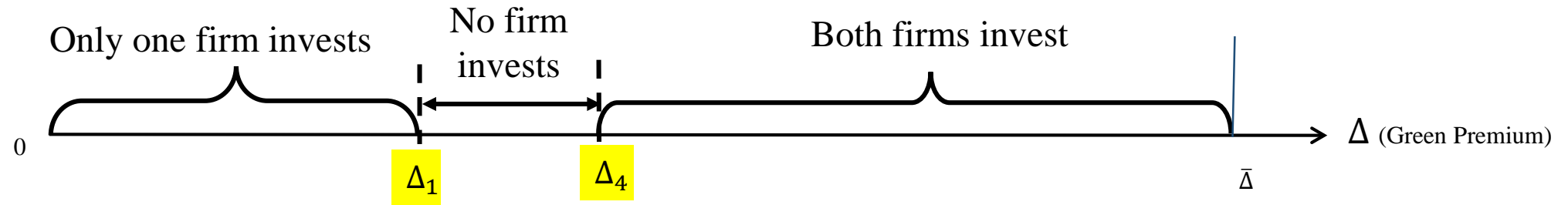
**Figure 1**

# Investment Equilibria

(a) With Ecolabel



(b) Without Ecolabel



(c) Mandatory Disclosure

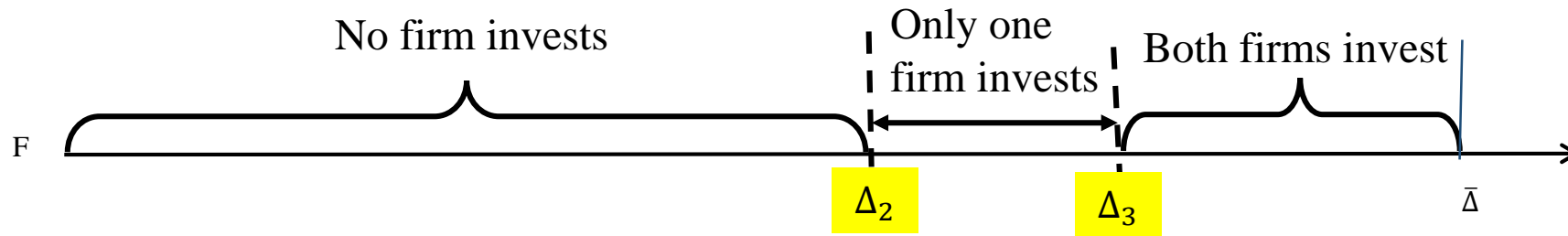
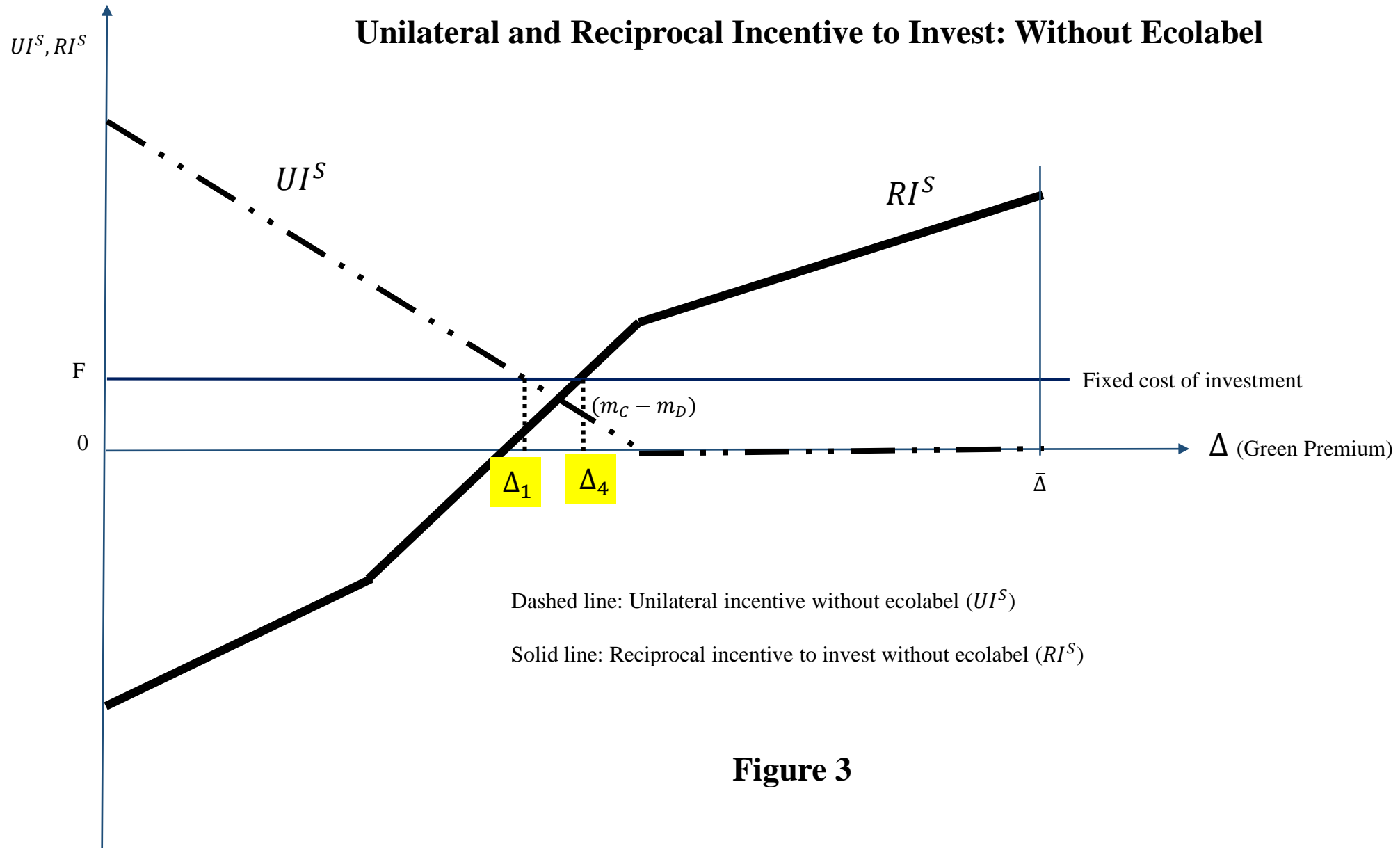


