Offshoring with Endogenous NGO Activism*

Sebastian Krautheim† Thierry Verdier‡

October 2014

Abstract

The process of globalization is characterized by an impressive growth of global value chains, as well as the proliferation of non-governmental organizations (NGOs) interacting with multinational firms. This paper presents a model of offshoring and NGO-firm interactions in which offshoring to a low-regulation country allows a monopolist to implement a 'dirty' technology undesired by consumers. Consumers can reduce the incentive for dirty production by financing an NGO monitoring the firm. NGO emergence and offshoring can arise as joint and interacting outcomes. For a range of trade costs, NGO emergence allows firms to capture gains from globalization, which would otherwise be unattainable. Somewhat paradoxically, NGO emergence can be at the expense of consumers possibly leading to welfare losses through offshoring.

JEL: *F23, F61, L31*

* We would like to thank participants of the LSE-NGO Workshop in London, the PSE Globalization Workshop in Paris, the ETSG 2012 in Leuven as well as seminar participants at the University of Nottingham, the University of Munich, the University of Rotterdam and the University of Cergy for useful discussions and comments. All remaining errors are our own. This is a revised version of CEPR Discussion Paper 9232 which was circulated under the title 'Globalization, Credence Goods and International Civil Society'.

† University of Passau.

‡ Paris School of Economics and CEPR.

1 Introduction

Over the last decades, globalization has been characterized by two remarkable features. The first one, widely noted by economists, is the impressive growth of global value chains, production networks and offshoring activities spanning across national borders. This process of increased internationalization of production was accompanied by a second phenomenon, somewhat less studied by economic scholars: the proliferation of non-governmental organizations (NGOs) and activist groups addressing a wide range of issues related to global production and sourcing decisions of multinational firms. In the United States and in Europe, such NGOs have triggered several well-publicized scandals involving pollution, child labor, hazardous working conditions, excessive working hours, and poor wages in factories supplying major global brands. In many cases these campaigns led to massive drops in demand and often forced firms to change their policies.

The increased internationalization of production has triggered strong NGO activity; NGO activity in turn affects the offshoring and technology decision of firms (Harrison and Scorse, 2010). Despite this inter-relation, the donor-NGO-firm interaction and the offshoring decision of the firm are analyzed in two mainly distinct literatures.

In this paper we contribute to both the NGO literature as well as to the offshoring literature and aim at closing the gap between them. We contribute to the NGO literature by suggesting a model of endogenous NGO emergence that simultaneously accounts for the NGO-donor interaction, the NGO-firm interaction as well as the linkages between the two. We contribute to the international economics literature on offshoring by introducing potential NGO emergence in response to offshoring into the problem of the firm. We link the two literatures by analyzing the mechanisms leading to the joint emergence of offshoring and NGO activism as a response to falling trade costs.

In our analysis we address the following questions: can falling trade costs ('globalization') trigger offshoring and NGO emergence simultaneously, and if so: under which conditions? When lower trade costs make offshoring more attractive for the firm, will an NGO supporting clean firms with a label and punishing dirty firms with a campaign dampen or foster offshoring? What determines the consumer's willingness to finance such an NGO? Are labels always good for firms and campaigns always bad? Why are NGOs active in some sectors but not in others? Can the firm exploit the fact that its offshoring decision may trigger the emergence of a donor-financed NGO to its own advantage? Is the offshoring decision of the firm always socially optimal?

To address these questions, we propose a North-South (Home-Foreign) model of offshoring where a monopolist serving consumers in Home faces a location choice (offshoring or domestic production) as well as a technology choice ('clean' or 'dirty') in a context of differences in regulation and enforcement in the two countries. Foreign has a general cost advantage for both types of technologies which captures standard gains from trade, e.g. due to different technologies or factor prices. Moreover, the dirty technology has a cost advantage compared to the clean one. If the firm offshores production to Foreign, it has to incur a transport cost to serve consumers in Home. Apart from the price of the consumption good, consumers care about externalities on third parties caused by the implementation of the dirty technology. In line with actual activist campaigns, we think of pollution in Foreign, poor labor standards, excessive and forced overtime, insufficient workplace security measures and the like leading to the cost advantage of the dirty technology.

We assume that Home is endowed with a perfect institutional regulatory capacity and that implementation of the clean technology is enforced by the government through regulation. In the baseline model, Foreign has no such enforcement capacity. When the firm offshores, consumers observe the technology used by the firm with some probability. When a firm is identified as dirty, demand drops to zero. An identified clean firm gets the full demand. Unidentified firms, however, face an intermediate demand which depends on the expectation consumers form about the probability of clean production. This implies that if the firm chooses dirty offshoring, it will always find it optimal to mimic a clean firm in all observables, i.e. setting the same price and quantity.

In response to offshoring, an NGO run by a motivated agent seeking to reduce the probability of dirty production can raise funds from consumers. Consumers are willing to donate to the NGO if they anticipate that the impact of the NGO on the technology choice of the firm is sufficiently strong to make the positive effect on utility from individual consumption large enough to compensate for the donation. While the objective of the NGO is only to reduce the probability of dirty production, consumers weigh this probability against the price of the consumption good. If the NGO can raise sufficient funds from consumers, it monitors the firm. If it is able to identify the technology used, it starts a campaign if the technology is dirty and provides a label if the technology is clean.

Our model highlights the ambiguous effects international integration can have on the production of credence goods. On the one hand, by locating production in a country with weak

governance structures, the firm is more likely to choose the dirty technology in order to benefit from the cost advantage, hoping that consumers do not find out. On the other hand, through the emergence of NGOs monitoring the firm, the globalization process triggers a countervailing reaction from within the civil society. More specifically, we show that, starting from an equilibrium where production takes place in Home, a sufficiently strong fall in trade costs can induce offshoring of the firm, which can in turn trigger NGO emergence turning the two into a joint phenomenon.

Thinking of globalization as a sequence of reductions in trade costs, the different corresponding equilibria in our model imply that if the NGO is sufficiently efficient, offshoring and NGO emergence occur simultaneously in response to lower trade costs. For lower NGO efficiencies, there is a range of trade costs for which the firm offshores without an NGO emerging. In these cases joint emergence of the two only occurs in equilibria with even lower trade costs, if it occurs at all.

We compare the offshoring decision of the firm in our model to the benchmark case without NGO emergence. We find that it is possible, in particular when the NGO is sufficiently efficient, that introducing the potential for NGO emergence increases the range of trade costs for which the firm decides to offshore. In our model we can therefore have a complementarity between the NGO and offshoring: anticipating NGO emergence makes offshoring more attractive to the firm.

This complementarity appears to contradict the popular view that NGOs and their campaigns hinder the internationalization of production as they harm offshoring firms. Our model is in line with this view in the sense that the ex-post effect of a campaign on a dirty firm is unambiguously negative as it reduces demand. From the ex-ante perspective the view of the firm is, however, more nuanced: the threat of facing a campaign when the dirty technology is implemented, reduces the attractiveness of the dirty technology. This in turn improves the expectations consumers form on the probability of clean production resulting in higher demand under offshoring. Combined with the unambiguously positive effect of labeling, this leads to the overall positive effect of NGO emergence on offshoring.[1]

As outlined above, NGO efficiency plays a key role in determining whether offshoring and

[1] This complementarity between NGO emergence and offshoring is directly related to the fact that in our model all consumers care about the type of technology used. One could think of a dynamic model with endogenized preferences where initially consumers do not care about the type of technology used and the NGO first needs to invest into awareness creation. This would likely lead NGO actions to reduce offshoring when the fraction of caring consumers is low, while we would expect the complementatiry to re-emerge when the fraction of caring consumers is sufficiently high. We briefly discuss such a setup in section 8.4.

NGO emergence occur jointly or not. Linking the parameters of our model determining NGO efficiency like the fundraising cost, increased transparency through the NGO action and transparency without an NGO, to real-world factors influencing these, our model is consistent with observing joint offshoring and NGO emergence in sectors producing final consumption products of well-known brands with large market size who tend to offshore through subcontractors. By the same token, based on the model, one would expect much fewer campaigns against producers of more homogeneous products or firms having their own production plants abroad. Although there are no established stylized facts on NGO campaigns in the literature, this is in line with most of the case study evidence on NGO campaigns.

Consumers in the model care both about the credence characteristic and the price of the consumption good. Offshoring unambiguously leads to lower prices, so that in principle a lower price could compensate consumers for a higher probability of facing a dirty product under offshoring. There is, however, a range of trade costs for which the firm offshores at the expense of consumers: while offshoring is profitable for the firm, consumers suffer from a reduced consumer surplus which for a subset of trade costs translates into a fall in total welfare. Due to the role of prices in consumer surplus, this effect dies out and is finally reversed for sufficiently low levels of trade costs.

This fall in social welfare as a consequence of offshoring is related to the first mover advantage the firm has when taking the offshoring decision. This leads to a somewhat paradoxical situation: consumers use the NGO as a tool to influence the firm technology choice if the firm offshores, but as the firm is the one to choose whether to offshore or not, it can use this first mover advantage to put consumers in a situation where they have to finance an NGO. There can therefore be a range of trade costs for which consumers would be better off if they could commit to not financing an NGO in response to offshoring, as this would lead the firm to produce in Home.

In an extension of the model, we analyze offshoring and international NGO activism when the national regulatory system in Foreign is endogenous and affected by the governments' investment in capacity-building. Specifically, we show that the emergence of an international NGO sector may actually crowd out the incentives of the government in Foreign to invest in local regulatory capacities that would fit the credence needs of Home consumers. The reason for this is simple. The Foreign government may underinvest in regulations, expecting sensitive consumers in Home to finance NGOs filling that gap. The existence of NGOs may therefore crowd out the implementation of effective public regulations in exporting countries. We show

that this crowding out is more likely to occur when Home consumers are more likely to ask for such regulatory mechanisms, the NGO technology is efficient at monitoring the multinational and when there are important efficiency gains from trade associated with offshoring.

Our paper contributes to the international economics literature on offshoring by introducing an NGO as a relevant player affecting the offshoring decision of the firm. In our analysis we abstract from asymmetric information and incomplete contracts as well as from the tradability of different tasks (Antràs and Helpman, 2004; Grossman and Rossi-Hansberg, 2008). Instead, we focus on the interaction between the offshoring decision of the firm (be it within the boundaries of the firm or at arm's length), the technology choice this involves and how these change when the possibility of the emergence of an activist NGO is taken into account by the firm.

Most of the existing literature on NGOs takes their existence as given and then either analyzes the fundraising dimension of NGO activity *or* the impact of NGO activity, e.g. on firms. Examples of the fundraising literature include work on non-profit competition for contractual aid projects (Chau and Huysentruyt, 2006), monopolistic competition on the donation market (Economides and Rose-Ackerman, 1993; Pestieau and Sato, 2006), fundraising competition between NGOs and diversion of funds (Castaneda et al., 2008 and Aldashev and Verdier, 2010), or NGOs as platforms in the two-sided "aid" market between donors and recipients when there are problems of moral hazard and adverse selection (Rowat and Seabright, 2006).

Starting with Baron (2001), a different strand of literature considers the role of *private politics* focusing on different aspects of the interaction between firms and activist groups (NGOs) and how these interactions affect corporate decisions on production, lobbying and advertising (Baron and Diermeier, 2007; Baron, 2012)[2], or patterns of competition and market structure outcomes (Aldashev, Limardi and Verdier, 2011). In contrast to our paper, this literature is mainly concerned about the actions of pre-existing agents and therefore does not focus on the endogenous emergence of NGOs.

The same holds true for the recent and related literature on Corporate Social Responsibility (CSR) surveyed by Kitzmueller and Shimshack (2012). The aspect of internationalization of production has not received much attention in either of these literatures (Kitzmueller, 2012, is a notable exception)[3]. Our paper therefore also contributes to the CSR literature by investigating the CSR investment of a firm (clean instead of dirty production) in the context of an international production choice under potential NGO emergence.

[2] See also Feddersen and Gilligan (2001), Innes (2006), Lyon and Salant (2013) and references therein.

[3] Another somewhat less related exception is Balboni and Balboni (2009).

Some papers (Baron, 2013; and Aldashev and Verdier, 2009) capture aspects of NGO emergence in the sense that they endogenize the number of NGOs depending on the donation decision. In these papers donors get direct utility from donating for a cause they support. Starting with Andreoni (1989) such preferences with a 'warm glow' or 'joy of giving' are widely used in the literature on charitable giving and NGO funding. One advantage of *not* using this assumption in our model is that we can analyze NGO funding driven by the effects an NGO has on utility through its equilibrium effects of the firm's technology and price setting choices: consumers anticipate the effect an NGO would have on the optimal choices of the firm and only donate if the emergence of an NGO triggers a firm reaction that makes consumers better-off. The second advantage is that a model without warm glow of giving allows us to carry out the welfare analysis, which would otherwise be conceptually problematic.

In our baseline model (as well as in an extension in the Appendix), we view the emergence of an NGO as coming from the initiative of an intrinsically motivated agent who needs to raise funds to set up the organization. In this sense, our work also relates to the literature on non-profit firms and motivated agents (Benabou and Tirole, 2003; Francois, 2003; Besley and Ghatak, 2005).

Our paper also connects to the literature on CSR and the private provision of public goods (Bagnoli and Watts, 2003; and Kotchen, 2006). The most related paper in this literature is Besley and Ghatak (2007), where perfectly competitive firms serve consumers who do and consumers who do not care about a public good which can be provided along with the production of a private good. They find that firms specialize in the production of the private good either with or without the public good component. Therefore in equilibrium, both caring and non-caring consumers are served. In their setup the level of provision of the public good through CSR is identical to the level obtained with voluntary contributions. They show that when firms cannot credibly commit to 'clean' production the equilibrium provision of the public good is lower. They also consider the possibility that the public good is directly provided by an NGO. We take a very different view on NGOs as the NGO in our model does not directly provide a public good, but rather incentivizes the firm to do so. In addition, they do not consider the international dimension of CSR provision and how it responds to falling trade costs.

Kitzmueller (2012) is related to our work as he takes the setup of Besley and Ghatak (2007) to the international level and considers the location choice of multinationals as well as the scope for government regulation in this context. In his model firms face a trade off between lower

production cost in the South and better transparency in CSR provision (increasing demand) in the North. This setting is similar to our modeling of lower production costs and lower probability of detection abroad. The focus in Kitzmueller (2012) is on the constraints informational asymmetries between locations put on national regulation and the scope for policy coordination. In our analysis we focus on the endogenous emergence of NGOs as a response to informational asymmetries and how this affects the offshoring decision of the firm.

In Immordino (2008) a multinational firm allocates a fraction of its production to the South where wages and labor standards are potentially different from the North. Some consumers care about labor standards and are informed about their application in the two countries. This exogenous fraction of caring informed consumers is interpreted as the manifestation of NGO activity. The focus of the analysis is on the game between the North and the South government in setting labor standards. This competition provides an incentive for mutual undercutting, possibly leading to a race to the bottom. This mechanism is the stronger the larger the number of informed caring consumers, leading to the result that there are welfare losses in both countries when more consumers care and are informed.

The remainder of the paper is structured as follows. Section 2 provides illustrative examples of major NGO campaigns and underlines the relevance of credence goods for consumer choices. Section 3 presents the baseline model. Section 4 describes the different stages of the strategic game consumers, the NGO and the firm play. Section 5 characterizes the firm's offshoring decision given NGO (non-) existence as well as the conditions for NGO emergence. Section 6 combines these, derives the equilibrium of the model and presents some of the main results. The welfare implications of offshoring are discussed in Section 7. Section 8 provides a discussion of several modeling choices, their plausibility, alternatives and possible effects on the outcomes. Section 9 extends the framework to endogenous regulatory capacity in Foreign. Section 10 presents conclusions and avenues for future research.

2 Globalization, NGOs and Credence Goods

The past decades have witnessed a large number of NGO campaigns and boycotts against multinational firms. The case of Nike's *sweatshops* in Indonesia is probably one of the most notorious examples. After NGOs triggered a wave of media attention and harsh criticism of the company's practice leading to weak consumer demand and retail oversupply, in May 1998 Nike's CEO Phil Knight admitted that *"the Nike product has become synonymous with slave*

wages, forced overtime, and arbitrary abuse. I truly believe the American consumer doesn't want to buy products made under abusive conditions." (quoted in Spar and La Mure, 2003, p. 91). Following this statement, the company raised the minimum ages of sneaker and apparel workers, adopted U.S. clean air regulations in all of its factories, and expanded monitoring and educational programs to workers. Harrison and Scorse (2010) empirically identify the causal effect of the campaigns on these changes in the production process.

The campaign of the Rainforest Action Network (RAN) against Staples, the leading retailer for paper and office supplies in the U.S. with more than 1000 retail outlets and annual sales of $11 billion, is another example. The campaign included an extensive research effort on the use of old growth rainforest woods in paper production for Staples as well as over 600 protests around the U.S. over a two year period. The campaign created large media attention and ultimately led Staples to agree to phase out the sale of products made from old growth rainforest wood and to obtain a minimum average of 30% of recycled paper products (O'Rourke, 2005).

An example for an ongoing campaign is the Detox Campaign by Greenpeace.[4] According to Greenpeace, a year long investigation on water pollution by textile factories in China preceded a campaign against major global brands like Nike, H&M, Lacoste and C&A. Confronting the companies with the threat of a negative PR campaign in 2011, a significant number of global brands agreed to eliminate hazardous chemicals across their entire supply chains by 2020.[5]

These examples show on the one hand that NGO campaigns can have a substantial impact on the technology choices of global companies. On the other hand they also illustrate that the information acquisition on behavior of foreign subsidiaries or subcontractors and their links to the final producer can be costly and time-consuming. The latter point underlines the informational asymmetry between firms and consumers which is at the core of our analysis.

The starting point of our analysis is the idea that the mounting leverage of NGOs on trade-related issues is directly related to the fact that consumers in developed northern economies are concerned about the *ethical* implications of their consumption. As these cannot be inferred from the final product they constitute credence characteristics. Different to experience goods where ex-ante unobservable properties are revealed ex-post, credence characteristics remain unobservable even after consumption implying an informational asymmetry between the firm and

[4] See http://www.greenpeace.org/international/en/campaigns/toxics/water/detox/intro/ (accessed September 2014) for information on the campaign provided by Greenpeace.

