WCERE 2018 - 6th World Congress of Environmental and Resource Economists

Gothenburg | Sweden 25-29 June 2018

Does a market with green goods voluntarily internalize externalities? Evidence from a lab experiment

Maria Eduarda Fernandes
eduarda.fernandes@ipleiria.pt
Affiliations:
School of Technology and Management, and CARME, Polytechnic Institute of Leiria, Portugal
ORCID: 0000-0001-6584-0412

Marieta Valente
mvalente@eeg.uminho.pt
Affiliations:
NIPE and EEG, Universidade do Minho, Portugal
ORCID: 0000-0003-3332-8945

Abstract

In a green goods market a combination of individual and corporate social responsibility may lead to the internalization of externalities. This economics experiment implements a market for green credence goods in the presence of externalities on other buyers and explores whether a combination of individual and corporate social responsibility may lead to the internalization of externalities. Under information asymmetry, we observe widespread false claims and an apparently pro-environmental market, when in reality green goods are sparingly sold. When a credible label is possible or when the information asymmetry is removed, the provision of actual green goods increases, but is roughly 20% to 25% of the market share. While this share is non-negligible, the niche market that ensues does not ensure that less environmentally damaging consumption options will be widespread, nor that social welfare will be maximized once the information asymmetry is removed.

Keywords: externalities; credence goods; labels; prosocial behavior; economics experiments.

1 Introduction

Traditionally, in economics, externalities are identified as market failure, given that producers or consumers do not internalize the total social costs or benefits in their production or consumption choices. However, as noted by Bénabou and Tirole (2010) there are several examples of socially responsible behavior whereby individuals and firms do indeed voluntarily consider externalities in their choices (respectively individual and corporate social responsibility). The focus of our paper is on one such example namely the market for green goods. Green goods are produced with often more expensive technologies but mitigate negative environmental impacts, that is reduce externalities. For green goods to be demanded, this means consumers need to care about the negative externality imposed on the environment and on others, and are voluntarily willing to pay for its alleviation. On the other hand, producers might also opt for greener technologies, as a reflection of corporate social responsibility (CSR). Therefore in the presence of a combination of individual and corporate social responsibility, it would be possible for the market to voluntarily reduce externalities through the supply and demand of green goods.

In most cases the environmental friendliness of green goods cannot be accessed beforehand by the consumer and in some cases, not even after the purchase, which means green goods often correspond to credence goods (Karl and Orwat 1999). Therefore, even if consumers have green preferences and green goods are being offered in the market, they are not able to make informed decisions. Also, firms may opt for less environmentally damaging technologies but may find it hard to be rewarded in the market for their CSR. One policy option to help market participants overcome this information asymmetry falls within the third wave of policy approaches to the externality market failure, namely an information

disclosure approach (OECD 2001). One concrete application of this principle is a labelling policy, whereby labelling corresponds to "any action by a firm, government, or third party that communicates product-specific information to end users" (Roe et al 2014, p. 408). The label thus refers to an unobservable characteristic, reducing the information asymmetry between the two sides of the market when truthful, as it provides information that the consumers would not have been able to acquire otherwise.

The present paper proposes an experimental design to explore to what extent buyers and sellers can voluntary internalize externalities in the market when green goods are possible. In the terminology of Bénabou and Tirole (2010), our experiment explores the scope of individual and corporate social responsibility to reduce externalities in a market for green goods. Specifically, the experimental design mimics a posted-offer competitive green goods market, simultaneously addressing two market failures, namely externalities and asymmetry of information in credence goods, as well as labels' capacity to overcome the latter consequences in the market.

Our proposed experimental framework includes the combination of prosocial behavior from sellers and buyers, within a competitive market with the possibility of cheating due to information asymmetry, as well as the optional inclusion of third party certification (eco-label), at a cost to firms. Three treatments are designed and implemented. The main treatment (TREATMENT CREDENCE) directly matches the market conditions of market for green goods, where goods produced with the least expensive technology generate externalities on others, whereas green goods do not cause negative impacts. However, the claims of "greenness" cannot be verified by buyers.

To overcome the information asymmetry, the second treatment includes third party certification (TREATMENT THIRD PARTY CERTIFICATION) that is a voluntary label. We can interpret this as a policy instrument that promotes more transparency in the market place. Given that this implies a use of resources in the certification, it is a second-best solution for society, but a necessary one.

The full information treatment (TREATMENT FULL-INFO) is a benchmark treatment. While it implements the theoretical first-best solution, it is not realistic in the case of credence goods, as opposed to the other two market setups in the two treatments above. However, we implement this treatment as it illustrates the limits of the voluntary internalization of externalities by market players, albeit in unrealistic conditions.

Previous laboratory experiments addressed these issues separately and represented markets for goods with externalities (e.g. Bartling et al. 2015; Plott 1983), while others treated the asymmetry information problem (Bougherara and Piguet 2009; Cason and Gangadharan 2002; Dejong et al. 1985; Etilé and Teyssier 2016; Miller and Plott 1985). We propose an experimental design that simultaneously represents both market failures and that includes design features aimed at providing a more realistic laboratorial framework for a green goods market. As a consequence, this design becomes more relevant for environmental policy purposes, assuring the parallelism condition as advocated by Smith (1982). As the main distinctive features of our experiment, we combine both market failures, as well as propose an effective implementation of an experimental credence good. Additionally, we use homegrown rather than the more commonly used induced values in the asymmetry information framework. These options reinforce the realistic framework aimed with our experimental design and allow us to establish how far homegrown prosocial preferences

can help address externalities in the market. This ultimately addresses the research question of whether voluntary approaches through green goods can lead to externality internalization.

The results from the experiment match theoretical expectations, which in turn contribute to validate our design as a framework to study externalities mitigation in the market. In TREATMENT CREDENCE the results corroborate the theoretical prediction for an inefficient market outcome in terms of social welfare. Furthermore, false claims as to the type of good are widespread in the market.

Additionally this experiment provides insights into the market dynamic between buyers and sellers and enriches our understanding of the scope of voluntary internalization of externalities. A priori we have no expectations as to the extent of prosocial preferences in this setup, nor about the outcome of the interaction between buyers and sellers. Other economics experiments, which we will explore later, show there is a non-negligible degree of prosocial preferences and behaviors, even in the controlled and less morally charged environment of an economics lab. While we partly confirm those observations, our evidence is less optimistically in support of voluntary approaches.

When the information asymmetry is solved under full information conditions or when a credible label is available, we find a higher share of green goods being traded and a positive effect for sellers of establishing a green reputation. However, the market share of actual green goods is not as high as other experiments would suggest (e.g. Bartling et al. 2015). This result is an indication about the limited power that people's concern (or not) about

consumption consequences on society in general, or on the environment in particular, can have to solve an externality problem.

Within this context our results call for more caution in assessing the potential of voluntary approaches to internalize externalities and for a less optimistic view of the power of individual and corporate social responsibility. We should not expect the market (participants) alone to solve this market failure.

The rest of the paper is organized as follows. In section 2 we discuss the approaches that have been used in the literature to study the above mentioned market failures, involving green goods: externalities (in section 2.1) and asymmetric information (in section 2.2), focusing on economics experiments that partially share our motivation. In section 3 we present our experimental design and procedures, and discuss how our design options differ from other experiments in the literature, contributing to a more realistic experimental design and ultimately to more externally valid results. Section 4 presents the main results across treatments and section 5 discusses the implications.

2 Overview of related economics experiments

Our experimental framework aims to mimic a market for green goods, simultaneously addressing two market failures, externalities and asymmetry of information. It also evaluates the ability of green goods market to internalize the negative externalities when the information asymmetry is withdrawn, namely by using a credible third party

certification. Different economics experiments partially address these issues and we review these contributions. Section 2.1 explores how previous experiments have addressed green and ethical goods and the considerations of externalities by market participants. Section 2.2 focuses on the information asymmetry problem implied by green credence goods.

2.1. Green goods and externalities

For there to be a market for green goods, with consumers demanding them and suppliers willing to offer them, either one of these market parts, or both, need to care about the externalities generated by the consumption of conventional "brown"[1] alternatives. We can thus assume that the demand for green goods stems from a consideration of prosocial motivations in consumption or production. In the terminology of Bénabou and Tirole (2010), there needs to be individual and corporate social responsibility in the marketplace.

