

Discussion Paper Series – CRC TR 224

Discussion Paper No. 426 Project B 04

The Market for Ethical Goods

Nicolas Bonneton¹

May 2023

¹ University of Mannheim, Email: nicolas.bonneton@gmail.com

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

The Market for Ethical Goods^{*}

Nicolas Bonneton †

May 5, 2023

Abstract

This paper examines the unintended consequences of the rapid growth of markets for ethically labeled goods such as "organic" or "child-labor-free", using a model of voluntary labeling. On the supply and demand side, agents act based on their intrinsic motivation and monetary payoff. After a positive demand shock, the expected social and environmental attributes of both labeled and unlabeled goods deteriorate. As such, even if producing ethically labeled goods help mitigates externalities, it can be optimal for policymakers not to intervene and sometimes to tax these goods. Sorting in this market is non-trivial: under plausible conditions, only the most and least intrinsically motivated agents strategically position themselves in the ethical segment of the market.

^{*}I wish to thank Emmanuelle Auriol, Roland Bénabou, Sylvain Chabé-Ferret, Marion Desquilbet, Andy Newman, Helena Perrone, Bernard Salanié, François Salanié, Christopher Sandmann, Nicolas Schutz, Jean Tirole, Nicolas Treich and seminar participants at several venues. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 669217 - MARKLIM), and financial support from the German Research Foundation (DFG) through CRC TR 224 (Project B04).

[†]University of Mannheim. E-mail: nicolas.bonneton@gmail.com.

1 Introduction

Growing concerns about sustainability have led to a surge in markets for ethically labeled goods. For instance, the global sales of certified organic products jumped from \$18 billion in 2000 to \$107 billion in 2019 (Willer et al., 2021). In financial markets, the volume of labeled green bonds issued has increased from \$10 billion in 2013 to \$500 billion in 2021 (The Climate Bonds Initiative).¹ Due to their potential role in mitigating externalities, such private initiatives have prompted policy-makers to offer support, such as the European Union's subsidy of 6.3 billion euros for the conversion and maintenance of organic agriculture over the period 2014-2020.

Motivating Evidence. Nonetheless, the rapid growth of markets for ethically labeled goods has had unintended consequences. Surveys reveal a shift in the motivations of those involved in their production, from ethical to monetary (Darnhofer et al. (2010) and Lamine and Bellon (2009)). Figure 1 reports that 73% of U.S. farmers who converted to organic farming before 1990 cited health or environmental reasons as their primary motivation, and only 36% of those who converted after 2000 did so. Instead, newer entrants report being primarily motivated by monetary gain.

Section 2 documents how this trend translates into a change in sustainable production methods among the entire population of French farmers (both organic and conventional). Specifically, I look at farmers' adoption of environmentally friendly and harmful practices, which are not directly regulated by organic certification standards. My findings suggest that early adopters of organic farming engage in more environmentally friendly practices than those who converted after the boom of organic agriculture. However, the latter still appear more environmentally friendly than those who have never converted. For example, among organic pioneers (those who converted before 2000) 6.9% are producing solar energy, compared to 2% among those who converted after 2010 and only 1% among conventional farmers.

Theory. Against this background, the theory developed in this paper focuses on the role of intrinsic motivation in shaping the market for ethically labeled goods, the welfare implications of its growth, and the scope for public intervention. To that end, I model a competitive credence-good market with many producers and consumers. Each producer can

¹Another example: The global sales of certified Fairtrade products grew from \$1 billion to \$9 billion between 2005 and 2017 (Fairtrade Foundation).

differentiate her production process along a vertical (unobservable) environmental or social attribute, which I refer to as quality. If a producer complies with an exogenous quality standard, she is entitled to stamp an ethical label on the goods she sells. At equilibrium, producers choose optimal quality levels; consumers decide whether to purchase a labeled or unlabeled good subject to rational expectations, and the price of the labeled (unlabeled) good is determined by market clearing.

The model includes two key features of markets for ethically labeled products. First, both consumer and producer decisions can be driven by intrinsic motivation in addition to monetary payoff. This contrasts with the existing theories that typically consider that only one side of the market (workers, consumers) is ethically motivated while the other side (firms, shareholders) is solely driven by the profit motive. Secondly, the label is imperfect in that it solely conveys a binary signal, despite a rich heterogeneity among firms' ethical practices.² This modeling specification is consistent with popular ethical labels such as "organic", "fair trade" or "child-labor-free". Labels are typically binary, as attested by the Directory Ecolabel Index: out of the 388 labels listed in the Directory, 291 are binary.

In equilibrium, an endogenous sorting of producers arises. The most intrinsically motivated producers make ethically labeled goods: Some bunch at the regulatory threshold while others, the most "virtuous" ones, overcomply with the labeling standard. Interestingly, practices are also heterogeneous in the unlabeled segment of the market: A fraction of producers conduct their business responsibly whenever it is not too expensive to do so, while the others maximize profit regardless of the human or environmental costs.

A seemingly paradoxical link ties the growth of the ethical segment of the market to the quality of both labeled and unlabeled products. Consider indeed an increase in the financial incentives to produce ethically labeled goods, such as a demand shock, a subsidy, or a green innovation. This improves the average quality, i.e., the overall market's social and environmental performance, as some previously unlabeled producers choose to upgrade their production process to obtain the label. However, due to a composition effect, the quality of both labeled and unlabeled goods deteriorates. First, the producers who adopt ethical practices after a rise in financial incentives are less intrinsically motivated and thus conduct their business less responsibly than the ones who were already producing labeled

 $^{^{2}}$ For a rationale of binary labels as an optimal rating system, see Hopenhayn and Saeedi (2019) and references therein.



Figure 1: Primary motivation for producing organic by conversion date.

Note: The data come from the Agricultural Resource Management Survey about organic production practices and costs. It contains a representative sample of 630 U.S. organic farmers who produce wheat, soybean or corn. In 2007, each farmer was asked: "Did you choose the organic production system used on this field primarily to a) adopt more environmentally friendly practices? b) protect health of family and community? c) increase farm income? d) for some other reason(s)?". The x-axis corresponds to the time at which farmers converted to organic agriculture, and the y-axis represents the percentage of farmers who answered a), b), and c), respectively.

goods. This weakens the expected quality of the labeled products and gives a theoretical foundation for the much-debated concept of "greenwashing": the development of ethical labels mechanically attracts firms that adopt pseudo-ethical conduct. Second, those who shift from unlabeled to labeled production are more intrinsically motivated than those who remain unlabeled. So only the least virtuous producers are left in the unlabeled segment of the market. This further reduces the expected quality of the unlabeled product.

How does such equilibrium sorting affect the optimal subsidy scheme? Although subsidizing the production of ethically labeled goods can enhance welfare by mitigating externalities that are not internalized by market participants, our theory reveals a countervailing force. Consumers' purchasing decisions are based on the expected quality difference between labeled and unlabeled products. Due to the above composition effect, this expected quality difference is strictly higher than the quality difference between marginal producers, which drives the demand for ethically labeled products above its level without informational frictions. Indeed, the marginal consumer would not be willing to buy the labeled good from the marginal producer at the equilibrium price. Overall, there is a trade-off between subsidizing the production of ethically labeled goods to improve the market's social and environmental performance and taxing it to alleviate issues related to information asymmetries. Even if producing ethically labeled goods helps mitigate externalities, it can be optimal for policymakers not to intervene and, when the externality differential is small, to tax these goods.

In the baseline model, I assume that the labeling standard constrains producers to adopt the most cost-efficient production methods. In Section 4, I relax this assumption and allow for less-than-perfect ethical standards. I begin by describing common ethical standards limitations and provide examples of how firms exploit regulatory loopholes. I then extend the model to a multi-task framework in the spirit of Holmstrom and Milgrom (1991). Some producers' actions are observable, while others are not, and the standard regulates only the observable actions. In this more general model, sorting into the production of ethical goods becomes non-trivial. For instance, when observable and non-observable actions are complementary, the most and least intrinsically motivated producers strategically position themselves in the market's ethical segment. The former overcomply with the labeling standard to be responsible and cost-efficient, while the latter engages in misleading labeling, taking advantage of the standard loopholes and being more socially and environmentally harmful than if they were producing unlabeled products. This result sheds light on a key drawback of the market for ethical goods, highlighting that even credible signals of costly moral actions can be deceitful. To ensure efficient sorting, i.e., to guarantee that the most motivated consumers buy from the most motivated producers, I show that the labeling standard must be designed to improve social and environmental performance, regardless of one's motive.

The rest of this section discusses the related literature. Section 2 presents motivating evidence in the context of organic farming. Section 3 lays out the model, analyzes market failures, and discusses public policy. Section 4 studies the sorting of producers when labeling standards are imperfectly designed or incomplete. Section 5 concludes. All proofs are gathered in the Appendix.

Related literature

This paper contributes to the growing literature on corporate social responsibility (CSR) (see Kitzmueller and Shimshack (2012) for a review).³ In my framework, two types of CSR coexist: "strategic CSR", which corresponds to a firm's ethical practices that result from profit maximization, and "not-for-profit CSR", which refers to a firm's ethical practices that are motivated by moral concerns, i.e., at the expense of profit. In contrast to Baron (2009) and most of the literature, I consider that consumers cannot distinguish between these two types of CSR. My results shed light on how strategic CSR free-rides on not-for-profit CSR and why the economic success of CSR goes hand in hand with the advent of firms' pseudo-ethical conduct. Wu et al. (2020) also considers an asymmetry of information between firms and consumers regarding CSR but with a different focus. In their model, a monopoly signals its type (moral value) through its observable ethical investment, and the consumer does not care about the monopoly's actions but solely its type.

Relatedly, this paper contributes to the stakeholder theory. Several contributions, including Tirole (2006), echoing Friedman et al. (1970), argue that allowing managers to adopt broader objectives than profit amplifies agency issues and may result in arbitrariness, waste, and corruption. More recently, Magill et al. (2015) and Fleurbaey and Ponthière (2023) theoretically show that simple management rules could in principle help firms maximize aggregate welfare. In line with the later development in the literature, I abstract from agency issues related to self-serving insider behavior and instead focus on the asymmetry of information between the firm's ethical practices and consumers.

Many firms claim to implement social and environmental initiatives, but it is difficult to determine whether these initiatives result from profit maximization strategies, such as threats from activists, the anticipation of new legal norms and marketing strategies, or if they arise from moral preferences on the supply side.⁴ Fioretti (2022) develops a structural model that quantifies firms' prosocial spending beyond profit maximization and finds that firms engage in prosocial spending beyond profit maximization, which substantially in-

³The European Commission (2002) defines CSR as "a concept whereby companies integrate social and environmental concerns in their business operations and their interaction with their stakeholders voluntarily."

⁴However, it is well established that some consumers are willing to pay for ethical attributes (see Cason and Gangadharan (2002); Teisl et al. (2002); Hainmueller et al. (2014); Lusk et al. (2005)).

creases welfare. Our theory complements Fioretti's article by demonstrating the economic implications of the presence of intrinsically motivated individuals on the supply side.

The idea that higher extrinsic rewards can attract less intrinsically motivated agents has drawn much attention in labor economics, both theoretically (e.g., Handy and Katz (1998), Delfgaauw and Dur (2007), Prendergast (2007)), and more recently empirically (see Deserranno (2019), Ashraf et al. (2019), Barfort et al. (2019), Dal Bó et al. (2013), Friebel et al. (2019)). I contribute to this strand of the literature by showing how such a phenomenon affects market prices and its consequences on sorting in other sectors. In addition, one key lesson of the existing papers is that low monetary incentives are sufficient to screen the most virtuous agents; Section 4 challenges this point.

A growing body of research, including Falk et al. (2020), Bartling et al. (2015), Hamman et al. (2010)), examines how markets affect individual ethical behavior. One key finding is that markets can erode social responsibility. However, less is known about the converse problem: how the presence of agents with socially responsible preferences affects, potentially negatively, the functioning of markets. This paper makes a step in this direction.