[5] Other examples of related campaigns can be found at http://www.cleanclothes.org/ (Clean Clothes Campaign, accessed September 2014) and on http://www.laborrights.org/ (International Labor Rights Forum, accessed September 2014).

consumers (Darby and Karni, 1973; Feddersen and Gilligan, 2001; Baksi and Bose, 2007). The relevance of such credence characteristics of goods is supported by survey evidence where a large fraction of consumers in developed countries reports preferences for ethical and environmentally friendly products along with a higher willingness to pay for such goods (O'Rourke, 2005; Loureiro and Lotade, 2005).[6] Recent evidence from field experiments shows that the preference for ethical goods affects real purchasing decisions (and not only survey-answers). Field experiments with fair trade and eco labels in U.S. department stores and in eBay auctions, showing that consumers express higher demand as well as a higher willingness to pay for labeled products (Hiscox and Smyth, 2011; Hainmueller and Hiscox, 2012).

3 Basic setup

3.1 Preference for 'clean' goods

We consider an economy (Home) with a continuum of individuals (of mass $L = 1$) consuming a bundle of normal goods and a credence good. The technology used ('clean' or 'dirty') in the production of the credence good may be unobservable for consumers. To capture this idea, we consider the following utility from consumption for a given consumer:

$$U = c_0 + IQ - \frac{Q^2}{2}.$$

c_0 is the level of consumption of a bundle of normal goods. In addition, consumers get utility from consumption of a differentiated good produced by a monopolist. The quantity of this good consumed is Q. Consumers do not only care about the physical properties of the good, but also about the technology used for production. Specifically, we assume that the valuation of the good is high when it was produced with a clean technology and low if it was produced with a dirty technology. This is reflected by the indicator variable I which equals one if the good is clean, and zero otherwise. Expected utility is then given by:

$$E(U) = c_0 + q^e Q - \frac{Q^2}{2} \tag{1}$$

[6] Whether the strong preference of such a large number of consumers found in most surveys is plausible or rather due to the fact that in surveys no actual purchasing decisions are taken into account is still a matter of debate (see e.g. O'Rourke (2005) and Vogel (2008)).

where q^e is the expectation formed by the consumer on the probability that the good was produced with a clean technology.[7] This implies that demand is given by:

$$Q = \max\{q^e - p, 0\}. \tag{2}$$

For analytical tractability it is convenient to assume that the expected probability of clean production enters equation (1) such that it only affects the intercept of the demand function but not the slope.

3.2 Technology of the firm

We model a monopolist that produces with a constant returns to scale technology. Depending on the choice of technology and location, the firm faces different variable costs of production. The firm can choose between a clean and a dirty technology. We assume that there is regulation in Home preventing the use of the dirty technology. For simplicity, we also assume that enforcement by the Home government is perfect and costless.[8] We denote the cost of production in Home by c_H.

When the firm chooses to locate production in Foreign, it faces lower input prices as well as imperfect enforcement. It can therefore choose between clean production with a marginal cost of $c < c_H$ and dirty production with a lower marginal cost of $c - \Delta$ with $\Delta \in (0, c)$.

For notational convenience, the marginal cost c includes both the cost of production and an additive trade cost of shipping the output to Home, where consumers are located. We model globalization as a decrease in these trade costs, which maps into a decrease in c.[9]

3.3 NGO: research, labels and campaigns

In our model a motivated agent can decide to create an NGO which monitors the technology used by the firm. The NGO interacts both with the firm (to reduce dirty output) and the consumer-donors (to raise funds). If the firm offshores and if in response to this an NGO emerges, it carries out research to identify the technology implemented by the firm. If this is successful, the NGO informs consumers about the firm type. In the case of a clean firm this can be thought

[7] Note that not even the consumption of the good reveals whether it was of the clean or dirty type.

[8] These assumptions are discussed in section 8.2.

[9] Assuming additive trade costs implies that the fall in trade costs is the same for both goods. The alternative assumption would be iceberg trade costs, which represent a fraction of the value of the good. Assuming additive trade costs appears appropriate in this context, as we consider goods that only differ in an unobservable credence characteristic. They should therefore also have identical shipping costs.

of as providing a label, in the case of a dirty firm this can be thought of as campaigning for a boycott. We simply model this by an increase in the probability that the type of the firm is revealed to consumers from π_0 (without NGO) to π_1 (with NGO) with $\pi_1 > \pi_0$.[10]

3.4 Cost of NGO emergence

While most consumers only care about whether the good they consume was produced with a clean or with a dirty technology, the motivated agent additionally cares about the possibility of dirty production in itself. When she faces a situation where the expected probability of dirty production $(1-q^e)$ is positive, she gets utility from 'fighting dirty production', i.e. from creating an NGO. This utility is independent of the actual level of the probability of dirty production: as soon as it is above zero, the motivated agent is ready to start monitoring the firm.[11]

NGO creation requires both a monetary fixed cost X and an effort fixed cost e. To avoid effects on aggregate welfare, we normalize the utility the motivated agent gets from fighting dirty production to be marginally higher than the fixed effort cost.[12] This implies that the motivated agent will set up an NGO as soon as there is some positive probability of dirty production and she is able to raise sufficient funds to cover the fixed cost X. In the appendix, we provide a more sophisticated microfounded model of the NGO entrepreneurial behavior where fundraising depends on two complementary inputs: cash and effort. We show that the model provides the same insights as the simplified version that we present here in the main text.

3.5 Donors and the NGO

Creating an NGO requires a fixed investment of X, which needs to be financed with fundraising. As consumer-donors are small, they would usually have an incentive to free ride on the donations of others, leading to zero donations to the NGO. A model of NGO financing needs to overcome this problem of free riding when the number of potential donors is large. The most common way to deal with the issue is to assume warm glow in preferences along the lines of Andreoni (1989,

[10] For the sake of analytical tractability, we keep the modeling of the NGO activity very stylized. Papers focusing exclusively on firm-NGO interactions provide a more detailed analysis of trade-offs concerning campaigns and labels (see e.g. Baron, 2012). These papers abstract, however, from NGO emergence and fund-raising as well as the interaction with offshoring decisions in the process of globalization, which are the focus of our analysis.

[11] We assume that any additional NGO monitoring the same firm would not lead to an additional increase in the detection probability. We therefore do not have to take a stand on the number of motivated agents in the economy, as for the firm we consider, at most one NGO will emerge.

[12] As the motivated agent running the NGO is small compared to the mass of consumers, aggregate welfare would not be affected, even if this assumption was relaxed. As we abstract from any intensive margin (monitoring intensity) choices of the NGO, all that matters is that the utility of fighting dirty production exceeds the effort cost e.

1990) (see Bagnoli and Watts, 2003; Baron, 2009; and references therein). In this case agents get utility from the mere fact that they are contributing to the provision of the public good.

While the warm glow assumption is a very parsimonious way to overcome the free riding problem, it remains an unresolved question how the utility from donating through warm glow should be accounted for in the welfare analysis. As we are also interested in the welfare effects of NGO emergence and offshoring, we use an alternative approach.

Atkinson (2009) analyzes the marketing strategies of NGOs and argues that NGOs aim at convincing donors that their donation does make a difference, e.g. by mapping each donor to a specific recipient - or at least to make the donor think that this is the case. We take a reduced form approach of this argument by assuming that through the costly process of activating donors, the NGO overcomes the free riding problem of funding by making donors 'feel important'.

Specifically, the NGO can pay costly resources to 'activate' a fraction ν of consumers. An activated consumer believes that her donation is crucial to the ability of the NGO to be active. The NGO requests a donation x from each donor. An activated donor is convinced to make a difference and is willing to donate up to the point where the donation equals the difference between expected individual consumer surplus with and without an NGO denoted by $E(W_1)$ and $E(W_0)$ respectively. This implies the following participation constraint for donors: $x \leq E(W_1) - E(W_0)$.

3.6 Timing

The timing in our model is as follows. In stage 1, the firm takes the location decision. In stage 2 the firm, consumer-donors and the motivated agent act simultaneously. The firm chooses the probability of clean production, consumers as well as the motivated agent form an expectation about the probability of clean production with an without an NGO. The motivated agent decides whether to run a fundraising campaign and consumers decide whether to donate to the NGO. In stage 3 nature chooses the type of the firm, which is revealed to consumers with probability π_1 if an NGO has emerged and with probability π_0 otherwise. In stage 4 firms set prices. In stage 5 consumers see whether prices of unidentified firms reveal their type and take purchase decisions. We solve the model by backward induction. The assumptions on the timing of events are discussed in Section 8.1.

4 Solving the model

4.1 Stage 5: demand

In stage 5 consumers observe prices set by firms. Equation (2) implies that for a product that is identified as clean demand is given by $Q_c = 1 - p$. By equation (2), demand for an unidentified product depends on the expectation consumers form on the probability of clean production q^e in stage 2. Demand for an identified dirty product is zero. Note that a dirty firm may not only be identified with probability π_i in stage 3, it may also be identified by consumers in stage 5, if its actions deviate from the actions consumers expect from an unidentified clean firm. The underlying signaling game is discussed in detail below.

4.2 Stage 4: price setting

In stage 4 firms set prices. By assumption, regulation and enforcement in Home are such that the firm always chooses the clean technology. In this case the firm gets the full demand from consumers but faces a higher production cost c_H and the optimal price set by the monopolist is given by:

$$p^H = \frac{1 + c_H}{2}.$$

In stage 4 there are identified clean, identified dirty and unidentified clean and dirty firms. We can therefore consider the optimal pricing of a clean and a dirty firm separately. A clean firm can be either identified as clean implying full demand or can be unidentified. In this case it faces the reduced demand depending on the expectation consumers have formed on its probability of clean production in stage 2. The optimal price set by a clean firm and the corresponding profits are therefore given by:

$$p_i^c = \begin{cases} \frac{1+c}{2} & \text{if identified} \\ \frac{q_i^e + c}{2} & \text{if not identified.} \end{cases} \tag{3}$$

When a dirty firm is identified as dirty, demand is zero regardless of the price it sets. What the optimal price of an unidentified dirty firm should be, however, is less obvious. The standard monopoly problem would suggest that the price should be set according to the same formula as the price of a clean firm but taking into account the lower cost of dirty production. The dirty firm, however, anticipates that in stage 5 consumers may interpret the deviation from the optimal price of a clean firm as a signal for being dirty.

We provide a formal analysis of the signalling game in the appendix. This formal analysis comprises both the price setting of the firm as well as the technology choice and its anticipation by consumers in stage 2. We therefore postpone the discussion of the signalling game and simply note at this stage that in the unique Perfect Bayesian Nash equilibrium (PBE) consistent with our equilibrium refinements the unidentified dirty firm sets the same price as the clean firm and the clean firm indeed sets its price according to equation (3).

The intuition for this result is that when the dirty firm deviates from the price a clean firm would set, consumers understand that this must be a dirty firm and demand drops to zero in stage 5. This mimicking behavior of an unidentified dirty firm to hide its type implies that while the firm benefits from the lower cost of dirty production, the cost advantage is not passed on to consumers. This introduces an asymmetry between the firm and consumers that will also play a role in the welfare analysis.

In stage 3 nature chooses the type of the firm.

4.3 Stage 2 - NGO

When the firm has decided to offshore production in step 1, in step 2 the NGO chooses the fraction of donors to activate and makes them a take it or leave it offer. Activated donors then decide whether or not to donate to the NGO and the NGO enters if it is able to raise sufficient funds.

The NGO minimizes fundraising costs φ subject to a fundraising constraint and a participation constraint for the donors:

$$\min_{\{\nu\}} \varphi(\nu)$$

subject to

$$\nu x - \varphi(\nu) \geq X \quad \text{and} \quad x \leq E(W_1) - E(W_0).$$

The second constraint implies that the gain of surplus the NGO gives to a donor must be larger or equal the amount donated.

For the maximum donation the NGO can extract from one particular individual, the participation constraint holds with equality: $x^* = E(W_1) - E(W_0)$. As activating additional donors is costly, the NGO will extract the maximum possible amount from each donor and choose the smallest possible fraction of consumers that still allows it to raise the funds X necessary for

operation. Defining ν^* as this minimum fraction of consumers to be activated, we can write:

$$\nu^* x^* - \varphi(\nu^*) = X.$$

For tractability, we consider the case of a linear NGO cost function of the type $\varphi(\nu) = y\nu$. The NGO optimality condition is then given by:

$$x^* \nu^* - y\nu^* = X.$$

The optimal fraction ν^* is:

$$\nu^*(c) = \frac{X}{x^* - y}. \tag{4}$$

Note that the maximum individual donation x^*, and therefore also ν^*, is a function of the cost level c. This implies that the willingness to donate varies for different levels of globalization.

4.4 Stage 2 - consumer-donors

As outlined above the maximum amount an activated consumer is willing to donate to the NGO is given by the difference in her consumer welfare with and without an NGO. Consumer welfare for the two cases $i \in \{0, 1\}$ is given by:

$$E(W_i) = q_i^e \pi_i \frac{(1-c)^2}{8} + (1 - \pi_i) \frac{(q_i^e - c)^2}{8}. \tag{5}$$

The first expression on the right hand side represents consumer surplus when the good is produced by an identified clean firm, weighted with the expected probability that the firm is clean q_i^e times the probability that it is identified π_i. The second term represents consumer surplus when the good is produced by an unidentified firm weighted with the probability that the firm in unidentified. Note that when the firm is identified as dirty (with expected probability of $(1 - q_i^e)\pi_i$), consumer surplus is zero. The difference between $E(W_1)$ and $E(W_0)$ therefore represents the gain in consumer surplus if an NGO emerges.

4.5 Stage 2 - firm

While the offshoring decision is taken by the firm in stage 1, in stage 2 it decides on the probability of clean production. If production takes place in home, the firm implements the clean technology, there is no asymmetric information on the firm type and no NGO emerges.

If the firm decides to offshore, it can choose a probability of clean production m_i. It does so as to maximize its expected profits under offshoring. In step 3 nature decides the firm type. Profits of a clean firm for given consumer expectations are given by:

$$\Pi_i^c(q_i^e) = \begin{cases} \frac{(1-c)^2}{4} & \text{if identified} \\ \frac{(q_i^e-c)^2}{4} & \text{if not identified.} \end{cases} \tag{6}$$

We will argue below that it is optimal for an unidentified dirty firm to mimic the price setting of an unidentified clean firm. It does, however, have lower costs so that profits of a dirty firm are given by:

$$\Pi_i^d(q_i^e) = \begin{cases} 0 & \text{if identified} \\ \frac{(q_i^e-c)^2}{4} + \Delta\frac{(q_i^e-c)}{2} & \text{if not identified.} \end{cases} \tag{7}$$

The last term in the second expression reflects the cost advantage an unidentified dirty firm has over an unidentified clean firm. With the above equations, we can write the expected firm profits for given consumer expectations as:

$$E(\Pi_i^c(q_i^e)) = \pi_i\frac{(1-c)^2}{4} + (1-\pi_i)\frac{(q_i^e-c)^2}{4} \tag{8}$$

$$E(\Pi_i^d(q_i^e)) = (1-\pi_i)\left[\frac{(q_i^e-c)^2}{4} + \Delta\frac{(q_i^e-c)}{2}\right]. \tag{9}$$

The firm chooses its probability of clean production m_i such that it maximizes expected profits. In this context, the three different cases of $m_i = 0$, $m_i \in (0,1)$ and $m_i = 1$ have to be distinguished.

The case where the firm chooses the dirty technology with certainty ($m_i = 0$) can never arise in equilibrium, as consumers would understand that the firm is dirty and decide not to buy the product. The opposite - and somewhat trivial - case where the firm uses the clean technology with probability $m_i = 1$ can arise for particular parametrizations, e.g. when the cost advantage of dirty production Δ is sufficiently low. In this case the firm always produces clean, gets full demand from consumers and no NGO emerges as no monitoring is necessary. This would be the standard case of gains from offshoring due to differences in production costs.

In this paper we are interested in the case where the gains from offshoring are more nuanced, as offshoring creates an incentive for the firm to use the dirty technology with some probability

$m_i \in (0,1)$. Such an interior solution exists if for some (endogenous) expectation of consumers $q_i^e \in (0,1)$ there exists an $m_i \in (0,1)$ for which the firm is indifferent between choosing the clean and the dirty technology, i.e. expected profits of clean and dirty production given by (8) and (9) are the same:

$$E(\Pi_i^c(q_i^e)) = E(\Pi_i^d(q_i^e)). \tag{10}$$

In an equilibrium where consumers form consistent expectations over the probability of clean production chosen by the firm (discussed below), this condition pins down the probability of clean production as:

$$m_i = q_i^e = \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi_i}{1-\pi_i} + c. \tag{11}$$

It can be seen from this expression that a high cost advantage of dirty production decreases the probability of clean production, while a high probability of being identified (possible through an NGO) increases the probability of clean production.

This directly implies the following parameter restriction which we assume to hold for the remainder of the analysis:

Assumption 1: $\frac{1-\pi_1}{\pi_1}\Delta > \frac{1-c}{2}$.

This condition assures that the expected gains from dirty production are sufficiently high for the firm to assign a positive probability to dirty production implying that for appropriate parameter values there can be dirty production despite the fact that consumers prefer clean production.

Two assumptions made above still remain to be justified. First, the mimicking of an unidentified dirty firm when setting its price and second, that consumers correctly anticipate the probability of clean production chosen by the firm ($q_i^e = m_i$).

Consumers interpret the price set by an unidentified firm as a signal about its type. Formally this implies that we are in the context of a signalling game in which firms have to take into account that consumers have beliefs about the price a clean firm would set. The anticipated outcome of this signalling game (in stage 4) is then taken into account by the firm when choosing the probability of clean production m_i in stage 2.