Following the framework of Lancaster (1966)'s characteristics theory of value, green goods, or ethical goods in general, combine a private with a public good attribute (Cornes and Sandler 1994; Kotchen 2006). For there to be demand for these goods, consumers need to care for the public good attribute to value it when choosing amongst market alternatives. In economics experiments, homegrown values may be used to studying green or ethical goods, by introducing a charitable cause or a third party in the design as the public good attribute. Using this kind of framework makes it unnecessary to induce higher preferences

[1] For simplicity in terms of exposition, we will use the term "green" good to mean a good that causes no externalities as opposed to a "brown" good that causes externalities. Reality is obviously more nuanced. Furthermore, the terms "green" and "brown" refer to environmental externalities, but both the experimental framework and analysis can be extended to more generically named "ethical" goods whereby the public good attribute is a broader social dimension.

for the higher quality good (the green good) since consumers will prefer it as long as they have prosocial preferences and producers decide to offer it in the market. When a specific type of good is sold, a third party is impacted positively (as in the case of a charitable cause that receives a certain donation) or negatively (as in the case of a negative externality on a third party) and consumers may or may not consider this in their utility maximization decisions. Rode et al. (2008), Valente (2015) and Bartling et al. (2015), for example, have taken this approach to social preferences in their experimental studies to study markets for green and ethical goods.

Both Rode et al. (2008) and Valente (2015) use a charitable donation in their experiments as a differentiating feature of the goods. Valente (2015) extends the analysis in Rode et al. (2008) to a simpler market setting but allowing the market interaction between buyers and sellers to endogenously determine the ethical dimension of the goods sold rather than exogenously defining how much the ethical cost is and who can sell those goods. The results in Valente (2015) show that ethical differentiation is valued by consumers and that sellers are able to charge higher prices successfully, adding arguments towards the existence of a market for ethical / green goods.

A different experimental approach to consider the specific case for a public bad is used by Bartling et al. (2017; 2015). In this case, an explicit externality for a "passive" third party is experimentally represented, consisting of another experimental subject unrelated to the market interactions. Aiming to study whether individual and corporate social responsibility is sustained within a market context, Bartling et al. (2015) design a market experiment involving the sale of two type of goods: the "unfair product", which generates externalities and is cheaper to produce, and the "fair good" that does not generate externalities but costs

more to produce. These authors study had a similar goal as our own although their experimental design differed in several aspects. Bartling et al. (2015) implement an experimental posted-offer market with excess supply, which means one seller per period does not sell an item, and where each third party is randomly matched with one of the units sold. We did not include these features, as detailed in section 3. Buyers' private value from the consumption of both fair and unfair goods is the same in Bartling et al. (2015), as well as in our experiment.

Plott (1983) introduces externalities in an experimental market lab. These are represented as a negative impact on the other subjects in the market when a brown good is purchased, a procedure we also adopt. However, differently from all the other studies just mentioned, it uses induced instead of homegrown values, where goods differing in their quality translate into different utility levels for consumers. That is to say consumers prefer the higher or better quality goods because their consumption yields more utility through the channel of monetary payoffs. From a social point of view, the higher quality good generates higher welfare gain than the purchase of a lower quality good, considering how monetary payoffs are parametrised. This option associates a higher quality as well as higher utility levels to the consumption of a green good than to the more conventional brown ones, which in reality is not usually the case. Therefore, in order to achieve the intended parallelism between our experimental study and the real market for green goods, we resort to homegrown values in our experiment.

2.2. Green goods as credence goods: the information asymmetry problem

Imperfect information about quality of goods and asymmetric information between producers and consumers imply market outcomes different from predicted for classic models of perfect competition. In some cases, consumers can learn eventually about true quality of the products through experience, or search (Nelson 1970). However, for credence goods, the solution is not as straightforward, because firms do not have the proper incentives to provide truthful information to consumers who are unable to verify the real good quality even after experience, as noted by Darby and Karni (1973). In terms of credence goods, two definitions are possible following Dulleck et al. (2011). One strand of literature follows from Darby and Karni (1973)'s definition of goods and services that consumers need but whose quality they cannot judge, as is the case with specialist services such as healthcare or mechanical repairs (e.g. Dulleck and Kerschbamer 2006). We focus instead on credence goods as goods that "have qualities which are expensive to judge even after purchase [… and] typical examples mentioned in this second strand of literature are goods vertically differentiated by process attributes (…) and typical assumption made is that consumers know what they want or need, but observe neither what they get nor the utility derived from what they get" (Dulleck et al. 2011, p. 527).

Credence goods may suffer from sellers' fraud considering the true quality of the good being sold, with cheap talk announcements in the market and, consequently, a predicted market equilibrium with only bad quality goods traded, i.e., Akerlof (1970)'s prediction of a "market for lemons". To overcome this problem, third-party certification and a credible

label policy is needed. Several authors have approached this issue from a theoretical perspective (e.g., Amacher et al. 2004; Baksi and Bose 2007; Baksi et al. 2016; Bonroy and Constantatos 2008; Cohen and Vandenbergh 2012; Dosi and Moretto 2001; Hamilton and Zilberman 2006; Karl and Orwat 1999; Krarup and Russell 2005; Mason 2006; Mason 2011; Roe and Sheldon 2007; Roe et al. 2014).

The experimental framework most commonly used to study information asymmetries considers an induced preferences approach (Dejong et al. 1985; Holt and Sherman 1990; Lynch et al. 1991; Miller and Plott 1985). Experimental studies referring specifically to green goods markets and eco-labelling have also been using the induced preferences approach to model the environmental attribute as credence or experience attribute. Consumers thus have an incentive to look for information to distinguish between the goods and this information can either be given by the seller, either in what is often referred to as a cheap talk label (since sellers can simply lie about the quality of the good) or in a certified manner. Cason and Gangadharan (2002), one of the most cited experimental studies motivated by green goods and labelling of environmental quality, also uses the induced preferences approach - an abstract setting with “supers” and “regulars”, which give different levels of utility, and have different costs on the production side. These authors conclude that allowing for reputations, further including cheap talk, and even further introducing certification increases the share of supers supplied relative to the baseline.

However, Cason and Gangadharan (2002) experimental design does not include spillovers from one consumer’s choice to the other (therefore the individual choice problem coincides with the social problem) neither externalities are considered (whereby the choice of one consumer benefits or harms other consumers). Therefore, although green goods motivate

this study (goods that cause less or no harm to the environment, in comparison to the brown goods), the authors do not include in their experiment the external costs they impose, as it was the case of the experiments mentioned on the previous section. Additionally, experience, not credence, goods are represented by Cason and Gangadharan (2002), as their experiment included end-of-round revelation, i.e., the type of good offered by all sellers was revealed at the end of each period. Bougherara and Piguet (2009) build on Cason and Gangadharan (2002) but include an outside option for buyers and variations of the labelling and certification schemes, using an experimental setting under conditions closer to the credence good nature. However, Bougherara and Piguet (2009) do not fully accomplish the objective of experimentally representing a market for credence goods as some end of period revelation still existed during the experimental sessions (although only after some periods).

Previous experimental studies did not simultaneously include the externalities and information asymmetry problems for the provision of green goods, i.e. the public good and the credence good characteristics. One notable exception in the experimental literature that addresses externalities and asymmetric information under homegrown preferences is Etilé and Teyssier (2016). The authors use a market where sellers can choose how much to give a charity out of the price (a positive externality) but also consider the possibility for asymmetric information. They design a baseline with credence good since donations associated with the units traded were not observed by participants. Additionally, these authors consider the impact of using a label in another treatment, whereby the label informs that at least a certain amount is (in fact) donated to charity. Etilé and Teyssier (2016) experimental design includes another treatment with a cheap talk label that can be used by sellers, even if a donation of at least a certain amount is not made, but with a positive

probability of being caught lying. As sellers can be identified, these authors also test a reputation effect. Etilé and Teyssier (2016) conclude efficiency gains in this market can be achieved through the introduction of a reliable label (third party certification), whereas when reputation it is not important, unsubstantiated claims produce a halo effect on consumers behavior.

3 Experimental framework and procedures

We study a market where two types of goods are available, namely brown goods generating an externality on others and green credence goods. The experimental design features and parameters are detailed in section 3.1. Section 3.2 describes the procedures adopted for the implementation of the experiment and section 3.3. briefly relates our design features to the literature.