The existing literature on ethical labels considers that some consumers are concerned about environmental and social issues, but agents on the supply side are solely driven by the profit motive (e.g. Amacher et al., 2004, Hamilton and Zilberman, 2006, Harbaugh et al., 2011).⁵ More generally, in economics, non-pecuniary preferences are typically modeled only on one side of the market, on the workers' or consumers' side (e.g., Bénabou and Tirole (2016) and Besley and Ghatak (2005)). One novelty of this paper is to model nonpecuniary preferences on both sides of the market. This allows us to study the interplay between consumer and producer intrinsic motivations and their welfare repercussions.

Sociologists were the first to compare the profile of recently converted organic farmers to those of organic pioneers (e.g., Buck et al. (1997) and Guthman (2004)). Qualitative studies in various countries find that most organic pioneers are primarily motivated by

⁵This literature has a different focus, and tackles the following questions: i) how firms strategically adapt to ethical labels - their incentives to invest in green technologies (Amacher et al. (2004)) as well as their incentives to fraud (Hamilton and Zilberman (2006)), ii) what are the differences between for-profit, public and nonprofit certifiers, and what are the welfare implications of their coexistence (Bottega and De Freitas (2009), Fischer and Lyon (2014), Li and van't Veld (2015)), iii) how ethical labels compare with their alternatives such as government regulation (e.g. Heyes and Maxwell (2004)), iv) how the market structure is affected by the introduction of an ethical label (see Bonroy and Constantatos (2015) for a review).

social and environmental concerns. Conversely, recently converted organic farmers appear to be more driven by profit. The latter are more likely to have common characteristics with conventional agriculture, such as mechanized production, vertical integration, and marketing strategies. Sociologists refer to this evolution as the *conventionalization thesis*; see Darnhofer et al. (2010) and Lamine and Bellon (2009) for reviews. Existing studies rely on small samples and self-reported data. In Section 2, I propose a complementary approach, using the adoption of green practices as a proxy for environmental concerns.

2 Motivating evidence

Organic food sales increase annually, far outstripping the growth rate of the overall food market.⁶ This has prompted many farmers to convert from conventional to organic agriculture, significantly increasing the number of organic farms. For example, in France, the number of organic farms has risen from 1,500 in 1995 to 34,000 in 2016 (see Figure 2). In this section, I document that the farmers who recently converted to organic agriculture appear to be less environmentally and socially concerned than the organic pioneers. Yet, they appear to be more environmentally friendly than the farmers who still use conventional production methods.

The data come from three sources: the French agricultural census of 2010, which contains information about production characteristics of the population of French farms (half a million observations); the Agence Bio data set, which includes the date of conversion and the certifier of each of the 36000 French organic farms; and a data set from INSEE, which includes French demographics and weather data at the "département" level (equivalent to a US county).

Indicators of farmers' adherence to organic principle. In contrast to the broad organic *principles*, organic *standards* have a relatively limited scope. The main goal of organic farming is to use the best environmental practices, preserve natural resources, and promote

⁶This rapid growth is mainly due to a gradual change in consumer demand. Consumers have become increasingly skeptical about the safety of conventional foods. A case in point is the crisis over dioxin-contaminated food and livestock diseases (such as Bovine Spongiform Encephalopathy). They also have become increasingly sensitive to environmental issues. Furthermore, the supply has been fostered by the development of organic production methods and subsidy programs (as the MAE 21 in the European Union).



Figure 2: Evolution of the number of French organic farms over time

animal as well as human wellbeing. However, organic standards mainly constrain farmers not to use chemical pesticides, synthetic fertilizers, and genetically modified organisms and limit the use of antibiotics for livestock.⁷

To get a sense of farmers' adherence to organic principles, I look at all the practices that are at the core of organic principles, which are available in the census but are not regulated by organic standards:⁸ *i*) producing renewable (solar or wind) energy, *ii*) beekeeping and *iii*) selling locally. Producing renewable energy helps combat greenhouse gas emissions and reduces our collective dependence on fossil fuels. As explained by Ashman et al. (2004), beekeeping plays an important role in creating and conserving biodiversity.⁹ Finally, selling locally preserves the environment as it reduces waste from packaging and reduces the fossil fuels used for transportation. In the same vein, I consider two practices that are antagonistic to organic principles but are not regulated by organic standards: *i*) force-feeding

⁸For more information regarding the census, refer to the questionnaire:

⁷A definition of organic principles can be found in Regulation (EC) No 834/2007 of the European Union, or on the website of the IFOAM (https://www.ifoam.bio/ principles-organic-agriculture-brochure). For more details about organic standards in the European Union, see (EC) No 889/2008.

http://agreste.agriculture.gouv.fr/IMG/pdf_questionnairemetropole.pdf

⁹The Food and Agriculture Organization of the United Nations (FAO) estimates that of the 100 crop species that provide 90% of food worldwide, 71 are pollinated by bees. According to the European food safety authority, over the past 15 years, there has been a severe loss of honey bees, making preservation a key environmental challenge.

animals and *ii*) producing tobacco.¹⁰ An extensive body of scientific evidence confirms that the practice of force-feeding is detrimental to animal welfare.¹¹ With respect to tobacco, according to Eurostat, smoking tobacco is the largest avoidable health risk in the EU, and its consequences are a major burden on healthcare systems.

Relationship between the date of conversion and intrinsic motivation. Table 1 reports, for the year 2010, the percentage of farmers who use the practices discussed above. The sample is divided into four groups: 1) the conventional farmers, 2) the farmers who converted to organic agriculture before 2000 ("organic pioneers"), 3) those who converted between 2000-2010, and 4) those who converted after 2010. Observe that organic pioneers are more likely to adopt practices in line with organic principles than those who recently converted to organic farming. Recent adopters are, however, more likely to adopt environmentally friendly practices than conventional farmers; the converse is true regarding "harmful" practices. For instance, 6.9% of organic pioneers are producing solar energy, while they are 2% among those who converted after 2010 and only 1% among conventional farmers. Likewise, organic pioneers trade on average 27% of their production locally; this percentage falls to 14% among the most recently converted farmers and to 7% among conventional farmers. It is noteworthy that this pattern holds for all indicators.

¹⁰Although the EU organic standards ban the practice of force-feeding, a farm can produce at the same time organic foods and non-organic meat using force-feeding.

¹¹In particular, force-feeding birds causes their liver to become diseased and swollen, pain and injury from feeding tube insertion, and increased mortality. For more detail, see the report of the European Union's Scientific Committee on Animal Health and Animal Welfare on Welfare Aspects of the Production of Foie Gras in Ducks and Geese, adopted on December 16, 1998.

% of farmers that:		Never converted		
	before 2000	2000-10	after 2010	
produce solar energy	6.7	3.6	2.0	1.0
produce wind energy	0.15	0.12	0.08	0.01
trade locally	27.1	19.6	13.9	7.2
engage in beekeeping	5.8	4.3	3.8	2.4
force-feed animals	1.7	3.8	3.7	3.8
produce tobacco	0.0	0.3	0.4	0.4
Number of farms	3300	10788	18604	500994

Table 1: Percentage of farmers adopting environmentally friendly/harmful practices as a function of the conversion date to organic farming

Controlling for farms and farmers' characteristics. Although these descriptive statistics exhibit a clear pattern, confounding variables unrelated to intrinsic motivations may partially explain why organic pioneers are more or less likely to adopt these practices. For instance, abundant sunshine can give a comparative advantage to producing solar energy and converting to organic farming (because the sun may reduce the need for fertilizers). Furthermore, the proximity of a densely populated and wealthy area can jointly raise the demand for local trade and organic products.¹² To account for these confounding factors, I now include a rich set of controls for observable farms and farmers' characteristics. I run, for each indicator, the following regression:

$$indicator_i = \alpha_0 + \alpha_1 organic_i + \alpha_2 date \ of \ conversion_i + \alpha_3 controls_i + \epsilon_i, \tag{1}$$

where $organic_i = 1$ if the farmer *i* produced organic food in 2016 and $organic_i = 0$ otherwise, and *date of conversion_i* is the year in which the farmer *i* converted to organic farming if she converted (1990 being normalized to 1). The controls include farmer characteristics (age, number of years of experience, gender, education), farm characteristics (size, kinds of animals raised, and plants cultivated), department fixed effect, weather (sun, rain), and demographics (population, average wage).

¹²Furthermore, farm's and farmer's characteristics differ significantly across the groups considered in table 1. Organic pioneers are relatively less likely to grow field crops, tend to be older, more educated, and, surprisingly, own larger farms than others (see Table 3 for summary statistics).

Table 2 presents the results of the OLS regression for the dependent variable trading locally, defined as the percentage of total production traded locally. The results indicate a positive and statistically significant relationship between organic farming and trading locally and a negative relationship between trading locally and the conversion date to organic farming. Including controls has little effect on the coefficients or standard deviations of the independent variables, suggesting that the correlations are robust to confounding variables. The tables for the remaining indicators are deferred to Appendix C; they all exhibit the same pattern.¹³

Dependant variable: Local	(1)	(2)	(3)	(4)
Organic	27.2***	27.1***	25.0***	24.7***
	(0.45)	(0.45)	(0.42)	(0.41)
Date of conversion	-0.51***	-0.56***	-0.52***	-0.53***
	(0.02)	(0.02)	(0.03)	(0.02)
Farmers' characteristics		yes	yes	yes
Farm characteristics			yes	yes
Location/weather				yes
Observations: 479,906				

Table 2: Trading locally and conversion to organic farming

Note: OLS regressions with constant; robust standard errors are in parentheses; p<0.1; p<0.05; p<0.05; p<0.01. The exhaustive list of controls is in Table 4. The dependent variable (Local) represents the percentage of the total production that is traded locally.

Discussion. This empirical exercise suggests that when producers' intrinsic motivations decrease, their adoption of sustainable production methods also decreases. Although some organic consumers prefer buying from producers who prioritize animal welfare, produce renewable energy, and sell locally, it should be noted that a change in producers' motivation can have more daunting consequences. For instance, intrinsic motivation can affect organic

¹³One exception is the coefficient of the conversion date for the force-feeding indicator, which is not statistically significant when all controls are included.

fraud and the extent to which farmers exploit regulatory loopholes common in the industry, such as replacing synthetic pesticides with their natural substitutes (as discussed in Section 4).

3 Baseline model

I study a competitive market with many producers and consumers. Each producer can vertically differentiate their production process according to (unobservable) environmental and social attributes, which I refer to as quality. If a producer complies with an exogenous quality standard, she is entitled to stamp an ethical label on the goods she sells. At equilibrium, producers choose optimal quality levels, consumers decide whether to purchase a labeled or unlabeled good subject to rational expectations, and the price of the labeled (unlabeled) good is determined by market clearing.

3.1 Supply

Consider a continuum of producers of mass one; each of them produces one unit of a good and takes prices as given. I denote by θ_p a producer's level of intrinsic motivation for producing ethically. The type θ_p is private information and is drawn from a cumulative distribution function F_p (density denoted f_p) on the support [0, 1]; F_p is differentiable and admits an increasing hazard rate.

Each producer selects a quality level $v \in \mathbb{R}^+$ which entails a monetary (or utility) quadratic cost $C(v) = \frac{1}{2}v^2$, and an internal reward $\theta_p v$. The quality level v represents the environmental and social benefits achieved during the production process. It can not be ascertained ex-ante nor ex-post by consumers (i.e., it is a credence attribute). Each producer can signal quality through a label to alleviate this information asymmetry. The labeling process is as follows: a good is labeled if and only if its quality v is greater than an exogenous threshold $\underline{v} \in [0, 1]$. For tractability reasons, we focus the analysis on \underline{v} that are not too large.¹⁴ Even though the good's quality is chosen from a continuum, the label solely conveys a binary signal: certified or not. To simplify the analysis, I also assume that the labeling process is implemented, for free, by a trustworthy third-party certifier.