We analyze this signalling game and its interaction with the technology choice (and the expectation thereof) in detail in Appendix D. We proceed in four steps. We first show that in

our setup all Perfect Bayesian Equilibria (PBEs) of the signaling game are pooling equilibria in which both firm types set the same price. In a second step, we characterize the equilibrium clean technology choice i.e. the probability the firm assigns to clean production in stage 2 $m(p') \in (0,1)$ consistent with the fact that stage 4 delivers a pooling PBE at some price p'. In the third step, we eliminate PBEs from the second stage using the intuitive criterion (Cho and Kreps, 1987) and the undefeated equilibrium concept (Mailath et al. 1993) as equilibrium refinements. This finally allows us to fully characterize the solution of the game in step four.

The result is that the unique Perfect Bayesian Nash equilibrium consistent with our equilibrium refinements has the following characteristics: first, consumers correctly anticipate the probability of clean production chosen by the firm, i.e. $q_i^e = m_i$. Second, when the firm type is not revealed, both a clean and a dirty firm set the monopoly price of a clean firm for a given q_i^e. This justifies the assumptions made above and fully characterizes stages 2-5 of the game.

4.6 Stage 1

In stage 1 firms simply compare profits from producing in Home to expected profits from offshoring and offshore if $E(\Pi_i^c) = E(\Pi_i^d) > \Pi^H$ with $\Pi^H = \frac{(1-c_H)^2}{4}$.[13]

5 Offshoring and NGO emergence

We can now start deriving the equilibrium of the model and analyze the determinants of offshoring and NGO emergence. To do so, we first consider the optimal firm choices taking NGO (non-) existence as given. We then determine under which conditions an NGO emerges and finally combine the two sides to derive the equilibrium.

5.1 Firm Offshoring Cutoffs

We start from the firm's perspective and characterize the conditions for offshoring with and without the existence of an NGO. Define $\bar{c}_i$ for $i \in \{0,1\}$ the cutoff cost levels below which the firm chooses production abroad when the NGO does not enter ($i = 0$) and when an NGO emerges in response to the offshoring decision ($i = 1$). Recall that the cost c includes trade costs. Different levels of c can therefore be interpreted as different levels of globalization. $\bar{c}_i$ is

[13] In the corner solution where the firm always chooses the clean technology, the condition is simply given by $E(\Pi_i^c) > \Pi^H$, which holds iff $c < c_H$. As discussed above, this case is ruled out by assumption 1.

pinned down by $\Pi_H = E(\Pi_i)(\bar{c}_i)$, which gives:

$$\frac{(1-c_H)^2}{4} = \pi_i \frac{(1-\bar{c}_i)^2}{4} + (1-\pi_i)\frac{(q_i^e - \bar{c}_i)^2}{4}. \tag{12}$$

Using equation (11) and solving the resulting quadratic equation for $(1-\bar{c}_i)^2$ delivers:

$$(1-\bar{c}_i)^2 = \frac{1-\pi_i}{\pi_i}\ 2\ \Delta^2\ \left[\left(1 + \frac{(1-c_H)^2}{1-\pi_i}\ \frac{1}{\Delta^2}\right)^{\frac{1}{2}}\ -1\right].$$

From this we have the first immediate result:

Proposition 1. *In an interior equilibrium where both technology types are chosen with a positive probability,*

i) The cutoff productivity level below which the firm chooses production abroad is given by:

$$\begin{aligned} \bar{c}_0 &= \bar{c}(c_H, \pi_0, \Delta) \text{ without the existence of an NGO} \\ \bar{c}_1 &= \bar{c}(c_H, \pi_1, \Delta) \text{ with the existence of an NGO} \end{aligned}$$

where $\bar{c}_i(c_H, \pi_i, \Delta)$ is the following function:

$$\bar{c}_i(c_H, \pi_i, \Delta) = 1 - \sqrt{\frac{1-\pi_i}{\pi_i}\ 2\ \Delta^2\ \left[\left(1 + \frac{(1-c_H)^2}{1-\pi_i}\ \frac{1}{\Delta^2}\right)^{\frac{1}{2}}\ -1\right]}.$$

ii) The function $\bar{c}_i(c_H, \pi_i, \Delta)$ is increasing in π_i, increasing in c_H, and decreasing in Δ.

Proof. See the appendix. □

Note that due to $\pi_1 > \pi_0$ we have $\bar{c}_1 > \bar{c}_0$. This implies that the internationalization decision of the firm does depend on whether it expects an NGO to emerge in response to offshoring or not. In our model the firm offshores for higher trade costs if it anticipates NGO emergence implying some complementarity between the two. This point is discussed in detail in Section 6.

The intuition for the remaining comparative statics is straightforward. First, a larger domestic cost c_H makes it more profitable for the firm to move abroad, making offshoring profitable for a larger range of trade costs both with and without an NGO. Second, a higher cost advantage of dirty production Δ increases the incentive to use the dirty technology. As consumers take this into account when forming their beliefs, demand is reduced, discouraging offshoring in the first

place. As a consequence the firm only offshores when trade costs are sufficiently low, translating into a lower cutoff cost level.

Figure 1 provides a graphical illustration of Proposition 1. The three solid lines represent expected profits in the cases of autarky $E(\pi_H)$, offshoring without NGO $E(\pi_0)$ and offshoring with NGO $E(\pi_1)$. The thin dashed lines depict the cutoff levels with and without an NGO $\bar{c}_1$ and $\bar{c}_0$ at the intersections of the expected offshoring profits with autarky profits.

-Figure 1 here-

5.2 NGO Emergence Cutoff

If the firm produces in Home, there is no asymmetric information between the firm and consumers and therefore no NGO emerges. When the firm decides to offshore in stage 1, the motivated agent decides on whether to raise funds from consumers or not. If she can raise sufficient funds, an NGO emerges.

Given that there is a cost of activating additional donors, but no cost of requesting additional funds from activated donors, the NGO will always extract the maximum donation from each donor. This maximum donation is given by the difference between expected consumer surplus with and without an NGO. Using equations (5) and (11) this maximum donation can be expressed as a function of the cost level c:

$$x^* = x^*(c, \pi_1, \pi_0, \Delta) = c\ (\pi_1 - \pi_0)\frac{(1-c)^2}{8} + \frac{(1-c)^4}{32}\frac{1}{\Delta}\left[\frac{\pi_1^2}{1-\pi_1} - \frac{\pi_0^2}{1-\pi_0}\right]\left(2 + \frac{1}{\Delta}\right). \quad (13)$$

This maximum donation is the gain in consumer surplus that an individual consumer has when the NGO enters. When the consumer is activated and faces the take it or leave it offer by the NGO, this is the maximum donation the NGO can extract from the donor. We can therefore state the following proposition:

Proposition 2. *In an interior equilibrium where both technology types are chosen with a positive probability,*

(i) The maximum donation the NGO can extract from an individual donor x^* is given by equation (13)*.*

(ii) Stronger globalization leads to an increase in the maximum donation when $\pi_0 > 1/10$ (i.e. $\frac{\partial x^*}{\partial c} < 0$).[14]

[14] Using numerical maximization, it can be shown that the much lower value of $\pi_0 > 0.019$ is a sufficient

(iii) A higher detection probability with an NGO π_1, a lower detection probability without an NGO π_0 and a lower cost advantage of dirty production Δ imply a higher maximum donation, (i.e. $\frac{\partial x^*}{\partial \pi_1} > 0$, $\frac{\partial x^*}{\partial \pi_0} < 0$ and $\frac{\partial x^*}{\partial \Delta} < 0$).

Proof. See the appendix. □

It is an important feature of our model that consumers do not donate due to a warm-glow motive, but due to the fact that NGO emergence can make them possibly better-off through its effect on the technology choice of the firm. This is why the difference in consumer surplus with and without an NGO determines the maximum donation. For most values of the initial detection probability π_0 falling trade costs lead to an increased willingness to fund an NGO. The same holds if NGO emergence strongly improves the detection probability (high $\pi_1 - \pi_0$). A high cost advantage of dirty production goes in the opposite direction. The intuition here is that a high Δ makes the use of the dirty technology more likely, which implies that the same increase in the detection probability induced by NGO emergence has a lower effect on the expected probability of clean production. This reduces the impact of an NGO on consumer surplus and therefore the willingness to fund the NGO.

With the linear cost function the marginal cost of activating an additional donor is constant. The NGO can therefore raise additional funds by activating more donors if the marginal cost of doing so is lower than the individual willingness to pay i.e. $x^* > y$. As the NGO needs to finance a fixed amount X, it can only enter if the funds it can raise when all consumers are activated are large enough, namely $x(c) - y \geq X$.

Define $\bar{c}_N$ as the cost level for which (conditional on offshoring of the firm) the NGO would be able to raise sufficient funds to start operating. With the linear cost function this is the case for $\nu^* = 1$. The NGO cutoff is therefore pinned down by:

$$E(W_1)(\bar{c}_N) - E(W_0)(\bar{c}_N) - y = x^*(\bar{c}_N, \pi_1, \pi_0, \Delta) - y = X. \tag{14}$$

and we have the following result:

Proposition 3. *In an interior equilibrium where both technology types are chosen with a positive probability, and assuming that $\pi_0 > 1/10$, we have:*

condition.

(i) The NGO cutoff for entry $\bar{c}_N$ is implicitly determined by the following equation :

$$x^*(\bar{c}_N, \pi_1, \pi_0, \Delta) = X + y. \tag{15}$$

(ii) A low fixed cost of entry X and low fundraising marginal cost y, a high NGO detection probability π_1, a low baseline detection probability π_0 as well as a low cost advantage of dirty production Δ increase the NGO cutoff, i.e. $\frac{\partial \bar{c}_N}{\partial \pi_1} > 0$, $\frac{\partial \bar{c}_N}{\partial \pi_0} < 0$ and $\frac{\partial \bar{c}_N}{\partial \Delta} < 0$.

Proof. See the appendix. □

The NGO cutoff $\bar{c}_N$ is the highest cost level for which - conditional on offshoring of the firm - the NGO can raise sufficient funds to operate. Equation (15) delivers an implicit solution for this cutoff cost level. Our model implies a direct link between the benefit consumers have from NGO emergence and the funding possibilities of the NGO. Anything that strengthens the favorable effect of the NGO on consumer surplus through affecting the technology choice of the firm improves the funding opportunities of the NGO and allows it to enter for higher cost levels. Therefore the comparative statics of the NGO cutoff with respect to π_1, π_0 and Δ correspond to the ones for the individual donation in Proposition 2.

-Figure 2 here-

Figure 2 provides a graphical illustration. The graph plots all the elements concerning the optimal choices of the firm (expected profits and cutoffs) from Figure 1 as thin dotted lines. The thick solid lines represent expected consumer surplus with an NGO $E(W_1)$ and without an NGO $E(W_0)$. Their difference gives the willingness to donate $x^*(c)$. It can be seen that for this parametrization the willingness to donate increases for lower values of c i.e. for stronger globalization. The vertical dashed line represents the resulting NGO cutoff $\bar{c}_N$.

An increase in the difference between the two solid lines implies an increased willingness to donate and therefore shifts the NGO cutoff level to the right. Such an increase can be triggered by an increase in π_1 (shifting the solid curve up) or a decrease in π_0 (shifting the dashed line down). A decrease in the cost advantage of dirty production Δ also increases the difference between the two curves by shifting up both of them, but the solid line by more. The intuition is that a lower Δ reduces the incentive for dirty production and therefore makes the firm more responsive to an increased detection probability.

6 Equilibrium

On the side of the NGO we now know for which cost levels the NGO enters. On the side of the firm we know the cutoff levels for offshoring conditional on NGO emergence. We can put these different elements together to derive the equilibrium of the model and to analyze under which conditions the joint emergence of the NGO and offshoring occurs.

Define $\bar{c}_g$ as the offshoring cutoff (where 'g' may stand for 'globalization'). This cutoff is given by $\bar{c}_g = \max\{\min\{\bar{c}_1, \bar{c}_N\}, \bar{c}_0\}$. There are therefore three cases to distinguish:

- *Case 1:* $\bar{c}_g = \bar{c}_1$ iff $\bar{c}_N \geq \bar{c}_1$. In this case, the NGO enters directly when the firm starts offshoring.

- *Case 2:* $\bar{c}_g = \bar{c}_N$ iff $\bar{c}_0 \leq \bar{c}_N \leq \bar{c}_1$. In this case, the firm would want to enter conditional on NGO emergence already for high cost levels, but here NGO emergence is the constraining factor.

- *Case 3:* $\bar{c}_g = \bar{c}_0$ iff $\bar{c}_N \leq \bar{c}_0$. In this case, the firm enters even without NGO emergence.[15]

To determine the offshoring cutoff $\bar{c}_g$, we need to determine in which of the three cases we are. To do so, we can use the fact that at $\bar{c}_N$ (and only there) we have $\nu^* = 1$. In which of the three cases we are depends on when (conditional on offshoring of the firm) the NGO is able to raise sufficient funds.

Proposition 4. *In an interior equilibrium where both technology types are chosen with a positive probability, assuming that $\pi_0 > 1/10$ and using $x^*(c, \pi_1, \pi_0, \Delta)$ as defined in equation (13) then,*

i) The globalization cutoff $\bar{c}_g$ is given by:

$$\bar{c}_g = \bar{c}_1 \quad \textit{if} \quad x^*(\bar{c}_1, \pi_1, \pi_0, \Delta) \geq X + y \qquad \textit{(Case 1)} \tag{16}$$

$$\bar{c}_g = \bar{c}_N \quad \textit{if} \quad x^*(\bar{c}_0, \pi_1, \pi_0, \Delta) \geq X + y \geq x^*(\bar{c}_1, \pi_1, \pi_0, \Delta) \qquad \textit{(Case 2)} \tag{17}$$

$$\bar{c}_g = \bar{c}_0 \quad \textit{if} \quad x^*(\bar{c}_0, \pi_1, \pi_0, \Delta) \leq X + y \qquad \textit{(Case 3).} \tag{18}$$

ii) Further, the globalization equilibrium is characterized in the following way:

iia) For $\bar{c}_g < c < c_H$, there is no NGO and no offshoring.

iib) When $\bar{c}_g \geq \bar{c}_0$ (cases 1 and 2) and $c \leq \bar{c}_g$ there is offshoring with NGO monitoring and the clean technology is used with probability q_1^e.

iic) When $\bar{c}_g \leq \bar{c}_0$ (case 3) we have the following: for $\bar{c}_N < c \leq \bar{c}_g$ there is offshoring without

[15] Note that the three cases are not mutually exclusive: $\bar{c}_g = \bar{c}_1 = \bar{c}_N$ is possible as well as $\bar{c}_g = \bar{c}_0 = \bar{c}_N$.

NGO and the clean technology is used with probability q_0^e; for $c \leq \bar{c}_N$, there is offshoring with NGO monitoring and the clean technology is used with probability q_1^e.

Proof. Proof see appendix. □

-Figure 3 here-

Figure 3 provides a graphical illustration of the three different cases. As before, the thin dotted lines represent expected profits with and without an NGO and under autarky from Figure 1. The thick dotted lines are the corresponding cutoffs $\bar{c}_1$ and $\bar{c}_0$.

The graph considers three different funding requirements for the NGO, implying three different NGO cutoffs. These are represented by the three solid vertical lines. In case 1 ($\bar{c}_{N1}$) the funding requirement is relatively low, so that $(\bar{c}_{N1}) > \bar{c}_1$. Conditional on offshoring of the firm, at $\bar{c}_{N1}$ the NGO would be able to raise sufficient funds to operate setting $\nu = 1$. But at this high cost level we have $\Pi_H > E(\Pi_1) > E(\Pi_0)$ so that it does not pay for the firm to offshore even in the presence of an NGO. In case 1 the globalization cutoff $\bar{c}_g$ is therefore at $\bar{c}_1$, as at this cost level the firm offshores and the NGO enters activating only a fraction $\nu < 1$ of potential donors.

In case 2 ($\bar{c}_{N2}$) the funding requirement for the NGO is higher and the difference between $\tilde{W}_1$ and $\tilde{W}_0$ is only large enough to finance the NGO for a cost level below $\bar{c}_1$. As in this cost range the firm only offshores if an NGO exists, the globalization cutoff is given by $\bar{c}_g = \bar{c}_{N2} < \bar{c}_1$.

In case 3 the funding requirement is so high that an NGO emerges only for very low cost levels. This implies that there is a range of trade costs for which the firm offshores without the NGO, while the NGO only emerges for even lower trade costs.

Proposition 4 implies that the model can generate NGO emergence and offshoring as joint phenomena depending on the level of trade costs. This is underlined by the need to distinguish cases 1 and 2: in these two cases offshoring and NGO emergence can only occur jointly with the one having to 'wait' for the other. Moreover, cases 1 and 2 imply that the move from an initial equilibrium with high trade costs and production in Home to an offshoring equilibrium with NGO emergence can be triggered by a fall in trade costs. This is in line with the stylized facts and examples of offshoring and NGO emergence as interrelated phenomena driven by falling trade costs and economic integration outlined in the introduction.

Before discussing case 3 in Proposition 4, we state the following corollary on the complementarity of NGO activity and offshoring:

Corollary to Proposition 4: *(i) When* $\min\{\bar{c}_1, \bar{c}_N\} > \bar{c}_0$ *(cases 1 and 2 in Proposition 4) there is complementarity between the NGO and offshoring i.e. the endogenous emergence of the NGO enables the firm to offshore at higher cost levels* c *than in a world without the possibility of NGO emergence. (ii) A better NGO detection technology* π_1 *and a lower cost advantage of dirty production* Δ *both induce offshoring for higher levels of trade costs.*

Proof. See the appendix. □

As consumers care about clean production, the firm faces a credibility problem. Offshoring creates asymmetric information on the technology implemented implying that there is an incentive to use the dirty technology. As consumers take this into account when forming expectations on the probability of clean production, demand is reduced. This makes it difficult for the firm to even benefit from the 'clean' efficiency gains from offshoring represented by the cost difference $c_H - c$ as it cannot credibly commit to clean production.

