

# Exit vs. Voice

Eleonora Broccardo

Università di Trento

Oliver Hart

Harvard University

Luigi Zingales\*

University of Chicago

First Version, July 2020, revised, December 2020

## Abstract

We study the relative effectiveness of exit (divestment and boycott) and voice (engagement) strategies in promoting socially desirable outcomes in companies. We show that in a competitive world exit is less effective than voice in pushing firms to act in a socially responsible manner. Furthermore, we demonstrate that individual incentives to join an exit strategy are not necessarily aligned with social incentives, whereas they are when well-diversified investors are allowed to express their voice. We discuss what social and legal considerations might sometimes make exit preferable to voice.

JEL: D02, D21, D23, D62, D64, H41, L21

Keywords: Exit, Voice, Social Responsibility, Divestment, Boycott, Engagement

\* We thank Anat Admati, Philippe Aghion, Lucian Bebchuk, Tim Besley, Patrick Bolton, John Campbell, Mathias Dewatripont, Alex Edmans, Tore Ellingsen, Daniel Green, Elhanan Helpman, Harrison Hong, Louis Kaplow, Augustin Landier, Christian Leuz, Eric Maskin, Ben Roth, Martin Schmalz, Dirk Schoenmaker, Andrei Shleifer, Yossi Spiegel, Kathy Spier, David Thesmar, Stefano Zamagni, Xingtang Zhang, seminar participants at Harvard, the NBER, the University of Chicago, University of Colorado Boulder, CUNY, Boston University, Cambridge Judge Business School, HEC, University of Zurich, Vienna University of Economics and Business, and the ESSO, FOM, ECGI, and Stony Brook International Game Theory conferences for their very useful comments. Eleonora Broccardo gratefully acknowledges financial support from the Department of Economics and Management at the University of Trento. Luigi Zingales gratefully acknowledges financial support from the Stigler Center at the University of Chicago.

## 1. Introduction

In recent years, companies have come under increasing stakeholder pressure to pursue environmental and social goals. In 2019, \$20.6bn flowed to funds that explicitly divest from “non-sustainable” companies, more than 10 times the level a decade earlier (CBInsights (2020)). A recent survey suggests that 38% of Americans are currently boycotting at least one company, up from 26% only a year ago.<sup>1</sup> In the last quarter of 2019, the term ESG (Environment, Social and Governance) was mentioned 357 times in earnings calls with CEOs versus only 49 times in the last quarter of 2016 (CBInsights (2020)).

At the same time a growing academic literature has argued that the usual presumption that firms should maximize profit or market value is no longer valid in a world where, as result of political failures either at the national or international level, externalities are not well-controlled.<sup>2</sup> In particular, Hart and Zingales (2017) show that, to the extent that a firm has a comparative advantage relative to individuals in producing a public good (or avoiding a public bad), a firm’s shareholders may wish it to pursue social goals at the expense of profit. Consumers and workers may also be willing to pay a price for a firm to act in a socially responsible way.

In this paper we analyze theoretically whether pressure by stakeholders—consumers, workers, shareholders—is likely to achieve a socially desirable outcome.<sup>3</sup> For concreteness we focus on the case of environmental harm caused by pollution, such as CO2 emissions. Using Hirschman’s (1970) terminology, we can describe stakeholders’ choices as *exit* versus *voice*. Investors or consumers can exercise their *exit* option by divesting from polluting companies or boycotting their products; alternatively, investors can use their *voice* by voting or engaging with management. (We focus on consumer boycotts, but argue that worker boycotts are conceptually similar.)

We consider a situation where the harm from a polluting firm is spread globally over many individuals. Under standard assumptions that agents are purely selfish, we are faced with a severe free rider problem: the direct benefit an agent receives from any exit or voice decision is negligible.

---

<sup>1</sup> <https://www.comparecards.com/blog/38-percent-boycotting-companies-political-pandemic-reasons/>

<sup>2</sup> See, for example, Baron (2007), Benabou and Tirole (2010), Edmans (2020), Elhauge (2005), Graff Zivin and Small (2005), Hart and Zingales (2007), Magill et al. (2015), Mayer (2018), Morgan and Tumlinson (2019), Schoenmaker and Schramade (2019), and Stout (2012).

<sup>3</sup> Our approach should not be confused with what Bebchuk and Tallarita (2020) call “stakeholderism.” Stakeholderism refers to a situation where, in making business decisions, corporate leaders take into account the well-being of stakeholders (rather than just shareholders). In contrast, we are interested in analyzing how various stakeholders (including shareholders) can persuade companies to act in a more socially responsible manner.

To explain social action, we assume – consistent with empirical evidence – that some investors and consumers are socially responsible in the sense that, when they make a decision, they put a positive weight  $\lambda$  on the well-being of others affected by the decision. Thus, the decision to boycott or divest is not based on purely moral consideration, but on the consequences that these actions have (hence, we call such agents consequentialists).

In our model each firm can choose to be clean or dirty. A dirty firm produces environmental damage equal to  $h$ .<sup>4</sup> A firm can avoid this damage by incurring an additional fixed cost  $\delta$  and becoming clean. Given our simple set-up, it is socially desirable for a firm to become clean if and only if  $h > \delta$ .

We start by computing a competitive free entry equilibrium of this economy in the absence of any environmental concerns. We then study how the equilibrium changes when environmental concerns become an issue, depending on the strategy adopted by socially responsible stakeholders.

The two exit strategies are very similar in their impact. Divestment and boycotts cause the market value of a dirty firm to fall, leading some value-maximizing managers to switch to the clean technology. However, as shown by Heinkel et al. (2001), this effect is attenuated given that selfish agents will partially offset the effects of divestment/boycotting by increasing their investment/purchases in companies shunned by socially responsible agents. The magnitude of the response depends on the slope of the demand curve, which is driven by agents' risk tolerance in the case of investors and by the utility of the good in the case of consumers.

When we consider the incentive to participate in an exit strategy, we find that only those agents with a social responsibility parameter  $\lambda$  above a cut-off will choose to exit (this cut-off depends on what others are doing). There is no simple relationship between the individual incentive to participate and the social incentive to create clean firms. Divesting or boycotting can lead to too little or too much exit from the perspective of a benevolent planner. However, under plausible assumptions about the distribution of  $\lambda$  in the population and the size of  $h$  relative to  $\delta$ , the unique equilibrium is where *no* agent divests or boycotts for consequentialist reasons.

---

<sup>4</sup> In this paper we assume  $h$  to be known. Uncertainty about  $h$  generates a risk management problem (analyzed in Andersson et al. (2016)). It also makes the correlation between an individual's degree of social responsibility  $\lambda$  and her expectation of  $h$  very important. With these qualifications the main gist of our analysis applies also to uncertain  $h$ .

We carry out our analysis under the assumption that exit decisions are common knowledge and agents can commit to them. As we explain in Section 6, in the absence of this assumption, both exit strategies become even less effective.

We then consider the voice strategy. Voice can in principle be exercised by several stakeholder groups, in the form of complaints (e.g., Gans et al. (2020)) or in the form of a vote. Here we focus on the unique ability shareholders have to express voice that stems from their possession of control rights or votes. As a starting point we abstract from any existing corporate governance rules and assume shareholders are presented with a binding vote on the choice of whether the firm they invest in should be clean or dirty. The only time an individual shareholder's vote matters is when she is pivotal. Thus, as in Hart and Zingales (2017), we assume that shareholders will vote as if they were pivotal. A pivotal shareholder trades off the net social benefit from the clean technology, weighted by the shareholder's social parameter  $\lambda$ , against her own capital loss resulting from the choice of that technology. The net social benefit equals the reduced pollution minus the cost of generating that reduction. If shareholders are well-diversified, the personal capital loss is negligible. Thus, as long as  $\lambda$  is positive, the first effect dominates and socially responsible shareholders vote in line with a benevolent planner's goal.

In practice, putting proposals up for a proxy vote is expensive and it will not be in the interest of atomistic investors to incur the cost of doing so. We argue that mutual funds can use engagement as a marketing strategy and show that socially responsible agents will be willing to invest in a Green Fund that is committed to promoting an environmental agenda.

Taken literally, our simple model suggests that if a majority of agents are even (slightly) socially responsible, shareholder voice dominates exit and voice by other stakeholders. In practice, there are several frictions, and other important factors, that might attenuate or reverse this result. We discuss them in Section 7. In spite of these factors, it remains true that, when agents choose voice, their individual incentives are aligned with the social incentive, whereas this is not the case when they choose exit.

There is a vast literature on socially responsible investment (SRI). Benabou and Tirole (2010), Kitzmueller and Shimshack (2012), and Christiansen et al. (2019) provide very useful overviews. On the divestment side, the first formal model is Heinkel et al. (2001). Our model of divestment is similar to theirs, but with the difference that they take as given that socially responsible investors refuse to hold shares of dirty companies, whereas we suppose that socially

responsible investors make the divestment decision based on the impact this decision has. Also our model incorporates boycotts and voice as well as divestment. Pastor et al. (2020) extend the Heinkel et al. (2001) model to derive an ESG factor in an equilibrium asset-pricing model when investors have a taste for ESG (for another paper along similar lines, see Pedersen et al. (2019)). They do endogenize the divestment decision, but under the assumption that investors are purely selfish<sup>5</sup>. Graff Zivin and Small (2005) and Morgan and Tumlinson (2019) suppose that investors value public goods and pay more for the shares of firms that bundle private and public goods; see also Aghion et al. (2020) and Bonnefon et al. (2019). However, each investor is selfish in that he values his consumption of the public good and not the utility from the public good accruing to others. Baron (2007), Chowdhry et al. (2019), and Gollier and Pouget (2014) consider the impact of divestment, but for the case of large as opposed to atomistic investors. Landier and Lovo (2020) study the social welfare effect of selected investment by an ESG fund that has some market power, while Oehmke and Opp (2020) and Green and Roth (2020) analyze optimal investment choices for large socially responsible investors who fund wealth-constrained entrepreneurs, exploring the complementarities between the actions of social investors and those of selfish investors.

There is also a smaller literature on consumer boycotting (see Kitzmueller and Shimshack (2012) for a survey). Boycotts can be seen as a way to redistribute surplus (see Baron (2001)), or as a way to induce companies to provide a public good (see Bagnoli and Watts (2003) and Besley and Ghatak (2007)). In Bagnoli and Watts (2003) and Besley and Ghatak (2007), each consumer is selfish in that he values his consumption of the public good and not the utility from the public good accruing to others.

There is also a vast literature on corporate social responsibility. This literature argues that companies can or should have a purpose beyond profit or value maximization, including to act in a socially responsible manner (e.g., Edmans (2020), Magill et al. (2015), Mayer (2018), Schoenmaker and Schramade (2019), and Stout (2012)). In contrast, we assume that some individuals are socially responsible and derive the consequences for corporate behavior, depending on the tools these socially responsible individuals have at their disposal.

Our work is related to, but different from, the literature on private politics (Baron (2003)). “Private politics refers to actions by private interest, such as activists and NGOs, that target private

---

<sup>5</sup> Admati and Pfleiderer (2009) consider a model where the threat by a large privately-informed shareholder to divest can put pressure on management to adopt a value-maximizing strategy, under the assumption that investors are purely selfish.

agents, typically firms, in the institution of public sentiment” (Abito et al. (2019)). The difference is that our agents are socially responsible, so they pursue the public interest, not just the private one.

The rest of the paper proceeds as follows. Section 2 describes our assumption on socially responsible investors and consumers. Section 3 presents the model. Section 4 analyzes the exit strategy, Section 5 the voice one. Section 6 covers robustness and extensions, and Section 7 includes discussion and qualifications. Section 8 concludes.

## 2. Socially Responsible Investors and Consumers

Responsible investing dates back at least as far as 1758, when the Philadelphia Yearly Meeting of the Society of Friends required its members to cease and desist from slaveholding (Brown (1988)). Consumer boycotting can be traced back even further to the vegetarianism of the Jain religion (Laidlaw (1995)). The rejection of slavery by the Quakers and of animal products by the Jains was on moral grounds, and thus did not lend itself to any economic calculus.<sup>6</sup> This original perspective survives in much of the contemporary socially responsible investment literature. From Heinkel et al. (2001) to Hong and Kacperczyk (2009), the early literature assumes that some investors simply do not want to own certain kinds of stocks. Such an approach is appropriate for “sinful” products, like tobacco, alcohol, or prostitution, but applies less well to social concerns that are less of a moral nature. Most investors are not morally against companies that emit CO<sub>2</sub>, they would just like these companies to emit less of it. Trinity Church was not morally against Walmart, it simply wanted Walmart not to sell assault weapons, and so on.