¹⁴To prove equilibrium uniqueness we will use that \underline{v} is in $\underline{v} \in [0, v^*]$, where v^* is the solution to $F_c^{-1}(F_p(v^*))\frac{f_p(v^*)}{1-F_p(v^*)} = 1.$

The producers of the labeled good are financially rewarded. They earn a price premium, denoted \hat{p} , which represents the price difference between the labeled good and the unlabeled one. The price premium will be endogenously determined by market forces. Without loss of generality, the costs and benefits common to producing labeled and unlabeled goods are set to zero. The objective of a producer of type θ_p amounts to the following:

$$\begin{array}{ccc} \theta_p & \underbrace{1}_{v \in \mathbb{R}^+} & \hat{p} + \theta_p v - C(v) & \text{ subject to } \{v \geq \underline{v}\} \\ & & & \\$$

If a producer of type θ_p chooses to produce the unlabeled good, then he selects the quality level $v = \theta_p$ that satisfies his first-order condition. Similarly, a producer who chooses to produce the labeled good selects a quality level $v = \max\{\underline{v}, \theta_p\}$. Thus, among the producers of the labeled good, the most intrinsically motivated overcomply with the standard while the others bunch at the regulatory threshold. Equalizing the payoffs associated with the production of the labeled and unlabeled good, it is immediate that there exists a unique cut-off type, denoted $\hat{\theta}_p$, such that a producer chooses to produce the labeled good if and only if his type is greater than $\hat{\theta}_p$. This feature of the model is consistent with empirical observations. For instance, as documented in Section 2, organic farmers tend to be more environmentally friendly than conventional ones.

The cut-off type $\hat{\theta}_p$ is a solution to the following equation:

$$\hat{p} = \frac{(\underline{v} - \hat{\theta}_p)^2}{2}.$$
(2)

Price and quality. Although the price premium is in equilibrium endogenously determined by market forces, it is insightful first to study how producers react to an exogenous change in the price premium.

Proposition 1 An exogenous increase in the price premium earned by labeled producers increases the average quality of the goods produced but reduces the average quality of both labeled and unlabeled goods. Formally, $\frac{\partial \mathbb{E}[v]}{\partial \hat{p}} \geq 0$, $\frac{\partial \mathbb{E}[v|l]}{\partial \hat{p}} \leq 0$, $\frac{\partial \mathbb{E}[v|u]}{\partial \hat{p}} \leq 0$.

Proof: See Appendix A.1.

To get the intuition of this seemingly paradoxical result, consider an increase of the price premium from \hat{p} to \hat{p}' , and let $\hat{\theta}_p$ and $\hat{\theta}'_p$ be the respective producer cut-off types. As

illustrated in Figure 3, the behavior of each producer with a type $\theta_p \in [0, \hat{\theta}'_p[\cup]\hat{\theta}_p, 1]$ remains unchanged. However, all producers with a type $\theta_p \in [\hat{\theta}'_p, \hat{\theta}_p]$ upgrade their quality from $v = \theta_p$ to $v = \underline{v}$ and shift from being unlabeled to labeled. This upgrade in quality leads to an increase in the average quality in the aggregate (represented by the hatched area in Figure 3). However, a composition effect reduces the average quality level of both labeled and unlabeled goods. Indeed, the new producers of the labeled good are less intrinsically motivated, bunch at the regulatory threshold, producing lower quality goods than the original producers of the labeled good. This implies that the average quality of the labeled good have shifted towards producing the labeled good, only the least motivated producers are left producing the unlabeled good, and so the average quality of the unlabeled good also goes down.



Figure 3: How the price premium affects the quality

The price premium, \hat{p} , is the difference between the price of the labeled good, p_l , and the price of the unlabeled one, p_u . Proposition 1 entails that there is a positive relationship between the price of an unlabeled product and its quality $\left(\frac{\partial \mathbb{E}[v|u]}{\partial p_u} \geq 0\right)$. This means that adverse selection, as in Akerlof (1970), occurs in the unlabeled segment of the market. Indeed, as in the market for lemons, due to a selection effect in the supply, an increase in price leads to a rise in the average quality of the goods traded. By contrast, an inverse relationship ties together the price and the quality of labeled products. Therefore, advantageous selection, as in Hemenway et al. (1990), occurs in the ethically labeled segment of the market.

Proposition 1 stems from producers' intrinsic motivation. If, as often assumed in the

literature, all producers were solely driven by profits, (i.e., $\theta_p = 0$ for all producers) they would bunch at the regulatory thresholds, $v = \underline{v}$ or v = 0, to maximize their profit. Thus, a change in price premium \hat{p} would not affect the two quality levels: $\mathbb{E}[v|l] = \underline{v}$ and $\mathbb{E}[v|u] = 0$.

3.2 Demand

There is a continuum of consumers of mass one, each consuming one unit of a good. I model the behavioral components of consumer preference as I have done for producers. I denote by θ_c a consumer's level of intrinsic motivation or ethical concern. I assume that the type θ_c is private information and is drawn from a differentiable cumulative distribution function $F_c(\theta_c)$.

The environmental and social features of the production process are credence attributes. Thus, the consumers cannot observe the quality of the goods they purchase. They, however, form rational expectations about the quality depending on whether the good is labeled or not. Let $\theta_c \mathbb{E}[v|l]$ be the benefit from consuming the labeled good where $\mathbb{E}[v|l]$ is its expected quality. A consumer of type θ_c chooses one of the following two options:

$$\theta_c \xrightarrow{l} \theta_c \mathbb{E}[v|l] - \hat{p}$$

$$u \to \theta_c \mathbb{E}[v|u]$$

As with the producers' problem, there exists a unique type, denoted $\hat{\theta}_c$, such that a consumer purchases the labeled good if and only if his type is greater than $\hat{\theta}_c$. The type $\hat{\theta}_c$ is a solution to the following equation:

$$\hat{\theta}_c(\mathbb{E}[v|l] - \mathbb{E}[v|u]) = \hat{p}.$$
(3)

3.3 Market equilibrium

At equilibrium, producers choose optimal quality levels, and consumers decide on whether to purchase the labeled good subject to rational expectations. Crucially, the price premium is such that the demand for the labeled good equals its supply. Each competitive equilibrium is then a solution of the following system:

$$\hat{p} = \frac{(\underline{v} - \theta_p)^2}{2} \tag{4a}$$

$$\hat{\theta}_c(\mathbb{E}[v|l] - \mathbb{E}[v|u]) = \hat{p} \tag{4b}$$

$$F_p(\hat{\theta}_p) = F_c(\hat{\theta}_c) \tag{4c}$$

$$\mathbb{E}[v|u] = \frac{1}{F_p(\hat{\theta}_p)} \int_0^{\hat{\theta}_p} \theta_p \, dF_p(\theta_p) \tag{4d}$$

$$\mathbb{E}[v|l] = \frac{1}{1 - F_p(\hat{\theta}_p)} \int_{\hat{\theta}_p}^1 \max\{\underline{v}, \theta_p\} \, dF_p(\theta_p) \tag{4e}$$

As shown previously, equations (4a) and (4b) define the producer and consumer cut-off types. Equation (4c) ensures that supply equals demand. Indeed, $F_p(\hat{\theta}_p)$ represents the total quantity of unlabeled goods produced, and $F_c(\hat{\theta}_c)$ represents the total quantity of unlabeled goods consumed. Finally, equations (4d) and (4e) give the relationships between the expected level of qualities and the producer's cut-off type.

Proposition 2 Suppose f_p is non-increasing. Then a competitive equilibrium $(\hat{\theta}_c, \hat{\theta}_p, \hat{p}, \hat{p}, \mathbb{E}[v|u], \mathbb{E}[v|l])$ defined by (4a), (4b), (4c), (4d) and (4e) exists and is unique.

Proof: See Appendix A.2.

To understand why assuming that f_p is non-increasing eases our analysis, recall that the distribution of producers' intrinsic motivation level determines quality levels through production choices. As shown by Proposition 1, an increase in the number of labeled goods produced (a reduction in $\hat{\theta}_p$) leads to the deterioration of the average quality of both labeled and unlabeled goods. Depending on the magnitude of the two effects, the quality difference between both goods, $\mathbb{E}[v|l] - \mathbb{E}[v|u]$, can be positively or negatively affected by a change in the cut-off type, and equation (4b) may admit multiple solutions. To gain insight into how the shape of the underlying distribution affects the quality difference, we can decompose it as follows:

$$\mathbb{E}[v|l] - \mathbb{E}[v|u] = \frac{1}{1 - F_p(\hat{\theta}_p)} \int_{\hat{\theta}_p}^1 \max\{\underline{v}, x\} dF_p(x) - \frac{1}{F_p(\hat{\theta}_p)} \int_0^{\hat{\theta}_p} x dF_p(x)$$
$$= \underbrace{E[\theta_p|\theta_p > \hat{\theta}_p] - E[\theta_p|\theta_p \le \hat{\theta}_p]}_{\Delta_p(\hat{\theta}_p)} + \frac{1}{1 - F_p(\hat{\theta}_p)} \int_{\hat{\theta}_p}^{\underline{v}} (\underline{v} - x) dF_p(x)$$
(5)

The difference in expectation, $\Delta_p(\hat{\theta}_p)$, is a known object from the signaling literature; its shape is characterized by the following lemma, by Jewitt (2004): If $f_p(\theta_p)$ is everywhere decreasing (increasing), then $\Delta_p(\hat{\theta}_p)$ is everywhere increasing (decreasing). If $f_p(\theta_p)$ has a unique interior maximum, then $\Delta_p(\hat{\theta}_p)$ has a unique interior minimum. This lemma sheds light on how changes in market shares can impact the quality differences between labeled and unlabeled goods.¹⁵ When the density is non-increasing, then a change in demand or subsidy, will impact relatively more the quality of the labeled good than the quality of the unlabeled good (through the signaling channel).

The second term on the right-hand side of (5) accounts for the producers who bunch at the regulatory threshold \underline{v} . This term is decreasing with respect to $\hat{\theta}_p$.¹⁶ Intuitively, the fact that a fraction of producers of the labeled good bunch at the regulatory threshold makes the average quality of labeled relatively less sensitive to change in demand or subsidy. In particular, if \underline{v} would be so high all producers of the labeled good choose to bunch at the regulatory threshold, then a change in market structure would only affect the quality of unlabeled goods.¹⁸

If f_p is increasing or admits a unique interior maximum, multiple equilibria can arise. For instance, the following two equilibria may coexist. In the first equilibrium E_1 , few labeled goods are produced, their qualities are alike, and only the most intrinsically moti-

¹⁵This holds because I have assumed that the cost function is quadratic. For the general convex cost function, the unconstrained quality level chosen by a producer of type θ_p is $v = C'^{-1}(\theta)$. Note that C'^{-1} is increasing because C is convex. Thus, I obtain the following equality:

$$\mathbb{E}[C'^{-1}(\theta_p)|\theta_p > \hat{\theta}_p] - \mathbb{E}[C'^{-1}(\theta_p)|\theta_p < \hat{\theta}_p] = \mathbb{E}[C'^{-1}(\theta_p)|C'^{-1}(\theta_p) > C'^{-1}(\hat{\theta}_p)] - \mathbb{E}[C'^{-1}(\theta_p)|C'^{-1}(\theta_p) < C'^{-1}(\hat{\theta}_p)].$$

As a result, one can apply Lemma Jewitt (2004) for the random variable $C'^{-1}(\theta_p)$.

$$\begin{aligned} \frac{\partial \frac{1}{1-F_{p}(\hat{\theta}_{p})} \int_{\hat{\theta}_{p}}^{\underline{v}} (\underline{v}-x) dF_{p}(x)}{\partial \hat{\theta}_{p}} &= \frac{f_{p}(\hat{\theta}_{p})}{(1-F_{p}(\hat{\theta}_{p}))^{2}} (\int_{\hat{\theta}_{p}}^{\underline{v}} (\underline{v}-x) dF_{p}(x) - (1-F_{p}(\hat{\theta}_{p}))(\underline{v}-\hat{\theta}_{p})) \\ &= -\frac{f_{p}(\hat{\theta}_{p})}{(1-F_{p}(\hat{\theta}_{p}))^{2}} \int_{\hat{\theta}_{p}}^{\underline{v}} 1 - F_{p}(x) dx \quad < 0.^{17} \end{aligned}$$

The second equality is obtained by integrating by parts:

$$\int_{\hat{\theta}_p}^{\underline{v}} (\underline{v} - x) dF_p(x) = (\underline{v} - x) F_p(x) |_{\hat{\theta}_p}^{\underline{v}} + \int_{\hat{\theta}_p}^{\underline{v}} F_p(x) dx = \int_{\hat{\theta}_p}^{\underline{v}} F_p(x) - F_p(\hat{\theta}_p) dx.$$

¹⁸This is for this reason we have assumed that the standard is not too high.

vated consumers buy labeled goods. By contrast, in the second equilibrium E_2 , the quality difference between both goods is high; thus, most consumers are willing to buy ethically. As the market shares of ethically labeled goods are usually small, E_1 can be considered the most relevant equilibrium. Furthermore, since the supply crosses the demand from above only in the second equilibrium, the equilibrium E_1 is stable while E_2 is not.¹⁹

From now on, we restrain our attention to cases in which the distribution of producers' level of intrinsic motivation admits a non-increasing density. This assumption entails that most producers primarily focus on profit maximization, while only a few are primarily motivated by social and environmental motives. It is arguably realistic in our context.