The NGO provides a partial solution to this commitment problem. It increases the detection probability and therefore the expected punishment of dirty production as well as the benefit of clean production, reducing the incentive of the firm to choose the dirty technology, which is the case for a higher π_1 and a lower Δ. Consumers take this into account, assign a higher probability to clean production and increase demand in the case that the good is not identified. NGO emergence therefore facilitates offshoring by the firm. This has a direct implication for the view of the firm on NGO campaigns. While from an ex-post perspective a dirty firm facing an NGO campaign makes zero profits, the threat of facing a campaign helps the firm from an ex-ante perspective to convince consumers that it has a strong incentive to choose the clean technology thereby increasing demand if its type is not revealed. Our model is therefore in line with the popular view that NGO campaigns harm the firms they target, but indicates at the same time a mechanism through which NGO campaigns can be beneficial for firms from an ex-ante perspective by changing firm behavior such that less campaigns occur.

The complementarity between NGO emergence and offshoring that we find in cases 1 and 2 is directly related to the fact that we start from a situation where all consumers care about clean production. In Section 8.4 we briefly discuss how a model could look in which the fraction of caring consumers, and therefore preferences, are endogenized.

In case 3 in Proposition 4 there is a range of trade costs for which the firm offshores without NGO emergence. This case is relevant for at least two reasons. First, it illustrates that despite

the complementarity between NGO emergence and offshoring, the two do not always have to occur together. If the cost of creating an NGO is too high or the impact on the firm too low, we can have cases of offshoring without NGO emergence. Second, when we think of globalization as a sequence of reductions in trade costs, we can have a pattern over time starting with production in Home, over only offshoring to offshoring plus NGO emergence in the three equilibria corresponding to the three levels of trade costs.

The model can also be used to think about sectoral differences, for example to address the question of why in some sectors offshoring triggers NGO emergence but does not do so in others. Depending on parameter values, the model can generate both patterns. In cases 1 and 2 offshoring and NGO emergence occur jointly, while in case 3 it is possible that the firm offshores without an NGO being present. In which case we are, depends on whether the value of the NGO cutoff $\bar{c}_N$ is to the left or the right of the firm's offshoring cutoff without an NGO $\bar{c}_0$. Proposition 3 implies (under the maintained assumption on the minimum value of π_0) that offshoring triggering NGO emergence is more likely in sectors in which an NGO has a high cost-efficiency (determined by the parameters y, π_1 and π_0).

Several features of NGO campaigns can be linked to these parameters. Many NGO actions are related to final consumption products of well-known brands with large market size, which is arguably the case - to different extents - for all examples discussed in Section 2. It appears plausible that consumers can be more easily convinced to donate if the campaign goes against a firm they know well, implying a lower y. Moreover, final products of big brands can be more easily liked to the brand than homogeneous products, an example being shirts with Mango, Bennetton and C&A labels found in the ruins of the 'Rana plaza' garment factory in Bangladesh in 2013 (high π_1). Offshoring often takes place through a very untransparent chain of subcontractors requiring an investigative effort to link production plants, e.g. in China or Bangladesh to US or European brands (low π_0). Our model is in line with these observations and can by the same token capture the pattern that there are much fewer campaigns against producers of more homogeneous products (high y, low π_1) or firms having their own production plants abroad (high π_0).

7 Welfare Analysis

We can now use our model to study the welfare effects of offshoring with NGO emergence. Due to the lower cost of clean production in Foreign, there is scope for standard welfare gains from

offshoring. We show that due to the misalignment of consumer and firm interests in the presence of credence goods, there can be a range of trade costs for which total welfare decreases when the firm decides to offshore.

Total welfare in our model is defined as the sum of expected firm profits and expected consumer surplus net of NGO fixed costs, which are incurred by consumers. We define net expected consumer surplus as:

$$E(W)_{net} = \begin{cases} W_H & \text{if production in Home} \\ E(W_1) - X & \text{if offshoring with NGO} \\ E(W_0) & \text{if offshoring without NGO.} \end{cases}$$

Expected consumer surplus $E(W)$ is defined in the same way, just without the NGO fixed costs X in the case of offshoring with an NGO.

Proposition 5. *In an interior equilibrium where both technology types are chosen with a positive probability and* $\min\{\bar{c}_1, \bar{c}_N\} > \bar{c}_0$ *(cases 1 and 2, where offshoring is associated with NGO entry), expected consumer surplus* $E(W)$ *and net expected consumer surplus* $E(W)_{net}$ *are non-monotonic in further trade integration (i.e. a reduction in the cost level c). In particular, in case 1 and 2 it is possible that for a range of costs below the globalization cutoff* $\bar{c}_g$*, expected and net expected consumer surplus* $E(W)$ *and* $E(W)_{net}$ *are below the autarky level of consumer surplus* W_H*.*

Proof. See the appendix. □

-Figure 4 here-

Figure 4 provides a numerical example for the pattern of consumer surplus in case 1. Again the thin vertical dashed lines represent the cutoff levels $\bar{c}_1$ and $\bar{c}_0$. The bold vertical dashed line is the NGO cutoff $\bar{c}_N$. As we have $\bar{c}_N > \bar{c}_1$, we are in case 1. The horizontal dashed line represents consumer surplus in autarky and the bold solid line depicts net expected consumer surplus $E(W)_{net}$ for the different values of c considered.

By definition, for all cost levels below $\bar{c}_1$ it is profitable for the firm to offshore. We can see from the graph that this is not in line with the interest of consumers. Upon offshoring at $\bar{c}_1$, net expected consumer surplus jumps below its autarky level and only gradually increases for lower cost levels. It is only to the left of the intersection of autarky surplus (dashed horizontal line) and net expected surplus that the lower prices make consumers gain from offshoring of the firm.

This implies that the firm offshores 'at the expense' of the consumers, who see their expected surplus decrease.

This negative effect on net expected consumer surplus can be decomposed into two components. The two dashed curves in the graph depict the function $E(W_1)$ and $E(W_1) - X$. We can see that at the cutoff $\bar{c}_1$ the expected surplus conditional on offshoring with an NGO $E(W_1)$ represented by the upper curve is below autarky welfare. In addition, the firm uses its first mover advantage and imposes the burden of financing the NGO onto consumers, so that $E(W_1) - X$ is even lower (lower dashed line).

While the latter effect is quite simple, the former is somewhat less obvious. Technically it stems from the fact that the first term on the r.h.s. of equation (5) includes a $q^e < 1$, while it is not present in the corresponding term in equation (8). The reason for this is that when a firm chooses the dirty technology, it is compensated for the risk of facing zero demand with probability π_i by the fact that it benefits from a cost reduction of Δ. This cost advantage, however, is not passed on to consumers, as a dirty firm mimics the price setting of the clean firm. This implies that the consumer is not compensated by a lower price for the possibility that her surplus drops to zero when the good is revealed to be dirty. This is reflected by the presence of the term q^e in the first term of the r.h.s. of equation (5).

-Figure 5 here-

Before we analyze how this misalignment of firm and consumer interests affects overall welfare effects of offshoring, consider Figure 5, which provides an illustration of case 2. In case 2 we have $\bar{c}_g = \bar{c}_N < \bar{c}_1$, so that the binding constraint is the possibility of NGO emergence. This is reflected by the fact that the bold dashed line is now to the left of $\bar{c}_1$. In this numerical example we obtained this by increasing the NGO funding requirement X.

In this example we still get a fall in net consumer surplus. The comparison to the upper dashed curve shows that this is due to the fact that consumers have to finance the NGO. If the bold dashed line representing the NGO cutoff was to the left of the point where autarky and net expected surplus intersect, there would be a positive jump in net consumer surplus upon offshoring.

Total welfare is given by the sum of net expected consumer surplus and firm profits. The discrete fall (or increase) in net expected consumer surplus also translates into the pattern observed for total welfare.

Proposition 6. *In an interior equilibrium where both technology types are chosen with a positive*

probability when $\min\{\bar{c}_1, \bar{c}_N\} > \bar{c}_0$ *(cases 1 and 2), expected total welfare is non-monotonic in further trade integration (i.e. a reduction in the cost level c). In particular, in case 1 and 2 it is possible that for a range of costs below the globalization cutoff* $\bar{c}_g$*, expected social welfare* W^* *is below its autarky level.*

Proof. See the appendix. □

-Figure 6 here-

Figure 6 illustrates the pattern in case 1, where offshoring takes place at $\bar{c}_1$. Like in Figure 4 we observe the downward jump in net expected consumer welfare given by the bold dashed curve. The dotted line above represents expected profits. As in case 1 offshoring takes place at $\bar{c}_g = \bar{c}_1$ with $E(\Pi_1)(\bar{c}_1) = \Pi_H$, there is no discontinuity at the cutoff. This in turn implies that the fall in net expected consumer surplus is reflected by a fall in total expected welfare (given by the solid line) at $\bar{c}_1$. Due to the fact that the firm offshores at the expense of consumers, also total expected welfare falls when the firm offshores at $\bar{c}_1$. This effect is compensated for lower values of c as on the one hand expected profits increase and on the other hand lower costs are passed on to consumers increasing expected consumer surplus.

-Figure 7 here-

Figure 7 illustrates an example of case 2 where the NGO cutoff $\bar{c}_g = \bar{c}_N$ is relatively close to the cutoff $\bar{c}_1$. In this case we still observe a discrete fall in net expected consumer surplus. At the same time there is a discrete increase in expected profits due to the fact that conditional on NGO existence, the firm would want to offshore already at $\bar{c}_1$. In the numerical example chosen here, the overall effect on expected welfare is still negative. For lower values of $\bar{c}_N$ (e.g. due to a higher NGO funding requirement X) the overall effect on expected welfare can also be positive.

As outlined above, the firm has the first mover advantage and decides whether to relocate production or not. When the firm takes the offshoring decision in stage 1, it anticipates whether this will lead to NGO emergence. From the viewpoint of the NGO this has a somewhat paradoxical implication. The aim of the NGO is to minimize dirty output. One way to achieve this would be to prevent the firm from offshoring in the first place by keeping the globalization cutoff $\bar{c}_g$ as low as possible. So if the NGO could commit to not entering for any cost level above $\bar{c}_0$, the firm could only offshore at $\bar{c}_0$. But as there is no such commitment device for the NGO, the firm can 'force the NGO into being' for cost levels below the NGO cutoff $\bar{c}_N$, as, conditional on offshoring, it will be optimal for the NGO to start working and for $c < \bar{c}_N$ it will also be able to raise sufficient funds to do so. In cases 1 and 2 the possibility of NGO emergence therefore

implies that for the cost range between $\bar{c}_0$ and $\bar{c}_1$, expected dirty output is higher than in a world without any possibility of NGO emergence.

The situation for consumers is somewhat different. They also suffer from the fact that the firm can use its first mover advantage taking into account that conditional on offshoring, consumers may be willing to finance an NGO. Consumer interests are, however, not perfectly aligned with the objective of the NGO. While the NGO only cares about dirty output, consumers also care about prices. Figures 4 and 5 illustrate cases where consumers lose from offshoring when trade costs fall such that the overall costs c fall from just above $\bar{c}_g$ to just below. In both cases, however, consumer surplus is above the autarky level for cost levels well above $\bar{c}_0$. In these situations, the price decreases due to the standard gains from offshoring more than compensate consumers for the presence of some dirty products.

8 Discussion

All the results from the baseline model are now established. Before turning to an extension with endogenous regulation in Foreign in Section 9, we provide a discussion of our main assumptions and of possible modeling alternatives.

8.1 Timing assumptions

Like in all games of the type considered here, timing plays a key role. In particular the result that the firm can offshore at the expense of consumers crucially depends on the first mover advantage of the firm when taking the offshoring decision. Of all possible timing assumptions the one where the firm first takes the offshoring decision and consumers and NGO only respond to this decision appears the most plausible to us. In reality most NGOs seem to emerge in response to an existing problem rather than to prevent a problem from arising in the first place. This is better captured by letting the firm decide on offshoring first taking into account how this decision will affect consumer choices.

8.2 Perfect and costless enforcement in North

A simplifying assumption in our model is that enforcement in North is perfect and costless. Relaxing these assumptions would strongly reduce tractability of the model. We conjecture that as long as enforcement in North is such that it is more cost-efficient than the NGO action (otherwise an NGO would also emerge in Home), the basic mechanics of the model are unaffected.

On the one hand offshoring would still increase the credence good problem and on the other the financial burden on consumers would increase with offshoring for the cases where NGO and offshoring occur jointly.

8.3 Only a fraction of consumers cares

The assumption that all consumers care about clean production is important in our model. In fact this assumption is complementary to the assumption that the credence good is provided by a single monopolist. Taken together, these assumptions generate tractability and allow us to focus the analysis on the joint emergence of NGOs and offshoring in the context of globalization.

There are two interesting possible extensions when considering consumer preferences for clean production. First, allowing for a fraction of consumers not to care about clean production and second, endogenizing this fraction.

The first extension of allowing for a fraction of consumers not caring about clean production, would render the assumption of a single monopolist questionable. If some consumers care and others do not, a natural modeling choice would be allowing for entry of at least one additional firm specializing in the provision of the dirty good. Since consumers care about the price and the credence characteristic, the markets would not be perfectly segmented and one would face a complex interaction in the price setting of the firms very likely rendering the model intractable. Despite this additional complexity, the factors determining whether an NGO emerges in response to offshoring should remain unchanged.

8.4 Endogenizing the fraction of caring consumers

The second extension would be endogenizing the fraction of caring consumers and therefore endogenizing preferences. This would require a dynamic framework. Besides this being beyond the scope of this paper, we would not expect this to affect the main results of the paper.

One qualification is, however, in order. In a dynamic model in which (most) consumers initially do not care about clean or dirty production, the NGO would first have to invest into awareness creation to build a sufficiently large stock of caring consumers. Once this point is reached, the main mechanisms outlined in our paper should apply. One interesting aspect of such a setup would, however, be that the result that the NGO provides a service to the firm would be more nuanced. It should still hold once the NGO has build up a sufficient stock of caring consumers. But it should not hold in the periods before this stock is reached: in these

periods, the firm would not want to commit to clean production as the demand for a dirty good would still be high enough. One would therefore expect such a setup to generate a much more nuanced and critical view of firms on NGOs than our setup with initially caring consumers generates. This interesting analysis would require a very different model from the one outlined here, so we leave it for future research.

8.5 Media and for-profits as providers of monitoring

In our model the NGO provides monitoring services in Foreign that cannot be provided by the Home government. As consumers demand such services, they could in principle also be provided by for-profit firms, similar to rating agencies in the financial sector. Not-for-profit organizations have several key advantages over for-profit firms. Surveys repeatedly found that NGOs are the most trusted institutions before government, business and media.[16] In addition, NGOs may have cost advantages over for-profit firms as they can largely rely on the help of volunteers, e.g. for verifying compliance of firms on the retail level. Moreover, NGOs are less likely than for-profit organizations to run the risk of being sued by the targeted firm, as they often have loose organizational structures that can be easily dissolved and recreated under new names. In fact, it seems hard to imagine that campaigns including a call for a consumer boycott of e.g. Nike, Staples Inc. or Shell could be run by for-profit firms, without enormous legal costs putting the profit objective at stake.

In reality the media play an important role in the context of campaigns. As the success of NGO campaigns depends strongly on media coverage, there can be a complementarity in the activities of the two. We do not model media activity explicitly. In a reduced-form way this complementing activity is represented by the strength of the increase in the detection probability. The impact of media activity that leaves the NGO efficiency unaffected is subsumed by the detection probability in the absence of an NGO π_0.

8.6 Extension to many firms

The model could easily be extended to a model with a large number of monopolists in a large number of sectors. In terms of interpretation this would have the nice feature that the - somewhat abstract - probability of clean production m_i would also represent the fraction of clean firms in the economy.

[16] See e.g. the Edelman Trust Barometer, http://www.edelman.com/insights/intellectual-property/2012-edelman-trust-barometer/about-trust/global-results/ accessed September 2014.

Introducing more than one firm in the same sector would render the analysis much more complex. For the case of monopolistic competition, the model would most likely lose its tractability. Strategic interactions between firms would be absent in such a setting and firms would still make monopoly profits, we would therefore expect the main mechanisms from our paper to be active also in such a setting.

In an oligopolistic setup there would be strategic interactions between firms possibly leading to interesting interactions between firms and the NGO. Such a framework would allow addressing interesting questions like which firm does the NGO optimally target (e.g. the market leader) and how does a switch in technology spill over from one firm to the other? A first attempt to do this is recent work by Lyon and Salant (2013). These interesting considerations are however, somewhat orthogonal to the interaction of NGO emergence and offshoring decision of firms in the context of globalization which are the focus of our paper.

In a setup with perfect competition, there would be no monopoly rents. The mimicking result would therefore break down and all firms would reveal their type. The informational asymmetry which is at the heart of our model would not exist. Besley and Ghatak (2007) analyze such a setup.

9 Endogenous public regulatory systems and NGOs

In our context, the rationale for the existence of NGO monitoring comes from the limits imposed on national regulatory powers by globalization. Governments can directly impose their own regulations (reflecting to some extent the preferences of their citizens as well as their capacity constraints) only within their sovereign jurisdictions. National asymmetries in these regulatory systems and the fact that international trade creates incentives for firms to locate in lower cost jurisdictions in order to satisfy the preferences of consumers in other jurisdictions is at the root of the demand for credible private regulation. Hence, the non-existence of efficient regulatory systems in southern countries with weak regulatory capacity induces the emergence of NGOs financed by northern consumers to monitor international trade in credence goods.

In this section, we show that when national regulatory capacities can be endogenously affected by governments, the causality can also go the other way around. Indeed the expectation of the emergence of international civil society organizations may actually crowd out incentives for southern governments to implement regulatory systems. The reason for this is simple. A southern government may just anticipate that by not investing into the necessary monitoring

capacity infrastructures, caring consumers in export markets will be ready to finance NGOs to play that role. When this happens, there is an incentive for the southern authorities to shift the cost of trade regulation to these consumers. The existence of NGOs as private international regulatory mechanisms may therefore crowd out the implementation of effective public regulations in exporting countries.