Some of the literature on socially responsible investment and consumption departs from the purely moral view. For example, Graff Zivin and Small (2005), Morgan and Tumlinson (2019), Bagnoli and Watts (2003), and Besley and Ghatak (2007) endogenize investor and consumer choice by assuming that an individual will value a share or good based on a combination of its private characteristics and the increased harm resulting from production. However, these authors assume that individuals consider only the personal disutility of the increased harm, ignoring the impact on others. As a result, in a large economy, there will be an extreme free rider

---

<sup>6</sup> In the *Wealth of Nations*, Adam Smith expressed skepticism that the Quakers would have voted to free their slaves if they had many slaves. But, according to Pack and Dimand (1996, p.268), “The Quakers of Philadelphia did make a substantial financial sacrifice when they freed their slaves.”

problem, leading to a large deviation between private and social optimality. Sugden (1982) convincingly argues that such a model is inconsistent with the evidence on charitable contributions. One way to mitigate the free rider problem is to introduce a “warm glow” effect, along the line of Andreoni (1989).<sup>7</sup> In a sense this is what Pastor et al. (2020) do in assuming an individual taste for green investment. However, in Pastor et al.’s approach, investors ignore their impact on others. For a recent paper in which moral individuals take into account their impact on others and act as consequentialists, see Schmidt and Herweg (2020).

In our model socially responsible individuals are altruistic in the sense that they put some weight on the utility of others. This assumption is uncontroversial for foundations that have an explicit social goal, such as the Gates Foundation. Yet, there is growing evidence in support of this assumption also for individual agents: see Andreoni and Miller (2002), Charness and Rabin (2002), Riedl and Smeets (2017), Brodback et al. (2019), and Bauer et al. (2020).<sup>8</sup> We adopt Hart and Zingales (2017)’s formulation: we assume that, in making a decision, an individual puts weight  $\lambda \in [0,1]$  on the welfare of those affected by the decision, where  $\lambda$  reflects her degree of social responsibility.<sup>9</sup> Consider, for example, the decision to wear a mask during a pandemic to protect others from the risk of being infected, when this decision is not mandatory. An individual will compare the private cost of her decision, say 10, with the social benefit of the decision, say 50, where the latter is weighted by  $\lambda$ . If  $\lambda > \frac{1}{5}$ , she will wear a mask, if not she will not.<sup>10</sup>

As in Hart and Zingales (2017), we assume that the socially responsible component enters at the time a decision is made, but not after the decision is made.<sup>11</sup> Assuming otherwise would lead to the paradoxical result that a pandemic raises people’s utility. To appreciate this point, go back to the mask example and suppose  $\lambda = 1/2$ . An individual with such a high  $\lambda$  will wear a mask, since

---

<sup>7</sup> Another way is to introduce reciprocal behavior along the lines of Sugden (1984).

<sup>8</sup> Andreoni and Miller (2002) and Charness and Rabin (2002) find support for such preferences in lab experiments. A preference for socially responsible investment has also been found in field experiments in situations where this preference yields lower expected returns (Bauer et al. (2020) and Riedl and Smeets (2017)). This preference is positively correlated with the degree of altruism (Brodback et al. (2018)). Such a preference is also consistent with the lower return of impact funds (Barber et al. (2020)).

<sup>9</sup> For similar formulations, see Acquatella (2020), Besley and Ghatak (2018), and Frydinger and Hart (2019). In contrast to Hart and Zingales (2017), we do not assume that an agent acts altruistically only when she feels responsible for a situation that has arisen; and we drop the (ad hoc) assumption that the impact on others is weighted by an investor’s shareholding.

<sup>10</sup> Consistent with our model, U.S. counties with higher civic capital (which can be interpreted as a higher  $\lambda$ ) wear masks more frequently and socially distance more; see Barrios et al. (2020).

<sup>11</sup> Acquatella (2020) and Frydinger and Hart (2019) make a similar assumption.

$-10 + 1/2(50) = 15 > 0$ . Yet, it is unreasonable to think that 15 is her final utility, because she would then be better off as a result of the pandemic. By contrast, if we assume (as we do in the rest of the paper) that the social responsibility component of the utility plays a role only in the decision-making process, but does not enter the final utility, then the final utility of the individual is  $-10$  and thus the pandemic reduces her utility.

Note that the only place in the analysis where including the socially responsible component in the utility function might change the results is in the calculation of the benevolent planner's solution in 3.4.

One interesting question is how broad is the group of people whose welfare enters a socially responsible individual's calculations: does it include people in one's neighborhood, the whole town, the whole country, or the whole world? The answer depends upon the socially responsible perspective of an individual and what she considers her relevant community. In this paper, we assume that the community includes everyone affected by the pollution. We return to this issue in Section 7.3.

### 3. The Model

#### *3.1 The case where pollution is not a problem*

Consider a three-date economy, as shown in Figure 1. There are three distinct groups: entrepreneurs, investors and consumers. At date 0 entrepreneurs can set up firms; they then leave the scene. Production decisions are made at date 1. Production and consumption take place at date 2. Entrepreneurs care only about date 0 money and have zero wealth. Investors care only about date 2 return. Investors and consumers are socially responsible but this does not affect the equilibrium in this subsection since at date 0 pollution is not an issue (and is not expected to be an issue).

There is a set-up cost  $F$  for each firm, and each firm has zero marginal cost up to a capacity constraint equal to one. After the set-up cost has been sunk, there is an additional fixed cost of production  $C$ , incurred at date 2. The expected value of  $C$  is zero, but  $C$  is uncertain. We suppose

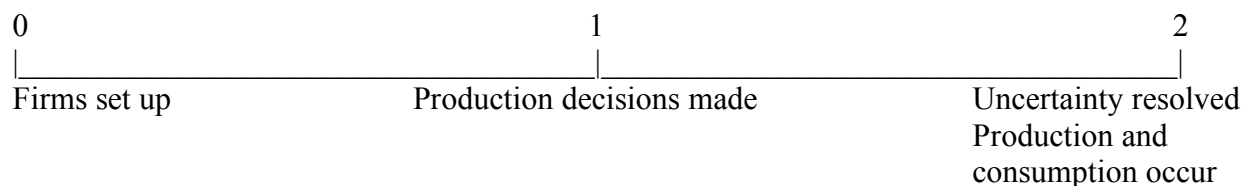
$$(3.1) \quad C = \varepsilon,$$

where  $\varepsilon$  is an aggregate shock, which is normally distributed with mean 0 and variance  $\sigma^2$ ;  $\varepsilon$  is realized at date 2. There is symmetric information throughout. We assume that the shock is an



aggregate one so that the limited risk bearing capacity of investors plays a role. However, there could also be an idiosyncratic component of the shock, which would explain why investors diversify their portfolios across firms.

Figure 1: Timeline



Entrepreneurs cover the date 0 set-up cost  $F$  by issuing shares to investors. Investors have an exponential utility function

$$(3.2) \quad U = -e^{-\gamma\omega},$$

where  $\omega$  is their final wealth. Investors hold the shares until date 2, when output is sold and profit is realized. However, at date 1 there can be some portfolio rebalancing.

We will study a competitive free entry equilibrium. In the basic economy, we normalize the number of investors and the number of consumers each to be one. Of course, a one investor, one consumer economy is not competitive. Therefore, in order to make the economy competitive, we will replicate it and take limits, as described below.

Assume that the product market consists of a homogenous good (whose origin can be easily determined, e.g., electricity). Suppose that the consumer's utility function is

$$(3.3) \quad U = \rho q - \frac{1}{2} \tau q^2 - pq$$

where the third term is the cost of buying  $q$  units of the good at price  $p$ . The maximization of this utility leads to the following demand curve,

$$(3.4) \quad p = \rho - \tau q, \quad q = \frac{\rho - p}{\tau}.$$

Output is sold in a competitive market at date 2. Production decisions are made at date 1. Each firm produces up to its capacity constraint of one since price exceeds the expected value of  $C$ , which is zero. Thus total supply equals  $N$ , where  $N$  is the number of firms set up at date 0, and equilibrium in the date 2 goods market is given by

$$(3.5) \quad N = \frac{\rho - p}{\tau}.$$

Each firm's date 2 profit is

$$(3.6) \quad \Pi = p - \varepsilon = \rho - \tau N - \varepsilon,$$

and expected profit is

$$(3.7) \quad \Pi = \rho - \tau N.$$

Consider the investor's date 0 portfolio decision. Assume that the investor can borrow and lend at a zero rate of interest. In a free-entry equilibrium the market value of each firm at date 0 must be  $F$  since otherwise firms would enter or exit. The total return for the investor at date 2 is therefore  $x\Pi - xF$ , where  $x$  is his investment level and we normalize the investor's initial wealth to be zero. This return has a certainty equivalent equal to

$$(3.8) \quad \text{CE} = x(\Pi - F) - \frac{1}{2}\gamma x^2 \sigma^2.$$

The investor's demand for shares at date 0 will be given by the  $x$  that maximizes this certainty equivalent. Thus,

$$(3.9) \quad x = \frac{\Pi - F}{\gamma \sigma^2}.$$

(3.9) provides the total demand for firms' shares. The total supply is equal to  $N$ . Hence, for the stock market to clear at date 0 we must have

$$(3.10) \quad \frac{\Pi - F}{\gamma \sigma^2} = N.$$

Using (3.7), we obtain

$$(3.11) \quad N = \frac{\rho - F}{\gamma \sigma^2 + \tau}.$$

This is the equilibrium number of firms that will set up at date 0.<sup>12</sup> From now on we assume  $\rho > F$ , so  $N > 0$ . For future reference, it is useful to derive the formula for the certainty equivalent at the optimal investment level  $x$ . This is obtained by substituting (3.9) into (3.8):

---

<sup>12</sup> We ignore the fact that the solution to (3.11) may not be an integer. This will become unimportant in the limit economy described below.

$$(3.12) \quad CE = \frac{1}{2} \frac{(\Pi - F)^2}{\gamma\sigma^2} = \frac{1}{2} \frac{(\rho - \tau N - F)^2}{\gamma\sigma^2}.$$

### 3.2 Replica economy

The economy as it stands is not competitive. To make it so we replicate the investor and consumer sectors  $r$  times and take limits as  $r \rightarrow \infty$ . In the replica economy there are  $r$  investors with the above investor preferences and  $r$  consumers with the above consumer preferences. (There is always an unlimited supply of entrepreneurs.) It is easy to see that the equilibrium number of firms will be  $Nr$ , where  $N$  is given by (3.11). For large  $r$  each investor, consumer and firm is small relative to the aggregate economy and so has little influence on market prices. In other words, for large  $r$  the economy is approximately perfectly competitive, and in the limit  $r = \infty$  it is perfectly competitive.<sup>13</sup>

In the equilibrium of the basic economy the single investor holds 100% of each of the  $N$  firms. In the replica economy we assume that each of the  $r$  investors holds  $1/r$  of each of the  $Nr$  firms, i.e., each investor is fully diversified.

In what follows we will have the replica or limit economy in mind even though we will not always be explicit about it. When we study the effects of individual divestment, boycott, and engagement decisions the replica economy will be particularly relevant.

### 3.3 Pollution Becomes a Problem at Date 1

Suppose that at date 1 pollution becomes a problem (to emphasize, this eventuality is unanticipated at date 0).<sup>14</sup> Operating with the existing technology (which we will now label dirty), each firm produces harm  $h > 0$  to the environment at date 2. We assume that the total harm from a single firm stays the same as the economy is replicated (replication simply makes the economy more competitive). We also suppose that this harm is spread over the whole population and so the harm an individual investor or consumer experiences from a single firm converges to zero as  $r \rightarrow \infty$ .<sup>15</sup>

<sup>13</sup> For details, see, e.g., Mas-Colell, Whinston, and Green (1995).

<sup>14</sup> We consider a rational expectations equilibrium in Section 6.