Market failures. In what follows, I show how the equilibrium interplay between sorting and quality affects the efficient functioning of the market for ethical goods. To that end, I compare market outcomes with the constrained efficient allocation. In such an allocation, the social planner chooses the percentage of producers (consumers) that make (consume) labeled goods as to maximize the aggregate surplus. It is characterized by two cut-off types $(\hat{\theta}_p^*, \hat{\theta}_c^*)$ that solve the following maximization problem:

$$(\hat{\theta}_p^*, \hat{\theta}_c^*) = \arg \max \{CS + PS + \alpha \mathbb{E}[v]\}$$
 subject to $(4c) - (4d) - (4e)$,

where CS and PS are the consumers and the producers aggregate surplus and $\alpha \mathbb{E}[v]$ accounts for the social and environmental impacts that are not taken into account by market participants. Consumers can promote values not shared or only partially shared by the social planner. For example, in laic countries, the welfare of religious people is considered (in the welfare function), but the presence of religious practice is not valued by itself. One can think of "halal" and "kosher" labels ($\alpha = 0$). In addition, when the externality occurs in another country, as is the case for "fair trade" and "child-labor-free", the government may internalize only partially the benefits of ethically labeled production (α small). The expression of CS and PS are as follow:

$$CS = \int_0^{\hat{\theta}_c} \theta \mathbb{E}[v|u] dF_c(\theta) + \int_{\hat{\theta}_c}^1 (\theta \mathbb{E}[v|l] - \hat{p}) dF_c(\theta),$$

$$PS = \int_0^{\hat{\theta}_p} (\frac{\theta^2}{2}) dF_p(\theta) + \int_{\hat{\theta}_p}^1 (\theta \max\{\theta, \underline{v}\} + \hat{p} - \frac{1}{2} \max\{\theta, \underline{v}\}^2) dF_p(\theta).$$

¹⁹In addition to a lack of stability in the equilibrium E_2 , there is counterintuitive comparative statics. For instance, if the demand goes up, then quantities traded fall. The next proposition compares the market equilibrium with the constrained efficient allocation.

Proposition 3 Suppose that there is no externality, $\alpha = 0$. Then the equilibrium market share of ethically labeled products exceeds its optimal level, that is, $1 - F_p(\hat{\theta}_p) > 1 - F_p(\hat{\theta}_p^*)$.

Proof: See Appendix A.3.

Consumers are not able to observe the environmental and social features associated with the goods offered on the market. Therefore, their purchasing decisions must be based on the expected quality difference between the labeled and the unlabeled goods. This creates a market inefficiency because such expected quality difference is strictly higher than the quality difference evaluated among marginal producers, that is, $\mathbb{E}[v|l] - \mathbb{E}[v|u] > \underline{v} - \hat{\theta}_p$. Indeed, as some intrinsically motivated producers overcomply with the standard, the expected quality of the labeled good is necessarily higher than the quality of the labeled good produced by the indifferent producers ($\mathbb{E}[v|l] > \underline{v}$). In the same vein, the expected quality of the unlabeled good is lower than the quality of the unlabeled good produced by the marginal producers ($\mathbb{E}[v|u] < \hat{\theta}_p$). This discrepancy between expected and marginal quality levels explains why, when there is no externality, the market share of ethically labeled products exceeds its optimal level in equilibrium. Put differently, the marginal consumer would not be willing to buy the labeled good from the marginal producer at the equilibrium price.

To better understand the forces at play, let's consider the total derivative of the aggregate surplus with respect to the producer cut-off type, evaluated at the equilibrium.²⁰

$$\frac{dW}{d\hat{\theta}_p} = \underbrace{\int_0^{\theta_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x)}_{\text{adverse selection (+)}} + \underbrace{\int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x)}_{\text{advantageous selection (+)}} + \underbrace{\alpha f_p(\hat{\theta}_p)(\hat{\theta}_p - \underline{v})}_{\text{externality (-)}}$$

The two first terms on the right-hand side are strictly positive (Proposition 1). Hence, there are therefore two forces that push the market share of ethically labeled products above its optimal level: advantageous selection in the labeled segment of the market, adverse selection in its unlabeled segment. It implies that, when there is no externality, a small increase in the producer cut-off type (or equivalently, a reduction in the market share of ethically labeled products) is socially desirable.

²⁰The computations are detailed in the Appendix A.3.

When there is no externality, $\alpha = 0$, Proposition 3 has straightforward implications regarding other market outcomes:

- (i) The quality of both labeled and unlabeled goods are inferior to their optimal level,
- (ii) The price premium earned by labeled producers is superior to its optimal level.

3.4 Public policy

An increase in the financial incentives to produce ethically labeled goods improves the overall market's social and environmental performances $\left(\frac{\partial \mathbb{E}[v]}{\partial \hat{\rho}} \ge 0\right)$, Proposition 1). Hence, it is unsurprising that subsidizing ethical labels has become a popular policy. A case in point is the Europe Union subsidy for conversion to organic farming. In 2016, 56 % of European Union organic land was granted a subsidy, receiving on average EUR 214/ha, for a cumulative amount of EUR 1.5 billion.²¹ In this section, I show that the subsidization of ethical labels has major drawbacks.

The following proposition summarizes the main policy lessons.

Proposition 4 The constrained efficient allocation can be decentralized by setting a (unique) per unit subsidy, $s^* \in \mathbb{R}$, on the labeled good. The optimal subsidy is strictly smaller than the Pigouvian subsidy, that is, $s^* < \alpha(\underline{v} - \hat{\theta}_p)$. If the externality, α , is small enough, then it is optimal to tax the labeled good, $s^* < 0$.

Proof: See Appendix A.4.

To make the marginal producers internalize their entire social and environmental impacts, the social planner must give a Pigouvian subsidy, $s = \alpha(\underline{v} - \hat{\theta}_p)$, to each producer of the labeled good (as proven in Appendix A.4). If this subsidy is implemented, the total derivative of the aggregate surplus with respect to the producer cut-off type, evaluated at the equilibrium, becomes:

$$\frac{dW(s = \alpha(\underline{v} - \hat{\theta}_p))}{d\hat{\theta}_p} = \underbrace{\int_0^{\hat{\theta}_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x)}_{\text{adverse selection (+)}} + \underbrace{\int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x)}_{\text{advantageous selection (+)}} > 0.$$

²¹Refer to European Union agricultural markets briefs No 13 of March 2019 for more details.

Although the Pigouvian subsidy effectively removes the externality channel, this equation shows that the two other forces remain. As a result, when the Pigouvian subsidy is implemented, a reduction in the market share of ethically labeled goods is socially desirable. Thus, the amount of the Pigouvian subsidy exceeds its optimal level. Overall, there is a trade-off between taxing the producers of the labeled good to alleviate adverse selection and advantageous selection and subsidizing them to improve the overall market's social and environmental performances. Depending on which effects dominate, it can be optimal to tax or subsidize the production of ethically labeled goods.

From an empirical perspective, to optimally set the subsidy level, one must estimate how much a producer attracted by a subsidy would improve her social and environmental impact. For the case of organic farming, the European Union could rely upon the large body of literature that estimates the environmental impact difference between organic and conventional agriculture.²² This corresponds to an empirical estimations of $\alpha(\mathbb{E}[v|l] - \mathbb{E}[v|u])$ from the model. A key insight is that this observable difference constitutes an upper bound of the social and environmental improvements that the producers attracted by a subsidy would achieve; indeed, as shown previously, $\alpha(\mathbb{E}[v|l] - \mathbb{E}[v|u]) > \alpha(\underline{v} - \hat{\theta}_p)$. As a result, econometricians need to account for equilibrium sorting to avoid overestimating the positive impacts of financial incentives on socially and environmentally friendly technology adoption. Furthermore, Proposition 3 shows that even accounting for positive bias, a Pigouvian subsidy exceeds the optimal level of subsidization.

4 Imperfect and incomplete standards

I have so far considered a unidimensional quality scale, v, and its associated convex cost function, C(v). Behind this modeling specification is the underlying assumption that the standard, \underline{v} , constrains the producers of the labeled good to adopt the most cost-efficient ethical production methods. In reality, this is often not the case.

Numerous headlines and academic papers have questioned the effectiveness of ethical labels. For example, in December 2016, the French magazine "60 Millions de consommateurs" published an issue on organic salmon, showing that metal contamination (mercury and arsenic) was higher for fresh organic salmon than for conventional salmon. One reason

²²See Seufert et al. (2012); van Elsen (2000); Crozier et al. (2010); Bengtsson et al. (2005).

for the imperfections of ethical labels is that certifiers' incentives are not always aligned with consumer preferences (see Bizzotto and Harstad (2020)). Additionally, the complexity of ethical standards is often constrained by consumer awareness. Research shows that consumers do not fully comprehend the meaning of ethical labels, such as the "organic" label. For example, Padel and Foster (2005) and Aarset et al. (2004) found that many consumers do not know what organic means or what benefits they can expect from organic products.²³ As a result, to simplify shopping, consumers need a unique and easy-to-remember standard, which significantly limits the scope of standard setting.

When ethical standards are imperfectly designed, firms can find creative ways to bypass the label requirement by exploiting legal loopholes. For example, the organic standards ban the use of chemical inputs. However, Allen and Kovach (2000) and Rosset and Altieri (1997) have shown that some organic farmers strategically replace chemical inputs by biological substitutes. This practice significantly reduces the environmental benefits of not using chemical inputs. More specifically, Crozier et al. (2010) finds that organic wine producers use an excessive amount of copper (approximately 6 kilos per hectare) as a substitute to synthetic pesticides and fertilizers. It harms the biomass of the land and is toxic to worms. Another case in point is standards for "child-labor-free" labels. The literature on the subject shows that the introduction of a "child-labor-free" label has an ambiguous impact on total welfare and even on the child-worker's welfare (see Baland and Duprez (2009), Chakrabarty and Grote (2009), and Basu et al. (2006)). Indeed, a displacement effect occurs: adult workers replace children in the labeled export sector, whereas children replace adults in the non-labeled industries, where generally working conditions are worse.

In this section, I study the sorting of producers when labeling standards are imperfectly designed or incomplete. To that end, I extend the model to a multi-task framework in the spirit of Holmstrom and Milgrom (1991) and Bénabou and Tirole (2016). Now, each producer chooses a bidimensional level of quality $v = (v_1, v_2) \in \mathbb{R}^2_+$; v_1 represents the contractible aspects of the quality, and v_2 its non-contractible aspects. Choosing a quality v entails a monetary (or utility) cost $C(v_1, v_2) = \frac{(v_1+v_2)^2}{2}$, and an internal reward $\theta_p V(v_1, v_2)$ where $V(v_1, v_2)$ is the social and environmental benefits achieved during the production process (assumed to be increasing and concave with respect to v_1 and v_2). Finally, the good

²³Furthermore, consumers often confuse organic with other labels, such as "free-range" or "pesticide safe" (Roitner-Schobesberger et al. (2008); Harper and Makatouni (2002)).

is labeled if and only if the contractible aspect of the quality v_1 is higher than an exogenous standard $\underline{v_1}$. Note that investing in non-contractible aspects of the quality can only come from ethical motives.

A producer of type θ_p now maximizes the following objective:

$$\theta_p \xrightarrow[(v_1,v_2) \in \mathbb{R}^2_+]{} \theta_p V(v_1,v_2) - C(v_1,v_2) \quad \text{subject to} \quad \{v_1 \ge \underline{v}_1\}$$

I denote $v_l(\theta_p)$ and $v_c(\theta_p)$ the bidimensional vector of quality chosen by a producer of type θ_p , conditional on producing the labeled good and unlabeled one respectively. The following proposition establishes a sufficient condition which ensures that ethically labeled products are made by the most ethical producers.