To illustrate our point, we extend our previous framework in the following way. Consider now that the firm (from Home) can offshore to Foreign where the government may invest at some cost K in some effective regulatory mechanism. Such capacity ensures that the offshoring firm has no choice but to produce in Foreign with the clean technology. Assume also that the Foreign government cares about total local employment which in the context of our constant return to scale technology is simply proportional to the firm local output. Formally the objective function of the Foreign government writes as:

$$W^S(Q, I) = \lambda Q - K \cdot I$$

where I is an indicator variable taking the value of one if the government invests in regulation and a value of zero otherwise. $\lambda > 0$ is a parameter reflecting the intensity with which the Foreign government cares about local production.[17] We consider the following timing. In an additional first stage 0, the Foreign government decides about investing or not in trade regulatory capacity, maximizing its expected objective function. After that, the game follows as before and is solved by backward induction.

When the Foreign government decides to invest (i.e. $I = 1$), the firm is facing perfect public monitoring in Foreign. Whenever it delocalizes production, it is then obliged to use the clean technology at local cost c. Given this, there is no demand by Home consumers for private regulation and therefore no funding for an NGO. Because $c_H > c$, the firm always delocalizes and produces the monopoly output level $Q^M = \frac{1-c}{2}$. The Foreign government objective function writes simply as:

$$W^R(c) = W^S(Q^M, 1) = \lambda \frac{1-c}{2} - K.$$

When the Foreign government does not invest, (i.e. $I = 0$), the game is exactly as before.

[17] λ could also reflect local wages, in which case the southern government cares about the wage bill generated by the trade activity of the firm.

When there is delocalization by the firm in Foreign, expected local production is given by

$$E(Q_i) = \pi_i q_i^e \frac{(1-c)}{2} + (1-\pi_i)\frac{(q_i^e - c)}{2} \tag{19}$$

with $\pi_i = \pi_0$ when there is no NGO and $\pi_i = \pi_1$ when there is the NGO. Correspondingly, the probability q_i^e of the firm choosing the clean technology is given by

$$q_i^e - c = \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi_i}{1-\pi_i}.$$

Equation (19) reflects the fact that with probability π_i, consumers learn about the true technology of the firm. When that happens to be the clean technology (with probability q_i^e), there is monopoly production $(1-c)/2$. On the other hand with probability $(1-\pi_i)$, nothing is revealed to consumers. As a consequence, realized output under both the clean and the dirty technology is given by $(q_i^e - c)/2$.

From this, one can write the expected utility of the Foreign government under a given monitoring probability π_i with $i \in \{0, 1\}$:

$$W^{NR}(\pi_i, c) = \lambda E(Q) = \lambda \pi_i q_i^e(\pi_i, c)\frac{(1-c)}{2} + \lambda(1-\pi_i)\frac{(q_i^e(\pi_i, c) - c)}{2}$$

with $q_i^e(\pi_i, c)$ given by

$$q_i^e(\pi_i, c) = c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi_i}{1-\pi_i}.$$

After substitution this rewrites as

$$W^{NR}(\pi_i, c, \Delta) = \lambda \pi_i \frac{(1-c)}{2}\left[c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi_i}{1-\pi_i}\right] + \lambda \frac{\pi_i}{\Delta} \cdot \frac{(1-c)^2}{2}. \tag{20}$$

Under our parameter restrictions and assumption 1, it can be shown that $W^{NR}(\pi_i, c, \Delta)$ is a decreasing convex function of the local cost c and is an increasing function of the monitoring probability π_i (see the appendix).

To highlight in the starkest way our mechanism of regulatory cost shifting, we assume that:

Assumption 2 : For all $c \in [\Delta, c_H]$, $W^{NR}(\pi_1, c, \Delta) > \lambda \frac{(1-c)}{2} - K > W^{NR}(\pi_0, c, \Delta) \geq 0.$

The first part of assumption 2 ensures that it is not optimal for the southern government to invest in regulatory capacity provided that there is firm offshoring and NGO entry (with

a monitoring probability π_1). The second part of assumption 2 conversely ensures that it is optimal to invest provided that there is firm delocalization and no NGO (with therefore a reduced probability of monitoring $\pi_0 < \pi_1$). In the appendix it is shown that for given values Δ and c_H, such an assumption can be satisfied when the NGO technology is good enough (i.e. π_1 high enough); the problem of observability of the firm technology by consumers is severe enough (π_0 small enough), and the cost of regulatory capacity K is low enough. We then get the following immediate result that simplifies the description of our equilibrium:

Proposition 7. *Under assumptions 1 and 2, for all local cost* $c \in [\Delta, c_H]$ *we have the following:*

i) When there is NGO entry, the Foreign government does not invest in trade regulatory capacity.

ii) When there is no NGO entry, the Foreign government invests in trade regulatory capacity.

We can now characterize completely the equilibrium of the game with endogenous Foreign regulatory capacity. Taking into account the equilibrium of the game without trade capacity in Foreign, we need to consider as before the three cutoff regimes as described in Section 6. We then get the following characterization of the globalization equilibrium:

Proposition 8. *Under assumptions 1 and 2, we have the following:*

i)When $\min\{\bar{c}_1, \bar{c}_N\} > \bar{c}_0$ *(cases 1 and 2, where offshoring occurs with NGO entry), the equilibrium is described in the following way: i1) When* $\min\{\bar{c}_1, \bar{c}_N\} < c < c_H$*, Foreign invests in regulatory capacity, there is no NGO and there is offshoring with the clean technology; i2) for* $c \leq \min\{\bar{c}_1, \bar{c}_N\}$*, Foreign does not invest in regulatory capacity, there is NGO entry and offshoring with the clean technology occurs with probability* q_1^e.

ii) When $\min\{\bar{c}_1, \bar{c}_N\} < \bar{c}_0$ *(case 3 where globalization can occur without NGO entry), the equilibrium is described in the following way: ii1) when* $\bar{c}_N < c < \bar{c}_0$*, Foreign invests in trade regulatory capacity, there is no NGO and there is offshoring with the clean technology; ii2) when* $c \leq \bar{c}_N$*, Foreign does not invest in trade regulatory capacity, there is NGO monitoring and offshoring with the clean technology with probability* q_1^e.

Proof. See the appendix. □

Proposition 8 basically states that the existence of international civil society mechanisms (NGOs) helps monitoring international trade flows in credence goods, but at the same time it may crowd out the incentives of exporting countries to invest in trade regulatory capacity. Moreover, this is most likely to happen when northern consumers are more likely to ask for

such regulatory mechanisms. This is the case when the severity of direct non observability of credence good characteristics is high (i.e. π_0 low), the NGO technology is efficient at dealing with the problem (i.e. π_1 is high) and when there are important potential gains from trade with globalization (i.e. when c is small enough). Indeed, whenever $c < \min\{c_g, \bar{c}_N\}$, Foreign does not invest in trade regulatory capacity, although it would be ready to do so without the possibility of NGO emergence. The reason is that southern governments with weak institutions have an incentive to shift the cost of international regulation of credence goods to other more caring Home consumers/donors. This suggests an interesting two-way relationship between the demand for private regulation on global markets and cross-country institutional asymmetries in terms of regulatory capacities for standards and norms that apply to traded credence goods.

Proposition 8 also has implications in terms of differential incentives for regulatory capacity investment across southern exporting countries. The closer the production cost in Foreign c is to the production cost in North c_H, the stronger the incentive of Foreign to invest into monitoring capacity and the less likely the emergence of private regulation by an NGO. By the same token, a large difference in c and c_H implies a stronger incentive for Foreign to free ride on NGO monitoring financed by donations from Home consumers.

Note also that our interpretation of local costs c includes trade costs. The previous discussion therefore implies that trade liberalization and easier market access through tariff reductions (i.e. a reduction of c) may also have consequences for the pattern of monitoring institutions of international trade in credence goods. Indeed, increased market access to northern consumers may dramatically alter investment in regulatory capacity in a given southern exporting country, as it also strengthens the incentives to shift the cost of regulation to Home consumers.

10 Conclusion

This paper brings together the economics and management literature on NGOs and private politics and the international economics literature on offshoring. We investigate how the offshoring decision and technology choice of a firm in the context of falling trade costs interact with the emergence of internationally active advocacy NGOs.

The starting point of our analysis is the observation that over the last three decades, the internationalization of production was accompanied by an increasing international advocacy NGO activity. This is reflected both in a large number of campaigns targeting multinational firms as well in the increasing number of labels granted by NGOs. Harrison and Scorse (2010) provide

evidence for the causal effect of NGO campaigns on the technology choice of a multinational firm.

We contribute to the literature on NGOs by suggesting a model with endogenous NGO emergence that models both the NGO-firm interaction and the NGO-consumer/donor interaction. We contribute to the literature on offshoring by introducing NGOs as an actor monitoring the technology choice of the firm into the firm's offshoring decision. We link the two literatures by introducing the potential of NGO emergence in response to offshoring and analyze how this affects the offshoring decision of the firm. We thereby explicitly model the interaction between the relocation of production to low-regulation countries and the emergence of international activist NGOs who run campaigns and provide labels to firms.

We find that, starting from production in Home, a fall in variable trade costs can trigger offshoring accompanied by NGO emergence. Thinking about globalization as a sequence of reductions in trade costs, the corresponding equilibria in our model are consistent with either the joint emergence of the NGO and of offshoring, but also with offshoring with later (or no) NGO emergence.

While in our model NGO emergence is unambiguously negative for a firm that has implemented the dirty technology, we find that the firm can benefit from the potential of NGO emergence from an ex-ante perspective. Potential NGO emergence increases the incentive of the firm to choose the clean technology and thereby increases its credibility in the eyes of consumers, which in turn increases demand. In this sense there can be a complementarity between NGO emergence and offshoring in our model, which means that allowing for NGO emergence makes the firm offshore in situations where it otherwise would not.

Moreover, we show that due to a misalignment of the firm's and consumer's interests, there is a range of trade costs for which the firm offshores at the expense of consumers. While expected firm profits increase, expected welfare of consumers decreases in these cases. This translates into a range of trade costs in which offshoring leads to reduced overall social welfare. In these cases the efficiency gains from offshoring cannot compensate the losses in consumer surplus when the firm offshores.

We discuss a number of possible extensions of our work. In our model all consumers care about the credence characteristic of the good. This simplifying assumption allows us to construct a model of NGO emergence that takes into account both the NGO-firm interaction, the NGO-donor interaction and how NGO emergence interacts with the offshoring decision of the firm.

In our view a promising avenue for future research is to construct a setup that endogenizes the number of caring consumers. This would require reducing complexity along other modeling dimensions altering the focus of the analysis. This should most probably be done in a dynamic setup. In such a setup the NGO would first build a 'stock' of caring consumers. We would then expect there to be a critical mass of consumers that needs to be reached before the firm considers cooperation with the NGO as better than the cost advantage of dirty production. Such a setup would likely have to abstract from endogenous NGO emergence, which is a key element of this paper, it could, however, provide a more nuanced view on the ex-ante perspective of the firm on NGO activity.

References

Aldashev, G., Limardi, M., and Verdier, T. (2011). Watchdogs of the invisible hand: NGO monitoring and industry equilibrium. *Mimeo, University of Namur*.

Aldashev, G. and Verdier, T. (2009). When NGOs go global: Competition on international markets for development donations. *Journal of International Economics*, 79(2):198–210.

Aldashev, G. and Verdier, T. (2010). Goodwill bazaar: NGO competition and giving to development. *Journal of Development Economics*, 91(1):48–63.

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97(6):1447–1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.

Antràs, P. and Helpman, E. (2004). Global Sourcing. *Journal of Political Economy*, 112(3):552-80.

Atkinson, T. (2009). Giving overseas and public policy. *Journal of Public Economics*, 93:647–653.

Bagnoli, M. and Watts, S. G. (2003). Selling to socially responsible consumers: Competition and the private provision of public goods. *Journal of Economics & Management Strategy*, 12(3):419–445.

Baksi, S. and Bose, P. (2007). Credence goods, efficient labelling policies, and regulatory enforcement. *Environmental and Resource Economics*, 37:411–430.

Balboni, A. and Balboni, F. (2009). A duopoly model of corporate social responsability and location choice. *University of Paris Dauphine, mimeo.*

Baron, D. P. (2001). Private politics, corporate social responsibility, and integrated strategy. *Journal of Economics & Management Strategy*, 10(1):7–45.

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.

Baron, D. P. (2012). The Industrial Organization of Private Politics. *Quarterly Journal of Political Science*, 7:135-174.

Baron, D. P. (2013). The Market for Activism. *mimeo.*

Baron, D. P. and Diermeier, D. (2007). Strategic activism and nonmarket strategy. *Journal of Economics & Management Strategy*, 16(3):599–634.

Benabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3):489–520.

Besley, T. and Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review*, 95(3):616–636.

Besley, T. and Ghatak, M. (2007). Retailing public goods: The economics of corporate social responsibility. *Journal of Public Economics*, 91(9):1645–1663.

Castaneda, M. A., Garen, J., and Thornton, J. (2008). Competition, contractibility, and the market for donors to nonprofits. *Journal of Law, Economics, and Organization*, 24(1):215–246.

Chau, N. H. and Huysentruyt, M. (2006). Nonprofits and public good provision: A contest based on compromises. *European Economic Review*, 50(8):1909–1935.

Cho, I-K and D.M. Kreps (1987). Signaling games and stable equilibria. *Quarterly Journal of Economics*, 102:179-221.

Darby, M. R. and Karni, E. (1973). Free competition and the optimal amount of fraud. *Journal of Law and Economics*, 16(1):67–88.

Economides, N. and Rose-Ackerman, S. (1993). Differentiated public goods: Privatization and optimality. In Ohta, H. and Thisse, J. F., editors, *Does economic space matter?: Essays in honour of Melvin L. Greenhut*, pages 111–132. St. Martin's Press, London.

Feddersen, T. J. and Gilligan, T. W. (2001). Saints and markets: Activists and the supply of credence goods. *Journal of Economics & Management Strategy*, 10(1):149–171.

Francois, P. (2003). Not-for-profit provision of public services. *The Economic Journal*, 113(486):C53–C61.

Gereffi, G. and Mayer, F. (2006). Globalization and the demand for governance. In Gereffi, G., editor, *The New Offshoring of Jobs and Global Development. ILO Social Policy Lectures*, pages 39–58. International Institute for Labour Studies, Geneva.

Grossman, G. and Rossi-Hansberg, E. (2008). Trading Tasks: A Simple Theory of Offshoring. *American Economic Review*, 98(5): 1978-1997.

Hainmueller, J. and Hiscox, M. J. (2012). The socially conscious consumer? Field experimental tests of consumer support for fair labor standards. *MIT Political Science Department Research Paper No. 2012-15.*

Harrison, A. and Scorse, J. (2010). Multinationals and anti-sweatshop activism. *The American Economic Review*, 100(1):247–273.

Hiscox, M. J. and Smyth, N. (2011). Is there consumer demand for fair labor standards? Evidence from a field experiment. *Working Paper SSRN 1820642.*

Immordino, G. (2008). Fairnes, consumer consciousness and the welfare of less developed countries. *Global Economy Journal*, 8(1).

Innes, R. (2006). A theory of consumer boycotts under symmetric information and imperfect competition. *Economic Journal*, 116, 355-381.

Kitzmueller, M. (2012). Is Better Information Always Good News? International Corporate Strategy and Regulation. *World Bank Policy Research Working Paper 6251.*

Kitzmueller, M. and Shimshack, J. (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature*, 50(1):51–84.

Kotchen, M. J. (2006). Green markets and private provision of public goods. *Journal of Political Economy*, 114(4):816–834.

Loureiro, M. and Lotade, J. (2005). Do fair trade and eco-labels in coffee wake up the consumer conscience? *Ecological Economics*, 53:129–138.

Lyon, T.P., and Salant, S. (2013) Linking Public and Private Politics: Activist Strategy for Industry Transformation. *mimeo.*

Mailath, G.G., Okuno-Fujiwara, M., Postlewaite, A. (1993). Belief-based refinements in signalling games. *Journal of Economic Theory* 60:241–276.

O'Rourke, D. (2005). Market movements. *Journal of Industrial Ecology*, 9(1):115–128.

Pestieau, P. and Sato, M. (2006). Limiting the number of charities. *Université catholique de Louvain, Center for Operations Research and Econometrics, CORE Discussion Paper 2006/74.*

Rowat, C. and Seabright, P. (2006). Intermediation by aid agencies. *Journal of Development Economics*, 79(2):469–491.

Spar, D. and La Mure, L. (2003). The power of activism: assessing the impact of ngos on global business. *California Management Review*, 45(3):78–101.

Vogel, D. (2008). Private global business regulation. *Annual Review of Political Science*, 11:261–282.

Appendix A: Proofs for the baseline model

For some derivations below it will be useful to note that combining equations (8) and (11) expected profits in an interior equilibrium without and with an NGO ($i \in \{0, 1\}$) can be expressed as:

$$E(\Pi_i) = \pi_i \frac{(1-c)^2}{4} \left[1 + \frac{\pi_i}{1-\pi_i} \frac{(1-c)^2}{4} \frac{1}{\Delta^2} \right]. \tag{21}$$

Moreover, the corresponding expression for consumer surplus follows from equations (5) and (11):

$$E(W_i) = \pi_i \frac{(1-c)^2}{8} \left[q_i^e + \frac{\pi_i}{(1-\pi_i)} \frac{(1-c)^2}{4} \frac{1}{\Delta^2} \right]. \tag{22}$$

Proof of Proposition 1

(i) is obvious.