3.1. Experimental design and parameters

We design a posted-offer market, with three buyers and three sellers, where two types of goods are available, a green and a brown good, which differ in their production costs and whether externalities are generated. In terms of the instructions, the framing is neutral and goods are referred to as simply A and B, green and brown good respectively. The difference between the two goods is presented in terms of differences in production costs and the imposition (or not) of external costs on others.

The production cost of a green good is of 50 (experimental points) whereas the brown good is 40, but if the firm does not sell it does not incur the production cost. Sellers have no supply restrictions, which imply that they can offer all the units being demanded in the market, although they can only produce one type of good (A or B) in each period. Buyers, on the other hand, buy at most one unit per period and receive the same payoff (100) for either a green or brown good, that is if a unit is purchased the buyer receives 100 points (minus the price paid), regardless of the type of good. Although not rational, buyers are allowed to accept a price for the goods that implies a loss (as the price set can be up to 110 points). They can choose from which producer to purchase their good and are aware that whenever a unit of a brown good is purchased a negative externality of 20 is caused, implying a 10 points loss on each of the other two buyers in the market.

At the beginning of the experiment, firms are given 400 experimental points, whereas buyers are given 20 experimental points in each round. While this is equivalent in terms of payoffs, we wanted buyers to be assured an income in each round even if they chose not to make a purchase (i.e., this is similar to an outside option). Parameters corresponding to the green and brown goods are summarized in Table 1.

[Table 1 here]

Our experimental setting represents a market for a credence good with negative externalities (TREATMENT CREDENCE). In this treatment firms have an incentive to provide misleading information to consumers (announcing a green good although effectively producing a brown good) because they know consumers are not able to distinguish the quality of the goods. This cheap-talk from sellers is only applicable for the green good, as

when announcing the brown good sellers have no incentive to produce a green good but the opposite is true. Therefore, when deciding to buy a brown good, buyers know they are imposing the externality on others, whereas when a green good is purchased the buyer is not sure if she is imposing this cost and the other two buyers do not know for sure if they will bear an externality cost. This information is revealed only at the end of the session, in the summary payoff screen. TREATMENT CREDENCE additionally allows for consumers' prosocial behavior and sellers' corporate social responsibility, in spite of the information asymmetry, as consumers can still choose to buy green and sellers can choose to actually produce green. On the contrary, this treatment also opens the possibility of "greenwashing", whereby false claims are made as to the true nature of goods, and consumers are misled (or allow themselves to be misled) into choosing these goods. However, in theory no green goods should be produced as these are more expensive to produce and buyers should not demand them given that their true quality is not verifiable.

An additional treatment explores the impact of using an optional seller label (the so-called eco-label in the case of green goods) to overcome the informational market failure (TREATMENT THIRD PARTY CERTIFICATION). This treatment is similar to TREATMENT CREDENCE except for the inclusion of the possibility of firms to pay a fixed cost of 2 points (regardless of selling or not, and per round) for a certification label for a green good. In this case, firms can announce green goods, brown goods and certified green goods. Participants know for sure the externalities imposed on each round when buying brown good or certified green goods but not in the case of a green good. Similarly to TREATMENT CREDENCE there is no type revelation in each round, and only at the end of the 20 periods are all payoffs computed and displayed.

Finally, in TREATMENT FULL-INFO the type of good proposed by sellers is truthfully revealed, so actual externalities imposed on others (if any) are also known with certainty at the end of each period (round) in the market. Therefore, this treatment allows us to evaluate whether or not there are social preferences in this context and serves as a benchmark for the results in the two main treatments.

Table 2 summarizes the main features and differences of the three treatments implemented.

[Table 2 here]

All treatments represent a market with twenty rounds including the three sequential stages described in Table 3. The appendix includes two screenshots from the Z-tree (Fischbacher 2007) interface for TREATMENT CREDENCE as an illustration. The instructions are included in the appendix for refereeing purposes.

[Table 3 here]

At the end of the twenty rounds actual payoffs, including externalities, are calculated and revealed to all participants - a summary display, consisting of the earnings from the twenty rounds, rather than per period information. Earnings from all the twenty rounds are converted at a conversion rate of 100 experimental points = £1.75, and privately announced.

3.2. Experimental procedures

This experiment was run in February 2015 at the Experimental Economics Laboratory of Royal Holloway, University of London (UK) using Z-tree (Fischbacher 2007). As per Table 4, 120 subjects took part and each subject participated in one treatment only in a fixed group and fixed role.

[Table 4 here]

Each experimental session started with a private reading of the instructions by all participants for both roles, an oral summary of the instructions and some practice questions to verify if all instructions were understood. Only after the actual experiment began, were participants informed of the role they would play for twenty rounds. Groups of three sellers and three buyers were randomly matched in each session and interacted in fixed groups and roles for twenty rounds per session.

At the end of the twenty-round market experiment, there was a short questionnaire concerning the participant and the experiment, which also allowed subjects to make any kind of comments about the session. Sessions lasted approximately one hour and thirty minutes and subjects earned on average £14.24, which were paid immediately at the end of the session.

3.3. Design features in relation to the literature

In this subsection we briefly highlight the main design features and how they differ from other economics experiments.

The first noteworthy design feature concerns how we implement externalities. We implement negative externalities as the impact on the other subjects that are part of the market, and whose payoff is affected by the choices of the other consumers. Etilé and Teyssier (2016), on the other hand, do not consider its impact on any agent involved in the experiment but only on the amount of donations NGOs would get from their decisions. We attempt to mimic a situation where a buyer's purchase decision impacts those in a similar situation, which means "browner" choices harm other individuals not far removed (as would be the case of a consumer in a developed country making choices that affect those in poorer countries), but close to her. This procedure also differs from that of Bartling et al. (2015) that implement externalities as the impact on a third party unrelated with the market setting (i.e., an experimental subject that does not participate in the market as buyer or seller). However, our choice is similar to the one adopted by Plott (1983) where each subject is hurt when others buy the goods and cause harm on others whenever engaged on trade (in our experiment, whenever a brown good is traded).

The second feature concerns the information asymmetry. The main experimental treatments we implement (CREDENCE and THIRD PARTY CERTIFICATION) mimic a market for credence goods with externalities in the sense that buyers cannot be sure whether they are in fact buying green or not, unless the good is certified in the latter treatment. In practice, in terms of design protocol, at the end of each period earnings are calculated for every buyer and seller, and externalities from announced brown goods computed. However, in the case of green goods, given their credence attribute, buyers are not sure of what they purchased, so potential externalities are not computed. Participants are reminded that a purchase of an announced green good may in fact be brown and only at the end of the experiment, total

externalities generated and suffered are computed. In this sense, our design differs from Plott (1983) and Bartling et al. (2015) that implement externalities but not the credence attribute, and Cason and Gangadharan (2002) who use experience goods (as there is end-of-round revelation) not credence. Our design also differs from Bougherara and Piguet (2009), that build upon the Cason and Gangadharan (2002) protocol, and try to implement credence goods, but continue to use end-of-round revelation, although only after some periods.

Third, we opt for homegrown rather than induced preferences. The same value is given in our experiment to green and brown goods as we opted to let the distinction between them be via the individual's utility function, that is whether the externality on others impacts the individual's utility. Like Rode et al. (2008), Valente (2015), Bartling et al. (2017; 2015) and Etilé and Teyssier (2016) we study goods that differ in their external impact and will thus be likely to be valued differently by different individuals rather than impose that their objective experimental value is different. This option distinguishes our work from that of Plott (1983), Cason and Gangadharan (2002) and Bougherara and Piguet (2009), for instance, as they opt for an induced preferences approach.

Fourth, we represent a competitive market, where sellers are free to propose goods that differ in their type and price, and able to offer all units demanded in the market, and buyers are not forced to accept a less valued good or a less attractive offer simply because there are not enough goods available for all buyers. On the contrary, Cason and Gangadharan (2002) restrict the number of items that can be sold, which implies in any given rounds there may not be goods available for certain buyers by the time they enter the market. Etilé and Teyssier (2016), on the other hand, opt to have a market where supply is able to meet all demand, but on the demand side impose that buyers buy a minimum amount each period

rather than allow them not to make a purchase, as we did by giving them a fixed endowment each period, even if not buying any good.

4 Results

In this section we present the results across the three treatments with particular emphasis on TREATMENT CREDENCE. We explore how introducing the possibility of reducing the information asymmetry, or removing it, impacts market outcomes, respectively in TREATMENTS THIRD PARTY CERTIFICATION and FULL-INFO.