<sup>15</sup> As an example, suppose that the environmental harm is the loss of beach space due to the rising sea level. Before pollution, there are  $B$  beach spaces available in the world. Given that there are  $Nr$  investors and  $Nr$  consumers, each individual is able to occupy a beach space for a fraction  $B/2Nr$  of the day. Imagine that a firm, emitting a certain number of CO<sub>2</sub> tons, causes the sea level to rise, reducing the number of beach spaces available by  $\alpha\%$ . If  $b$  represents an individual's utility from a full day at the beach, and utility is linear in beach consumption, then total

A firm can avoid polluting by incurring an additional fixed cost  $\delta$  at date 1; this fixed cost comes out of date 2 profits. We call the firms that decide to pay this cost “clean”. Thus, the cost of a clean firm is

$$(3.13) \quad C^C = \delta + \varepsilon,$$

while the cost of a dirty firm is as before

$$(3.14) \quad C^D = \varepsilon.$$

We assume that

$$(3.15) \quad \delta < F.$$

(3.15) ensures that a firm prefers to install the clean technology rather than closing down.

If all investors and consumers are purely selfish, the existence of pollution will not change any production or investment decision significantly when  $r$  is large. The reason is that, since the pollution impact of any production and investment decision on each individual converges to zero as  $r \rightarrow \infty$ , nobody internalizes the pollution externalities (as in Pastor et al. (2020)). As we will see shortly, this is not the case when people are socially responsible. In this case, the outcome depends upon the strategy adopted by socially responsible investors and consumers. Before analyzing this, however, we need to consider what a benevolent planner would do, so that we have an appropriate benchmark.

### 3.4 Benevolent Planner’s Response to Environmental Damage

As a benchmark, we derive a benevolent planner’s solution in a world *where all investors and consumers are purely selfish*.<sup>16</sup> The number of firms  $N$  that entrepreneurs have set up at date 0 is given at date 1. However, a benevolent planner can dictate what technology—clean or dirty—each firm should adopt at date 1, that is, she can choose the proportion of clean firms  $\phi = \frac{n_c}{N}$ . Assume that this is the *only* instrument at the planner’s disposal. That is, the planner chooses  $\phi$  and then lets the date 1 stock market and the date 2 product market clear. The question is at what level will she set  $\phi$ .

---

utility falls from  $Bb$  to  $(1 - \alpha) Bb$  for large  $r$ , regardless of the size of  $r$ . Hence, the damage caused by the firm is  $\alpha Bb$ , which is  $h$  in our model.

<sup>16</sup> The solution is the same under the assumption that investors and consumers are socially responsible but the socially responsible component does not enter their final utility. See the discussion in Section 2.

We suppose that the planner's objective is to maximize the sum of investor and consumer surplus, net of the harm imposed by pollution. In the Appendix we show that the solution is very simple. If  $\delta > h$ , that is, the cost of avoiding pollution is bigger than the cost of pollution itself, the planner will want all firms to use the dirty technology ( $\phi=0$ ), while, if  $\delta < h$ , that is, the cost of avoiding pollution is less than the cost of polluting, the planner will want all firms to become clean ( $\phi=1$ ).

#### 4. Exit Strategies

##### *4.1 Divesting*

We now analyze what happens when there is no planner (or government) and social action is left to individual investors or consumers. As in Section 2, we assume that, in making a decision, an individual puts weight  $\lambda \in [0,1]$  on the welfare of those affected by the decision, where  $\lambda$  reflects her degree of social responsibility. For simplicity we suppose that the distribution of  $\lambda$  in the population is the same for investors and consumers. The distribution has finite support  $[\lambda_1, \dots, \lambda_n]$ , where  $\lambda_1 < \dots < \lambda_n$ , with associated strictly positive probabilities  $\pi_1, \dots, \pi_n$ .<sup>17</sup> (Here  $\lambda_1$  could be zero.)

We will study equilibrium in the limit economy where  $r = \infty$ , but in order to analyze individual exit decisions we will take limits as  $r \rightarrow \infty$ . We consider first the strategy of divestment by investors. Assume that a fraction  $\mu$  of investors announce at date 1 that they will hold shares only in clean firms; we will see below that only investors with a  $\lambda$  above a particular cut-off will choose to divest. We suppose that investors' announcements are visible and that investors can commit to their divestment decisions (we return to the visibility and commitment issue in Section 6). Firms observe these announcements, and then decide whether to stay dirty or become clean. We want to characterize a (Nash) equilibrium. To this end we derive the product market and capital market equilibrium under the assumption that a fraction  $\mu$  of investors divest. Then, we check that a fraction  $\mu$  of investors do indeed want to divest. In this subsection we assume that there is no consumer boycott.

---

<sup>17</sup> To avoid the replica economy being stochastic, the reader can imagine that each  $\lambda$  type is represented in the replica economy exactly according to its frequency.

We suppose that at date 1 firms are run by value-maximizing managers. One can imagine that (before there were any environmental concerns) initial entrepreneurs designed an incentive scheme to encourage managers to maximize market value at date 1 in order to obtain the highest valuation at date 0 (there could be some unmodeled agency problems). Note that initial entrepreneurs are not well-diversified and so they want to maximize the value of their own company, not the joint value of all companies, as the common ownership literature suggests (see Azar et al. (2018))<sup>18</sup>.

Value maximization implies that in an equilibrium where both clean and dirty firms operate they must have the same value  $V$ , otherwise there would be switching.<sup>19</sup> Let  $n_c$  be the number of clean firms and  $n_d = N - n_c$  the number of dirty firms. Note that the mix of clean and dirty firms has no effect on the date 2 product market equilibrium since each firm will supply at its capacity constraint of one whether it is clean or dirty.

For divestors, the analogy of (3.9) is

$$(4.1) \quad x = \frac{\Pi - \delta - V}{\gamma\sigma^2},$$

since C firms yield expected profits  $\Pi - \delta$ , rather than  $\Pi$ , and cost  $V$ . Since divestors represent a mass  $\mu$  of investors, their demand for clean firms is

$$(4.2) \quad \mu x = \mu \left( \frac{\Pi - \delta - V}{\gamma\sigma^2} \right).$$

The rest of the market will not invest in clean firms since they are less profitable, but equally expensive. Hence, (4.2) represents the total demand for clean firms, and we must have

---

<sup>18</sup> We consider the possibility of socially responsible entrepreneurs in Section 6. In this paper we do not discuss how incentive contracts can affect the ESG decision of managers; on this, see Davies and Van Wesep (2018).

<sup>19</sup> An interesting question is whether a purely selfish investor could take advantage of the fact that clean and dirty firms have the same price, but different expected profitability, by short selling one and using the proceeds to invest in the other. The feasibility of this strategy depends on whether socially responsible investors are willing to lend shares to short sellers and whether they are willing to accept borrowed shares as “bona fide” clean shares. In our model, where socially responsible investors care about their impact, the answer to both questions is negative. A socially responsible investor, who accepts a lower return for a greater cause, would be foolish to lend his shares to a speculator who undoes his strategy without fully compensating him. The same is true for an investor buying lent shares.

$$(4.3) \quad \mu \left( \frac{\Pi - \delta - V}{\gamma \sigma^2} \right) = n_c .$$

Similarly, the demand for dirty firms will be given by

$$(4.4) \quad (1 - \mu) \left( \frac{\Pi - V}{\gamma \sigma^2} \right),$$

which must be equal to  $n_d$ :

$$(4.5) \quad (1 - \mu) \left( \frac{\Pi - V}{\gamma \sigma^2} \right) = n_d .$$

Adding (4.3) and (4.5) yields

$$(4.6) \quad \Pi - V - \mu \delta = N \gamma \sigma^2 .$$

We know from (3.10) that  $N \gamma \sigma^2 = \Pi - F$ , and therefore

$$(4.7) \quad V = F - \mu \delta .$$

Substituting back into (4.3), we obtain

$$(4.8) \quad n_c = \mu \left( \frac{\Pi - F - \delta(1 - \mu)}{\gamma \sigma^2} \right) \\ = \mu N - \frac{\mu \delta (1 - \mu)}{\gamma \sigma^2} .$$

At this point it is helpful to provide some intuition. To understand (4.7), note that divestment leads to a fall in the demand for dirty firms' shares, causing  $V$  to fall. If  $V$  fell by  $\delta$ , clean firms would have the same net return as dirty firms previously, while dirty firms would have a higher net return.

As a result, the demand for shares would exceed the supply. Hence  $V$  must fall by less than  $\delta$ , indeed by  $\mu\delta$  according to (4.7)<sup>20</sup>.

(4.7) also throws light on (4.8). If  $V$  fell by  $\delta$ , the demand for clean firms' shares would be in proportion to the number of divestors since divestors would invest as much as before. However, since  $V$  falls less, the demand for clean shares is lower and the number of clean firms is less than proportional to the number of divestors (see also Heinkel et al. (2001) on this). Indeed  $n_c$  is quadratic in  $\mu$ . It is also clear that  $n_c$  is increasing in  $\gamma\sigma^2$ , reflecting the fact that the impact of divestment will be greater if the environment is riskier or investors are more risk averse since divestment will cause share prices to fall more.

(4.8) implies that the marginal impact of divestment is increasing in  $\mu$ . If  $N < \frac{\delta}{\gamma\sigma^2}$ , we have a corner solution: the number of clean firms  $n_c=0$  in a neighborhood of  $\mu=0$  and, for low  $\mu$ , the marginal impact of  $\mu$  on  $n_c$  is zero. Note that a corner solution will occur if  $\gamma\sigma^2$  is sufficiently low. Under these conditions it is an equilibrium for no investor to divest: starting at  $\mu=0$ , nondivestors will absorb any divested stock with minimal price impact and as a result no firms will become clean.

Conversely,  $n_c > 0$  if  $N > \frac{\delta(1-\mu)}{\gamma\sigma^2}$ . From now on we assume that we are at an interior solution for any  $\mu > 0$ , that is,

$$(4.9) \quad N > \frac{\delta}{\gamma\sigma^2}.$$

We next determine whether an investor wants to divest when (4.9) holds. As a first step, we compare the certainty equivalent of a divestor with the certainty equivalent of a nondivestor. We then bring in the environmental impact of divestment.

Since nondivestors invest only in high return dirty firms, their payoff is given by

$$(4.10) \quad x\Pi + (x_0 - x)V - x_0F,$$

---

<sup>20</sup> Note that  $V = F - \mu\delta > 0$  given (3.15). In other words, a value-maximizing firm prefers to adopt the clean technology rather than closing down.



where  $x_0$  is their date 0 investment.

The certainty equivalent of (4.10) is

$$(4.11) \quad x(\Pi - V) + x_0(V - F) - \frac{1}{2}\gamma x^2\sigma^2,$$

and the  $x$  that maximizes (4.11) is

$$(4.12) \quad x = \frac{\Pi - V}{\gamma\sigma^2}.$$

Substituting (4.12) and (3.9) (with  $x=x_0$ ) into (4.11) and using (4.7), we obtain the following expression for the CE of a nondivestor:

$$(4.13) \quad CE_{nd} = \frac{1}{2\gamma\sigma^2}(\Pi - F + \mu\delta)^2 - \mu\delta \frac{\Pi - F}{\gamma\sigma^2}.$$

Carrying out the same exercise for a divestor yields

$$(4.14) \quad CE_d = \frac{1}{2\gamma\sigma^2}(\Pi - F - \delta + \mu\delta)^2 - \frac{\mu\delta}{\gamma\sigma^2}(\Pi - F).^{21}$$

Thus by divesting an investor loses

$$(4.15) \quad CE_{nd} - CE_d = \frac{\delta}{2\gamma\sigma^2}(2\Pi - 2F - \delta(1 - 2\mu)).$$

Note that the right-hand side of (4.15) is decreasing in  $\gamma\sigma^2$ , reflecting the fact that when  $\gamma\sigma^2$  is low, the price of the risky asset will be close to its expected return and so the investor does not lose much from not investing in it. We see that a low  $\gamma\sigma^2$  reduces the (total) impact of divestment, but also reduces the cost of divesting. We will see later that the latter effect may outweigh the former.