Proposition 5 If for all $\theta_p \in [0,1]$, $V(v_l(\theta_p)) > V(v_u(\theta_p))$, then a producer chooses to produce the ethically labeled good if and only if his type is greater than a cut-off type $\hat{\theta}_p$.

Proof: See Appendix A.5

In words, $V(v_l(\theta_p)) > V(v_u(\theta_p))$ means that a producer of type θ_p would have a better social and environmental impact by producing the ethically labeled good, rather than by producing the unlabeled good. If that is true for all types, then one can rewrite the equilibrium as in the baseline model and show that all the previous results hold.

If the sufficient condition from Proposition 5 is not satisfied, the sorting of producers becomes more complex and possibly highly inefficient. To illustrate this claim, I study two cases. To simplify the analysis I assume in what follows that the distributions of types are drawn from a uniform distribution.

4.1 Loopholes in ethical standards

Let V take a Leontief form $V = \min\{v_1, v_2\}$. This specification captures that producing ethically may require more than just complying with the standard. For instance, $1 - v_1$ can be interpreted as the use of synthetic fertilizers and $1 - v_2$ as the use of its biological substitutes. **Proposition 6a** Suppose that, $V = \min\{v_1, v_2\}$ with $\underline{v}_1 \leq \frac{1}{2}$. Then, there exist two distinct cut-off types, $\hat{\theta}_{p1} < \hat{\theta}_{p2}$, such that a producer of type θ_p makes the labeled good if and only if $\theta_p \in [0, \hat{\theta}_{p1}] \cup [\hat{\theta}_{p2}, 1]$.

Proof: See Appendix A.6.

Let $U(\theta_n)$ denote the utility difference between producing the labeled and unlabeled good for a producer of type θ_p . A producer of type θ_p chooses to produce the labeled good if and only if $\overline{U}(\theta_p) \geq 0$. The function $\overline{U}(\theta_p)$ and the quality produced $V(\theta_p)$ are plotted in figure 4a and 4b respectively. Observe that the producers who are the most willing to produce the labeled good are both the least and the most motivated ones. The producer who is the least willing to produce the ethically labeled good has an intermediate type, $\theta_p = \frac{4}{3} \underline{v_1}$. As in the baseline model, the most motivated producers overcomply with the standard (if $\theta \geq 2\underline{v}_1$, $V = v_1 = v_2 = \underline{v}_1$). However, as shown in appendix A.5, the quality chosen by producers with a type $\theta_p \in [0, \frac{4}{3}\underline{v}]$ would be higher if they remained unlabeled. Complying with the standard \underline{v}_1 and investing in v_2 is too costly for these producers. Thus, as illustrated in 4a, they solely comply with the standard $(v_1 = \underline{v}_1$ and $v_2 = 0$, so V = 0). By contrast, when producing the unlabeled good, these producers adopt modest but efficient, socially and environmentally friendly practices $(v_1 = v_2 = \frac{\theta_p}{4} > 0)$. As a result, among the least motivated producers, the most motivated ones relatively prefer to produce unlabeled goods. Finally, note that there is a mismatch between motivated producers and motivated consumers: the most ethical consumer also buys from the least intrinsically motivated producer.



Figure 4: Equilibrium quality and sorting with $V = \min\{v_1, v_2\}$ and $\underline{v} \leq \frac{1}{2}$

4.2 Cost-efficient alternatives

Now let $V(v_1, v_2) = v_1 + \frac{4}{3}v_2$. As an example, v_1 can be interpreted as the number of children the company fired and v_2 as the number of school scholarships offered. This specification implies that the most cost-efficient ethical production methods may not be part of the standard.

Proposition 6b Suppose that, $V(v_1, v_2) = v_1 + \frac{4}{3}v_2$. Then, there exists a non empty set $\Omega \subset]0,1]$, such that for all $\underline{v}_1 \in \Omega$, there exist two distinct cut-off types, $\hat{\theta}_{p1} < \hat{\theta}_{p2}$, such that a producer of type θ_p makes the labeled good if and only if $\theta_p \in [\hat{\theta}_{p1}, \hat{\theta}_{p2}]$.

$$\begin{array}{c|cccc} unlabel & label & unlabel \\ \hline 0 & \hline \theta_{p1} & & \hline \theta_{p2} & 1 \end{array}$$

Proof: See Appendix A.7.

Here the least and the most motivated producers are the least willing to produce the labeled good.²⁴ Interestingly, the most motivated producers choose to produce the unlabeled good to adopt the most cost-efficient ethical production methods ($\underline{v}_1 = 0$ and $v_2 = \theta_p$, so $V = \frac{4}{3}\theta_p > \underline{v}_1$). This is at the expense of selling at a lower price and not being socially

²⁴The more willing to produce it is a producer with an intermediate type $\theta_p = \frac{9}{16}v_1$.

rewarded. Just as in the previous case, there is a mismatch between motivated producers and motivated consumers.

5 Conclusion

This paper indicates that the development of markets for ethically labeled goods negatively affects the qualities of both ethically labeled and unlabeled goods. This is because recent adopters of ethical labels are typically less environmentally and socially concerned than the early adopters. However, they are more so than those still using conventional production methods. As documented in Section 2, the burgeoning organic industry is a good example of such a phenomenon.

This equilibrium sorting impedes the efficient functioning of the market, causing the market share of labeled goods to exceed what it would be without the presence of informal frictions. Therefore, policymakers should subsidize ethically labeled goods at a lower rate than the standard Pigouvian subsidy. Taxing these goods is socially desirable whenever producing them does not entail a significant positive externality. The underlying idea is that capping the profitability of ethically labeled firms curbs the arrival of opportunistic corporations. In my model, marginal consumers use the pseudo-ethical conduct of certain firms as an excuse to stop buying ethical goods. Note that this phenomenon can be fostered by psychological motives (refer to Exley (2016) and Exley (2020) for experimental evidence of such a pattern of behavior). This gives an additional reason to cap the profitability of ethically labeled firms.

This paper also shows that sorting in these markets is non-trivial and possibly highly inefficient. For instance, when firms can strategically get around the social and environmental constraints imposed by the labeling standard, ethically labeled goods are exclusively produced by the most and the least virtuous agents. In such cases, the most ethical consumers also buy from the least ethical corporations. To avoid this, the labeling standard must be designed so that, irrespective of one's motive, adopting it improves social and environmental performances.

Throughout this paper, I have considered that each firm is owned and managed by a single individual. Although this modeling specification fits well with the organic farming industry, it does not account for large firms.²⁵ The validity of my results for large corporations hinges on whether or not there is positive assortative matching within firms, among managers/owners/workers concerning their social and environmental preferences, as in Besley and Ghatak (2005). If such positive sorting arises, one can interpret each producer from the model as a representative agent of a large firm.

Appendix

A Proofs of the propositions

A.1 Proof of Proposition 1

The producer's cut-off type is given by:

$$\hat{p} = \frac{(\underline{v} - \hat{\theta}_p)^2}{2}.$$

Hence, when the price increases, more producers are willing to produce the labeled good, that is $\frac{\partial \hat{\theta}_p}{\partial \hat{p}} \leq 0$. Then, using the Leibniz integral rule with respect to $\hat{\theta}_p$ we obtain:

$$\frac{\partial \mathbb{E}[v|u]}{\partial \hat{\theta}_p} = \frac{f_p(\hat{\theta}_p)}{F_p(\hat{\theta}_p)} (\hat{\theta}_p - \mathbb{E}[v|u]) \ge 0,$$

$$\frac{\partial \mathbb{E}[v|l]}{\partial \hat{\theta}_p} = \frac{f_p(\hat{\theta}_p)}{1 - F_p(\hat{\theta}_p)} (\mathbb{E}[v|l] - \underline{v}) \ge 0,$$

And as $\mathbb{E}[v] = \int_0^{\hat{\theta}_p} x \, dF_p(x) + \int_{\hat{\theta}_p}^{\underline{v}} \underline{v} \, dF_p(x) + \int_{\underline{v}}^1 x \, dF_p(x),$
$$\frac{\partial \mathbb{E}[v]}{\partial \hat{\theta}_p} = f_p(\hat{\theta}_p) (\hat{\theta}_p - \underline{v}) \le 0.$$

It follows that $\frac{\partial \mathbb{E}[v]}{\partial \hat{p}} = \frac{\partial \mathbb{E}[v]}{\partial \hat{\theta}_p} \frac{\partial \hat{\theta}_p}{\partial \hat{p}} \ge 0$. Using the same reasoning, $\frac{\partial \mathbb{E}[v|l]}{\partial \hat{p}} \le 0$ and $\frac{\partial \mathbb{E}[v|u]}{\partial \hat{p}} \le 0$.

A.2 Proof of Proposition 2

First, we show that the equilibrium is interior. Suppose, to the contrary, that each producer makes the labeled good. Then, the least motivated producer, $\theta_p = 0$, by choosing to produce

²⁵According to Kitzmueller and Shimshack (2012), more than one-third of large firms have social or environmentally friendly labels.

the unlabeled good would get 0 instead of $-\frac{1}{2}\underline{v}^2 + \hat{p}$, so to sustain a pooling equilibrium \hat{p} needs to be strictly positive. On the demand side, the least motivated consumers, $\theta_c = 0$, by consuming the unlabeled good would get 0 instead of $-\hat{p}$, which leads to a contradiction. It is also impossible that all produce the unlabeled good. As $\underline{v} \leq 1$, the most motivated producers, $\theta_p > \underline{v}$, always make the labeled good (for any price $\hat{p} \geq 0$). Hence, there is no pooling equilibrium.

Secondly, we prove existence and uniqueness. Note that the price premium, the expected qualities, and the consumer cut-off type can be expressed as a function of $\hat{\theta}_p$. Thus, the equilibrium defined by (4a), (4b), (4c), (4d), and (4e) boils down to the following equation with one unknown $\hat{\theta}_p$:

$$\phi(\hat{\theta}_p) \equiv F_c^{-1}(F_p(\hat{\theta}_p))(\mathbb{E}[v|l] - \mathbb{E}[v|u]) - \frac{1}{2}(\underline{v} - \hat{\theta}_p)^2 = 0$$

To apply the intermediate value theorem, we need to show that ϕ is strictly increasing. Using the decomposition of the net quality derived in (5) we obtain the following derivative of ϕ :

$$\begin{split} \phi'(\hat{\theta}_p) &= F_c^{-1}(F_p(\hat{\theta}_p))] \frac{\partial \Delta_p}{\partial \hat{\theta}_p} + \frac{\partial F_c^{-1}(F_p(\hat{\theta}_p))}{\partial \hat{\theta}_p} \Delta_p - \frac{f(\hat{\theta}_p)F_c^{-1}(F_p(\hat{\theta}_p))}{(1 - F_p(\hat{\theta}_p))^2} [\int_{\hat{\theta}_p}^{\underline{v}} (1 - F(x))dx] \\ &+ \frac{\partial F_c^{-1}(F_p(\hat{\theta}_p))}{\partial \hat{\theta}_p} \frac{1}{1 - F(\hat{\theta}_p)} \int_{\hat{\theta}_p}^{\underline{v}} [F_p(x) - F_p(\hat{\theta}_p)]dx + (\underline{v} - \hat{\theta}_p). \end{split}$$

The lemma by Jewitt (2004) ensures that $\frac{\partial \Delta_p}{\partial \hat{\theta}_p} \geq 0$. Additionally, as F_p and F_c are increasing, $\frac{\partial F_c^{-1}(F_p(\hat{\theta}_p))}{\partial \hat{\theta}_p} \geq 0$. Consequently, the following inequality holds.

$$\phi'(\hat{\theta}_p) > (\underline{v} - \hat{\theta}_p) - \frac{f_p(\hat{\theta}_p)F_c^{-1}(F_p(\hat{\theta}_p))}{(1 - F_p(\hat{\theta}_p))^2} [\int_{\hat{\theta}_p}^{\underline{v}} (1 - F_p(x))dx]$$

Since $\hat{\theta}_p \leq \underline{v}$, and the cumulative distribution function $F_p(\theta_p)$ has an increasing hazard rate and $F_c^{-1}(F_p(\hat{\theta}_p))$ is increasing, the following inequality holds.

$$\phi'(\hat{\theta}_p) > (\underline{v} - \hat{\theta}_p) - \frac{f_p(\underline{v})F_c^{-1}(F_p(\underline{v}))}{(1 - F_p(\underline{v}))(1 - F_p(\hat{\theta}_p))} [\int_{\hat{\theta}_p}^{\underline{v}} (1 - F_p(x))dx]$$

I have assumed that \underline{v} is sufficiently small. Formally, we need \underline{v} such that $\underline{v} \in [0, v^*]$, where v^* is solution to $F_c^{-1}(F_p(v^*))\frac{f_p(v^*)}{1-F_p(v^*)} = 1$. This allows us to simplify the inequality as follows:

$$\phi'(\hat{\theta}_p) > (\underline{v} - \hat{\theta}_p) - \frac{1}{1 - F_p(\hat{\theta}_p)} [\int_{\hat{\theta}_p}^{\underline{v}} (1 - F_p(x)) dx] = \int_{\hat{\theta}_p}^{\underline{v}} 1 - \frac{1 - F_p(x)}{1 - F_p(\hat{\theta}_p)} dx > 0$$

The above expression is strictly positive because $1 - F_p(\hat{\theta}_p) > 1 - F_p(x)$ for all x in $\in [\hat{\theta}_p, \underline{v}]$. Therefore $\phi' > 0$. To conclude, observe that $\phi(0) = -\frac{1}{2}\underline{v}^2 < 0$ and $\phi(\theta_p = \underline{v}) > 0$. Thus, the intermediate value theorem ensures that the equilibrium exists and is unique.