First, we explore market outcomes in terms of the share of green and brown goods and calculate social welfare for each treatment. Then we focus on market outcomes but as perceived by market participants given that production decisions may not correspond to announcement decisions. Therefore, false claims and greenwashing in TREATMENT CREDENCE are explored. Then, in section 4.2. and 4.3. we focus on the supply and demand side of the market respectively. Finally, we discuss the main results for each treatment and how they allow us to understand the potential role of the market for green goods to help internalize externalities.

4.1. Overview of results across treatments

Table 5 presents a summary of the share of green goods in each treatment. In terms of social welfare, only goods that were produced with green technologies are relevant, so we start by presenting an overview of actual green goods proposed and purchased, that is of actual market outcomes. Later in section 4.1.2. we explore how actual green goods diverge from claims of greenness to show that each market looks in fact different from the perspective of consumers. For the following analysis we will pool market/ period observations.

[Table 5 here]

4.1.1. Actual market outcomes

The shares of actual green goods proposed and traded are in

Table 5. TREATMENT CREDENCE includes two sessions with three markets each, thus six groups of three buyers and sellers. The proposals are for the twenty rounds for a total of 360 units that can be purchased. We observe that 15.6% (56) of producer proposals are for actual green goods.

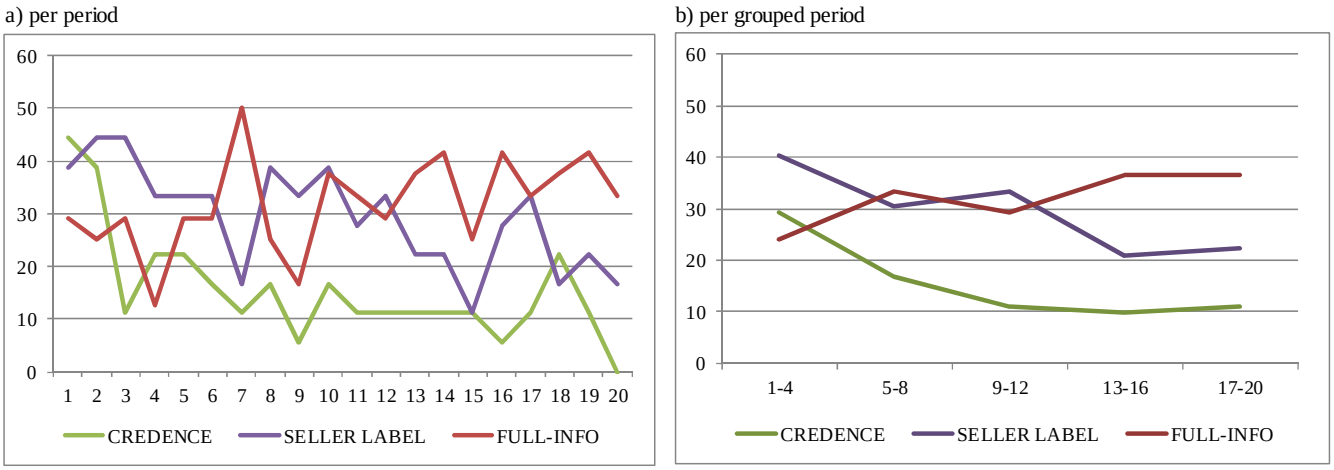
Similarly to the TREATMENT CREDENCE, in TREATMENT THIRD PARTY CERTIFICATION there are six groups resulting in 360 proposals and potential units sold. There are 106 green goods proposed out of the 360 potential, corresponding to a 29.4%. A first comparison of these results shows that producers seem more willing to propose the production of green goods than in TREATMENT CREDENCE – as expected. When comparing announcements for green goods in TREATMENT CREDENCE and TREATMENT THIRD PARTY CERTIFICATION we find statistically significant differences following a non-parametric rank-sum test (Wilcoxon-Mann-Whitney test $z = -3.12$, $p<0.01$).

In TREATMENT FULL-INFO, of the 480 producer proposals, 31.9% are for greens goods while the remainder are for brown goods. In terms of announcements of actual green goods in TREATMENT FULL-INFO and TREATMENT CREDENCE, we find statistically significant differences following a non-parametric rank-sum test that compares the announcements in each market per period (Wilcoxon-Mann-Whitney test $z = 4.65$, $p<0.01$). As for the comparison of TREATMENT FULL-INFO with TREATMENT THIRD PARTY CERTIFICATION, there are no statistically significant differences (Wilcoxon-Mann-Whitney test $z = -1.06$, $p=0.29$).

We now explore market outcomes, i.e. the share of actual green goods purchased in each treatment. This is illustrated in Figure 1a over the 20 periods and in Figure 1b with

averages calculated over four periods to smooth variations. Both in TREATMENT CREDENCE and TREATMENT THIRD PARTY CERTIFICATION there is a decreasing share of green goods. Only TREATMENT FULL-INFO shows relatively high green good market shares.

Figure 1 Share of actual green purchases



In all three treatments the share of actual green goods is not high (compared to other expressions of prosocial behavior in the experimental literature). In TREATMENT FULL-INFO in terms of demand behavior, we observe that 87 of the potential 480 units sold are of green goods (18.1%). In TREATMENT CREDENCE, only 10.3% of units were actually green. When comparing number of green goods sold in each market in TREATMENTS FULL-INFO and CREDENCE over the 20 periods, we find statistically significant differences following a non-parametric rank-sum test (Wilcoxon-Mann-Whitney test $z = 1.81$, $p<0.10$). Also the overall share of actual greens is significantly different across both treatments following a Fisher independence test (Fisher exact test 2-tailed *p*-value < 0.01).

In the TREATMENT THIRD PARTY CERTIFICATION, the overall share of green is 25.8%. As for the comparison of number of greens in each market and period, there are statistically significant differences between TREATMENT THIRD PARTY CERTIFICATION and CREDENCE (Wilcoxon-Mann-Whitney test $z = 2.87$, $p<0.10$) but no differences in comparison with FULL-INFO (Wilcoxon-Mann-Whitney test $z = 1.27$, $p=0.21$).

To evaluate the different market outcomes observed depending on treatment, for each treatment the theoretical net social benefit is calculated taking into consideration the experimental setup in Table 3 and the parameters in Table 1. Then, we take the actual number of green and brown goods purchased in each treatment and calculate the net social benefit observed tocompare it with the social optimum case (all greens). The results are in Table 6.

[Table 6 here]

In TREATMENT CREDENCE the actual social benefit is 81.2% of the social optimum (i.e., the net social benefit benchmark of all greens). Fully removing the information asymmetry yields social benefit equivalent to 83.1% of social optimum (in TREATMENT FULL-INFO) and a partial information scenario results in 84% of social optimum (in TREATMENT THIRD PARTY CERTIFICATION). As such social welfare maximization is not achieved in the presence of negative externalities and credence goods. While the result in TREATMENT CREDENCE is not surprising, neither the possibility of reducing the information asymmetry via a label nor full information substantially increases observed net social benefit. The remaining inefficiency can be attributed to how individuals value the negative externality, which does not appear to be enough to counter its impact.

4.1.2. Market outcomes as perceived by market participants

Given the information asymmetry as to the type of good in TREATMENT CREDENCE and THIRD PARTY CERTIFICATION, unless the green good is certified buyers and the other sellers are not sure if the good was produced with green technology. Therefore what they observe can differ from what is actually produced. A striking result in

Table 5 is that market participants throughout the experiment observe roughly 70% to 77% of good proposals to be green. Also, in terms of goods purchased, the shares are 63.6% and 76.2%. To the uniformed observer, these markets are highly pro-environmental.

In TREATMENT THIRD PARTY CERTIFICATION, a few non-certified green goods are indeed green (

Table 5). Also in TREATMENT CREDENCE, in spite of the information asymmetry 10.3% of all traded goods were produced with green technology. Within the setup of this experiment these two cases can be interpreted as examples of corporate social responsibility in that sellers incur higher production costs, even though the market cannot verify it.

We focus more closely in TREATMENT CREDENCE in Figure 2 and Figure 3. Given the information asymmetry, in theory no green goods should be offered or purchased. However as Figure 2 illustrates, out of the 18 units announced and potentially traded each round, a significant share is for actual green goods. In terms of claimed quality of the goods in the market, "greens" represent 77.8% of the market announcements in period 1 and 66.7% in period 20 (average of 69.4% over the 20 periods). However, only 44.4% of initial announcements are for actual green goods in period 1 and in the last period no good proposed is in fact green (average actual green share of 15.6% over the 20 periods). Therefore, while the market supply appears to be green (the corresponding share is in dashed green in Figure 3a), it is in fact falsely announced as green (the actual green share is in solid green in Figure 3b), which is a tactic that is often referred to as greenwashing.