An investor will compare the loss in (4.15) with the effect her divestment has on the environment and on other people's utilities. To evaluate this effect we compute it for the replica economy and then take limits as  $r \rightarrow \infty$ . In the replica economy there are  $r$  investors,  $\mu r$  of whom divest;  $r$  consumers; and  $Nr$  firms set up in the free entry equilibrium, of which  $n_c r$  choose to

---

<sup>21</sup> Note that (4.9) implies  $n_c > 0$ , which in turn, given (4.3), implies  $\Pi - F - \delta + \mu\delta > 0$ .

become clean at date 1, where  $n_c$  is given by (4.8). The effect of one investor's divestment decision is composed of three elements: the impact on other investors, the impact on consumers, and the impact on the environment. Investors are optimizing and so, by the envelope theorem, a small change in the market value of firms caused by one investor divesting will have a second order effect on other investors. Consumers will be unaffected because total supply equals  $N$ , independent of the mix of clean and dirty firms. Thus, we are left with the effect on the environment.

Currently  $\mu r$  investors are divesting. If one investor stops divesting  $\mu$  changes from  $\mu$  to  $\mu - \frac{1}{r}$ , i.e.,  $\Delta\mu = -\frac{1}{r}$ . The number of clean firms changes from  $n_c r$  to  $(n_c + \frac{\partial n_c}{\partial \mu} \Delta\mu)r$ , plus some second order terms. That is, as  $r \rightarrow \infty$ , the change in the number of clean firms is

$$(4.16) \quad \frac{\partial n_c}{\partial \mu} \left(-\frac{1}{r}\right)r = -\frac{\partial n_c}{\partial \mu} = -N + \frac{\delta(1-2\mu)}{\gamma\sigma^2},$$

where we use (4.8). So the damage created by the investor's decision not to divest is  $\left[ N - \frac{\delta(1-2\mu)}{\gamma\sigma^2} \right] h$ , which the investor weights by her socially responsible parameter  $\lambda$ . She then compares this to the expression in (4.15). We may conclude that an investor will be willing to stay divested if

$$(4.17) \quad \frac{\delta}{2\gamma\sigma^2} (2\Pi - 2F - \delta(1-2\mu)) \leq \lambda h \left[ N - \frac{\delta(1-2\mu)}{\gamma\sigma^2} \right],$$

which can be rewritten, using (3.10), as

$$(4.18) \quad (\lambda h - \delta) \left( N - \frac{\delta}{\gamma\sigma^2} \right) + \frac{\mu\delta}{\gamma\sigma^2} (2\lambda h - \delta) \geq \frac{\delta^2}{2\gamma\sigma^2}.$$

Note that the left-hand side (LHS) is increasing in  $\lambda$ , while the right-hand side (RHS) is constant, from which we conclude that there is a cut-off: only investors with  $\lambda$  above a critical value will divest. It is easy to show that the cutoff is decreasing in  $\mu$ .

We can use (4.18) to characterize a divestment equilibrium for the limit economy. In the following recall that  $\lambda = \lambda_i$  with probability  $\pi_i$ .

Definition 1. A divestment equilibrium for the limit economy ( $r = \infty$ ) is a  $0 \leq \mu^* \leq 1$ , where  $\mu^*$  represents the fraction of investors who divest, such that one of the following holds:

- (1)  $\mu^*=0$ , and the LHS of (4.18) is less than or equal to the RHS at  $\mu = 0$ ,  $\lambda = \lambda_n$ .
- (2)  $\mu^*=1$ , and the LHS of (4.18) is greater than or equal to the RHS at  $\mu = 1$ ,  $\lambda = \lambda_1$ .
- (3)  $\mu^* = \sum_{j=i+1}^n \pi_j$  for some  $i = 1, \dots, n - 1$ , and the LHS of (4.18) equals the RHS at  $\mu = \mu^*$  for some  $\lambda_i < \lambda^* < \lambda_{i+1}$ .
- (4)  $\sum_{j=i+1}^n \pi_j < \mu^* < \sum_{j=i}^n \pi_j$  for some  $i=1, \dots, n$ , and LHS of (4.18) equals the RHS at  $\mu = \mu^*$ ,  $\lambda = \lambda_i$ .

To understand this definition, note that in (1) no-one divests and nobody wants to divest. In (2) everyone divests and everyone wants to divest. In (3) the cut-off is such that only those whose  $\lambda$  strictly exceeds  $\lambda^*$  want to divest and the fraction of them is  $\mu^*$ . (4) is like (3) except that the fraction  $\mu^*$  of divestors is made up of those who strictly want to divest ( $\lambda > \lambda_i$ ) and those who are indifferent ( $\lambda = \lambda_i$ ).

(4.18) has a number of implications for the nature of equilibrium. First, as we shall see below, it can fail to hold when  $h > \delta$  but can hold when  $h < \delta$ . In other words the private incentive to divest is not aligned with the social incentive to create clean firms. Second, since the cutoff is decreasing in  $\mu$ , as  $\mu$  increases the fraction of investors with a  $\lambda$  above the cut-off rises. This suggests that there can be multiple equilibria (something that we will verify below). Third, it is easy to show that the cut-off is increasing in  $\gamma \sigma^2$  if  $\mu > 1/2$ : in other words, keeping  $\mu$  fixed, divestment becomes *less* attractive in an environment that is riskier or where investors are more risk-averse (the impact of divestment rises but the cost rises more).

Proposition 1 states that a divestment equilibrium always exists.

**Proposition 1:** A divestment equilibrium exists.

Proof:

We use a fixed point argument. For each  $\lambda \geq 0$ , define the correspondence  $G(\lambda) = \{1\}$  if  $\lambda < \lambda_1$ ,  $G(\lambda) = \{\sum_{j=i+1}^n \pi_j\}$  if  $\lambda_i < \lambda < \lambda_{i+1}$  ( $i=1, \dots, n-1$ ),  $G(\lambda) = [\sum_{j=i+1}^n \pi_j, \sum_{j=i}^n \pi_j]$  if  $\lambda = \lambda_i$  ( $i=1, \dots, n$ ),  $G(\lambda) = \{0\}$  if  $\lambda > \lambda_n$ . For each  $\mu$ , let  $\lambda(\mu)$  be the unique value of  $\lambda$  such that the LHS

of (4.18) equals the RHS. (Here  $\lambda(\mu)$  could exceed 1.) Now consider the correspondence  $\varphi$  from  $[0,1]$  into itself, where  $\varphi(\mu) = G(\lambda(\mu))$ . It is easy to see that  $\varphi$  is upper hemicontinuous and convex-valued and so by Kakutani's fixed point theorem there exists  $\mu^*$  such that  $\mu^* \in \varphi(\mu^*)$ . It is easy to check that  $\mu^*$  is a divestment equilibrium. Q.E.D.

In the following three propositions we will focus on the case where  $\lambda_n h < \delta$ . Note that if  $\lambda_n h > \delta$  we would expect to see the most socially responsible investors (of whom there are an infinite number in the limit economy) approaching dirty firms and individually paying them to become clean, which does not seem very realistic (see also the discussion of Coasian bargaining in Section 6.2). Proposition 2 covers this leading case, and also a second case where  $\lambda_n h < (\frac{3}{4})\delta$ , that is, the most socially responsible investors are not willing to pay for  $\frac{3}{4}$  of the cost of turning a firm clean.

**Proposition 2:**

- (1) Suppose that  $\lambda_n h < \delta$ . Then  $\mu = 0$  is an equilibrium.
- (2) Suppose that  $\lambda_n h < (\frac{3}{4})\delta$ . Then  $\mu = 0$  is the unique equilibrium.

Proof:

Proposition 2(1) follows from the fact that the LHS of (4.18) is negative for all  $\lambda$  if  $\mu = 0$ . Proposition 2(2) follows from the fact that the second term of the LHS of (4.18) is negative whenever  $2\lambda h < \delta$ ; and the second term is less than the RHS if  $2\lambda h > \delta$  (it is easy to show the latter when  $\mu = 1$ ,  $\lambda = \lambda_n$ , and hence the latter must be true for all  $\mu, \lambda$  since the second term is monotonically increasing in  $\mu, \lambda$ ). Since the first term of the LHS is negative, the LHS is less than the RHS for all  $\mu$  and  $\lambda$ . Q.E.D.

One implication of Proposition 2 is that there can be too little divestment when  $h > \delta$ . When  $h > \delta$  the social optimum is  $n_c = N$  (see Section 3.4), and so we want all socially responsible investors to divest. (Even if they do, according to (4.8) we have

$n_c = \pi N - \frac{\pi\delta(1-\pi)}{\gamma\sigma^2} < \pi N < N$ .) Yet, if  $\lambda_n h < (\frac{3}{4})\delta$ ,  $\mu = 0$  is the only equilibrium: there is no divestment at all.

Charness and Rabin (2002) have conducted experiments that suggest that a median  $\lambda$  of less than 0.25 is plausible. The next proposition shows that as long as a majority of investors have  $\lambda \leq 1/4$ , and  $h < 2\delta$ , the unique equilibrium is  $\mu = 0$ .

**Proposition 3:**

Assume that  $\lambda_n h < \delta$ . Suppose that a majority of investors have  $\lambda \leq 1/4$ , that is,  $\sum_{i \in \{\lambda_i \leq \frac{1}{4}\}} \pi_i \geq \frac{1}{2}$ .

Suppose also that  $h < 2\delta$ . Then the unique equilibrium is  $\mu = 0$ .

Proof:

It follows from the assumptions of the proposition that the second term of the LHS of (4.18) is less than the RHS if  $\mu \leq 1/2$ . Since the first term of the LHS is negative, it follows that (4.18) cannot hold if  $\mu \leq 1/2$ : the only possible equilibrium if  $\mu \leq \frac{1}{2}$  is  $\mu = 0$  (and  $\mu = 0$  is an equilibrium by Proposition 2). What about an equilibrium with  $\mu > \frac{1}{2}$ ? Since the majority of investors have  $\lambda \leq 1/4$ , (4.18) must then hold for some investor with  $\lambda \leq 1/4$ . But that is impossible since  $\lambda \leq 1/4 \Rightarrow 2\lambda h < \delta$ , which in turn implies that the LHS of (4.18) is negative. Q.E.D.

Propositions 2 and 3 are consistent with Bill Gates’s view that “Divestment, to date, probably has reduced about zero tonnes of emissions.”<sup>22</sup>

It is worth noting that Propositions 2 and 3 generalize to the case where some investors divest for moral reasons. For the sake of brevity, we focus on the extension of Proposition 3 and show that only moral investors will divest. In the following it is supposed that the distribution of  $\lambda$  among the consequentialists is the same as in the overall population.

<sup>22</sup> <https://www.ft.com/content/21009e1c-d8c9-11e9-8f9b-77216ebe1f17>.

**Proposition 4:** Assume that  $\lambda_n h < \delta$ . Suppose that a fraction  $\bar{\mu}$  of investors divest for moral reasons. Assume that the investors who are consequentialists and who have  $\lambda \leq 1/4$  are in a

majority, that is,  $(1 - \bar{\mu}) \sum_{i \in \{\lambda_i \leq \frac{1}{4}\}} \pi_i \geq \frac{1}{2}$ . Suppose also that  $h < 2\delta$ . Then the unique equilibrium is

$$\mu = \bar{\mu}.$$

Proof:

By the argument in the first part of the proof of Proposition 3, the only equilibrium with  $\mu \leq 1/2$  is  $\mu = \bar{\mu}$  (and this is an equilibrium). But if  $\mu > 1/2$ , (4.18) must hold for some investor with  $\lambda \leq 1/4$ . By the argument in the second part of the proof of Proposition 3, this is impossible. Q.E.D.

Propositions 2-4 show that under plausible assumptions no divestment (by consequentialists) will occur in equilibrium. However, we do not want to suggest that this is always the case. The next proposition provides sufficient conditions for divestment to occur among consequentialists.

**Proposition 5:**

- (1) Suppose that  $\delta \simeq N\gamma\sigma^2$ ,  $\pi_n > 1/2$ ,  $\lambda_n h < \delta$ ,  $\lambda_n h \simeq \delta$ . Then there exists an equilibrium with  $\mu > 0$  (as well as an equilibrium with  $\mu = 0$ ).
- (2) Suppose that  $\lambda_n > 0$  and  $h$  is sufficiently large. Then every equilibrium satisfies  $\mu > 0$ .