To prove *(ii)*, note that $\bar{c}_i(c_H, \pi_i, \Delta)$ is determined by equation:

$$\begin{aligned} \frac{(1-c_H)^2}{4} &= \pi_i \frac{(1-\bar{c}_i)^2}{4} + (1-\pi_i) \frac{(q_i^e - \bar{c}_i)^2}{4} \\ &= (1-\pi_i) \left[\frac{(q_i^e - \bar{c}_i)^2}{4} + \Delta \frac{(q_i^e - \bar{c}_i)}{2} \right] \end{aligned}$$

where

$$(q_i^e - \bar{c}_i) = \frac{1}{\Delta} \cdot \frac{(1-\bar{c}_i)^2}{2} \cdot \frac{\pi_i}{1-\pi_i}$$

Substitution provides then the alternative condition characterizing $\bar{c}_i(c_H, \pi_i, \Delta)$:

$$\frac{(1-c_H)^2}{4} = \pi_i \frac{(1-\bar{c}_i)^2}{4} + \frac{(1-\bar{c}_i)^4}{16} \frac{1}{\Delta^2} \frac{\pi_i^2}{(1-\pi_i)}. \tag{23}$$

which gives

$$\bar{c}_i(c_H, \pi_i, \Delta) = 1 - \sqrt{\frac{1-\pi_i}{\pi_i} \; 2 \; \Delta^2 \; \left[\left(1 + \frac{(1-c_H)^2}{1-\pi_i} \frac{1}{\Delta^2} \right)^{\frac{1}{2}} - 1 \right]}$$

The comparative statics immediately follows from the fact that: a) the r.h.s. of (23) is decreasing in $\bar{c}_i$, b) the r.h.s. of (23) is increasing in π_i and is decreasing in Δ, c) the l.h.s. of (23) is decreasing in c_H. **QED.**

Proof of Proposition 2

(i) follows directly from the participation constraint of the donor together with the expressions for consumer surplus with and without an NGO given by equation (22).

(ii) to show that $\pi_0 > 0.1$ is a sufficient condition for $\frac{\partial x^*}{\partial c} < 0$, we start by rewriting equation (13) as:

$$x(c)^* = \frac{(1-c)^2}{8} \left[c(\pi_1 - \pi_0) + (1-c)^2 T_1 \right] \tag{24}$$

with T_1 being an exogenous constant:

$$T_1 = \frac{1}{4\Delta}\left[\frac{\pi_1^2}{1-\pi_1} - \frac{\pi_0^2}{1-\pi_0}\right]\left(2+\frac{1}{\Delta}\right).$$

The partial derivative is then:

$$\begin{aligned}\frac{\partial x^*}{\partial c} &= -\frac{1-c}{4}\left[c(\pi_1-\pi_0)+(1-c)^2T_1\right] + \frac{(1-c)^2}{8}\left[(\pi_1-\pi_0)-2(1-c)T_1\right] && (25)\\ &= \frac{(1-c)^2}{2}\left[(\pi_1-\pi_0)\left(\frac{1}{4}-\frac{c}{2(1-c)}\right)-(1-c)T_1\right]. && (26)\end{aligned}$$

Next, note that as $\pi_1 > \pi_0 > 0$, we can express π_1 as $\pi_1 = \rho\pi_0$ with $\rho > 1$. We can therefore write:

$$\frac{\pi_1^2}{1-\pi_1} - \frac{\pi_0^2}{1-\pi_0} = \pi_0\left[\frac{\rho\pi_1}{1-\pi_1} - \frac{\pi_0}{1-\pi_0}\right] > \pi_0\left[\frac{\pi_1}{1-\pi_1} - \frac{\pi_0}{1-\pi_0}\right] > \pi_0(\pi_1-\pi_0). \quad (27)$$

Using (26), the definition of T_1 and (27) we have:

$$\frac{2}{(1-c)^2}\frac{\partial x^*}{\partial c} = (\pi_1-\pi_0)\left(\frac{1}{4}-\frac{c}{2(1-c)}\right)-(1-c)\frac{1}{4\Delta}\pi_0\left[\frac{\rho\pi_1}{1-\pi_1}-\frac{\pi_0}{1-\pi_0}\right]\left(2+\frac{1}{\Delta}\right) < 0. \quad (28)$$

By (27) a sufficient condition for the inequality in (28) to hold is:

$$\left(\frac{1}{4}-\frac{c}{2(1-c)}\right) - \pi_0\frac{(1-c)}{4\Delta}\left(2+\frac{1}{\Delta}\right) < 0.$$

Rearranging delivers the following sufficient condition for $\frac{\partial x^*}{\partial c} < 0$:

$$\underbrace{\frac{c}{2(1-c)}}_{T_2} + \underbrace{\pi_0\frac{(1-c)}{4\Delta}\left(2+\frac{1}{\Delta}\right)}_{T_3} > \frac{1}{4}. \quad (29)$$

Note that both T_2 and T_3 are positive and only T_3 depends on π_0. Furthermore, note that $T_2 > \frac{1}{4}$ if $c > \frac{1}{3}$. This implies that $c > \frac{1}{3}$ is a sufficient condition for $\frac{\partial x^*}{\partial c} < 0$.

For $c \leq \frac{1}{3}$, π_0 has to be sufficiently high for the inequality to hold. We now establish a sufficient condition on π_0. For analytical tractability, we ignore the positive term T_2 and derive a sufficient condition for $T_3 > 1/4$. (The numerical analysis outlined below takes the term T_2 into account.) Note that $\frac{\partial T_3}{\partial \Delta} < 0$ so that a higher Δ makes it less likely that condition (29) is satisfied. As even under dirty production the cost of production cannot be less than zero, we

must have $\Delta \leq c$. This implies that:

$$\pi_0 \frac{(1-c)}{4\Delta}\left(2+\frac{1}{\Delta}\right) \geq \pi_0 \frac{(1-c)}{4c}\left(2+\frac{1}{c}\right). \tag{30}$$

We can therefore state the following sufficient for condition (29) to hold:

$$\pi_0 > \frac{c}{1-c}\,\frac{1}{2+1/c}. \tag{31}$$

Note that the r.h.s. increases monotonically in c implying that the condition requires higher values of π_0 for higher values of c. Recall that $c > 1/3$ is a sufficient condition for $\frac{\partial x^*}{\partial c} < 0$. This implies that the highest relevant value for condition (31) is $c = \frac{1}{3}$. This directly implies that $\pi_0 > 0.1$ is a sufficient condition for $\frac{\partial x^*}{\partial c} < 0$ to hold over the full range of $c \in (0,1)$. This proves *(ii)*.[18]

(iii) follows directly from partially differentiating equation (13) with respect to π_1, π_0 and Δ, which delivers $\frac{\partial x^*}{\partial \pi_1} > 0$, $\frac{\partial x^*}{\partial \pi_0} < 0$ and $\frac{\partial x^*}{\partial \Delta} < 0$. **QED**.

Proof of Proposition 3

(i) follows from the definition of the NGO cutoff level c_N.

To prove *(ii)*, differentiate (15) for any variable $\alpha = \pi_1, \pi_0, \Delta$. Thus,

$$\frac{d\bar{c}_N}{d\alpha} = -\frac{\frac{\partial E(W_1)(\bar{c}_N)}{\partial \alpha} - \frac{\partial E(W_0)(\bar{c}_N)}{\partial \alpha}}{\frac{\partial E(W_1)(\bar{c}_N)}{\partial \bar{c}_N} - \frac{\partial E(W_0)(\bar{c}_N)}{\partial \bar{c}_N}} = -\frac{\frac{\partial x^*(\bar{c}_N)}{\partial \alpha}}{\frac{\partial x^*(\bar{c}_N)}{\partial \bar{c}_N}}.$$

By proposition 2, we have $\frac{\partial x^*(\bar{c}_N)}{\partial \bar{c}_N}$ is negative (under the maintained sufficient condition of $\pi_0 > 0.1$). This implies that:

$$sign\left[\frac{d\bar{c}_N}{d\alpha}\right] = sign\left[\frac{\partial x^*(\bar{c}_N)}{\partial \alpha}\right]. \tag{32}$$

It now follows directly from proposition 2 *(iii)* that $\frac{\partial \bar{c}_N}{\partial \pi_1} > 0$, $\frac{\partial \bar{c}_N}{\partial \pi_0} < 0$ and $\frac{\partial \bar{c}_N}{\partial \Delta} < 0$. **QED.**

[18] A stronger sufficient condition can be found numerically substituting Δ by c in condition (29), solving for π_0 and numerically determining the maximum. Our algorithm delivers $\pi_0^{max} = 0.01894899$, so that we can state $\pi_0 > 0.019$ as a sufficient condition on numerical grounds. This procedure delivers a much lower condition as it allows us to include the term T_2 in the analysis, which had been dropped in the algebraic proof.

Proof of Proposition 4

By (14) at the NGO cutoff, we have $x^*(\bar{c}_N, \pi_1, \pi_0, \Delta) = X + y$. Given that $x^*(c, \pi_1, \pi_0, \Delta)$ is decreasing in c, case 1 occurs if and only if $\bar{c}_N \geq \bar{c}_1$ which is equivalent to $x^*(\bar{c}_1, \pi_1, \pi_0, \Delta) \geq X + y$. Similarly case 2 if and only if $\bar{c}_0 \leq \bar{c}_N \leq \bar{c}_1$, which is equivalent to $x^*(\bar{c}_1, \pi_1, \pi_0, \Delta) \leq X + y \leq x^*(\bar{c}_0, \pi_1, \pi_0, \Delta))$. Finally case 3 occurs if and only if $\bar{c}_N \leq \bar{c}_0$ which is equivalent to $x^*(\bar{c}_0, \pi_1, \pi_0, \Delta) \leq X + y)$. **QED.**

Proof of the Corollary to Proposition 4

Note that in cases 1 and 2 the cost level at which offshoring takes place is the globalization cutoff $\bar{c}_g = \min\{\bar{c}_1, \bar{c}_N\}$. To prove *(i)*, we need to show that $\bar{c}_g > \bar{c}_0$. As we consider case 1 and 2, we always have $\bar{c}_N > \bar{c}_0$. The fact that $\bar{c}_1 > \bar{c}_0$ follows from equation (21) together with the fact that at the cutoffs we must have $\Pi_H(c_H) = E(\Pi_1(\bar{c}_1))$ and $\Pi_H(c_H) = E(\Pi_0(\bar{c}_0))$. As the l.h.s. of the two expressions is the same and as $\pi_1 > \pi_0$, it follows that $\bar{c}_1 > \bar{c}_0$. This proves *(i)*.

To prove *(ii)* we need to show that both cutoff levels $\bar{c}_N$ and $\bar{c}_1$ increase in π_1 and decrease in Δ. It follows from the fact that the l.h.s. of the condition $\Pi_H(c_H) = E(\Pi_1(\bar{c}_1))$ is constant, and the r.h.s. increases in π_1 (decreases in Δ), that an increase in the detection probability π_1 (a decrease in Δ) must lead to an increase in the cutoff level $\bar{c}_1$. Moreover, we know from Proposition 3 that for the NGO cutoff level $\bar{c}_N$ we have $\frac{\partial \bar{c}_N}{\partial \pi_1} > 0$ and $\frac{\partial \bar{c}_N}{\partial \Delta} < 0$. This proves *(ii)*. **QED**.

Proof of Proposition 5

In case 1 we have $c_g = \bar{c}_1$. By definition, at $\bar{c}_1$, we have $E(\Pi_1)(\bar{c}_1) = \Pi_H$. Note that in autarky consumer surplus is just half of profits: $W_H = \frac{1}{2}\Pi_H$. To show the discrete fall in expected consumer surplus, it is sufficient to show that $E(W_1)(\bar{c}_1) < \frac{1}{2}E(\Pi_1)(\bar{c}_1) = \frac{1}{2}\Pi_H$. This follows directly from equation (5) (for $i = 1$) and the fact that in an interior solution we have $q_i^e < 1$. The discrete fall in net expected consumer surplus then follows directly from the fact that if an NGO exists we have $E(W)_{net} = E(W_1) - X$. This proves *(i)*.

In case 2 we have $c_g = \bar{c}_N$ with $\bar{c}_N \in [\bar{c}_0, \bar{c}_1]$. In case 2 outsourcing takes place only with NGO emergence, so that we have a discrete fall in expected consumer surplus upon outsourcing iff $E(W_1)(\bar{c}_N) < W_H$. We know from *(i)* that this is the case when $\bar{c}_N = \bar{c}_1$. As the function $E(W_1)(c)$ is continuous in c, the difference $E(W_1)(c) - W_H$ must be negative for all values of $\bar{c}_N \in (\tilde{c}, \bar{c}_1]$ with $\tilde{c}$ close enough to $\bar{c}_1$. The same argument holds for net expected consumer

surplus $E(W)_{net}$ which falls upon outsourcing as long as $E(W_1)(\bar{c}_N) - X < W_H$, which is the case for all $c_N \in (\tilde{c'}, \bar{c}_1]$ with $\tilde{c} > \tilde{c'}$ as long as $X > 0$. **QED.**

Proof of Proposition 6

For case 1, the result simply follows from noting that total welfare is the sum of net expected consumer surplus and firm profits. The former component is described in Proposition 6. Expected profits are either given by $E(\Pi_i)$ in equation (21) or by autarky profits Π_H. In case 1 the firm offshores at the cutoff $\bar{c}_1$ at which we have $E(\Pi_1)(\bar{c}_1) = \Pi_H$ this implies that upon offshoring there is no discrete jump in expected profits. As Proposition 6 implies that there is a discrete negative jump in net expected consumer surplus, we must have a negative jump in the sum of the two, namely total welfare.

As for case 2, note that $E(\Pi_1)$ decreases continuously in c and $E(\Pi_1)(\bar{c}_1) = \Pi_H$. Therefore, in case 2, where offshoring takes place at $\bar{c}_g = \bar{c}_N < \bar{c}_1$, we must have $E(\Pi_1)(\bar{c}_g) > \Pi_H$. This implies that in case 2 we always have a discrete upward jump of expected profits upon offshoring. The fact that $E(\Pi_1)$ decreases continuously in c also implies that the magnitude of the upward jump is the lower the closer the globalization cutoff $\bar{c}_g = \bar{c}_N$ is to $\bar{c}_1$. As it goes to zero for $\bar{c}_g$ going to $\bar{c}_1$, Proposition (ii) implies that in case 2 when $\bar{c}_g$ is sufficiently close to $\bar{c}_1$ there is a discrete fall in total welfare. **QED.**

Appendix B: A Micro-founded model of NGO behavior

In this appendix, we provide a less extreme microfounded model of an NGO entrepreneur and we generalize to that setting the result that NGO entry occurs when the local cost of production c is below a cutoff level $\bar{c}_N$.

Consider an intrinsically motivated agent that contemplates creating an NGO in case there is a positive probability of production with a dirty technology. Assume that the motivated agent cares only about the possibility of dirty production in itself. When she faces a situation where the probability of dirty production $(1 - q_i^e)$ is positive, she gets utility V from 'fighting dirty production' i.e. from creating an NGO. This utility is independent of the actual level of the probability of dirty production: as soon as it is above zero, the motivated agent is ready to start monitoring the firm.

NGO creation requires both a monetary fixed cost X and an effort level $e \in [0, e_{\max}]$ with a linear disutility of effort ke. We impose $V - ke_{max} > 0$ which implies that the motivated

agent is willing to create an NGO as soon as sufficient funds can be raised. The technology of fundraising requires two inputs: money and effort and is described in the following way. With fundraising expenditures m and some effort level e, the motivated agent succeeds to activate $\nu(m,e)$ donors who feel convinced to "make a difference" when donating. The function $\nu(m,e)$ has the usual properties:

$$\begin{aligned} \text{for all } (m,e) &\in \mathbb{R}^+ \times [0, e_{\max}], \nu(m,e) \in [0,1] \ \nu(0,e) = 0, \nu(m,0) = 0, \nu'_m(m,e) \geq 0, \nu'_e(m,e) > 0, \\ \nu''_{mm}(m,e) &< 0, \nu''_{ee}(m,e) < 0, \nu''_{me}(m,e) \geq 0, \lim_{m \to 0} \nu'_m(m,e) = +\infty; \lim_{m \to \infty} \nu'_m(m,e) = 0. \end{aligned}$$

Now the problem of the NGO entrepreneur is:

$$\max_{e,m,x} V - ke$$

under the constraints:

$$\nu(m,e)x - m \geq X \tag{33}$$

$$x \leq x^*(c, \pi_1, \pi_0, \Delta) = E(W_1) - E(W_0). \tag{34}$$

The first constraint is the fund raising constraint to have the capacity to create the NGO. The second constraint implies that the surplus gain the NGO gives to a donor must be larger or equal to the amount donated. The necessary conditions for m, e and x are simply given by:

$$\begin{aligned} -k + \mu \nu'_e(m,e)x &= 0 \\ \mu \left[\nu'_m(m,e)x - 1 \right] &= 0 \\ \mu \nu(m,e) - \xi &= 0 \end{aligned}$$

where $\mu \geq 0$ and $\xi \geq 0$ are the lagrange multipliers of the constraints (33) and (34). Simple inspection shows that $\mu > 0$ and $\xi > 0$. Therefore, (33) and (34) are binding and the solution has to satisfy:

$$\begin{aligned} x &= x^*(c, \pi_1, \pi_0, \Delta) \\ \nu(m,e)x - m &= X \\ \nu'_m(m,e)x &= 1. \end{aligned}$$

Now the third equation defines $m^*(e,c) > 0$. Substitution in the second equation provides the solution for the NGO effort e^*, solution of:

$$\Gamma(e,c) = \nu(m^*(e,c),e)x^*(c,\pi_1,\pi_0,\Delta) - m^*(e,c) = X.$$

This defines the solution of the NGO entrepreneur when $e^* \leq e_{\max}$. Note that because of the envelope theorem applied to m, we have:

$$\Gamma'_e(e,c) = \nu'_e(m^*(e,c),e)x^*(c,\pi_1,\pi_0,\Delta) > 0 \text{ and } \Gamma'_c(e,c) = \nu(m^*(e,c),e)\frac{\partial x^*}{\partial c}(c,\pi_1,\pi_0,\Delta) < 0.$$

Thus, the NGO will enter if and only if:

$$\Gamma(e_{\max},c) = \nu(m^*(e_{\max},c),e_{\max})x^*(c,\pi_1,\pi_0,\Delta) - m^*(e_{\max},c) > X.$$