A.3 Proof of Proposition 3

The aggregate welfare is $W = CS + PS + \alpha \mathbb{E}[v]$ where

$$CS = \int_{\hat{\theta}_c}^{1} (\theta \mathbb{E}[v|l] - \hat{p}) dF_c(\theta) + \int_{0}^{\hat{\theta}_c} \theta \mathbb{E}[v|u] dF_c(\theta),$$

$$PS = \int_{\hat{\theta}_p}^{1} (\theta \max\{\theta, \underline{v}\} + \hat{p} - \frac{1}{2} \max\{\theta_p, \underline{v}\}^2) dF_p(\theta) + \int_{0}^{\hat{\theta}_p} \frac{\theta^2}{2} dF_p(\theta).$$

The monetary transfers cancel out, hence, the derivative of the aggregate welfare with respect to the cut-off type $\hat{\theta}_p$ is:

$$\begin{aligned} \frac{dW}{d\hat{\theta}_p} &= \int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x) + \int_0^{\hat{\theta}_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x) - \frac{d\hat{\theta}_c}{d\hat{\theta}_p} f_c(\hat{\theta}_c) \hat{\theta}_c(\mathbb{E}[v|l] - \mathbb{E}[v|u]) \\ &+ f_p(\hat{\theta}_p) \frac{(\underline{v} - \hat{\theta}_p)^2}{2} + \alpha f_p(\hat{\theta}_p) (\hat{\theta}_p - \underline{v}). \end{aligned}$$

Due to market clearing $f_p(\hat{\theta}_p) = \frac{d\hat{\theta}_c}{d\hat{\theta}_p} f_c(\hat{\theta}_c)$ because $F_p(\hat{\theta}_p) = F_c(\hat{\theta}_c)$. Therefore, the derivative of the welfare function can be rewritten as follow:

$$\begin{aligned} \frac{dW}{d\hat{\theta}_p} &= \int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x) + \int_0^{\hat{\theta}_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x) \\ &+ f_p(\hat{\theta}_p)(\frac{(\underline{v} - \hat{\theta}_p)^2}{2} - \hat{\theta}_c(\mathbb{E}[v|l] - \mathbb{E}[v|u]) + \alpha(\hat{\theta}_p - \underline{v})) \end{aligned}$$

We have shown in the proof of Proposition 2 that the equilibrium is such that:

$$\frac{1}{2}(\underline{v}-\hat{\theta}_p)^2 - \hat{\theta}_c(\mathbb{E}[v|l] - \mathbb{E}[v|u]) = 0.$$

Therefore, the derivative of the aggregate welfare evaluates at the equilibrium cut-off type $\hat{\theta}_p$ amounts to:

$$\frac{dW}{d\hat{\theta}_p} = \int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x) + \int_0^{\hat{\theta}_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x) + f_p(\hat{\theta}_p)\alpha(\hat{\theta}_p - \underline{v})$$

We know that $\frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} > 0$ and $\frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} > 0$ (Proposition 1). As a result, if $\alpha = 0$, the derivative of the aggregate welfare evaluates at the equilibrium cut-off type, is strictly positive. It means that, in equilibrium, a small increase in the cut-off type is socially desirable. By contrast, a reduction of the cut-off type cannot be optimal: it would create a deadweight loss and foster market failures (adverse and advantageous selection). Consequently, if $\alpha = 0$, the equilibrium market share of labeled goods exceeds its optimal level.

A.4 Proof of Proposition 4

Each labeled producer is given an amount $s \in \mathbb{R}$. The equilibrium condition becomes:

$$\phi(\hat{\theta}_p) + s = 0.$$

Thus, the optimal subsidy s^* is implicitly defined by $s^* + \phi(\hat{\theta}_p^*) = 0$, where $\hat{\theta}_p^*$ is the constraint efficient cut-off type. As $\hat{\theta}_p^* \in [0, 1]$, ϕ continuous and $\phi' > 0$ (shown in the proof of Proposition 1), then the intermediate value theorem ensures that the optimal subsidy exists and is unique.

The subsidy, s, does not appear in the formulation of the total welfare. Indeed, monetary transfers cancel out. However, it features in the equilibrium condition: $\phi(\hat{\theta}_p) + s = 0$. Therefore, the derivative of the total welfare becomes:

$$\frac{dW}{d\hat{\theta}_p} = \int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x) + \int_0^{\theta_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x) + f_p(\hat{\theta}_p)(\alpha(\hat{\theta}_p - \underline{v}) + s).$$

If $\alpha = 0$, then, as shown in the proof of Proposition 3, an increase in the cut-off type is welfare enhancing. As $\phi' > 0$, The higher is $\hat{\theta}_p^*$; the smaller is s^* . Therefore, taxing the labeled good, $s^* < 0$, is optimal. By continuity, this result holds for any α small enough.

If $s = \alpha(\underline{v} - \hat{\theta}_p)$, then the total derivative of the aggregate surplus with respect to the producer cut-off type, evaluated at the equilibrium, becomes:

$$\frac{dW}{d\hat{\theta}_p} = \int_0^{\hat{\theta}_c} x \frac{d\mathbb{E}[v|u]}{d\hat{\theta}_p} dF_c(x) + \int_{\hat{\theta}_c}^1 x \frac{d\mathbb{E}[v|l]}{d\hat{\theta}_p} dF_c(x) > 0.$$

As a result, the derivative of the aggregate welfare is strictly positive. Thus, the amount of the Pigouvian subsidy exceeds its optimal level.

A.5 Proof of Proposition 5

Let $U(\theta_p)$ represent the utility difference between producing the labeled good and unlabeled good for a producer of type θ_p . A producer of type θ_p chooses to produce the labeled good if and only if $\overline{U}(\theta_p) \ge 0$.

$$\bar{U}(\theta_p, v_l(\theta_p), v_u(\theta_p)) = \theta_p[V(v_l(\theta_p)) - V(v_u(\theta_p))] - C(v_l(\theta_p)) + C(v_u(\theta_p)) + \hat{p}$$

Using the chain rule, we obtain the following derivative:

$$\frac{d\bar{U}}{d\theta_p} = \frac{\partial U}{\partial \theta_p} + \frac{\partial U}{\partial v_l} \frac{\partial v_l}{\partial \theta_p} + \frac{\partial U}{\partial v_u} \frac{\partial v_u}{\partial \theta_p}.$$

According to the envelope theorem, the last two terms are zero. Or,

$$\frac{d\bar{U}}{d\theta_p} = \frac{\partial U}{\partial \theta_p} = V(v_l(\theta_p)) - V(v_u(\theta_p))$$

This established the desired result:

$$\frac{d\bar{U}}{d\theta_p} \gtrless 0 \Leftrightarrow V(v_l(\theta_p)) \gtrless V(v_u(\theta_p)).$$

The above inequality ensures that if for all $\theta_p \in [0, 1]$, $V(v_l(\theta_p)) > V(v_u(\theta_p))$, then $\overline{U}(\theta_p)$ is strictly increasing, which ensures that there exists a unique cut-off type $\hat{\theta}_p$.

A.6 Proof of Proposition 6a

Step 1: characterization of $\overline{U}(\theta_p)$. Suppose that $V(v_1, v_2) \equiv \min\{v_1, v_2\}$. If a producer makes the unlabeled good, she chooses $v_1 = v_2$ (this is due to the extreme complementarity of the Leontief function). Replacing v_1 with v_2 in the producer's objective and taking the first order condition, we obtain that $v_1 = v_2 = \frac{\theta_p}{4}$.

Let's consider the case in which the producer produces the labeled good. First, if $\theta_p \geq 4\underline{v}_1$, the producer is not constrained by the minimum quality standard, so $v_1 = v_2 = \frac{\theta_p}{4}$. By contrast, if the producer's type is such that $\theta_p \leq 4\underline{v}_1$, then the producer selects the corner solution, $v_1 = \underline{v}_1$. Plugging $v_1 = \underline{v}_1$ into the producer's objective and taking the first-order derivative with respect to v_2 , we obtain the following first-order condition: $v_2 = \theta_p - \underline{v}_1$. Therefore, i) if $\underline{v}_1 \leq \theta_p \leq 2\underline{v}_1$, then $v_2 = \theta_p - \underline{v}_1$; ii) if $\theta_p \leq \underline{v}_1$, then $v_2 = \theta_p - \underline{v}_1 < 0$, so v_2 is a corner solution, $v_2 = 0$; if $2\underline{v}_1 \leq \theta_p \leq 4\underline{v}_1$, then $v_2 = \theta_p - \underline{v}_1 > \underline{v}_1$, so v_2 is a corner

solution, $v_2 = \underline{v}_1$. Plugging the value of v_1 and v_2 into the producer's objective, we obtain the following expression of $\overline{U}(\theta_p)$:

$$\bar{U}(\theta_p) = \begin{cases} \hat{p} & \text{if } \theta_p \ge 4\underline{v_1} \\ \hat{p} - 2\underline{v_1}^2 + \theta_p \underline{v_1} - \frac{\theta_p^2}{8} & \text{if } 2\underline{v_1} \le \theta_p \le 4\underline{v_1} \\ \hat{p} - \underline{v_1}\theta_p + \frac{3}{8}\theta_p^2 & \text{if } \underline{v_1} \le \theta_p \le 2\underline{v_1} \\ \hat{p} - \frac{\underline{v_1}^2}{2} - \frac{\theta_p^2}{8} & \text{if } \theta_p \le \underline{v_1} \end{cases}$$

Step 2: Equilibrium price premium. Note that $\bar{U}(\theta_p)$ is continuous, decreasing on the support $[0, \frac{4}{3}\underline{v}_1]$ and increasing on the support $[\frac{4}{3}\underline{v}_1, 1]$. Also, observe that $\bar{U}(0) = \bar{U}(2\underline{v}_1) = \hat{p} - \frac{1}{2}\underline{v}_1^2 < \bar{U}(1)$. Therefore, if the price premium is such that $\bar{U}(\frac{4}{3}\underline{v}_1) = \hat{p} - \frac{2}{3}\underline{v}_1^2 < 0 < \bar{U}(0) = \hat{p} - \frac{1}{2}\underline{v}_1^2$, then the equilibrium is of the form: there exist two distinct cut-off types, $\hat{\theta}_{p1} < \hat{\theta}_{p2}$, such that a producer of type θ_p produces the labeled good if and only if $\theta_p \in [0, \hat{\theta}_{p1}] \cup [\hat{\theta}_{p2}, 1]$.

$$\mathbb{E}[V|u] = \frac{1}{2}(V(v_u(0)) + V(v_u(2\underline{v}_1))) = \frac{1}{2}(0 + \frac{2\underline{v}_1}{4}) = \frac{\underline{v}_1}{4},$$
$$\mathbb{E}[V|l] = \underline{v}_1.$$

At the equilibrium, the supply equals the demand. Therefore, $\hat{\theta}_c = \hat{\theta}_p = 2\underline{v}_1$. However, observe that:

$$\hat{\theta}_{c}[\mathbb{E}[V|l] - \mathbb{E}[V|u]] - \hat{p} = 2\underline{v}_{1}[\underline{v}_{1} - \frac{\underline{v}_{1}}{4}] - \frac{1}{2}\underline{v}_{1}^{2} = \underline{v}_{1}^{2} > 0.$$

As a result, for $\hat{p} = \frac{1}{2}\underline{v}_1^2 - \epsilon$, the demand for the labeled good exceeds it supply. Thus, the price premium must increase until the supply equals its demand. Consequently, the equilibrium price premium is such that $\hat{p} > \frac{1}{2}\underline{v}_1^2$. Thus, $\overline{U}(0) > 0$. We need also to consider the case (b), $0 < \underline{v}_1 < \frac{1}{4}$. In this case, the expected quality of the labeled good is:

$$\mathbb{E}[V|l] = \frac{1}{1 - 2\underline{v}_1} [2\underline{v}_1^2 + (1 - 4\underline{v}_1)\frac{(\underline{v}_1 + \frac{1}{4})}{2}] = \frac{1}{8 - 16\underline{v}_1}$$

Plugging the expected quality of the labeled good into the consumer's equilibrium condition, we obtain the same result as above.