Proof:

(1) The second term of the LHS of (4.18) is easily seen to be strictly greater than the RHS when  $\mu = \pi_n$ ,  $\lambda = \lambda_n$  given that  $\pi_n > 1/2$  and  $\lambda_n h \simeq \delta$ . Since the first term of the LHS is approximately zero, the LHS is strictly greater than the RHS at  $\mu = \pi_n$ ,  $\lambda = \lambda_n$ . However, the LHS is strictly less than the RHS at  $\mu = 0$ ,  $\lambda = \lambda_n$ . By continuity there must exist  $0 \leq \mu^* \leq \pi_n$  such that the LHS equals the RHS at  $\mu = \mu^*$ ,  $\lambda = \lambda_n$ . Since the LHS is increasing in  $\lambda$ , no-one with  $\lambda < \lambda_n$  wants to divest while investors with  $\lambda = \lambda_n$  are indifferent. Thus  $\mu^*$  is an equilibrium.

(2) Choose  $h$  large enough so that the LHS of (4.18) strictly exceeds the RHS at  $\mu = 0$ ,  $\lambda = \lambda_n$ . Then  $\mu = 0$  cannot be a divestment equilibrium. But we know from Proposition 1 that a divestment equilibrium exists. Hence it must satisfy  $\mu > 0$ . Q.E.D.

Note that although Proposition 5 provides a sufficient condition for divestment to occur, it does not follow that divestment is socially desirable. The assumptions of Proposition 5(1) can hold if  $\lambda_n=1$  and  $h$  less than but close to  $\delta$ . In other words, Proposition 5(1) shows that there can be too much divestment.

#### 4.2 Consumer Boycotts

In this section we ignore the possibility of divesting and focus on a different form of exit: consumer boycotts. We will show that under plausible assumptions a consequentialist will choose not to boycott.

For boycotts to be possible, we need to assume that consumers know the technology behind the good they buy: they can tell whether the good is produced by a clean firm or a dirty firm. We suppose that boycotting decisions are common knowledge and that consumers can commit to them (but see Section 6). As in previous sections we suppose that a boycott is not anticipated at date 0 when firms are set up, but only becomes a factor at date 1. Thus,  $N$  is predetermined at date 1 and is given by (3.11).

Consider the replica economy where there are  $r$  consumers. As above, we start by assuming that a fraction  $\theta$  of consumers will boycott the dirty product and then derive the equilibrium value of  $\theta$ . Boycotters buy only clean items at a price  $p_c$ . The other consumers are either indifferent about what they buy (if  $p_c = p_d$ ) or buy only dirty items (if  $p_d < p_c$ ). We will see that  $p_d < p_c$ . Thus, a fraction  $\theta$  of the demand will be for clean products and a fraction  $(1 - \theta)$  for dirty products.

Consider an equilibrium with  $n_c$  clean firms and  $n_d$  dirty firms, where  $n_d = N - n_c$ . The equilibrium in the output market requires that

$$(4.19) \quad \theta \left( \frac{\rho - p_c}{\tau} \right) = n_c, \quad (1 - \theta) \left( \frac{\rho - p_d}{\tau} \right) = n_d,$$

where  $p_c$  and  $p_d$  are the prices of clean and dirty goods, respectively.

Solving these equations yields,

$$(4.20) \quad p_c = \frac{\theta \rho - \tau n_c}{\theta},$$

$$p_d = \frac{(1 - \theta) \rho - \tau n_d}{1 - \theta}.$$

In an interior equilibrium the expected date 1 profit of clean and dirty firms must be the same, since otherwise the lower profit firms would have a lower market value (since investors must be induced to hold the shares), and a dirty firm would have the incentive to become clean or vice versa. Hence,

$$(4.21) \quad \Pi_c = p_c - \delta = \Pi_d = p_d.$$

Substituting the value of  $p_c$  and  $p$  we have

$$(4.22) \quad \frac{\theta\rho - \tau n_c}{\theta} - \delta = \frac{(1-\theta)\rho - \tau n_d}{1-\theta}$$

and using  $n_d = N - n_c$  we can rewrite this as

$$(4.23) \quad n_c = \theta N - \frac{\delta\theta(1-\theta)}{\tau}$$

$$n_d = (1 - \theta)N + \frac{\delta\theta(1-\theta)}{\tau}.$$

Note that the first equation in (4.23) is very similar to (4.8). The impact of boycotting is similar to the impact of divesting. Boycotting will be effective when either the mass of boycotters is close to 1 or the cost of the clean technology is small relative to slope of the demand curve. As for divestors, boycotters impact the equilibrium level of clean firms less than proportionally.

If  $N < \frac{\delta}{\tau}$ , we have a corner solution: the number of clean firms  $n_c = 0$  in a neighborhood of  $\theta = 0$  and, for low  $\theta$ , the marginal impact of  $\theta$  on  $n_c$  is zero. Note that this will be the case when the slope of the demand curve is low. Under these conditions it is an equilibrium for no consumer to divest: starting at  $\theta = 0$ , nonboycotting consumers will absorb any goods boycotters shun with minimal price impact and as a result no firms will become clean.

For small  $\theta$ , we have an interior solution with a positive number of clean firms ( $n_c > 0$ ) if and only if  $N > \frac{\delta}{\tau}$ . From now on, we assume

$$(4.24) \quad N > \frac{\delta}{\tau}.$$



Suppose that one of the consumers stops boycotting. When she was boycotting dirty products, she was maximizing her utility  $\rho q - \frac{1}{2}\tau q^2 - p_c q$ , yielding  $q = \frac{\rho - p_c}{\tau}$ . This purchase generates a utility of  $(\rho - p_c)\frac{\rho - p_c}{\tau} - \frac{1}{2}\tau\left(\frac{\rho - p_c}{\tau}\right)^2 = \frac{1}{2\tau}(\rho - p_c)^2$ . When she stops boycotting she maximizes  $\rho q - \frac{1}{2}\tau q^2 - p_d q$  and so her utility becomes  $\frac{1}{2\tau}(\rho - p_d)^2$ . Thus, the change is

$$(4.25) \quad \frac{1}{2\tau}\left[(\rho - p_d)^2 - (\rho - p_c)^2\right] = \frac{1}{2\tau}\left[(2\rho - p_d - p_c)(p_c - p_d)\right].$$

At the same time, the consumer bears a cost of not boycotting due to her internalizing a fraction of social welfare. As in the divestment case the effect of her stopping her boycott on other consumers' and investors' utility is zero by the envelope theorem. But there is a negative effect on the environment equal to  $h\frac{\partial n_c}{\partial \theta}$ , which will have weight  $\lambda$  in her utility function. Thus, a boycott is sustainable if and only if

$$(4.26) \quad \frac{1}{2\tau}(2\rho - p_d - p_c)(p_c - p_d) \leq \lambda h \frac{\partial n_c}{\partial \theta}.$$

We can rewrite this as

$$(4.27) \quad \frac{1}{2\tau}\left(2\rho - \left[\rho - \frac{\tau n_c}{\theta} + \rho - \frac{\tau n_d}{1-\theta}\right]\right)\left(\frac{\tau n_d}{1-\theta} - \frac{\tau n_c}{\theta}\right) \leq \lambda h \left[N - \frac{\delta}{\tau}(1 - 2\theta)\right],$$

where we use (4.20). After some manipulation and the use of (4.23), this can be simplified to

$$(4.28) \quad (\tau N - \delta)(\delta - \lambda h) \leq 2\theta\delta\left(\lambda h - \frac{\delta}{2}\right) - \frac{\delta^2}{2}.$$

The following definition and propositions parallel the material in the divestment section, and we state them without discussion or proof.

**Definition 6.** A boycott equilibrium for the limit economy ( $r = \infty$ ) is a  $0 \leq \theta^* \leq 1$ , where  $\theta^*$  represents the fraction of consumers who boycott, such that one of the following holds:

- (1)  $\theta^* = 0$ , and the LHS of (4.28) is less than or equal to the RHS at  $\theta = 0$ ,  $\lambda = \lambda_n$ .
- (2)  $\theta^* = 1$ , and the LHS of (4.28) is greater than or equal to the RHS at  $\theta = 1$ ,  $\lambda = \lambda_1$ .

(3)  $\theta^* = \sum_{j=i+1}^n \pi_j$  for some  $i = 1, \dots, n-1$ , and the LHS of (4.28) equals the RHS at  $\theta = \theta^*$  for some  $\lambda_i < \lambda^* < \lambda_{i+1}$ .

(4)  $\sum_{j=i+1}^n \pi_j < \theta^* < \sum_{j=i}^n \pi_j$  for some  $i=1, \dots, n$ , and LHS of (4.28) equals the RHS at  $\theta = \theta^*$ ,  $\lambda = \lambda_i$ .

**Proposition 7:** A boycott equilibrium exists.

**Proposition 8:**

(1) Suppose that  $\lambda_n h < \delta$ . Then  $\theta = 0$  is an equilibrium.

(2) Suppose that  $\lambda_n h < (\frac{3}{4})\delta$ . Then  $\theta = 0$  is the unique equilibrium.

**Proposition 9:**

Assume that  $\lambda_n h < \delta$ . Suppose that a majority of consumers have  $\lambda \leq 1/4$ , that is  $\sum_{i \in \{\lambda_i \leq \frac{1}{4}\}} \pi_i \geq \frac{1}{2}$ .

Suppose also that  $h < 2\delta$ . Then the unique equilibrium is  $\theta = 0$ .

In the next proposition, it is supposed that the distribution of  $\theta$  among the consequentialists is the same as in the overall population.

**Proposition 10:**

Assume that  $\lambda_n h < \delta$ . Suppose that a fraction  $\bar{\theta}$  of consumers boycott for moral reasons. Assume that the boycotters who are consequentialists and who have  $\lambda \leq 1/4$  are in a majority, that is,  $(1 - \bar{\theta}) \sum_{i \in \{\lambda_i \leq \frac{1}{4}\}} \pi_i \geq \frac{1}{2}$ .

Suppose also that  $h < 2\delta$ . Then the unique equilibrium is  $\theta = \bar{\theta}$ .

**Proposition 11:**

(1) Suppose that  $\delta \simeq \tau N$ ,  $\pi_n > 1/2$ ,  $\lambda_n h < \delta$ ,  $\lambda_n h \simeq \delta$ . Then there exists an equilibrium with  $\theta > 0$  (as well as an equilibrium with  $\theta = 0$ ).

(2) Suppose that  $\lambda_n > 0$  and  $h$  is sufficiently large. Then every equilibrium satisfies  $\theta > 0$ .

The most important takeaway from this section is that what we said about divestment is true of boycotts as well. Under plausible assumptions consequentialists will not boycott, but sometimes very altruistic consumers may boycott even when this is socially inefficient ( $h < \delta$ ).

### *4.3 Labor Boycott*

Our simple model does not have any labor costs, let alone the possibility of workers boycotting a firm. Yet, in a competitive labor market the effect of a labor boycott is very similar to that of the consumer boycott we analyzed in Section 4.2. Purely selfish workers work for any firm, while socially responsible workers boycott dirty firms. The resulting equilibrium would be similar to that in Section 4.2, with workers in dirty firms being paid more than workers in clean firms and the equilibrium level of clean firms depending on the slope of the labor demand curve. Indeed, Nyborg and Zhang (2013) provide evidence that workers in socially responsible firms are paid less.

The situation is different if a firm has some market power. Consider, for instance, a case where there is one monopsonist and many workers. The monopsonist has the choice to stay dirty and be able to hire only from a smaller pool of workers or pay the cost  $\delta$  and be able to hire all workers. As we discuss in the working paper version, when the market is not competitive, if the pool of boycotters is large enough, not only will boycotters be able to turn the firm clean, but they will be able to do so without bearing any cost.

## 5. Voice

Socially responsible agents can also engage with management. Voice can be exercised by several stakeholder groups but we focus on the unique ability shareholders have to exercise voice given that they possess control rights. We start by assuming that environmental proposals are exogenously put up for a proxy vote. Later we will consider the role of mutual funds in facilitating engagement. We assume that all investors continue to buy dirty as well as clean firms, i.e., they do not discriminate. To put it another way they do not both engage and divest. In this section, we ignore consumer boycotts.