This relationship determines implicitly an NGO entry cutoff $\bar{c}_N$ such that $\Gamma(e_{\max},\bar{c}_N) = X$. The envelope theorem gives also:

$$\frac{\partial \bar{c}_N}{\partial \alpha} = -\frac{\frac{\partial x^*}{\partial \alpha}}{\frac{\partial x^*}{\partial \bar{c}_N}} \text{ for all parameters } \alpha = \pi_1, \pi_0, \Delta.$$

Therefore the comparative statics of Proposition 3 are preserved. **QED.**

Appendix C: Endogenous public regulatory systems and NGOs

Proof that $W^{NR}(\pi_i,c,\Delta)$ is decreasing convex in $c \in [\Delta, c_H]$:

Differentiation of $W^{NR}(\pi_i,c,\Delta)$ in equation (20) provides:

$$\begin{aligned}
\frac{\partial W^{NR}(\pi_i,c,\Delta)}{\partial c} &= -\frac{\lambda\pi_i}{2}\left[c + \frac{1}{\Delta}\cdot\frac{(1-c)^2}{2}\cdot\frac{\pi_i}{1-\pi_i}\right] + \lambda\pi_i\frac{(1-c)}{2}\left[1 - \frac{1}{\Delta}\cdot(1-c)\cdot\frac{\pi_i}{1-\pi_i}\right] \\
&\quad -\lambda\frac{\pi_i}{\Delta}\cdot(1-c) \\
\frac{\partial^2 W^{NR}(\pi_i,c,\Delta)}{\partial c^2} &= -\lambda\pi_i\left[1 - \frac{1}{\Delta}\cdot(1-c)\cdot\frac{\pi_i}{1-\pi_i}\right] + \lambda\pi_i\frac{(1-c)}{2}\left[\frac{1}{\Delta}\cdot\frac{\pi_i}{1-\pi_i}\right] + \lambda\frac{\pi_i}{\Delta} \\
&= \lambda\pi_i\frac{1-\Delta}{\Delta} + \frac{\lambda}{\Delta}\cdot(1-c)\cdot\frac{\pi_i^2}{1-\pi_i} + \lambda\pi_i\frac{(1-c)}{2}\left[\frac{1}{\Delta}\cdot\frac{\pi_i}{1-\pi_i}\right] > 0.
\end{aligned}$$

Thus $\frac{\partial W^{NR}(\pi_i,c,\Delta)}{\partial c}$ is increasing in c. Also $\frac{\partial W^{NR}(\pi_i,1,\Delta)}{\partial c} = -\frac{\lambda\pi_i}{2} < 0$. Hence

$$\frac{\partial W^{NR}(\pi_i, c, \Delta)}{\partial c} < 0$$

and $W^{NR}(\pi_i, c, \Delta)$ is decreasing convex in $c \in [\Delta, c_H]$. **QED.**

Proof that assumptions 1 and 2 can be satisfied for a non empty set of parameters $(c_H, \Delta, \pi_1, \pi_0, K)$:

Fix the two parameters Δ and c_H with $\Delta < c_H$. Consider then the function $\Theta(\pi_i, c, \Delta) = W^{NR}(\pi_i, c, \Delta) - \lambda\frac{(1-c)}{2} + K$. As $W^{NR}(\pi_i, c, \Delta)$ is convex in c, $\Theta(\pi_i, c, \Delta)$ is also convex in c. It is a simple matter to see that for $\pi_1 \leq \overline{\pi}_1 = 2\Delta/(1+\Delta)$, assumption 1 is satisfied for all $c \in [\Delta, c_H]$.

Moreover,

$$\begin{aligned}
\Theta\left(\overline{\pi}_1, c, \Delta\right) &= \lambda\frac{\Delta}{(1+\Delta)}(1-c)\left[c + \frac{(1-c)^2}{(1-\Delta)}\right] + \lambda\frac{(1-c)^2}{(1+\Delta)} - \lambda\frac{(1-c)}{2} + K \\
&= \lambda\frac{\Delta}{(1+\Delta)}(1-c)\left[c + \frac{(1-c)^2}{(1-\Delta)}\right] + \lambda\frac{(1-c)}{(1+\Delta)}\left[1 - c - \frac{1+\Delta}{2}\right] + K.
\end{aligned}$$

We then observe that:

$$1 - c - \frac{1+\Delta}{2} > 1 - c_H - \frac{1+\Delta}{2} > 1 - c_H - \frac{1+c_H}{2} = \frac{1-3c_H}{2}.$$

Thus, for $\Delta < c_H < 1/3$, one has that for all $c \in [\Delta, c_H]$, $\Theta\left(\overline{\pi}, c, \Delta\right) > 0$. Hence for π_1 close enough to $\overline{\pi}$, one can ensure that for $\Delta < c_H < 1/3$,

$$\text{for all } c \in [\Delta, c_H]\,,\ W^{NR}(\pi_1, c, \Delta) > \lambda\frac{(1-c)}{2} - K.$$

Similarly,

$$\begin{aligned}
\Theta(\pi_i, c, \Delta) &= \lambda\pi_i\frac{(1-c)}{2}\left[c + \frac{1}{\Delta}\cdot\frac{(1-c)^2}{2}\cdot\frac{\pi_i}{1-\pi_i}\right] + \lambda\frac{\pi_i}{\Delta}\cdot\frac{(1-c)^2}{2} - \lambda\frac{(1-c)}{2} + K \\
&< \lambda\pi_i\frac{(1-c)}{2}\left[c + \frac{1}{\Delta}\cdot\frac{(1-c)^2}{2}\cdot\frac{\pi_i}{1-\pi_i}\right] + \lambda\frac{\pi_i}{\Delta}\cdot\frac{(1-c)^2}{2} - \lambda\frac{(1-c_H)}{2} + K.
\end{aligned}$$

Hence in such a case when π_0 close enough to 1/10, one has:

$$\begin{aligned}\Theta(\pi_0, c, \Delta) &< \lambda \frac{1}{10} \frac{(1-c)}{2} \left[c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\frac{1}{10}}{1-\frac{1}{10}} \right] + \lambda \frac{\frac{1}{10}}{\Delta} \cdot \frac{(1-c)^2}{2} - \lambda \frac{(1-c_H)}{2} + K + \varepsilon \\ &< \lambda \frac{(1-c)}{20} \left[c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{18} \right] + \frac{\lambda}{\Delta} \cdot \frac{(1-c)^2}{20} - \lambda \frac{(1-c_H)}{2} + K + \varepsilon\end{aligned}$$

with ε some positive small number. Now

$$\lambda \frac{(1-c)}{20} \left[c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{18} \right] + \frac{\lambda}{\Delta} \cdot \frac{(1-c)^2}{20} - \lambda \frac{(1-c_H)}{2} + K + \varepsilon < 0$$

when

$$c_H < 1 - \frac{2K}{\lambda} - \frac{(1-c)}{10} \left[c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{18} \right] - \frac{1}{\Delta} \cdot \frac{(1-c)^2}{10} - \varepsilon.$$

Denote

$$\sigma(\Delta) = \frac{2K}{\lambda} + \max_{c \in [0,1]} \left[\frac{(1-c)}{10} \left[c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{18} \right] + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{10} \right].$$

Then for $c_H < 1 - \sigma(\Delta) - \varepsilon$, and π_0 close enough to 1/10, for all $c \in [\Delta, c_H]$, $\Theta(\pi_0, c, \Delta) < 0$. Note that a non empty interval $[\Delta, c_H]$ with $c_H < 1 - \sigma(\Delta) - \varepsilon$ exists when $\Delta < 1 - \sigma(\Delta) - \varepsilon$ or $\sigma(\Delta) < 1 - \Delta - \varepsilon$. Now $\sigma(\Delta) < \frac{2K}{\lambda} + \frac{1}{10}\left[1 + \frac{1}{\Delta} \cdot \frac{1}{18}\right] + \frac{1}{\Delta} \cdot \frac{1}{10} < 1 - \Delta - \varepsilon$ when $\frac{K}{\lambda}$ is small enough and Δ larger than some $\overline{\Delta}$.

It follows finally that for $\overline{\Delta} < \Delta < c_H < \min\{1/3; 1 - \sigma(\Delta) - \varepsilon\}$, $\pi_1 < \overline{\pi}$ but close enough to $\overline{\pi}$ and π_0 close enough to 1/10, one has:

$$\text{for all } c \in [\Delta, c_H], \ \Theta(\pi_1, c, \Delta) > 0 > \Theta(\pi_0, c, \Delta)$$

and both assumptions 1 and 2 can be satisfied for a non empty set of parameters.

Proof of Proposition 8:

(i) Consider first the case where $\min\{\bar{c}_1, \bar{c}_N\} > \bar{c}_0$ (complementarity between globalization and NGO entry).

(i1) When $\min\{\bar{c}_1, \bar{c}_N\} < c < c_H$, we know that under this cost configuration, the firm does not offshore if there is no NGO. Hence without southern regulatory capacity, there is no offshoring. At the same time assumption 2 implies that $\lambda \frac{(1-c)}{2} - K > \lambda \frac{(1-c_H)}{2} - K > 0$. It is optimal for Foreign to invest in regulation capacity to induce the firm to offshore internationally. Consequently there is international offshoring with the clean technology.

(i2) When $c \leq \min\{\bar{c}_1, \bar{c}_N\}$, without southern trade regulatory investment, the firm does offshore and there is NGO entry. In such a case, from assumption 2 and Proposition 8, the southern government prefers not to invest in efficient regulatory institutions. There is therefore international offshoring, NGO entry but the firm chooses the clean technology only with probability q_1^e.

(ii) Consider now the case where $\min\{\bar{c}_1, \bar{c}_N\} < \bar{c}_0$ (case 3, where offshoring can occur without NGO entry). It is very similar to previous one except that one substitutes $\bar{c}_N$ for $\bar{c}_1$.

(ii1) follows from the fact that under this cost configuration, the firm offshores without NGO entry. Proposition 8 implies that it is then in the interest of Foreign to invest in regulatory capacity to optimize its objective function.

(ii2) when $c \leq \bar{c}_N$, without trade regulatory investment, the firm does offshore and there is NGO entry. According to Proposition 8, the southern government then prefers not to invest in efficient regulatory institutions. There is therefore offshoring but the firm chooses the clean technology only with probability q_1^e. **QED.**

Appendix D: The signaling game

In the main text only the case is considered where, under uncertainty regarding the firm type, the firm sets the monopoly price if clean and mimics this behavior if dirty. We argue in the main text that this mimicking is optimal for a dirty firm as deviating from the optimal price of a clean firm reveals that the firm is dirty. This implies that there is a signal about the firm type contained in the price set by the firm. In this section we analyze this signalling game more formally.

Recall the sequencing of the game: in stage 2 of the game, the firm chooses a probability of clean production $m_i \in [0, 1]$. In stage 3 nature then determines the type of technology used $\theta \in \{c, d\}$. With probability π_i, the type θ is revealed. In this case there is no uncertainty and there is no role for an NGO to be played.

With probability $(1 - \pi_i)$ the type of the firm is not revealed to consumers. In this case the firm sets a price $p(\theta) \in [0, \infty)$. In stage 5 consumers pick an action $s \in \{0, 1\}$ which is not to buy or to buy the product. Consumers share a common belief function $\mu_i(\theta|p)$ that assigns a probability of $\mu_i(\theta|p)$ to type θ of the firm observing the price p. Consumers receive a utility $u(s, \mu_i(\theta|p), p(\theta))$ given by equation (1).

Consider the case where beliefs are such that consumers take the firm to be dirty (consumers choose $s = 0$) whenever they observe a price $p \neq p'$. This implies that the belief function takes the following form:

$$\mu_i(\theta = c|p) = \begin{cases} q_i^e & \text{if } p = p' \\ 0 & \text{if } p \neq p' \end{cases}$$

where q^e is the expectation consumers form on the probability of facing a clean firm, when the firm type is not revealed. In this case a clean firm would set a price $p = p'$ iff expected profits profits of p' are positive: $\Pi_{ic}^u(p', \mu(c|p')) > 0$ (where the superscript u stands for uncertainty).

The analysis in this appendix is valid for the case with and without an NGO ($i = 1$ and $i = 0$) and takes (non) existence of the NGO as given. For the ease of notation we therefore drop the index i where it causes no confusion. Moreover, as firms maximize expected profits in all cases, also the expectations operator is omitted.

We now proceed in four steps. We first show that in our setup all Perfect Bayesian Equilibria (PBEs) are pooling equilibria in which both firm types set the same price. In a second step, we characterize the equilibrium clean technology choice i.e. the probability the firm assigns to clean production in stage 2 $m(p') \in (0, 1)$ of the firm consistent with the fact that the second stage delivers a perfect Bayesian pooling equilibrium at some price p'. In the third step, we eliminate PBEs from the second stage using the intuitive criterion and the undefeated equilibrium concept as equilibrium refinements. This finally allows us to fully characterize the two stage solution of the game in step four.

Step 1: We first consider the situation where the type of the firm is not revealed to consumers in stage 3. This occurs with probability $(1 - \pi)$ conditional on the firm playing mixed strategies in the first stage, i.e. $m \in (0, 1)$. In this case the price setting of the firm potentially signals its type.

Define the set of prices p' that satisfy $\Pi_c^u(p', \mu(\theta = c|p')) > 0$ as Φ. Note that under uncertainty in a pooling equilibrium both types of firms face the same demand Q. We can therefore write $\Pi_d^u = \Pi_c^u + x$ with $x = Q\Delta > 0$ and $\Delta > 0$. As in any PBE we must have $\Pi_c^u > 0$, we must also have $\Pi_d^u > 0$. As demand for an identified dirty firm is always zero, it follows that in any PBE the dirty firm will always set the same price as a clean firm. We can therefore conclude that all PBE are pooling equilibria. Moreover, under uncertainty, a dirty firm has higher profits than a clean firm as long as demand is positive. As the firm takes the

belief function of consumers $\mu(\theta|p)$ and therefore the value of q^e as given, we can plot profits of an unidentified clean and dirty firm for the full range of possible PBE represented by the set of prices leading to positive profits of a clean firm.

-Figure 8 here-

Figure 8 illustrates the payoffs of firms under uncertainty about the technology choice for a given q^e (note that the general shape of the profit functions plotted remains the same for different values of q^e). We have seen above that any PBE is a pooling equilibrium in which both types of firms charge the same price. The range of pooling PBEs is between $p_{\min}$ and $p_{\max}$, which depicts the prices for which both firms make positive profits. Due to the cost advantage Δ, profits of an unidentified dirty firm are higher than profits of an unidentified clean firm for the full range of prices. As a dirty firm faces lower costs, its optimal price p_d^* is lower than the optimal price of a clean firm p_c^*.

Step 2: Before eliminating PBEs in step 3, we now consider stage 2 of the game where the firm chooses $m \in [0,1]$, the probability of clean production, taking into account that when it plays mixed strategies($m \in (0,1)$) and its type is not revealed to consumers $(1-\pi)$ all PBEs of the signaling game are pooling equilibria.

As defined in the main text, under uncertainty, demand is given by $q^e - p$, where q^e is the *expectation* consumers form on the probability the firm assigns to clean production. Assume all agents have the (correct) expectation that in stage 4 of the game, there is a pooling equilibrium path where both types of firms $\theta \in \{c, d\}$ play price p'. Then $q^e = \mu(\theta = c|p')$ represents the belief of consumers to face a clean firm along the pooling equilibrium path when observing the price p'. Note that by definition at this pooling equilibrium one must have $q^e \in (0,1)$.

Moreover, the probability of clean production m chosen by the firm must reflect the a-priori beliefs of consumers $\Pr(\theta = c)$ of facing a clean firm at the beginning of the second stage of the game before firms announce prices.

In the first stage, the technology choice of the firm depends on the expected profits of the two options, which are given by:

$$\Pi_c(p', q^e) = \pi \frac{(1-c)^2}{4} + (1-\pi)(q_i^e - p')(p' - c) \tag{35}$$

$$\Pi_d(p', q^e) = (1-\pi)(q^e - p')(p' - c + \Delta). \tag{36}$$

The firm will choose $m = 1$ whenever $\Pi_c(p', q^e) > \Pi_d(p', q^e)$ and will choose $m = 0$ in the opposite case of $\Pi_d(p', q^e) > \Pi_c(p', q^e)$. So the only case in which the firm chooses a mixed strategy $m \in (0, 1)$ is when we have $\Pi_c(p', q^e) = \Pi_d(p', q^e)$.

Along any pooling equilibrium path at price p' in the second stage of the game, the expectation $q^e = \mu(\theta = c|p')$ of consumers satisfy Bayes Law implying that the common belief function is given by:

$$\begin{aligned}
\mu(\theta = c|p') &= \frac{\Pr(p' \mid \theta = c) \cdot \Pr(\theta = c)}{\Pr(p' \mid \theta = c) \cdot \Pr(\theta = c) + \Pr(p' \mid \theta = d) \cdot \Pr(\theta = d)} \\
&= \frac{1 \cdot \Pr(\theta = c)}{1 \cdot \Pr(\theta = c) + 1 \cdot \Pr(\theta = d)} = \Pr(\theta = c) = m_i
\end{aligned}$$

as along the pooling equilibrium path of the second stage both types of firms $\theta \in \{c, d\}$ play price p', and therefore $\Pr(p' \mid \theta = c) = \Pr(p' \mid \theta = d) = 1$. It follows that along any Bayesian Nash Perfect equilibrium with pooling at price p' in the second stage, one should have $m = \mu(\theta = c|p') = q^e < 1$. Hence in such an equilibrium, the firm should play mixed strategies in term of technology choice in stage 2 of the game. For this to happen and have expectations and actual firm choices coinciding, one should have $\Pi_c(p', q^e) = \Pi_d(p', q^e)$ with $m = q^e$. This provides the following equilibrium value of m consistent with a second stage pooling equilibrium at price p':

$$m = p' + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi}{1-\pi}.$$

this is possible only for

$$p' < 1 - \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi}{1-\pi}.$$

Step 3: We now consider equilibrium refinements that can help to reduce the number of possible PBEs that range from $p_{\min}$ to $p_{\max}$.