Figure 2 Announcements and purchases of green and brown goods in TREATMENT CREDENCE

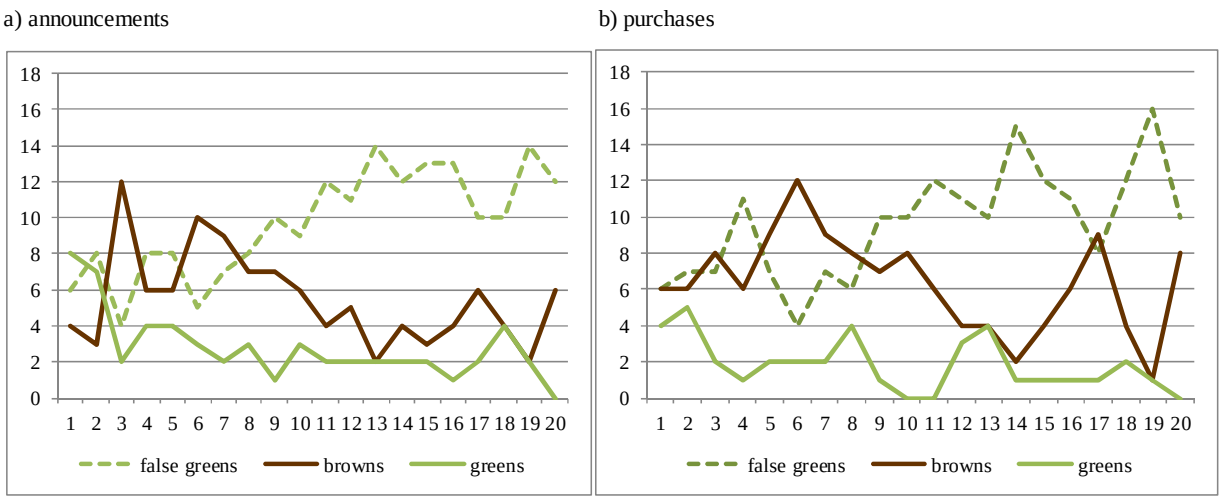
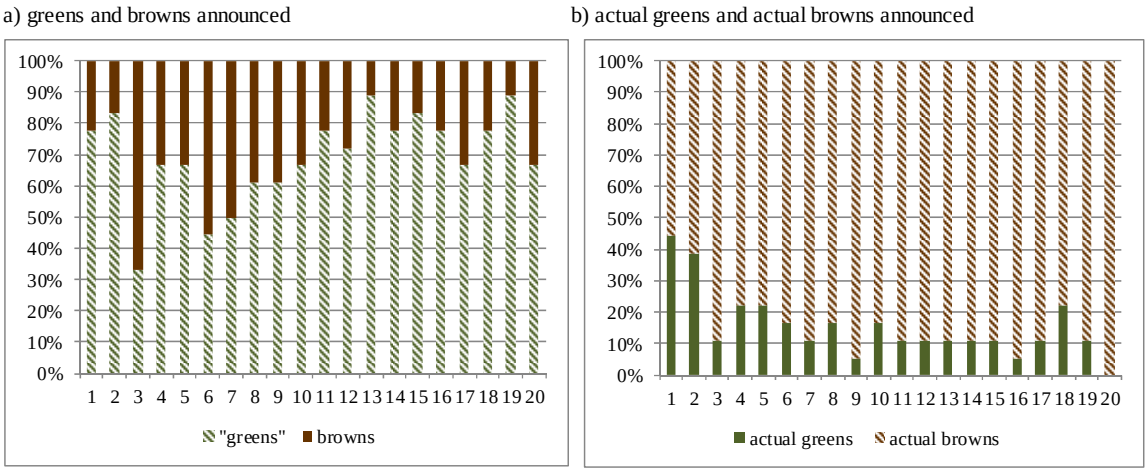


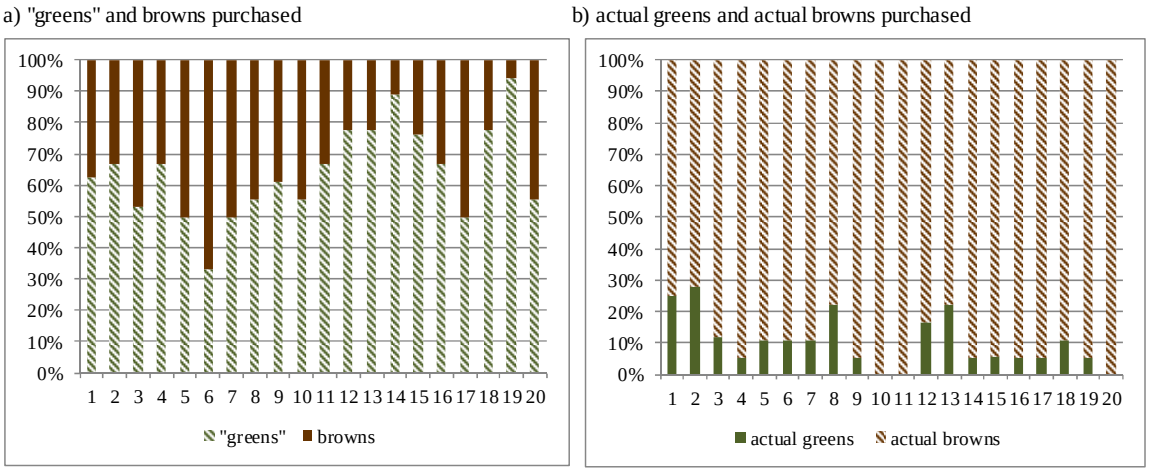
Figure 3 Announcements of green and brown goods in TREATMENT CREDENCE



Note: announcements refer to one unit per seller, regardless of whether each seller makes a sale or not.

As consequence of the false claims made by sellers protected by the information asymmetry, many consumers choose to buy a green good (as in Figure 4a), which in reality turns out to be majorly brown (Figure 4b). The contrast between how consumers perceive the market in Figure 4a and the actual reality in Figure 4b attests to the existence of greenwashing in this market as while we know that the green share in this market is small, for consumers this market appears supportive of green goods.

Figure 4 Purchases: share of green and brown goods in TREATMENT CREDENCE



4.2. Behavior on the supply side and reputation effects

In all three treatments, subjects are paired into fixed groups, so in each market buyers can follow the behavior of each of the sellers throughout the 20 periods. It is thus possible for reputation effects to emerge. To capture reputation, specifically green reputation of sellers, we compute for each period the running sum up to that period (inclusive) of the times a seller proposed a green good in the market (*green reputation*). In TREATMENT FULL-INFO this corresponds to goods that were produced with green technology as well, whereas in the other treatments these are goods announced as green. This is in fact how the goods in the market are perceived by buyers and sellers. Then to account for the relative green reputation of a seller, we calculate how this value relates to the reputation of the greenest competitor in terms of proposals (*relative green reputation*). A positive *relative green reputation* implies the seller has proposed to produce green goods more often than

competitors. On the contrary, a negative *relative green reputation* implies that at least one other seller in the market has a greener reputation.

Table 7 presents the results from a panel data Logit model that regresses the probability of selling in a particular period on this relative reputation measure. In both cases, we control for the *proposed price*, how the proposed price compares relative to the lowest price in the market (*lowest competitors' price*), whether the proposed good is green (*green good proposed*) and whether at least one competitor is proposing a green good (*other green goods proposed*). In TREATMENT THIRD PARTY CERTIFICATION we include a binary variable for whether the seller opted for certification (*certification*).

[Table 7 here]

In regressions (1) – (4) the probability of selling in a given period depends, as expected, negatively on the price proposed and positively on the competitors' price. All else being equal, proposing a green good increases the probability of selling in TREATMENT CREDENCE and TREATMENT FULL-INFO. While the coefficient of this variable is not statistically significant in TREATMENT THIRD PARTY CERTIFICATION, whether the seller chooses to certify its proposed green good increases the probability of selling (model 3). Therefore in a context where certification is possible, simply offering a green good is not sufficient to sell. Specifically in the case of TREATMENT CREDENCE, the coefficient of the binary variable corresponding to at least a competitor proposing a green good, decreases the likelihood of selling, but has no statistically significant impact on other treatments.