### *5.1 Shareholders' Voting*

Suppose there are  $\nu N$  clean firms and  $(1-\nu)N$  dirty firms at date 1. Assume that a vote takes place on whether one of the dirty firms should become clean. As mentioned in the Introduction, we follow Hart and Zingales (2017) in assuming that each shareholder votes as if she were pivotal, since this is the only time her vote matters.

Given  $\nu N$  clean firms, the date 1 stock market equilibrium will be as follows. The gross return of a clean firm is  $\delta$  less than that of a dirty firm. Thus, in order to ensure that investors stay invested in both kinds of firms, and have the same overall demand for shares as before, we must have  $V_c = F - \delta$  and  $V_d = F$ . Applying (3.12) and (3.10), we see that the certainty equivalent of each fully diversified investor is

$$(5.1) \quad CE = \frac{(\Pi - F)^2}{2\gamma\sigma^2} - \nu\delta \frac{(\Pi - F)}{\gamma\sigma^2},$$

where the second term reflects the capital loss caused by a fraction  $\nu$  of the firms the investor owns becoming clean.

Suppose that one additional dirty firm becomes clean as a result of the vote. This will cause CE to change by  $\frac{\partial CE}{\partial \nu} \Delta \nu$ . But in the replica economy  $\Delta \nu = \frac{1}{Nr}$ , and so the change in CE converges to zero as  $r \rightarrow \infty$ .

It is important to emphasize the importance of our assumption that investors are well-diversified. As a result the capital loss incurred by any one firm becoming clean will be shared equally by  $r$  investors, and each individual's capital loss will converge to zero as  $r \rightarrow \infty$ . We discuss the case of concentrated ownership in Section 5.3.

We see that the impact of the vote on each investor's personal utility is negligible. The remaining effect of bringing about an extra clean firm consists of two elements: the impact on the environment and the impact on the wealth of other investors (the effect on consumers is zero, since the supply of output remains at  $N$ ). The impact on the environment is

$$(5.2) \quad h \frac{\partial(rn_c)}{\partial \nu} \Delta \nu = hrN\Delta \nu = h$$

because  $n_c = \nu N$  and  $\Delta \nu = \frac{1}{Nr}$ .

The impact on other investors' wealth is no longer zero (the envelope theorem does not apply since firms that are pressed to choose clean by engaged shareholders are not maximizing value), but is now given by

$$(5.3) \quad \frac{\partial[(r-1)CE]}{\partial v} \Delta v = \frac{(r-1)}{Nr} \frac{\partial CE}{\partial v} = -\frac{(r-1)}{Nr} \delta \frac{(\Pi-F)}{\gamma\sigma^2},$$

which converges to

$$(5.4) \quad -\delta \frac{(\Pi-F)}{N\gamma\sigma^2}$$

as  $r \rightarrow \infty$ .

Summing the two effects, weighting them by  $\lambda$ , and using the fact that by (3.10)  $N = \frac{\Pi-F}{\gamma\sigma^2}$ , yields the conclusion that bringing about an extra clean firm is desirable for the investor if

$$(5.5) \quad \lambda(h - \delta) > 0.$$

We may conclude that an investor with  $\lambda > 0$  will vote "clean" (that is, to make the dirty firm clean) if and only if  $h > \delta$ , which is the same criterion used by the planner. Hence, as long as the majority of investors are a little socially responsible, voting will deliver the social optimum. If the majority of investors are purely selfish, however, and a vote requires majority support to be effective, engagement by socially responsible shareholders will have no impact.

**Proposition 12:**

Suppose that a majority of investors have  $\lambda > 0$ . Then majority rule will deliver a socially efficient outcome.

*5.2 Engaging through an intermediary*

Putting proposals up for a proxy vote is expensive. It will not be in the interest of atomistic investors to incur the cost of doing so, and management is unlikely to take the lead. By contrast, mutual funds can use engagement as a marketing strategy (O'Leary and Valdmanis (2020)). A Green Fund can sell its ability to put socially responsible proposals on the ballot as a feature of their fund and pay the cost out of the management fee. The question then is whether atomistic socially responsible investors are willing to put their money in the Green Fund.

As a benchmark we consider the case where the impact of engagement by a Green Fund is linear, that is, if a fraction  $\nu$  of investors' wealth is placed with the fund, a fraction  $\nu$  of firms will adopt the clean technology.

Suppose that the Green Fund controls a fraction  $\nu$  of wealth at date 1. That leads to  $\nu N$  clean firms and  $(1-\nu)N$  dirty firms. Consider an investor's decision to divest from the Green Fund. Her current certainty equivalent is given by (5.1). If she divests, her certainty equivalent changes by  $\frac{\partial CE}{\partial \nu} \Delta \nu$ , where  $\Delta \nu = -\frac{1}{r}$ . So, as in the case of voting, the change in CE converges to zero as  $r \rightarrow \infty$ .

The remaining effects of divesting from the Green Fund are also exactly the same as with voting. The impact on the environment is given by (5.2) and the impact on other investors' wealth by (5.3). Thus, a socially responsible investor will choose to divest from the Green Fund if and only if  $h < \delta$ .

Our conclusion is that all the socially responsible investors will invest in the Green Fund if  $h > \delta$  and none will invest if  $h < \delta$ . Hence, socially responsible mutual funds can help to achieve the social optimum.

The presence of intermediaries greatly relaxes the informational burden that the voice option imposes on investors. While it is hard to imagine that investors are sufficiently informed to vote on all ballot propositions, institutional investors have a fiduciary duty to exercise their voting rights. Thus, socially responsible investors can express their choices by picking the intermediary with the right "ideology", in the language of Bolton et al. (2020).

Note that a socially responsible fund can engage in ways other than just through voting. For example, Dimson et al. (2015), Barko et al. (2018), and Naaraayanan et al. (2020) show the effectiveness of behind-the-scenes engagement by socially responsible funds. One of the strategies that intermediaries use behind the scenes is to threaten to withhold their votes in board elections if management does not accede to their wishes. If credible, this strategy can induce management to compromise because they fear for their jobs. Such a strategy, however, is difficult for individual investors to monitor.

### 5.3 Nonatomistic Investors

The engagement result derived in Section 5.1 depends heavily on investors being well diversified, so that the impact of their decisions on the value of their own portfolio is infinitesimal. This is not true anymore if an investor is a significant stakeholder. In that case, a socially responsible investor will weigh the net social benefit of a vote for clean (multiplied by  $\lambda$ ) against the reduced return on her portfolio due to the drop in profitability of the clean firm (multiplied by the investor's stake in that firm). Thus, significant investors (even if socially responsible) may vote dirty when this is socially inefficient.

In the value maximizing approach to corporate governance, large shareholders are often thought to be beneficial because they reduce the agency costs produced by the separation of ownership and control (Shleifer and Vishny (1997)). In contrast, in the socially responsible approach to corporate governance, large shareholders reduce the level of provision of public goods by firms and to that extent are undesirable.

Note that one downside of investors being well-diversified is that they have an incentive to persuade firms to take advantage of their monopoly power in the product market and raise consumer prices. This idea has received a great deal of attention recently (see Azar et al. (2018)). In order to guard against this we would advocate that only ESG proposals should be voted on, not those concerning prices and quantities. More important, we do not fear open votes to collude because these will attract the attention of the antitrust authorities.

#### 5.4 Takeovers

A natural question to ask is whether takeovers affect engagement. As Hart and Zingales (2017) show (see also Elhauge (2005)), takeovers can undermine social action to turn companies clean, creating an “amoral drift.” Here we briefly sketch the argument.

Suppose that engagement leads a company to choose clean (provisionally). This means that its market value will be  $V_c = F - \delta$ . A (purely selfish) bidder could make an unconditional tender offer for the company at a price  $V_c < p < V_d = F$ , at the same time announcing that, if more than 50% of the shares are tendered, he plans to freeze-out nontendering shareholders at a price  $V_c < p' < p$ . Even a socially responsible investor will tender. The reason is that given that she has a very small shareholding the chance that her tender decision will be pivotal is negligible. Furthermore, by not tendering she receives  $p'$  if the bid succeeds as opposed to  $p$ ; while if the bid

fails she owns shares worth  $V_c$  rather than receiving  $p$  (she could always buy back her shares). Thus tendering is a dominant strategy. Since everyone tenders the bid succeeds and the bidder makes a profit of  $V_d - p$ . This is true even if a majority of the investors are socially responsible and would have voted against the bid if given the chance. For further details, see Hart and Zingales (2017).

There is an asymmetry here. It is unlikely that a socially responsible bidder will buy a dirty company and turn it clean. The reason is that the bidder will have to pay at least  $V_d$  to persuade shareholders to tender (at a lower price it would be profitable for someone, e.g., management to make a counteroffer), which means he loses  $V_d - V_c$  on the transaction. There is an environmental gain of  $h$ , but this is weighted by  $\lambda$ . Thus only if  $V_d - V_c = \delta < \lambda h$  will he proceed. (In Propositions 2-4 and 8-10, we have assumed that  $\lambda_n h < \delta$ .) In contrast, dispersed shareholders will vote for the company to become clean if  $h > \delta$ .

One important qualification to the above is that, as a result of a number of legal decisions in recent years and the existence of poison pills, it has become hard to take over a U.S. company if the majority is against the bid. These developments serve to mitigate the amoral drift, and make it less likely that takeovers will interfere with socially responsible engagement.

### *5.5 Voice by Other Stakeholders*

So far, we have considered engagement only by shareholders. But other groups can engage with and influence management even though they do not have votes. For example, workers (or consumers) might form a coalition and bargain with the firm over environmental strategy (this would obviously be easier in the presence of a union). If workers are shielded from the costs of reducing pollution, it is sufficient that they have a positive  $\lambda$  for them to vote as a benevolent planner would. However, if they have to absorb part or all of the cost of abating pollution, workers will not vote in a socially desirable way. The reason is that workers are not well diversified and hence a significant part of the cost of turning a firm clean falls on them. Hence, they will vote clean only if their degree of social responsibility  $\lambda$  is high. In other words, they will behave as large shareholders.



## *5.6 Conclusions on Voice*

While there is no reason that the incentives to boycott or divest should be aligned with a benevolent planner's choices, shareholders' incentives to approve a proposal are aligned with what a benevolent planner would do. Yet, this is true only of well-diversified shareholders. Shareholders who are not well diversified or other stakeholders will be more reluctant to approve a clean technology given that they bear a disproportionate fraction of the cost.

## 6. Robustness and Extensions

### *6.1 Visibility and Commitment*

So far we have assumed that individuals can commit to their strategy (be it divestment, or boycotting) and that this strategy is common knowledge. In practice, it is difficult for individuals to communicate and commit to their strategy. Here technology and institutions might make a difference.

In our model firms are assumed to be aware that some investors (consumers) plan to divest (boycott). But how do they know this? One way is for investors or consumers to make announcements. In a pre-internet world, the authors of this paper could have announced that they would divest, but it would have been hard for anyone to know about it. In contrast, large institutions and companies could easily publicize their divestment and boycott decisions. Today, thanks to social media, this difference has become smaller, facilitating the announcement of divestments and boycotts.

Even today it is difficult to verify whether someone has carried through their announced strategy, given the variability of demand (see Ashenfelter et al. (2007)). Verification is important because there is a commitment issue. At date 1, some investors could announce that they will divest. This announcement might, if believed, be sufficient to push some companies to switch to clean. But, after having achieved their goal, the divestors will be tempted to sell the clean companies and buy the more profitable dirty ones, which trade at the same price. If this behavior is anticipated, divestment will become ineffective.

The same problem arises in the case of boycotts. Some consumers may announce that they will buy only clean products, causing some companies to install a clean technology. But once this

is done what ensures that consumers do not renege on their promise and buy cheaper dirty products?

These commitment problems can exist even in the presence of intermediaries. Suppose that investors invest through a mutual fund, e.g., Fidelity. Fidelity might have a fund that plans to invest only in clean companies and another fund that plans to invest only in dirty companies. A socially responsible investor might put all her money in the Fidelity Green Fund. Seeing how much wealth has been invested in the Green Fund, some companies may elect to become clean. But once companies have made this decision what is to stop investors from switching their money from the Green Fund to the Dirty Fund?<sup>23</sup>

The commitment problem is stark in our setting because we study a one-shot game: firms make their production decisions at date 1, then investors and consumers make their investment and consumption decisions, then the world ends. Reality is more complex and commitment may be easier to establish in a repeated setting.