Suppose now that the price premium is such that $\hat{p} > \frac{2}{3}\underline{v_1}^2$. Then, the market is covered: each producer makes the labeled good, and each consumer consumes the labeled good. It is impossible because the consumer with the lowest type, $\theta_c = 0$, always prefers to buy the unlabeled good (for any $\hat{p} > 0$); thus, $\hat{p} < \frac{2}{3}\underline{v_1}^2$. This concludes the proof.

A.7 Proof of Proposition 6b

Step 1: characterization of $\bar{U}(\theta_p)$. Suppose that $V(v_1, v_2) = v_1 + hv_2$ with h > 1. If a producer makes the unlabeled good, then she chooses $v_1 = 0$ (because v_2 is more productive). Taking the first-order derivative with respect to v_2 , we obtain that $v_2 = h\theta_p$. If the producer makes the labeled good, then she chooses $v_1 = \underline{v}_1$ (if she wants to over-comply, she puts her additional effort into v_2). Plugging $v_1 = \underline{v}_1$ into the producer's objective and taking its derivative with respect to v_2 , we obtain the following first-order condition: $v_2 = h\theta_p - \underline{v}_1$. If $\theta_p \geq \frac{v_1}{\overline{h}}$, then the producer chooses $v_2 = h\theta_p - \underline{v}_1$. If $\theta_p < \frac{v_1}{\overline{h}}$, $h\theta_p - \underline{v}_1 < 0$, she chooses the corner solution $v_2 = 0$. Plugging the value of v_1 and v_2 into the producer's objective, we obtain the following expression of $\overline{U}(\theta_p)$:

$$\bar{U}(\theta_p) = \begin{cases} \hat{p} + (1-h)\underline{v}_1\theta_p & \text{if } \theta_p \ge \frac{v_1}{h} \\ \hat{p} + \theta_p \underline{v}_1 - \frac{1}{2}\underline{v}_1^2 - \frac{1}{2}(\theta_p h)^2 & \text{if } \theta_p \le \frac{v_1}{h} \end{cases}$$

Step 2: equilibrium price premium. Note that $\bar{U}(\theta_p)$ is continuous, increasing on the support $[0, \frac{v_1}{h^2}]$ and decreasing on the support $[\frac{v_1}{h^2}, 1]$. Also, observe that $\bar{U}(0) = \hat{p} - \frac{1}{2}\underline{v}_1^2 < \bar{U}(1) = \hat{p} + (1-h)\underline{v}_1$. Indeed, this inequality is equivalent to $2(h-1) > \underline{v}_1$, which is satisfied when $h = \underline{v}_1$. Therefore, if the price premium is such that $\bar{U}(1) = \hat{p} + (1-h)\underline{v}_1 < 0 < \bar{U}(\frac{v_1}{h^2}) = \hat{p} + \frac{\underline{v}_1^2}{2}(\frac{1}{h^2} - 1)$, then the equilibrium is of the form: there exist two distinct cut-off types, $\hat{\theta}_{p1} < \hat{\theta}_{p2}$, such that a producer of type θ_p produces the unlabeled good if and only if $\theta_p \in [0, \hat{\theta}_{p1}] \cup [\hat{\theta}_{p2}, 1]$.

First, suppose that each producer makes the labeled good, that is, $\hat{p} > (h-1)\underline{v}_1$. As in case 1, it is not possible because the least motivated consumer, $\theta_c = 0$, always prefers to buy the unlabeled good (for any $\hat{p} > 0$). Secondly, suppose that each producer makes the unlabeled good, that is, $\hat{p} < \frac{v_1^2}{2}(1-\frac{1}{h^2})$. Then, the expected quality of the unlabeled good is $\mathbb{E}[V|u] = \frac{1}{2}(0+h^2) = \frac{h^2}{2}$. Note that the producer with a type, $\theta_p = \frac{v}{h^2}$, is willing to produce the labeled good for a price premium: $\hat{p} \ge \frac{v_1^2}{2}(1-\frac{1}{h^2})$; and the most motivated consumer, $\theta_c = 1$, is willing to buy the labeled good of quality \underline{v}_1 (the quality that would be produced by producer $\theta_p = \frac{v}{h^2}$) if $\hat{p} < \underline{v}_1 - \frac{h^2}{2}$. Consequently, we can discard pooling equilibria if the following condition holds: $\frac{v_1^2}{2}(1-\frac{1}{h^2}) < \underline{v}_1 - \frac{h^2}{2}$ as we have assumed that $h = \underline{v}$, this condition is equivalent to $\underline{v} < \frac{1}{2}(1+\sqrt{3})$. Therefore, there is no pooling equilibria. Thirdly, if the price premium is such that $\frac{1}{2}\underline{v}_1^2 > \hat{p} > (h-1)\underline{v}_1$, then the equilibrium is of the form: there exists a unique cut-off type, $\hat{\theta}_p$, such that the producer is labeled if and only if her type is greater than $\hat{\theta}_p$. If the price premium is at its lowest level $\hat{p} = (h-1)\underline{v}_1$, $\bar{U}(\theta_p) = 0$ admits two solutions: $\theta_p = 1$ and $\theta_p = \frac{2-h}{h}$. Note that the expected quality of the labeled good is $\mathbb{E}[V|l]) = \underline{v}_1$, and the expected quality of the unlabeled good is $\mathbb{E}[V|u]) = \frac{1}{2}(0 + \hat{\theta}_p h^2) = \frac{(2-h)h}{2}$. The consumer's equilibrium condition is: $\hat{\theta}_c(\mathbb{E}[V|l]) - \mathbb{E}[V|u]) - \hat{p} = 0$, which can be rewritten, $\frac{1}{2}(4 - 3h)h = 0$. However, as we have assumed that h > 4/3, this equality is not satisfied the price is too high. Equilibrium forces are going to push the price premium downward. As a result, the price premium is such that $\bar{U}(1) = \hat{p} + (1-h)\underline{v}_1 < 0 < \bar{U}(\frac{v_1}{h^2}) = \hat{p} + \frac{v_1^2}{2}(\frac{1}{h^2} - 1)$; this concludes the proof.

B Motivating evidence

	Date of conversion			Never converted	
	before 2000	2000-10	after 2010	never	
number of farms	3300	10788	18604	500994	
farm size in ares (median)	4300	3576	3429	3184	
farmer age (median)	52	47	47	51	
number of year at the head of the farm (median)	24	16	16	20	
female farmer (%)	15	18	18	22	
access to internet (%)	68	74	64	46	
(A) indicators (%):					
solar energy	6.7	3.6	2.0	1.0	
wind energy	0.15	0.12	0.08	0.01	
local trade	27.1	19.6	13.9	7.2	
beekeeping	5.8	4.3	3.8	2.4	
force-feeding	1.7	3.8	3.7	3.8	
tobacco	0.0	0.3	0.4	0.4	
(B) main production of the farm (%):					
cattle: meat	10	8	12	11.5	
cattle:meat and milk	1	1	1.5	2	
cattle:milk	16.5	10.5	8	9.5	
chicken and pig	5.5	5	3.5	5.5	
field crops	14	14.5	16	23.5	
fruits and permanent crops	7	8	6	3.5	
mixed livestock and polyculture	18	14	11.5	11	
sheep ad goat	10	9	10	10.5	
vegetable and flower	6	5	3	2.5	
wine	10	22	19	13	
(C) farmer's education (%):					
no degree	21	17.5	21	36	
middle school degree	31	29	28.5	28	
vocational education	11	11.5	14.5	15	
high school	20.5	21	18.5	11	
bachelor degree	7.5	9	7	4	
master degree	9	12	9	4.5	

Table 3: Summary statistics

Table 4: List of variables

Variable	Definition
LOCAL	percentage of product sold in local markets
SOLAR	1 if produce solar energy, 0 otherwise
BEE	1 if beekeeping, 0 otherwise
WIND	1 if produce wind energy, 0 otherwise
GAVAGE	1 if practice force-feeding, 0 otherwise
TABAC	1 if produce tobacco, 0 otherwise
dep_pop	population in departements
dep_wage	median wage per departements
dep_sun	number of hours of sun per year in the departement
dep_rain	number of cm of water of rain per departement.
INTERNET	1 if the farm has access to internet
dep	1:100 number of the departement
COMPT	1 if the farmer is using an accounting software, 0 otherwise
CEXSEX	1 if the farmer is a female, 0 otherwise
UTATOT	total labor force (hours worked per year)
SAU	size of the farm (ares)
AGE	age of the farmer
EXP	number of year of the farmer at the head of the farm
EDUC	education of the farmer:
	no degree
	vocational education
	middle school degree
	high school degree
	bachelor degree
	master degree
PROD	main type of production:
	cattle: meat and milk
	cattle: milk
	chiken and pig
	Field crops
	fruit and permanent crops
	mixed livestock and polyculture
	sheep and goat
	vegetable and flowers
	wine

Dependent variable: Wind	(1)	(2)	(3)	(4)
Organic	2.11e-3***	2.05e-3***	2.13e-3***	2.16e-3***
	(4.32e-4)	(4.32e-4)	(4.33e-4)	(4.34e-4)
Date of conversion	-6.96e-5*	$-7.24e-5^{*}$	-8.24e-5**	-8.20e-5**
	(2.82e-5)	(2.82e-5)	(2.83e-5)	(2.83e-5)
Farmers' characteristics		yes	yes	yes
Farm characteristics			yes	yes
Location/weather				yes
Number of observations: 479,906				

Table 5: Wind energy and conversion to organic farming

Note: OLS regressions with constant; robust standard errors are in parentheses; p<0.1; p<0.05; p<0.01. The exhaustive list of controls is in Table 4. The variable Wind is equal to one if the farmer produces wind energy, and zero otherwise.

Table 6: Solar energy and conversion to organic farming

Dependent variable: Solar	(1)	(2)	(3)	(4)
Organic	5.87e-2***	5.67e-2***	5.35e-2***	5.15e-2***
	(2.40e-3)	(2.42e-3)	(2.42e-3)	(2.41e-3)
Date of conversion	-2.08e-3***	-2.11e-3***	-2.01e-3***	-1.96***
	(1.57e-4)	(1.57e-4)	(1.58e-4)	(1.57e-4)
Farmers' characteristics		yes	yes	yes
Farm characteristics			yes	yes
Location/weather				yes
Number of observations: 479,906				

Note: OLS regressions with constant; robust standard errors are in parentheses; p<0.1; p<0.05; p<0.01. The exhaustive list of controls is in Table 4. The variable Solar is equal to one if the farmer produces solar energy, and zero otherwise.