How a seller fares relative to competitors in terms of green reputation has a positive impact in TREATMENT FULL-INFO and TREATMENT THIRD PARTY CERTIFICATION. Having proposed

more green goods than competitors up to a particular period increases the probability of selling in that period. However, given the information asymmetry in TREATMENT CREDENCE, this reputational measure has no impact on buyers' choices, as expected.

4.3. Behavior on the demand side and prices paid for green goods

It is also relevant to analyze buyers' behavior in terms of how they value a good's greenness. We present in Table 8 a random effects panel regression of the price paid by each buyer as a function of the type of good (variable *green good purchased* from the perspective of buyers) and in TREATMENT THIRD PARTY CERTIFICATION a dummy variable captures whether the purchased good is certified (*green good certified*). Additionally we include a linear time variable.

[Table 8 here]

In all three treatments, there is a statistically significant and positive coefficient on the dummy that indicates if the good purchased is green. Green goods are thus more highly priced relative to brown goods. Additionally there is a price premium for being certified green in TREATMENT THIRD PARTY CERTIFICATION. Concerning the period variable, there is a significant decrease trend in the price paid in models (1) and (3). This is consistent with the results in Bartling et al. (2015) for example.

In all treatments green goods accrue a price premium relative to brown goods. This is corroborated when comparing prices. Goods that are announced as green have an average price of 51 experimental points and brown goods of 46.2 in TREATMENT CREDENCE. There

are statistically significant differences between the prices of green goods and brown goods (Wilcoxon-Mann-Whitney test $z = 6.15$, $p<0.01$). On the contrary, in TREATMENT THIRD PARTY CERTIFICATION the differences are not significant (Wilcoxon-Mann-Whitney test $z = 0.53$, $p=0.600$). However when certified green goods are compared to the other goods in the market, the prices are higher (Wilcoxon-Mann-Whitney test $z = 9.43$, $p<0.01$), which indicates that the market attributes a price premium to certification.

Furthermore, we can compare prices for green goods across treatments. When comparing green good prices in TREATMENT CREDENCE with each of the other treatment, there is evidence that those are lower namely comparing with certified green goods in TREATMENT THIRD PARTY CERTIFICATION (Wilcoxon-Mann-Whitney test $z = -7.44$, $p<0.01$) and comparing with green goods in TREATMENT FULL-INFO (Wilcoxon-Mann-Whitney test $z = -3.62$, $p<0.01$). So a consequence of the information asymmetry in TREATMENT CREDENCE is that green goods accrue a lower price in the market than when the asymmetry can be partially removed (or is removed all together). This is consistent with the fact that buyers cannot be sure of the true quality of the good and thus prices cannot reflect the good's true quality.

4.4. Discussion of results

Several of the results from this experiment conform to theoretical expectations. However there are surprising results that expand our knowledge as to the degree to which green preferences and goods can help mitigate externalities.

In TREATMENT CREDENCE given the information asymmetry, false claims about the greenness of goods are abundant, when in fact actual green goods produced are few. In terms of the perception of market greenness, in each round, markets appears quite environmentally concerned. This observation attests to the pervasiveness of greenwashing taking place on the part of sellers, and to buyers allowing themselves to be deceived. In fact, even in this setup, proposing a "green" good increases the probability of selling in a given period. An additional result is that prices for green goods are lower in this treatment, so the pervasiveness of green choices may be explained by the lower prices.

Furthermore, while that low share of actual green goods is not surprising, the fact that it is non-zero is. Given that buyers cannot ascertain the true type of the good, the fact that some goods are actually produced with green technology is counterintuitive. This corresponds to CSR given that producer could have simply lied about the true quality. As consequence, in spite of the information asymmetry some sellers opt for more expensive technologies that do not impose externalities.

These are good news both from the supply and from the demand side of the market for the potential of voluntary approaches to help solve an externality problem. Therefore voluntary approaches can be a part of the solution to solving externalities.

On the contrary, the results from the other treatments limit that potential. With no information asymmetry, or with the option to remove it, we observe a share of green goods produced of just around 1/5 to 1/4 of the market. In TREATMENT THIRD PARTY CERTIFICATION there is also the appearance of pervasive green behavior. Statistically, prices of green and brown goods do not differ, however certified green goods do accrue a price

premium. Therefore, the introduction of an optional eco-label partially helps market participants internalize externalities. However, false green claims and purchases are still abundant. This is not surprising given that prices of certified goods are higher.

Finally, TREATMENT FULL-INFO serves as a benchmark as it presents the unrealistic first-best solution. With no information asymmetry, no false claims can be made and buyers need not be wary of uncertified green goods or discount them in the market. The results from this treatment thus provide a reference in terms of the potential of the voluntary internalization of externalities. Given the share of 18.1% of green purchases overall, we can conclude that there appears to be a limit to how far green preferences and CSR can solve the externality market failure.

5 Conclusions

Green and ethical goods are now available in consumer markets as alternatives to more conventional goods and offer a pro-socially motivated consumer the option to care for a public good cause through her consumption decisions. However without an eco-label it is not possible for the consumer to check both before and after the purchase if the claims really are legitimate and the purchased item is not causing externalities on the environment. In economics these types of goods are said to have a credence attribute. The current paper, using an experimental approach, explores precisely this credence attribute in a context where consumers are given the possibility to care about the externalities caused by their purchase decisions and sellers can produce goods that do not cause externalities.

Theoretically, both features individually imply that social welfare maximization should not be achieved and this is in fact observed in the experiment reported here.

We study a market where brown goods cause an externality on others and where green goods have a credence attribute, and allow for individuals to choose according to their preferences, rather than induce preferences. Also, we do not restrict supply as has been done in similarly motivated experiments. As a consequence all buyers in the market are able to buy from just one firm, and are not forced or restricted to buy from the remaining offers available.

Our main research question concerns the simultaneous impact of the information asymmetry (credence feature) and the existence of externalities on others when a brown good is consumed. Furthermore, while information asymmetries may prevent the social optimum from being achieved, consumers need to care about the external costs of their actions on others. A standard economic assumption is for consumers or firms not to voluntarily internalize externalities. A seller label or a full information setting may help overcome the inefficiencies in TREATMENT CREDENCE, but only as long as market participants value not causing externalities on others.

As expected theoretically we observe that in TREATMENT CREDENCE the share of actual greens sold is the lowest of all three treatments. There is a marginal improvement in the market outcomes of the other two treatments (TREATMENT THIRD PARTY CERTIFICATION and TREATMENT FULL-INFO) relative to TREATMENT CREDENCE. Social welfare maximization is however not achieved, as it would imply that all market participants value the burden placed on others from their consumption choices. Indeed the green market share

is only roughly 1/5 to 1/4 in this experiment when the information asymmetry is partially or totally removed.

Therefore theoretical predictions are verified and we observe the expected results. However we find a level of actual green consumption below that of other experiments (e.g., Bartling et al. 2015 report 44% of market share for the good that does not impose externalities) and believe this is driven by non-trivial design options, but which are aligned with more realistic market conditions. As we mentioned, we choose features that closely resemble the type of good that motivated our experiments. As such, our protocol in the main treatment mimics credence goods by not revealing information about the actual externalities caused, until the end of the experiment when payments need to be disclosed. Also, as in Plott (1983) we implement externalities of consumption as the impact on other buyers, namely similar participants who are asked to make similar decisions within the same group, which means the decisions are reciprocal. Finally, we do not introduce supply restrictions which could have led to "forced" green consumption or at least not being free to buy the preferred choice. All of these features combined may have reduced prosocial behavior throughout the experiment in comparison to other experiments in the literature, but we believe that has in turn made both our design and our results more realistic.

Moreover, the reduced market share for green goods observed within our experimental setting is in line with studies on consumers' attitude-behavior gap (e.g. Boulstridge and Carrigan 2000; Devinney et al. 2010). In fact, in spite of widespread support in surveys for pro-environmental consumption choices, actual purchases make green goods still part of a niche market. These results conform with a recent EU survey whereby 26% of respondents indicated that they bought green goods (European Commission 2013), even though support

for green goods is much higher (89% agree that "buying environmentally-friendly products can make a difference to the environment", p. 6).