Visibility can also help with commitment. Even today, if the authors of this paper announce that they will divest from oil companies, it would be hard for anyone to check.<sup>24</sup> In contrast, the decision by the Harvard University endowment to achieve greenhouse gas neutrality by 2050 can easily be verified.<sup>25</sup> In a similar fashion, on June 26, 2020, Unilever announced that it would not advertise on Facebook or Twitter for the rest of the year, citing hate speech and divisive content on the platforms.<sup>26</sup> Unilever's action can easily be verified, and so Unilever is likely to stick to this commitment.

## 6.2 Coasian Solution

A natural question to ask is whether Coasian bargaining could achieve a socially efficient outcome without the need for divestment or voice. To understand how this might work, suppose  $h > \delta$  and consider a situation where there are some agents with

---

<sup>23</sup> One way in which a mutual fund can help increase the level of commitment is by offering only “clean” products, increasing the cost for investors to switch.

<sup>24</sup> However, someone who makes a personal decision to divest or boycott may incur a personal cost if they deviate from this decision, which can help to sustain commitment. See, e.g., Ederer and Stremitzler (2017). Note that commitment is not an issue in the literature that assumes that people divest or boycott for moral reasons.

<sup>25</sup> See <https://www.thecrimson.com/article/2020/4/22/endowment-ghg-neutral-2050/>

<sup>26</sup> See <https://www.wsj.com/articles/unilever-to-halt-u-s-ads-on-facebook-and-twitter-for-rest-of-2020-11593187230>.

$\lambda$  slightly above  $\frac{1}{4}$  (in a large economy, there will be many of them). Then a coalition of four of them could get together and approach a dirty firm with the following offer: We will pay you  $\delta$ ; in return you agree to become clean. The cost  $\delta$  is split equally among the four. Each agent should be prepared to do this since  $\lambda h - \delta/4 > 0$  and the firm should agree since it is no worse off.

The main difficulty with this solution is that it is not clear who should be in the coalition and who should not. That is, each agent would like other agents with  $\lambda$  above  $1/4$  to form the coalition and pay the  $\delta/4$ : there is a classic free-rider problem. Thus the coalition may not form.

Somewhat related free-rider problems arise in the case of divestment. We have endogenized these and have found that there is often too little divestment. Note that voice operates very differently. If a majority of shareholders vote to make a company clean then the decision is binding on the minority: they have to bear their proportionate share of the cost. A shareholder vote is thus very similar to a vote by citizens to levy a tax to pay for a public good. In both cases the free-rider problem is avoided.

### *6.3 Rational Expectation Equilibrium*

We have assumed that the harm at date 1 is unanticipated at date 0. Relaxing this assumption does not change the analysis very much. If all investors and consumers are purely selfish, it does not change it at all since the date 1 market value of firms will be independent of  $h$ , and so the incentives of entrepreneurs to set up firms at date 0 will be unaffected. On the other hand, if divestment or boycotting by socially responsible investors or consumers is anticipated to occur at date 1, then this reduces the date 1 and date 0 market value of firms, and so the equilibrium number of firms will be lower. The same is true if it is anticipated that successful engagement will cause firms to choose the clean technology since this reduces future profitability.<sup>27</sup>

### *6.4 Social Entrepreneurs*

---

<sup>27</sup> Given this, founders of firms at date 0 may try to make engagement more difficult at date 1 through provisions in the corporate charter, for example, by putting in super-majority provisions or a dual-class voting structure.

We have assumed that the entrepreneurs at date 0 who set up companies are interested only in money. But suppose that some of them are socially responsible. If they anticipate the harm at date 1 might they try to influence their firms to act in a socially responsible way? In our setting this is difficult to achieve. The reason is that we have assumed free entry, that is an infinite supply of entrepreneurs. Even if some are socially responsible many are likely not to be. As a result, in the free entry equilibrium the market value of a firm that does not encourage social responsibility will be  $F$  and the market value of a firm that does encourage social responsibility will be below  $F$ . Since we have assumed that entrepreneurs have zero wealth they will not be able to finance the latter. In effect competition drives out good behavior (on this, see Aghion et al. (2020), Dewatripont and Tirole (2020) and Shleifer (2004)).

Note that the situation is different at date 1. At this point the entry cost  $F$  is sunk and so firms earn rents. Therefore, firms have the ability to choose clean without being driven out of the market.

## 7. Discussion

### *7.1 Awareness and Social Pressure Campaigns*

The biggest limitation of our analysis is that we take social preferences as given. As a result, we miss an important benefit of exit campaigns: their effect on social preferences. The fur-free campaign launched by the Humane Society is not just the announcement of a boycott, but an attempt to shame fur users, i.e., to lead purely selfish individuals to behave as if they were socially responsible consumers.<sup>28</sup>

When it comes to informing and changing people's preferences the exit strategy is superior to the voice one. A successful information campaign can keep the relevant piece of news in the media for an extended period of time. A corporate vote is not so newsworthy to begin with. The media feel compelled to cover it at most twice, when the vote is announced and when the votes are counted. By contrast, an exit campaign is newsworthy every time a famous person/institution joins the exiters. Thus, exit is more effective at communicating the news.

Exit is also more effective at pressuring people into behaving socially, even if their  $\lambda$  is equal to zero. It is not only peer pressure that operates, but also the pressure to join a growing and

---

<sup>28</sup> <https://www.humanesociety.org/all-our-fights/going-fur-free>.

potentially successful movement (Thaler and Sunstein (2008)). Both these forces help a highly motivated minority to achieve successes it would never be able to achieve through a voice strategy. Consistent with this idea, corporate boycotts succeed mostly by affecting a target's reputation in the media, not the demand for their product (King (2011)).

For these reasons, a highly-motivated minority might find exit a more successful strategy than voice. Yet, there is no guarantee that the ability of an exit strategy to succeed is linked to the social desirability of its goal. Thus, extending the model to incorporate information and social pressure is unlikely to change the fundamental result that voice is more aligned to social incentives than exit.

### *7.2 Best Strategy Contingent on Size*

If a majority of investors are socially responsible, voice will deliver the socially desirable outcome. Even when socially responsible investors are not a majority, as long as they can concentrate their holdings in a subset of firms in which they represent a majority (without affecting much the diversification of their portfolio), they can have an impact that is proportional to their size. Thus, they can do better than an exit strategy, which will have an impact less than proportional to the mass of socially responsible consumers/investors (see (4.8)).

Problems start to arise when socially responsible consumers/investors represent such a small minority that they cannot reach a majority of the votes in any single firm without significantly reducing the diversification of their portfolio. In such situations, the exit strategy may be more effective.

### *7.3 Community of Reference*

We have assumed that the group toward which people feel responsible is the entire world. In practice, people are more likely to internalize the impact they have on their community than on the world at large. This local bias in social responsibility might explain some of the observed trends in corporate governance. Until the 1970s companies were owned very locally. Even during the 1990s, Huberman (2001) documents a bias in favor of owning local companies. A locally concentrated ownership favors an internalization of the externalities produced by firms, especially if production and distribution are also locally concentrated. From the 1980s we have witnessed two important trends: the globalization of firms and the indexation of individual portfolios. The

combination of these two trends have led firms to become more asocial, i.e. to ignore most of the externalities they produce. We can interpret the rise of the ESG movement as a reaction to this increasing asociality of firms.

## 8. Conclusions

This paper is an attempt to analyze the welfare implications of two traditional strategies aimed at shaping corporate outcomes: exit and voice. To make the problem tractable we have made a number of simplifying assumptions: identical firms with zero marginal cost up to a capacity constraint, a linear demand curve, constant absolute risk aversion, normal distribution, etc. We have also studied the three principal socially responsible strategies, divestment, boycotting and engagement, separately, without considering how they might interact with each other. Subject to these limitations, we find that in a competitive world exit is less effective than voice in pushing firms to act in a socially responsible manner. Our conclusion is consistent with Kruger et al.'s (2020) survey of institutional investors, which finds that such investors consider engagement, rather than divestment, to be the better approach for addressing an externality such as climate risk. Furthermore, we show that individual incentives to join an exit strategy are not necessarily aligned with the social incentive, while they are when investors are allowed to express their voice.

We have derived these results under the best possible scenario for the exit strategies: investors and consumers who can announce their strategies to the world and commit to them. If we relax these assumptions exit becomes even less effective.

One question raised by our paper is why social engagement is relatively rare in spite of all its desirable properties. In some cases, engagement is infeasible because somebody owns a majority of the votes, such as Mark Zuckerberg with Facebook, or the company is privately held, such as Koch Industries. We think that an important additional factor resides in the current U.S. proxy system, which tends to limit shareholders' ability to influence corporate policy. The restrictions reflect a fear that individual shareholders are activists in the sense that they put a lot of weight on a single issue (e.g., their utility is  $-Nh$ ). If instead individuals are socially responsible (in the way we define), this fear is unfounded. Individual shareholders have the incentive to vote on issues in a socially optimal way and their engagement can lead to more efficient outcomes.

Another question is what comparative advantage firms have vis-à-vis the government in addressing externalities (e.g., Egorov and Harstad (2017) and Besley and Persson (2020)). After

all, our voice option is not very different from single-issue referenda, common in Switzerland and California (see Matsusaka (2020)). The corporate solution has three advantages. First, a referendum-imposed regulation has – by necessity– to be general, creating potentially large deadweight costs. A firm-by firm solution is more flexible and cost effective. Second, in the investment world there are monetary incentives for mutual funds to cater to investors’ preferences, which are not present in the political world. Mutual funds can pay for the cost of setting up a proxy, in a way that political parties cannot. Last but not least, in the United States companies can spend massively to influence the outcomes of referenda (as Uber and Lyft did recently in California) and their spending is constitutionally protected. By contrast, shareholders can choose to limit such spending. Thus, shareholder voice has the chance of being less prone to capture than political voice.

## References

- Abito, Joe, Miguel, D. Besanko and D. Diermeier. 2019. *Corporate Reputation and Social Activism: Strategic Interaction, Firm Behavior, and Social Welfare*, Oxford University Press.
- Acquatella, Angie. 2020. "Optimal Reimbursement Contracts for Physician Administered Treatments." Harvard. Unpublished paper.
- Admati, Anat and Paul Pfleiderer. 2009. "The 'Wall Street Walk' and Shareholder Activism: Exit As a form of Voice." *Review of Financial Studies*, 22:7, 2645-2685.
- Aghion, Philippe, Roland Bénabou, Ralf Martin and Alexandra Roulet. 2020. "Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?" CEP Discussion Paper No 1684.
- Andersson, Mats, Patrick Bolton & Frédéric Samama (2016). "Hedging Climate Risk". *Financial Analysts Journal*, 72:3, 13-32,
- Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97 (6): 1447–1458.
- Andreoni, James, and John Miller. 2002. "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica*, 70 (2), 737–753.
- Ashenfelter Orley C., Ciccarella Stephen, and Shatz Howard J. 2007. "French Wine and the U.S. Boycott of 2003: Does Politics Really Affect Commerce?" NBER Working Paper # 13258.
- Azar, José, Martin C. Schmalz, and Isabel Tecu. 2018. "Anticompetitive Effects of Common Ownership." *The Journal of Finance* 73(4): 1513-1565.
- Bagnoli, Mark, and Watts Susan G. 2003. "Selling to Socially Responsible Consumers: Competition and The Private Provision of Public Goods." *Journal of Econ. & Management Strategy* 12 (3): 419–45.
- Barber, Brad, Adair Morse, and Ayako Yasuda. 2020. "Impact Investing." *Journal of Financial Economics*, forthcoming.
- Barko, Tamas, Cremers, K. J. Martijn, and Renneboog, Luc. 2018. "Shareholder Engagement on Environmental, Social, and Governance Performance". CentER Discussion Paper No. 2017-040.
- Baron, David P. 2001. "Private Politics, Corporate Social Responsibility, and Integrated Strategy." *Journal of Economics and Management Strategy*, 10(1): 7-45.
- Baron, David P. 2003. "Private Politics." *Journal of Economics & Management Strategy* 12 (1): 31–66.
- Baron, David P. 2007. "Corporate Social Responsibility and Social Entrepreneurship." *Journal of Economics & Management Strategy* 16 (3): 683–717.
- Barrios, John, Efraim Benmelech, Yael V. Hochberg, Paola Sapienza, and Luigi Zingales. 2020. "Civic Capital and Social Distancing during the Covid-19 Pandemic." *Journal of Public Economics*.