	(1)	(0)	(9)	(4)
Dependent variable: Bee	(1)	(2)	(3)	(4)
Organic	4.9e-2***	3.77e-2***	2.9e-2***	2.7e-2***
	(3.33e-3)	(3.37e-3)	(2.86e-3)	(2.85e-3)
Date of conversion	-1.00e-3***	-0.86e-3***	$-0.4e-3^{*}$	-0.36e-3*
	(2.21e-4)	(2.20e-4)	(1.86e-4)	(1.86e-4)
Farmers' characteristics		yes	yes	yes
Farm characteristics			yes	yes
Location/weather				yes
Number of observations: 479,906				

Table 7: Beekeeping and conversion to organic farming

Note: OLS regressions with constant; robust standard errors are in parentheses; p<0.1; p<0.05; p<0.01. The exhaustive list of controls is in Table 4. The variable Bee is equal to one if the farmer engages in beekeeping, and zero otherwise.

Table 8: Force feeding and conversion to organic farming

Dependent variable: Force feeding	(1)	(2)	(3)	(4)
Organic	-5.6e-3**	-5.33e-3***	-8.3e-3***	-6.3e-3***
	(1.9e-3)	(1.9e-3)	(1.8e-3)	(1.77e-3)
Date of conversion	$2.4e-4^{*}$	1.8e-4	$3.0e-4^{**}$	1.36e-4
	(1.2e-4)	(1.2e-4)	(1.1e-4)	(1.1e-4)
Farmers' characteristics		yes	yes	yes
Farm characteristics			yes	yes
Location/weather				yes
Number of observations: 479,906				

Note: OLS regressions with constant; robust standard errors are in parentheses; p<0.1; p<0.05; m<p<0.01. The exhaustive list of controls is in Table 4. The variable Force feed is equal to one if the farmer force-feeds animals, and zero otherwise.

Dependent variable: Tobacco	(1)	(2)	(3)	(4)
Organic	-5.1e-3***	-5.1e-3***	-6.1e-3***	-6.4e-3***
	(1.48e-3)	(1.47e-3)	(1.50e-3)	(1.49e-3)
Date of conversion	$2.19e-4^*$	$2.15e-4^{*}$	$2.50e-3^{*}$	$1.83e-3^{*}$
	(9.69e-5)	(9.7e-5)	(9.8e-5)	(9.72e-5)
Farmers' characteristics		yes	yes	yes
Farm characteristics			yes	yes
Location/weather				yes
Number of observations: 479,906				

Table 9: Tobacco and conversion to organic farming

Note: OLS regressions with constant; robust standard errors are in parentheses; p<0.1; p<0.05; $m^*p<0.01$. The exhaustive list of controls is in Table 4. The variable Tobacco is equal to one if the farmer produces tobacco, and zero otherwise.

References

- Aarset, B., Beckmann, S., Bigne, E., Beveridge, M., Bjorndal, T., Bunting, J., McDonagh, P., Mariojouls, C., Muir, J., Prothero, A., et al. (2004). The european consumers' understanding and perceptions of the "organic" food regime: The case of aquaculture. *British* food journal, 106(2):93–105.
- Akerlof, G. A. (1970). The market for" lemons": Quality uncertainty and the market mechanism. The quarterly journal of economics, pages 488–500.
- Allen, P. and Kovach, M. (2000). The capitalist composition of organic: The potential of markets in fulfilling the promise of organic agriculture. Agriculture and human values, 17(3):221–232.
- Amacher, G. S., Koskela, E., and Ollikainen, M. (2004). Environmental quality competition and eco-labeling. *Journal of Environmental Economics and Management*, 47(2):284–306.
- Ashman, T.-L., Knight, T. M., Steets, J. A., Amarasekare, P., Burd, M., Campbell, D. R., Dudash, M. R., Johnston, M. O., Mazer, S. J., Mitchell, R. J., et al. (2004). Pollen

limitation of plant reproduction: ecological and evolutionary causes and consequences. *Ecology*, 85(9):2408–2421.

- Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. (2019). Losing prosociality in the quest for talent? sorting, selection, and productivity in the delivery of public services. *American Economic Review (Forthcoming)*.
- Baland, J.-M. and Duprez, C. (2009). Are labels effective against child labor? Journal of Public Economics, 93(11):1125–1130.
- Barfort, S., Harmon, N. A., Hjorth, F., and Olsen, A. L. (2019). Sustaining honesty in public service: The role of selection. *American Economic Journal: Economic Policy*, 11(4):96–123.
- Baron, D. P. (2009). A positive theory of moral management, social pressure, and corporate social performance. Journal of Economics & Management Strategy, 18(1):7–43.
- Bartling, B., Weber, R. A., and Yao, L. (2015). Do markets erode social responsibility? The Quarterly Journal of Economics, 130(1):219–266.
- Basu, A. K., Chau, N. H., and Grote, U. (2006). Guaranteed manufactured without child labor: the economics of consumer boycotts, social labeling and trade sanctions. *Review* of Development Economics, 10(3):466–491.
- Bénabou, R. and Tirole, J. (2016). Bonus culture: Competitive pay, screening, and multitasking. Journal of Political Economy, 124(2):305–370.
- Bengtsson, J., Ahnström, J., and WEIBULL, A.-C. (2005). The effects of organic agriculture on biodiversity and abundance: a meta-analysis. *Journal of applied ecology*, 42(2):261– 269.
- Besley, T. and Ghatak, M. (2005). Competition and incentives with motivated agents. *The American economic review*, 95(3):616–636.
- Bizzotto, J. and Harstad, B. (2020). The choice of certifier in endogenous markets.
- Bonroy, O. and Constantatos, C. (2015). On the economics of labels: how their introduction affects the functioning of markets and the welfare of all participants. *American Journal of Agricultural Economics*, 97(1):239–259.

- Bottega, L. and De Freitas, J. (2009). Public, private and nonprofit regulation for environmental quality. *Journal of Economics & Management Strategy*, 18(1):105–123.
- Buck, D., Getz, C., and Guthman, J. (1997). From farm to table: The organic vegetable commodity chain of northern california. *Sociologia ruralis*, 37(1):3–20.
- Cason, T. N. and Gangadharan, L. (2002). Environmental labeling and incomplete consumer information in laboratory markets. *Journal of Environmental Economics and Management*, 43(1):113–134.
- Chakrabarty, S. and Grote, U. (2009). Child labor in carpet weaving: impact of social labeling in india and nepal. *World Development*, 37(10):1683–1693.
- Crozier, P., Bidaut, F., Dubus, C., Peres, G., Cluzeau, D., and Chossod, R. (2010). Comparaison de trois modes de production en bourgogne viticole: Premiers resultats.
- Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, 128(3):1169–1218.
- Darnhofer, I., Lindenthal, T., Bartel-Kratochvil, R., and Zollitsch, W. (2010). Conventionalisation of organic farming practices: from structural criteria towards an assessment based on organic principles. a review. *Agronomy for sustainable development*, 30(1):67–81.
- Delfgaauw, J. and Dur, R. (2007). Signaling and screening of workers' motivation. *Journal* of Economic Behavior & Organization, 62(4):605–624.
- Deserranno, E. (2019). Financial incentives as signals: experimental evidence from the recruitment of village promoters in uganda. American Economic Journal: Applied Economics, 11(1):277–317.
- Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. The Review of Economic Studies, 83(2):587–628.
- Exley, C. L. (2020). Using charity performance metrics as an excuse not to give. Management Science, 66(2):553–563.

- Falk, A., Neuber, T., and Szech, N. (2020). Diffusion of being pivotal and immoral outcomes. The Review of Economic Studies, 87(5):2205–2229.
- Fioretti, M. (2022). Caring or pretending to care? social impact, firms' objectives, and welfare. Journal of Political Economy, 130(11):2898–2942.
- Fischer, C. and Lyon, T. P. (2014). Competing environmental labels. Journal of Economics & Management Strategy, 23(3):692–716.
- Fleurbaey, M. and Ponthière, G. (2023). The stakeholder corporation and social welfare. Journal of Political Economy.
- Friebel, G., Kosfeld, M., and Thielmann, G. (2019). Trust the police? self-selection of motivated agents into the german police force. *American Economic Journal: Microeconomics*, 11(4):59–78.
- Friedman, M. et al. (1970). A friedman doctrine: The social responsibility of business is to increase its profits. The New York Times Magazine, 13(1970):32–33.
- Guthman, J. (2004). The trouble with 'organic lite'in california: a rejoinder to the 'conventionalisation'debate. *Sociologia ruralis*, 44(3):301–316.
- Hainmueller, J., Hiscox, M. J., and Sequeira, S. (2014). Consumer demand for the fair trade label: Evidence from a multi-store field experiment. *Review of Economics and Statistics*, (0).
- Hamilton, S. F. and Zilberman, D. (2006). Green markets, eco-certification, and equilibrium fraud. Journal of Environmental Economics and Management, 52(3):627–644.
- Hamman, J. R., Loewenstein, G., and Weber, R. A. (2010). Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, 100(4):1826–46.
- Handy, F. and Katz, E. (1998). The wage differential between nonprofit institutions and corporations: getting more by paying less? *Journal of comparative economics*, 26(2):246– 261.

- Harbaugh, R., Maxwell, J. W., and Roussillon, B. (2011). Label confusion: The groucho effect of uncertain standards. *Management science*, 57(9):1512–1527.
- Harper, G. C. and Makatouni, A. (2002). Consumer perception of organic food production and farm animal welfare. *British Food Journal*, 104(3/4/5):287–299.
- Hemenway, D. et al. (1990). Propitious selection. Quarterly Journal of Economics, 105(4):1063–1069.
- Heyes, A. G. and Maxwell, J. W. (2004). Private vs. public regulation: political economy of the international environment. *Journal of Environmental Economics and Management*, 48(2):978–996.
- Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7:24–52.
- Hopenhayn, H. and Saeedi, M. (2019). Optimal quality ratings and market outcomes.
- Jewitt, I. (2004). Notes on the shape of distributions. Technical report, Mimeo, Oxford University, Summer.
- Kitzmueller, M. and Shimshack, J. (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature*, 50(1):51–84.
- Lamine, C. and Bellon, S. (2009). Conversion to organic farming: a multidimensional research object at the crossroads of agricultural and social sciences. a review. Agronomy for sustainable development, 29(1):97–112.
- Li, Y. and van't Veld, K. (2015). Green, greener, greenest: Eco-label gradation and competition. *Journal of environmental economics and management*, 72:164–176.
- Lusk, J. L., Jamal, M., Kurlander, L., Roucan, M., and Taulman, L. (2005). A meta-analysis of genetically modified food valuation studies. *Journal of Agricultural and Resource Economics*, pages 28–44.
- Magill, M., Quinzii, M., and Rochet, J.-C. (2015). A theory of the stakeholder corporation. *Econometrica*, 83(5):1685–1725.

- Padel, S. and Foster, C. (2005). Exploring the gap between attitudes and behaviour: Understanding why consumers buy or do not buy organic food. *British food journal*, 107(8):606– 625.
- Prendergast, C. (2007). The motivation and bias of bureaucrats. American Economic Review, 97(1):180–196.
- Roitner-Schobesberger, B., Darnhofer, I., Somsook, S., and Vogl, C. R. (2008). Consumer perceptions of organic foods in bangkok, thailand. *Food policy*, 33(2):112–121.
- Rosset, P. M. and Altieri, M. A. (1997). Agroecology versus input substitution: a fundamental contradiction of sustainable agriculture. *Society & Natural Resources*, 10(3):283–295.
- Seufert, V., Ramankutty, N., and Foley, J. A. (2012). Comparing the yields of organic and conventional agriculture. *Nature*, 485(7397):229–232.
- Teisl, M. F., Roe, B., and Hicks, R. L. (2002). Can eco-labels tune a market? evidence from dolphin-safe labeling. *Journal of Environmental Economics and Management*, 43(3):339– 359.
- Tirole, J. (2006). The theory of corporate finance. Princeton university press.
- van Elsen, T. (2000). Species diversity as a task for organic agriculture in europe. Agriculture, Ecosystems & Environment, 77(1):101–109.
- Willer, H., Trávníček, J., Meier, C., and Schlatter, B. (2021). The world of organic agriculture 2021-statistics and emerging trends.
- Wu, Y., Zhang, K., and Xie, J. (2020). Bad greenwashing, good greenwashing: Corporate social responsibility and information transparency. *Management Science*, 66(7):3095– 3112.