In the broader study of sustainable development, sustainable consumption and individual choices can play a major role (EU Commission 2014; OECD 2001, 2002). It is often argued that information provision as to the environmental attributes of goods can foster sustainable consumption, notably "consumers should be empowered to make informed choices through better information on green credentials of different products" (EU Commission 2014, p.7). Our experimental results do not challenge this approach but caution as to the limits of voluntary individual and corporate social responsibility. Also, given the results concerning greenwashing and false claims as to the greenness of goods, we can go as far as saying that from a societal perspective it is worth for green claims made in the market to be frequently audited or else we all run the risk of assuming the market is addressing externalities, when in fact it is not.

Plott (1983) challenges the belief that given that individuals exhibit some forms of prosocial behavior, markets will not ignore externalities, or alternatively "people are aware, sensitive, and concerned about others so why should they behave in such an atomistic fashion?" (p. 106). Furthering this conclusion, our experimental results show that voluntary measures relying on consumer choices and potentially concern for others, even if supported by eco-label solutions, are not enough to provide a social optimum equilibrium. The market share for the green goods still remains residual, even when the asymmetric information market failure is solved with reliable certification. This suggests that more conventional policy approaches are needed to address the externality problem.

In conclusion, in this paper we propose a novel and more realistic experimental design to studying green goods market and assess to what extent individual and corporate social responsibility can voluntarily address the externality market failure. The results point to the limited role that this type of voluntary approaches can play in tackling externalities. A natural extension of this experimental setup would be to test different policy instruments to address the externality problem.

References

Akerlof, G. A. (1970). The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics, 84*(3), 488-500.

Amacher, G. S., Koskela, E., & Ollikainen, M. (2004). Environmental quality competition and eco-labeling. *Journal of Environmental Economics and Management, 47*(2), 284-306, doi:10.1016/S0095-0696(03)00078-0.

Baksi, S., & Bose, P. (2007). Credence Goods, Efficient Labelling Policies, and Regulatory Enforcement. *Environmental and Resource Economics, 37*(2), 411-430, doi:10.1007/s10640-006-9032-0.

Baksi, S., Bose, P., & Xiang, D. (2016). Credence Goods, Misleading Labels, and Quality Differentiation. *Environmental and Resource Economics*, 1-20.

Bartling, B., Valero, V., & Weber, R. (2017). On the scope of externalities in experimental markets. *Experimental Economics*, doi:10.1007/s10683-017-9549-9.

Bartling, B., Weber, R. A., & Yao, L. (2015). Do Markets Erode Social Responsibility? *The Quarterly Journal of Economics, 130*(1), 219-266, doi:10.1093/qje/qju031.

Bénabou, R., & Tirole, J. (2010). Individual and Corporate Social Responsibility. *Economica, 77*(305), 1-19, doi:10.1111/j.1468-0335.2009.00843.x.

Bonroy, O., & Constantatos, C. (2008). On the use of labels in credence goods markets. *Journal of Regulatory Economics, 33*(3), 237-252, doi:10.1007/s11149-008-9058-z.

Bougherara, D., & Piguet, V. (2009). Market behavior with environmental quality information costs. *Journal of Agricultural & Food Industrial Organization, 7*(2).

Boulstridge, E., & Carrigan, M. (2000). Do consumers really care about corporate responsibility? Highlighting the attitude-behaviour gap. *Journal of Communication Management, 4*(4), 355–368.

Cason, T. N., & Gangadharan, L. (2002). Environmental Labeling and Incomplete Consumer Information in Laboratory Markets. *Journal of Environmental Economics and Management, 43*(1), 113-134.

Cohen, M. A., & Vandenbergh, M. P. (2012). The potential role of carbon labeling in a green economy. *Energy Economics, 34, Supplement 1*(0), S53-S63, doi:10.1016/j.eneco.2012.08.032.

Cornes, R., & Sandler, T. (1994). The comparative static properties of the impure public good model. *Journal of Public Economics, 54*(3), 403-421.

Darby, M. R., & Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics, 16*(1), 67-88.

Dejong, D. V., Forsythe, R., & Lundholm, R. J. (1985). Ripoffs, Lemons, and Reputation Formation in Agency Relationships: A Laboratory Market Study. *The Journal of Finance, 40*(3), 809-820, doi:10.1111/j.1540-6261.1985.tb05006.x.

Devinney, T. M., Auger, P., & Eckhardt, G. M. (2010). *The Myth of the Ethical Consumer* Cambridge University Press.

Dosi, C., & Moretto, M. (2001). Is Ecolabelling a Reliable Environmental Policy Measure? *Environmental and Resource Economics, 18*(1), 113-127.

Dulleck, U., & Kerschbamer, R. (2006). On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature, 44*(1), 5-42.

Dulleck, U., Kerschbamer, R., & Sutter, M. (2011). The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition. *American Economic Review, 101*(2), 526-555, doi:doi: 10.1257/aer.101.2.526.

Etilé, F., & Teyssier, S. (2016). Signaling Corporate Social Responsibility: Third□Party Certification versus Brands. *The Scandinavian Journal of Economics, 118*(3), 397-432.

EU Commission (2014). *Towards a circular economy: A zero waste programme for Europe* (Vol. COM(2014) 398 final). Brussels,: European Commission.

European Commission (2013). *Flash Eurobarometer 367 - Attitudes of Europeans towards building the single market for green products*: Conducted by TNS Opinion & Social at the request of Directorate-General for the Environment.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171-178.

Hamilton, S. F., & Zilberman, D. (2006). Green markets, eco-certification, and equilibrium fraud. *Journal of Environmental Economics and Management, 52*(3), 627-644.

Holt, C., & Sherman, R. (1990). Advertising and product quality in posted-offer experiments. *Economic Inquiry, 28*(1), 39-56, doi:10.1111/j.1465-7295.1990.tb00802.x.

Karl, H., & Orwat, C. (1999). Economic Aspects of Environmental Labelling. In H. Folmer, & T. Tietenberg (Eds.), *The International Yearbook of Environmental and Resource Economics 1999/2000 - a survey of current issues* (pp. 107-170). Cheltenham: Edward Elgar Publishing.

Kotchen, M. J. (2006). Green Markets and the Private Provision of Public Goods. *Journal of Political Economy, 114*(4), 816-834.

Krarup, S., & Russell, C. S. (2005). *Environment, information and consumer behaviour*. Cheltenham: Edward Elgar Publishing.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *The Journal of Political Economy, 74*(2), 132-157.

Lynch, M., Miller, R. M., Plott, C. R., & Porter, R. (1991). Product quality, informational efficiency, and regulations in experimental markets. *Research in Experimental Economics, 4*, 269-318.

Mason, C. F. (2006). An economic model of ecolabeling. *Environmental Modeling and Assessment, 11*(2), 131-143.

Mason, C. F. (2011). Eco-Labeling and Market Equilibria with Noisy Certification Tests. *Environmental and Resource Economics, 48*(4), 537-560, doi:10.1007/s10640-010-9402-5.

Miller, R. M., & Plott, C. R. (1985). Product quality signaling in experimental markets. *Econometrica: Journal of the Econometric Society*, 837-872.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy, 78*(2), 311-329.

OECD (2001). *Policies to Enhance Sustainable Development* (OECD Environment & Sustainable Development). Paris: OECD.

OECD (2002). Towards sustainable consumption: An economic conceptual framework. *Working Party on National Environmental Policy ENV/EPOC/WPNEP(2001)12/FINAL - OECD Environmental Directorate*. Paris.

Plott, C. R. (1983). Externalities and corrective policies in experimental markets. *The Economic Journal, 93*(369), 106-127.

Rode, J., Hogarth, R. M., & Le Menestrel, M. (2008). Ethical differentiation and market behavior: An experimental approach. *Journal of Economic Behavior & Organization, 66*(2), 265-280.

Roe, B., & Sheldon, I. (2007). Credence Good Labeling: The Efficiency and Distributional Implications of Several Policy Approaches. *American Journal of Agricultural Economics, 89*(4), 1020-1033, doi:10.1111/j.1467-8276.2007.01024.x.

Roe, B., Teisl, M., & Deans, C. (2014). The Economics of Voluntary Versus Mandatory Labels. *Annual Review of Resource Economics, 6*(1), 407-427, doi:10.1146/annurev-resource-100913-012439.

Smith, V. L. (1982). Microeconomic Systems as an Experimental Science. *The American Economic Review, 72*(5), 923-955.

Valente, M. (2015). Ethical Differentiation and Consumption in an Incentivized Market Experiment. *Review of Industrial Organization, 47*(1), 51-69, doi:10.1007/s11151-015-9455-2.