- Bauer, Rob, Tobias Ruof, and Paul Smeets. 2020. "Get real! individuals prefer more sustainable investments." Working Paper.
- Bebchuk, Lucian A. and Tallarita, Roberto. 2020. "The Illusory Promise of Stakeholder Governance." *Cornell Law Review*, December.
- Bénabou, Roland, and Jean Tirole. 2010. "Individual and Corporate Social Responsibility." *Economica* 77 (305): 1–19.
- Besley, Timothy, and Maitreesh Ghatak. 2007. "Retailing Public Goods: The Economics of Corporate Social Responsibility." *Journal of Public Economics* 91 (9): 1645–63.
- Besley, Timothy, and Maitreesh Ghatak. 2018. "Prosocial Motivation and Incentive." *Annual Review of Economics* 10:411–38.
- Besley, Timothy, and Torsten Persson. 2020. Escaping the Climate Trap? Values, Technologies, and Politics. Unpublished paper.
- Bolton, Patrick, Li, Tao, Ravina, Enrichetta, and Rosenthal, Howard. 2020. "Investor Ideology." *Journal of Financial Economics*, August: 320-352.
- Bonnefon, Jean-Francois, Landier, Augustin, Sastry, Parinitha and Thesmar, David. 2019. "Do Investors Care About Corporate Externalities? Experimental Evidence." HEC Research Paper FIN-2019-1350.
- Brodback, Daniel, Nadja Guenster, and David Mezger. 2019. "Altruism and Egoism in Investment Decisions." *Review of Financial Economics* 37 (1): 118–48.
- Brown, Ira V. 1988. "Pennsylvania's Antislavery Pioneers, 1688–1776." *Pennsylvania History: A Journal of Mid-Atlantic Studies* 55(2): 59-77.
- CBINSIGHTS. 2020. State of Fintech: Investment & Sector Trends to Watch, <https://www.cbinsights.com/research/report/fintech-trends-q4-2019/>
- Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *The Quarterly Journal of Economics* 117 (3): 817–69.
- Chowdhry, Bhagwan, Shaun William Davies, and Brian Waters. 2019. "Investing for Impact." *The Review of Financial Studies* 32 (3): 864–904.
- Christensen, Hans B., Luzi Hail, and Christian Leuz. 2019. "Adoption of CSR and Sustainability Reporting Standards: Economic Analysis and Review." NBER Working Paper # 26169.
- Davies, Shaun William, and Edward Dickersin Van Wesep. 2018. "The Unintended Consequences of Divestment." *Journal of Financial Economics* 128 (3): 558–75.
- Dewatripont, Mathias and Jean Tirole. 2020. The Morality of Markets and the Nature of Competition. Unpublished paper.

- Dimson, Elroy, Oğuzhan Karakaş, and Xi Li. 2015. "Active Ownership." *Review of Financial Studies* 28 (12): 3225–68.
- Ederer, Florian, and Alexander Stremitzler. 2017. "Promises and Expectations." *Games and Economic Behavior*, 106(C): 161-178.
- Edmans, Alex. 2020. *Grow the Pie: How Great Companies Deliver Both Purpose and Profit*. Cambridge University Press.
- Egorov, Georgy, Bård Harstad. 2017. "Private Politics and Public Regulation," *Review of Economic Studies* (2017) 84, 1652–1682
- Elhauge, Einer. 2005. "Sacrificing Corporate Profits in the Public Interest." *New York University Law Review*. 80(3): 733–869.
- Frydlinger, David, and Oliver D. Hart. 2019. "Overcoming Contractual Incompleteness: The Role of Guiding Principles." NBER Working Paper # 26245.
- Gans, Joshua S., Avi Goldfarb Mara Lederman. 2020. "Exit, Tweets, and Loyalty," *American Economic Journal: Microeconomics*, forthcoming.
- Gollier Christian, and Sébastien Pouget. 2014. "The "Washing Machine": Investment Strategies and Corporate Behavior with Socially Responsible Investors", TSE Working Paper, n.14-457, January 2014.
- Graff Zivin, Joshua, and Arthur Small. 2005. "A Modigliani-Miller Theory of Altruistic Corporate Social Responsibility." *Topics in Economic Analysis & Policy* 5 (1): 1-19.
- Green, Daniel, and Benjamin N. Roth. 2020. "The Allocation of Socially Responsible Capital." Unpublished Paper.
- Hart, Oliver, and Luigi Zingales. 2017. "Companies Should Maximize Shareholder Welfare Not Market Value." *Journal of Law, Finance, and Accounting* 2 (2): 247–75.
- Hartzmark, Samuel M., and Abigail B. Sussman. 2019. "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows." *The Journal of Finance* 74 (6): 2789–2837.
- Heinkel, Robert, Alan Kraus, and Josef Zechner. 2001. "The Effect of Green Investment on Corporate Behavior." *The Journal of Financial and Quantitative Analysis* 36 (4): 431-449.
- Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge: Harvard University Press.
- Hong, Harrison, and Marcin Kacperczyk. 2009. "The Price of Sin The Effects of Social Norms on Markets." *Journal of Financial Economics*, 93: 15-36.
- Huberman, G. (2001). Familiarity breeds investment, *Review of Financial Studies* 14(3): 659–680.

- King, Brayden G. 2011. The Tactical Disruptiveness of Social Movements Sources of Market and Mediated Disruption in Corporate Boycotts,” *Social Problems*, Vol. 58, No. 4, pp. 491-517
- Kitzmueller, Markus, and Jay Shimshack. 2012. “Economic Perspectives on Corporate Social Responsibility.” *Journal of Economic Literature* 50 (1): 51–84.
- Krueger, Philipp, Zacharias Sautner, and Laura T. Starks. 2020. “The Importance of Climate Risks for Institutional Investors.” *The Review of Financial Studies* 33(3): 1067-1111.
- Laidlaw, James. 1995. *Riches and Renunciation. Religion, Economy, and Society among the Jains*, Clarendon Press, Oxford.
- Landier Augustin and Stefano Lovo. 2020. “ESG Investing: How to Optimize Impact?” Unpublished paper.
- Magill, Michael, Martine Quinzii, and Jean-Charles. Rochet. 2015. “A Theory of the Stakeholder Corporation.” *Econometrica* 83(5): 1685–1725.
- Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*, Oxford University Press.
- Matususaka, John. 2020. *Let the People Rule: How Direct Democracy Can Meet the Populist Challenge*. Princeton University Press.
- Mayer, Colin. 2018. *Prosperity: Better Business Makes the Greater Good*. Oxford University Press.
- Morgan, John, and Justin Tumlinson. 2019. “Corporate Provision of Public Goods.” *Management Science* 65 (10): 4489–4504.
- Naaraayanan, S. Lakshmi, Kunal Sachdeva, Varun Sharmak. 2020. “The Real Effects of Environmental Activist Investing.” Unpublished paper.
- Nyborg, Karine, and Tao Zhang. 2013. “Is corporate social responsibility associated with lower wages?” *Environmental and Resource Economics* 55(1): 107-117
- Oehmke, Martin, and Marcus M. Opp. 2020. “A Theory of Socially Responsible Investment.” Unpublished paper.
- O’Leary and Valdmanis. 2020. *Accountable: The Rise of Citizen Capitalism*, Harper Business.
- Pack, Spencer J. and Robert W. Dimand, 1996, “Slavery, Adam Smith's Economic Vision And The Invisible Hand,” *History of Economic Ideas*, Vol. 4, No. 1/2 (1996), pp. 253-269
- Pastor, Lubos, Robert F Stambaugh, and Lucian A. Taylor. 2020. “Sustainable investing in equilibrium.” *Journal of Financial Economics*, forthcoming.
- Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski. 2019. “Responsible Investing: The ESG-Efficient Frontier.” *SSRN Electronic Journal*.

- Riedl, Arno, and Paul Smeets. 2017. "Why Do Investors Hold Socially Responsible Mutual Funds?" *The Journal of Finance* 72 (6): 2505–50.
- Schmidt, Klaus and Fabian Herweg. 2020. "Prices versus Quantities with Morally Concerned Consumers." Unpublished paper.
- Schoenmaker, Dirk and Willem Schramade. 2019. *Principles of Sustainable Finance*, Oxford Un. Press.
- Shleifer, Andrei. 2004. "Does Competition Destroy Ethical Behavior?" *American Economic Review*, 94 (2): 414-418.
- Shleifer, Andrei, and Robert W Vishny. 1997. "A Survey of Corporate Governance." *Journal of Finance* 52 (2): 737-783.
- Stout, Lynn. 2012. *The Shareholder Value Myth*. San Francisco: Berrett-Koehler Publishers, Inc.
- Sugden, Robert. (1982). "On the Economics of Philanthropy." *The Economic Journal* 92(366): 341-350.
- Sugden, Robert, 1984. "Reciprocity: The Supply of Public Goods through Voluntary Contributions." *Economic Journal*, 94(376): 772-787.
- Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*, Penguin Books

### Appendix: The Benevolent Planner Solution

Each firm produces one unit whether it is clean or dirty. As a result, the product market equilibrium and consumer surplus are independent of  $\phi$ . To derive investor utility we need to compute the investors' return at date 1 after the planner sets the proportion of clean firms at  $\phi$  and investors freely re-optimize their investment choices.

Let the equilibrium prices of the two types of firms be  $V_c$  and  $V_d$ . The gross return of a clean firm is  $\delta$  less than that of a dirty firm. Thus, in order to ensure that (purely selfish) investors stay invested in both kinds of firms at date 1 we must have

$$(A1) \quad V_c = V_d - \delta.$$

The return of an investor at date 2 is

$$(A2) \quad x(\Pi - V_d) + x_0 [\phi V_c + (1 - \phi)V_d] - x_0 F,$$

where  $x$  is her date 1 portfolio holding. The first term reflects the fact that the net return on her investment is  $\Pi - V_d = \Pi - \delta - V_c$ . In the second and third terms  $x_0$  is the portfolio holding chosen at date 0 (given by (3.9)). The second term reflects the fact that a fraction  $\phi$  of the firms the investor owns have become clean, and the third term is the original cost of the date 0 investment.

The certainty equivalent of this return is

$$(A3) \quad CE = x(\Pi - V_d) + x_0 [\phi V_c + (1 - \phi)V_d] - x_0 F - \frac{1}{2} \gamma x^2 \sigma^2,$$

and so the investor's date 1 choice of  $x$  will satisfy

$$(A4) \quad x = \frac{\Pi - V_d}{\gamma \sigma^2}.$$

The condition for date 1 stock market equilibrium is  $x = N$ , which combined with (3.10) yields

$$(A6) \quad V_d = F.$$

Thus,

$$(A7) \quad \begin{aligned} CE &= \frac{(\Pi - F)^2}{2\gamma\sigma^2} - \phi\delta x_0 \\ &= \frac{(\Pi - F)^2}{2\gamma\sigma^2} - \phi\delta \frac{(\Pi - F)}{\gamma\sigma^2}. \end{aligned}$$

By choosing  $\phi N$  clean firms, the planner will cause the total amount of pollution to be  $(1-\phi)Nh$ .

Thus, the planner will maximize

$$(A8) \quad \left[ \frac{(\Pi - F)^2}{2\gamma\sigma^2} - \phi\delta \frac{(\Pi - F)}{\gamma\sigma^2} - (1-\phi)Nh \right]$$

with respect to  $\phi$ . Recall that  $\Pi = \tau - \rho N$ , which is independent of  $\phi$ . We obtain a bang-bang solution (either  $\phi=0$  or  $\phi=1$ ) depending on whether

$$(A9) \quad \delta \frac{(\Pi - F)}{\gamma\sigma^2} > or < Nh.$$

Using (3.10), this boils down to

$$(A10) \quad \delta > or < h.$$