Figure 2



# Unilateral and Reciprocal Incentive to Invest: Without Ecolabel



**Figure 3**

$UI^E, UI^S$

### Unilateral Incentive to Invest: Ecolabel vs. Without Ecolabel

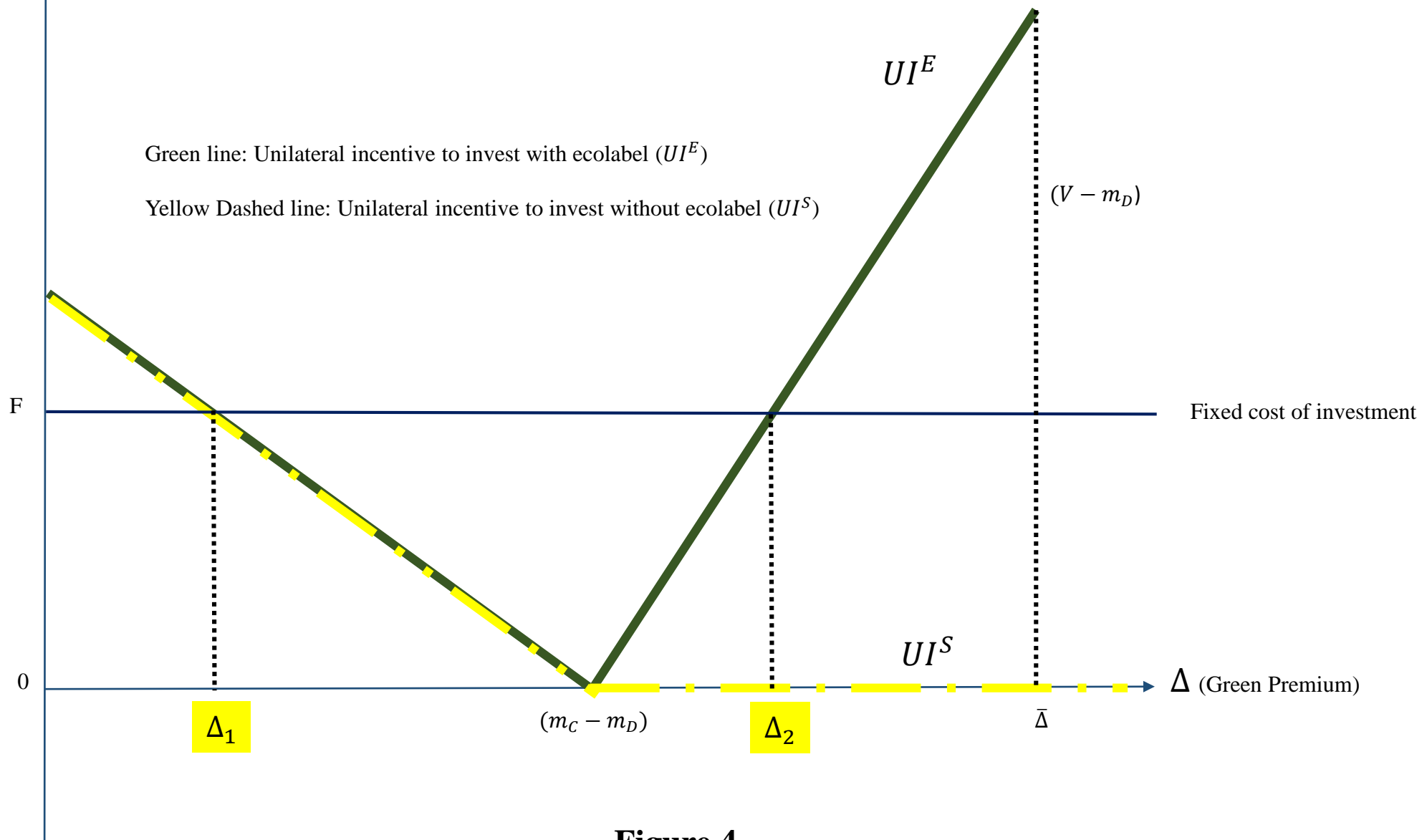


Figure 4

# Reciprocal Incentive to Invest: Ecolabel