Table 1 Parameters for the green and brown goods (in experimental units)

	green	brown
(-) production cost	50	40
(-) externalities	0	20
(+) benefit	100	100
(=) net social benefit	50	40
(-) certification cost (if applicable)	2	

Table 2 Treatment features

	CREDENCE	THIRD PARTY CERTIFICATION	FULL-INFO
Truthful announcements			
brown good	✓	✓	✓
green good			✓
certified green good		✓	
Certification		✓	
Externalities known with certainty			
brown good	✓	✓	✓
green good	?	?	✓
certified green good	-	✓	-

Legend: "-" not applicable; "✓" present; "?" uncertain from the buyer's perspective

Table 3 Stages within each round

Stage 1:	**Firms' proposals**
All treatments:	firms choose to produce green or brown goods and set a price for the good.
Additionally for TREATMENT CREDENCE and TREATMENT THIRD PARTY CERTIFICATION:	firms decide the type of good to be announced (which may differ from the one being produced).
Additionally for TREATMENT THIRD PARTY CERTIFICATION:	when choosing to produce a green good, firms are asked if they wish to pay for a certification of the good.
Stage 2:	**Market proposals and buyer decision**
All treatments:	proposals of each seller concerning the price and (announced) type of good are revealed. Each buyer can choose from whom to purchase one unit or whether not to make a purchase.
Stage 3:	**Information about market outcomes**
All treatments:	participants are informed about the choices and all transactions made in the market. At this stage no decision has to be made. Producers learn about the price, announced type of good and sales of the other firms, as well as their own. Period and cumulative earnings are calculated and revealed to sellers. Buyers learn about the price, announced type of good and sales of each

producer. Period and cumulative earnings, as well as externalities, are calculated and revealed to buyers as far as is possible in each period.

TREATMENT FULL-INFO: all relevant information is known with certainty and revealed in stage 3 (including externalities imposed).

TREATMENT CREDENCE and TREATMENT THIRD PARTY CERTIFICATION: participants are informed their payoffs may be lower if other buyers bought good B, consequently imposing an external cost on the other buyers in the same group. Externalities known for sure are calculated at the end of each period but actual externalities are only revealed at the end of the experimental session.

Table 4 Sessions overview

Treatments	# groups	# subjects	# sellers	# buyers
CREDENCE	2 sessions *3 groups	36	18	18
THIRD PARTY CERTIFICATION	2 sessions *3 groups	36	18	18
FULL-INFO	2 sessions *4 groups	48	24	24
Total	20	120	60	60

Table 5 Share of green goods in proposals and purchases

Treatment	Actual Proposals	Actual Market share	Perceived proposals	Perceived Market Share
CREDENCE	15.6%	10.3%	69.4%	63.6%
THIRD PARTY CERTIFICATION	29.4%	25.8%	76.4%	76.2%
Only certified greens	25.3%	24.2%		
FULL-INFO	31.9%	18.1%		

Table 6 Social welfare: benchmarks and actual outcomes

TREATMENT CREDENCE	# greens	# browns	net social benefit
benchmark: all greens	360	0	18000
actual market outcome	37	319	14610
TREAT. THIRD PARTY CERTIFICATION	**# greens non-certified**	**# greens certified**	**# browns**	**net social benefit**
benchmark: all greens non-certified	360		0	18000
actual market outcome	6	87	266	15116
TREATMENT FULL-INFO	**# greens**	**# browns**	**net social benefit**
benchmark: all greens	480	0	24000
actual market outcome	87	390	19950

Table 7 Panel data Logit model for the probability of seller selling at least one unit

	TREAT CREDENCE	TREAT THIRD PARTY CERTIFICATION		TREAT FULL-INFO
VARIABLES	(1)	(2)	(3)	(4)
Proposed price	-0.304***	-0.135***	-0.173**	-0.334***
	(0.0434)	(0.0459)	(0.0685)	(0.0549)
Lowest competitors' price	0.332***	0.175***	0.198***	0.331***
	(0.0549)	(0.0491)	(0.0703)	(0.0696)
Green good proposed (1:yes 0:no)	0.913**	-0.0654	-0.485	0.964***
	(0.465)	(0.591)	(0.535)	(0.317)
Other green goods proposed (1:yes 0:no)	-1.005***	-0.416	-0.229	0.00440
	(0.371)	(0.702)	(0.675)	(0.224)
Relative green reputation	-0.0976	0.200***	0.203***	0.175***
	(0.0933)	(0.0391)	(0.0393)	(0.0414)
Certification (1: yes 0:no)			1.612**	
			(0.762)	
Constant	-0.185	-0.576	0.0637	1.006
	(1.484)	(0.946)	(1.000)	(1.467)
Observations	360	360	360	480
Number of unique subjects	18	18	18	24

Notes: dependent variable: dummy variable for seller selling or not in a given period; Robust standard errors in parentheses; levels of statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 8 Random effects panel regression of price paid by each buyer

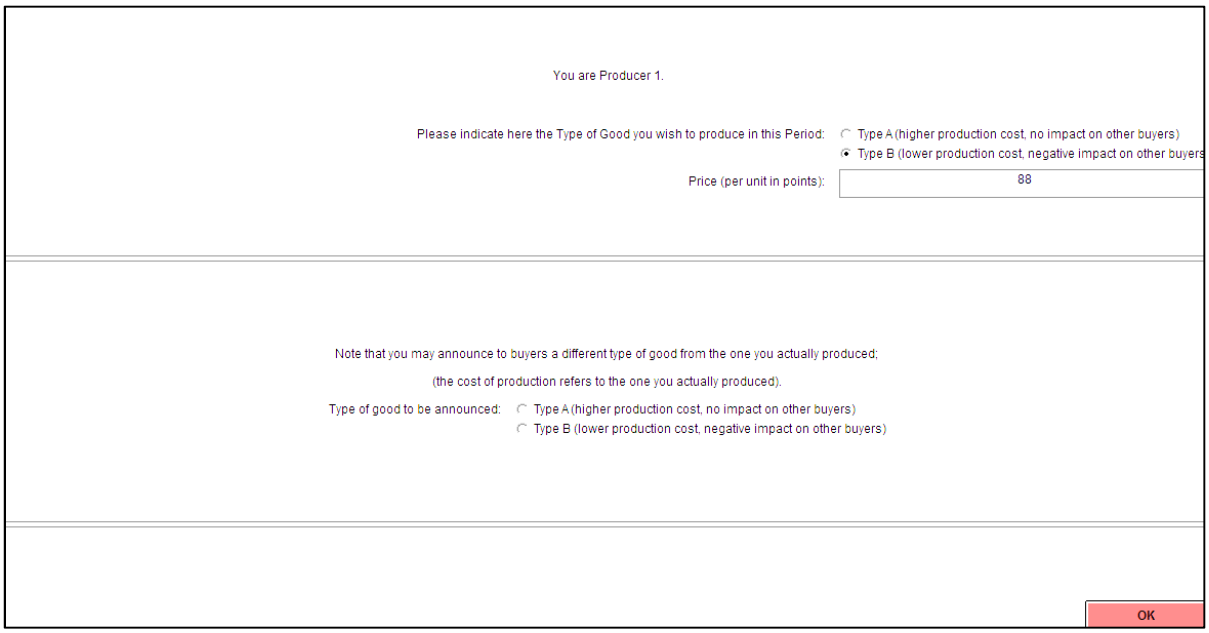
	TREAT CREDENCE	TREAT THIRD PARTY CERTIFICATION	TREAT FULL-INFO
	(1)	(2)	(3)
Green good purchased (1: yes 0: no)	6.278***	5.111***	8.520***
	(0.958)	(1.980)	(1.266)
Green good purchased is certified (1: yes; 0: no)		7.394*	
		(4.346)	
Period	-0.621***	-0.538	-0.299*
	(0.158)	(0.386)	(0.173)
Constant	51.81***	49.90***	47.44***
	(2.542)	(3.417)	(2.294)
Observations	356	359	477
Number of unique subjects	18	18	24

Notes: dependent variable: price paid for a good by a buyer; Standard errors in parentheses clustered by group; levels of statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Appendix

Figure A. Screenshots for the experiment

A1. TREATMENT CREDENCE: Input screens for sellers (Stage 1)



A2. TREATMENT CREDENCE: Input screens for buyers (Stage 2)