In the range between p_d^* and p_c^* the intuitive criterion introduced by Cho and Kreps (1987) can be applied as an equilibrium refinement. This refinement assumes that players do not try to switch into a strictly dominated equilibrium. Intuitively, the player sending the signal (the firm) 'argues' as follows: "you (the consumer) know that I know that you will buy from me if I set the price $p = p'$. If now I deviate from the this price, I propose a new equilibrium to you. If this equilibrium is such that only a clean firm is better off in this new equilibrium, you should consider moving there." This implies that any off-equilibrium price $p \neq p'$ should be interpreted as an attempt to move to another equilibrium.

It turns out that in the range between p_d^* and p_c^* the intuitive criterion can be applied. In this range, increasing the price towards p_c^* makes the clean firm better off, but decreases profits of the dirty firm. All PBE between p_d^* and p_c^* can therefore be excluded on the grounds of the intuitive criterion (including $p' = p_d^*$).

For the remaining equilibria, we have to refer to a stronger equilibrium refinement introduced by Mailath et al.(1993): the concept of the undefeated equilibrium. Consider two equilibria characterized by p_1' and p_2' with $p_{\min} < p_1' < p_2' < p_d^*$. The assumption imposed on out-of-equilibrium beliefs is such that firms always set the price p_2' if equilibrium 2 (characterized by p_2') defeats equilibrium 1 (characterized by p_1'). This is the case if the firm is strictly better off in equilibrium 2 than in equilibrium 1. Figure 8 nicely illustrates that this is the case for both firm types for all equilibria characterized by prices $p_{\min} < p' < p_d^*$. All these equilibria are therefore defeated by $p' = p_c^*$ for a clean firm and by $p' = p_d^*$ for a dirty firm.

Note that in the equilibrium $p' = p_d^*$, the intuitive criterion discussed above kicks in as according to the intuitive criterion, a clean firm should deviate from this equilibrium. We have therefore excluded all equilibria to the left of $p' = p_c^*$. Note that by the same argument as above, all equilibria to the right of p_c^* are defeated by $p' = p_c^*$ for both firm types.

Overall, this implies that $p' = p_c^*$ is the only PBE that is undefeated *and* consistent with the intuitive criterion. This implies that under uncertainty the optimal price set by a clean firm and mimicked by a dirty firm is given by

$$p' = \frac{q^e + c}{2}.$$

Step 4: Given our selected pooling price equilibrium in step 3 in the second stage of the game and the fact that through Bayes Law $q^e = \mu(\theta = c|p') = m$, the Nash Perfect Equilibrium of the two stage game is given by the two following equations:

$$\begin{aligned} p' &= \frac{m+c}{2} \\ m &= p' + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{4} \cdot \frac{\pi}{1-\pi}. \end{aligned}$$

This gives:

$$
\begin{aligned}
m^* &= c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{2} \cdot \frac{\pi}{1-\pi} \\
p^* &= c + \frac{1}{\Delta} \cdot \frac{(1-c)^2}{4} \cdot \frac{\pi}{1-\pi}.
\end{aligned}
$$

Assumption 1 ensures that $m^* < 1$.

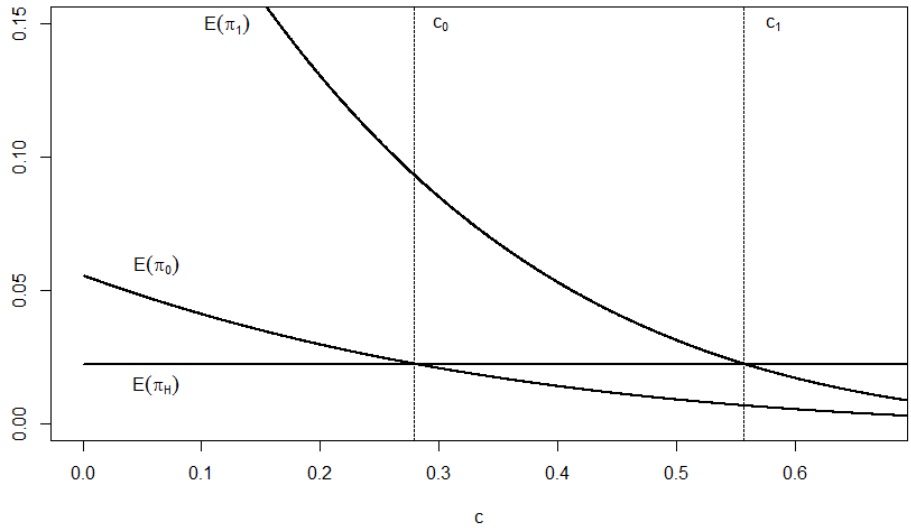

Figure 1: Expected profits as a function of offshoring cost c for the cases of production in Home $E(\pi_H)$, offshoring without NGO $E(\pi_0)$ and offshoring with NGO $E(\pi_1)$ (solid line). Dotted lines: offshoring cutoffs with and without an NGO ($\bar{c}_1$ and $\bar{c}_0$).

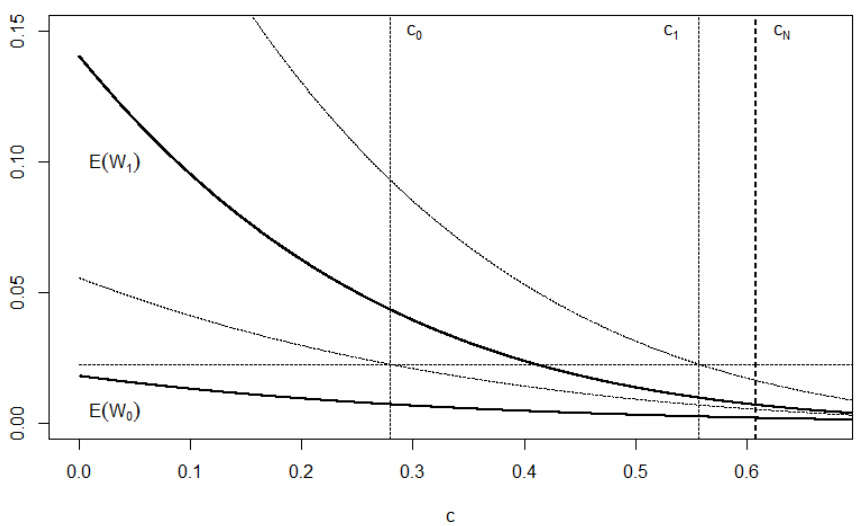

Figure 2: Pattern of expected consumer surplus with NGO $E(W_1)$ and without NGO $E(W_0)$ (solid lines) as a function of the cost level c. Thin dotted lines: expected profits and firm cutoffs from Figure 1. Vertical dashed line: NGO cutoff $\bar{c}_N$.

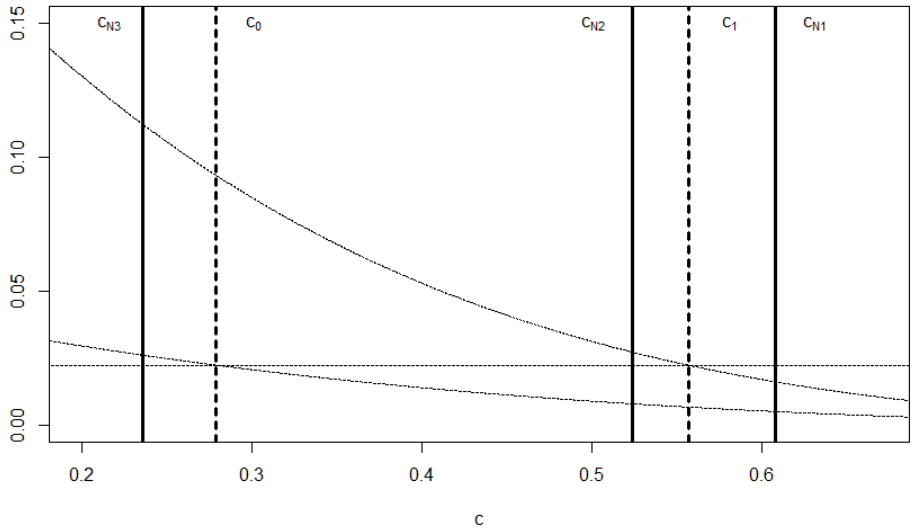

Figure 3: Three possible cases: depending on the parametrization, the NGO cutoff can be above, between or below the firm cutoffs $\bar{c}_1$ and $\bar{c}_0$. Thin dotted lines: expected profits with and without and NGO and under autarky from Figure 1. Thick dotted lines: offshoring cutoffs $\bar{c}_1$ and $\bar{c}_0$. Three solid vertical lines: cases 1, 2 and 3 with low, medium and high NGO funding requirements implying three different NGO cutoffs $\bar{c}_{N1}$, $\bar{c}_{N2}$ and $\bar{c}_{N3}$.

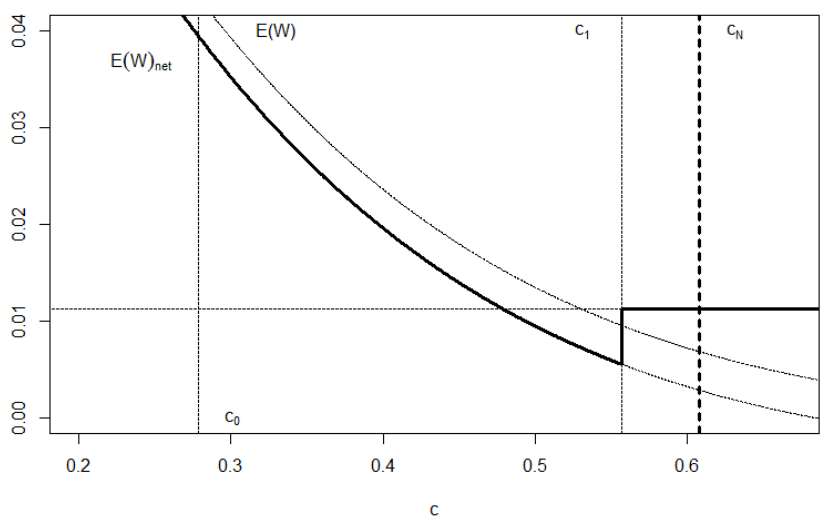

Figure 4: Pattern of expected consumer surplus $E(W)_{net}$ (thick solid line) for different levels of offshoring costs c in case 1. Offshoring takes place at the cost level $\bar{c}_g = \bar{c}_1$ and can take place at the expense of the consumers, who face a non monotonic pattern of $E(W)_{net}(c)$ with a range of offshoring cost levels leading to expected consumer surplus below the autarky level (horizontal dashed line). Thin vertical dashed lines: cutoff levels $\bar{c}_1$ and $\bar{c}_0$. Bold vertical dashed line: NGO cutoff $\bar{c}_N$.

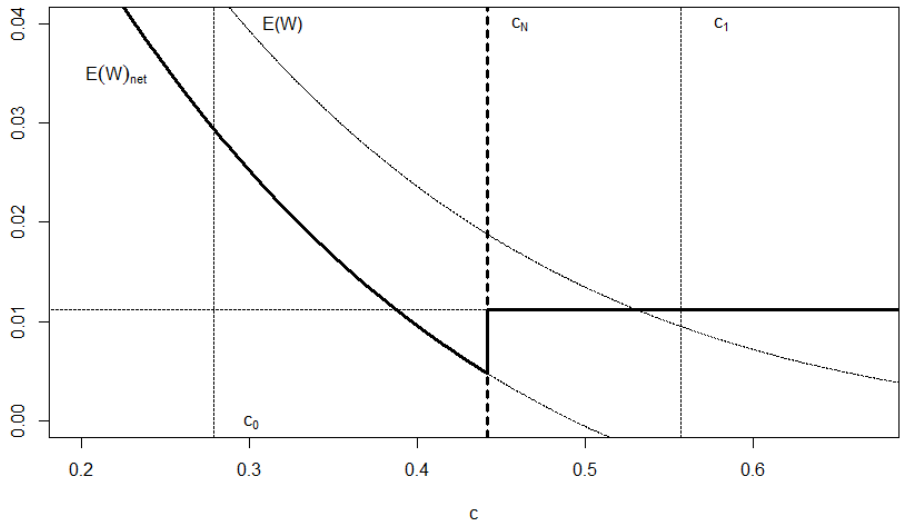

Figure 5: Pattern of expected consumer surplus $E(W)_{net}$ (thick solid line) for different levels of offshoring costs c in case 2. Offshoring takes place at the cost level $\bar{c}_g = \bar{c}_N$ and can take place at the expense of consumers, who face a non monotonic pattern of $E(W)_{net}(c)$ with a range of offshoring cost levels leading to expected consumer surplus below the autarky level (horizontal dashed line). Thin vertical dashed lines: cutoff levels $\bar{c}_1$ and $\bar{c}_0$. Bold vertical dashed line: NGO cutoff $\bar{c}_N$.

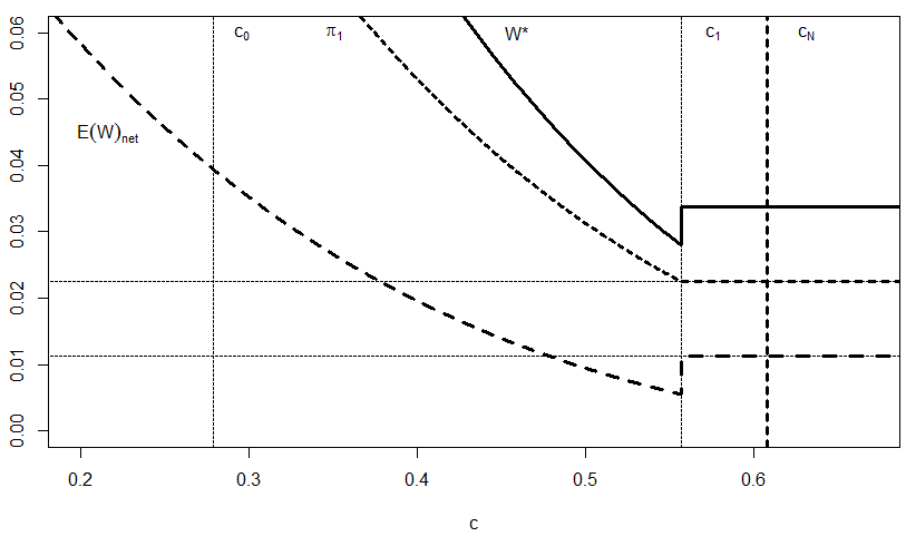

Figure 6: The solid line depicts total expected welfare W^* as the sum of expected consumer surplus $E(W)_{net}$ and firm profits under offshoring π_1 for different levels of the offshoring cost c ind case 1. As in case 1 expected profits are continuous, the discontinuity in expected consumer surplus translates into a corresponding discontinuity in expected total welfare.

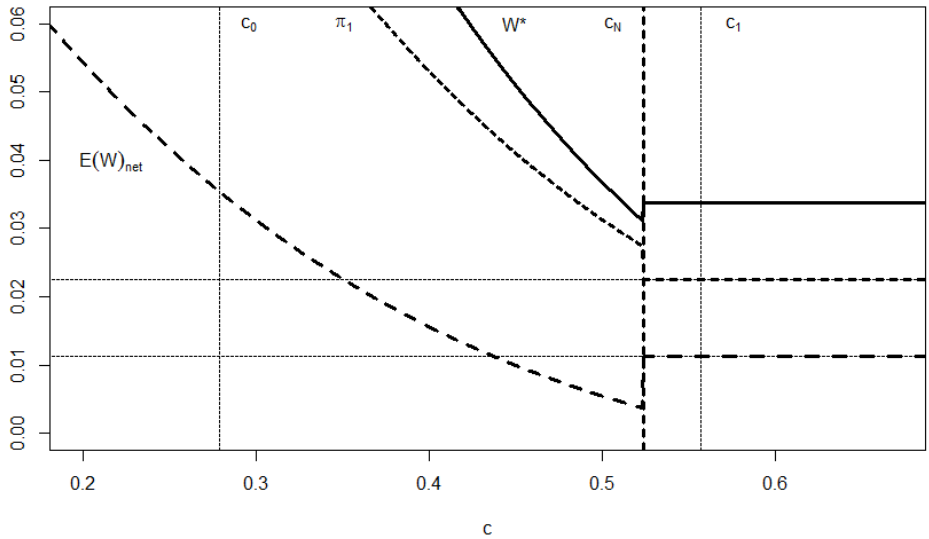

Figure 7: Numerical example in which the fall in total expected welfare W^* is preserved also in case 2 where $\bar{c}_g = \bar{c}_N$.

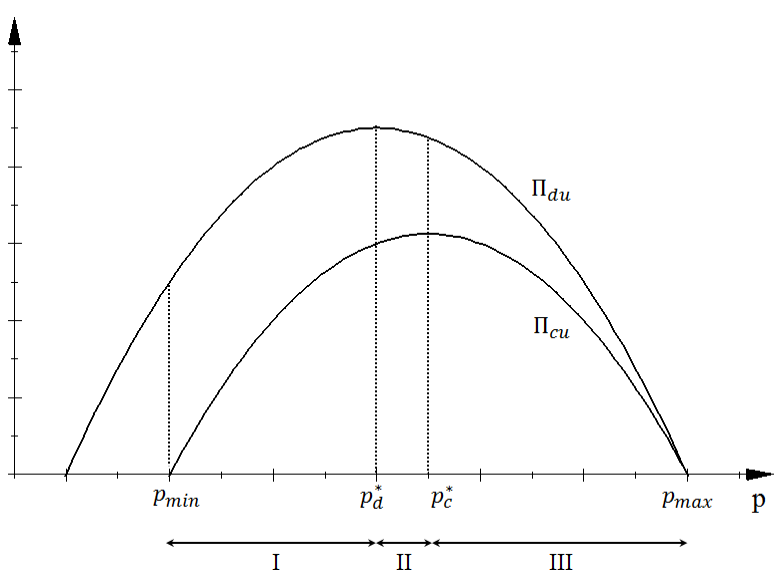

Figure 8: Expected profits of unidentified clean and dirty firm types for a given q^e. The range of pooling PBEs reaches from $p_{\min}$ to $p_{\max}$.