vs. Without Ecolabel

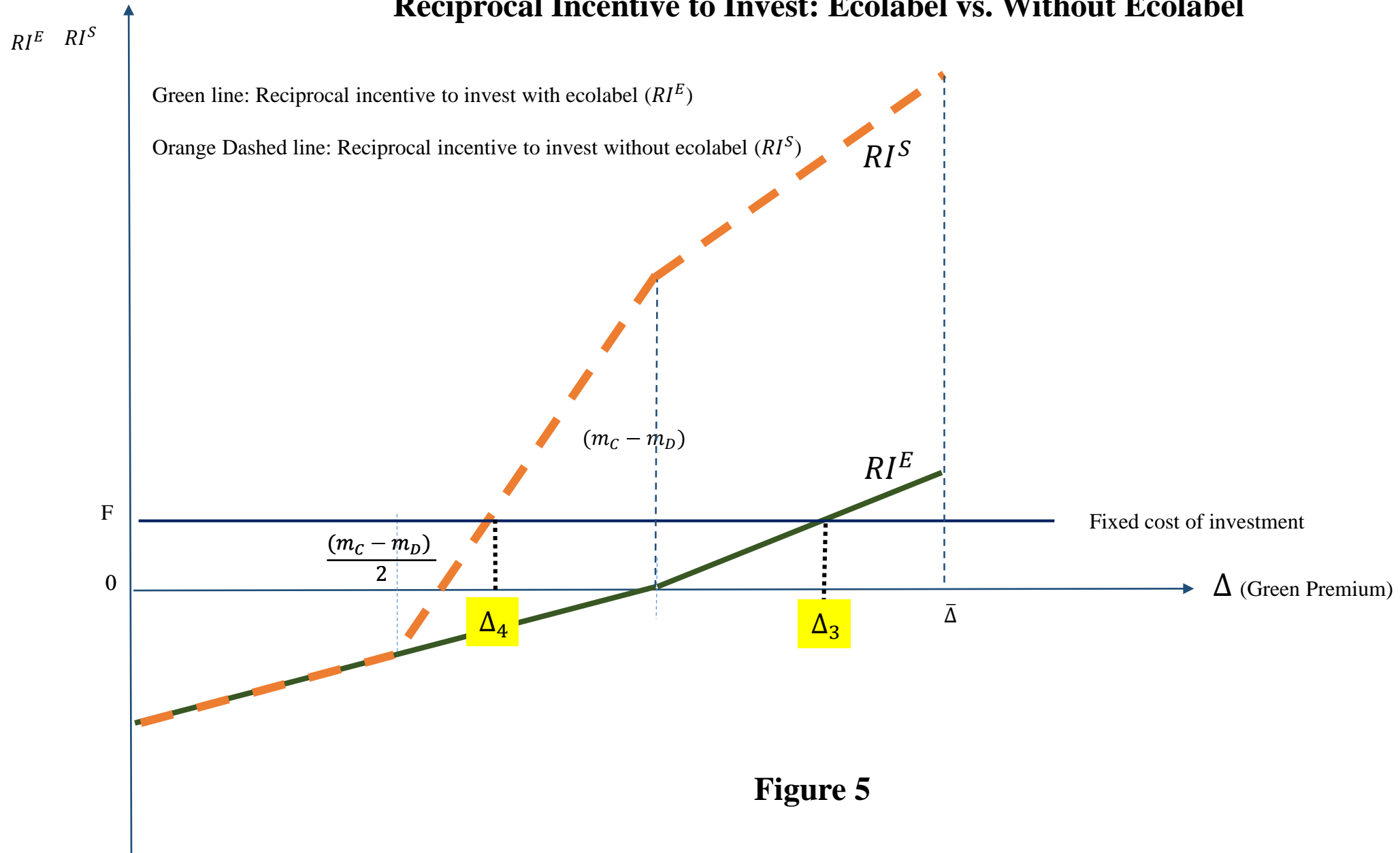
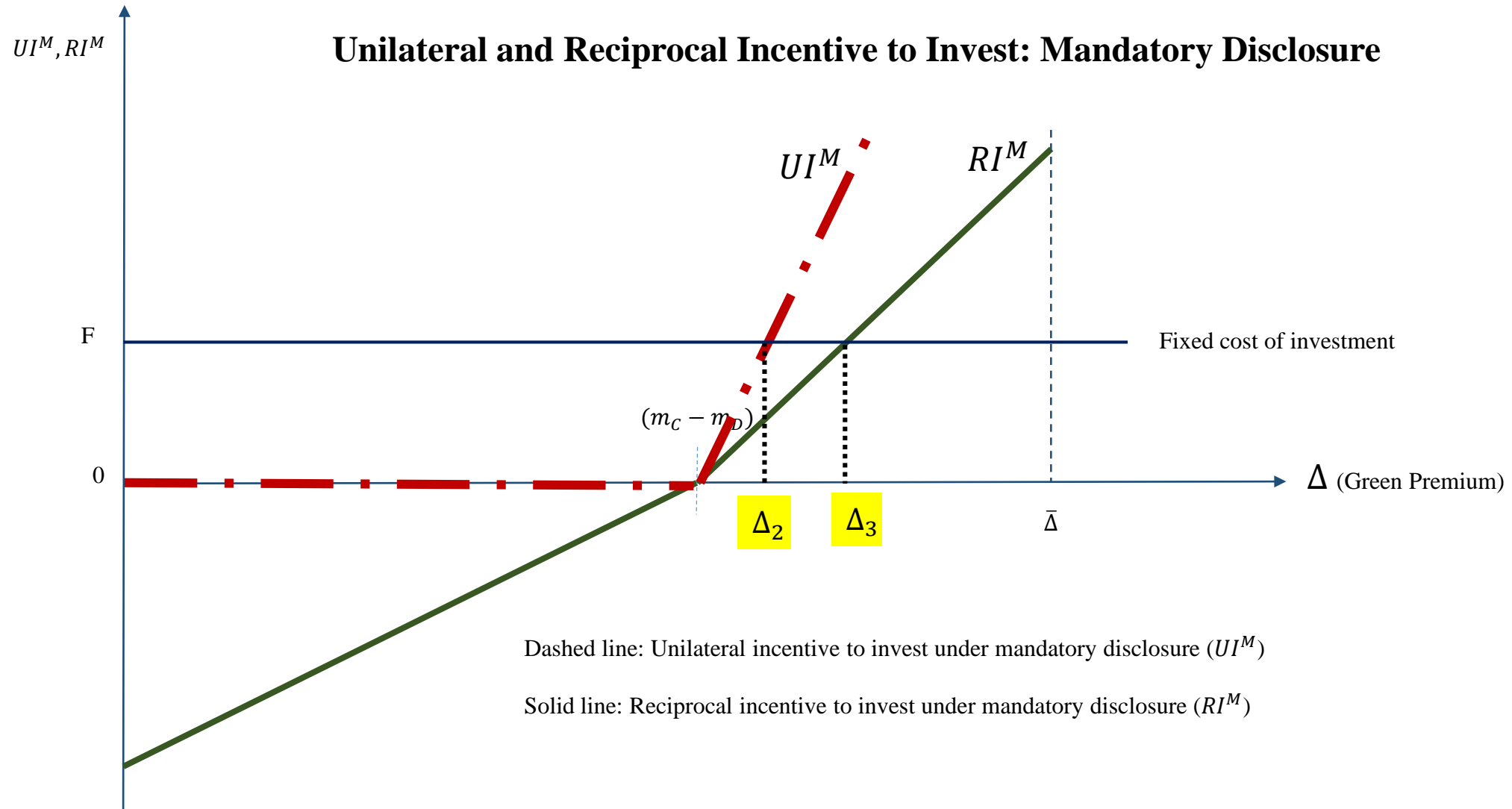


Figure 5



**Figure 6**