Trade, Voting, and ESG Policies: Theory and Evidence

John Duffy Daniel Friedman Jean Paul Rabanal
Olga A. Rud*

October 23, 2024

Abstract

We model the interaction between shareholder trading and voting on ESG policies under different sets of preferences and test the model's equilibrium predictions in the laboratory. The model suggests, and laboratory results confirm, that low policy costs favor policy adoption while intermediate costs lead to a lower policy adoption rate under dispersed preferences than under polarized preferences. Observed share prices are higher than equilibrium predictions when the policy is adopted and are close to the fundamental value of the firm had the policy not been adopted. We conclude that ESG policy adoption may be less burdensome to shareholders than predicted.

Keywords: shareholder voting, social responsibility, ESG, experimental economics
JEL: D74, G34, G11, Q50, C92

*John Duffy (duffy@uci.edu) is at UC Irvine and ISER Osaka University; Daniel Friedman (dan@ucsc.edu) is at UC Santa Cruz and University of Essex; Jean Paul Rabanal (jeanpaulrab@gmail.com) and Olga Rud (olga.rud@gmail.com) are at the University of Stavanger. We arae grateful for comments and feedback received from Te Bao, Doug Davis, Oliver Hart, Dorothea Kubler, David Levine, Nadya Malenko, Sebastién Pouget, Klaus Ritzberger, Paul Smeets, Simon Weidenholzer and Donald Wittman. Funding for this project was provided by the UC Irvine School of Social Sciences. The experimental protocol was approved by the UC Irvine Institutional Review Board, HS #20118378.

1 Introduction

Environmental, Social and Governance (ESG) policies have become increasingly important to many investors in deciding whether to acquire shares in a firm.[1] Indeed, it is now mainstream to consider climate change and environmental sustainability, employee satisfaction and workplace diversity, corporate policies regarding political lobbying, customer privacy and cybersecurity – more generally "corporate social responsibility policies"– in the analysis of a firm's business performance, in addition to fundamental factors such as the firm's free cash flow. Not surprisingly, this shift in thinking by some away from the "Friedman doctrine" of maximizing shareholder value (Friedman, 1970) has provoked much debate, especially given the difficulty of measuring ESG compliance and the potential risks of inaction.[2] Recent discussions have often emphasized value creation for shareholders while simultaneously advancing ESG goals. However, at some point profitable opportunities from ESG policies are likely be exhausted, either due to technological barriers that limit profitability, or to decreasing returns as an increasing number of firms and funds focus on ESG goals. As that process unfolds, firms and funds may question whether to undertake the ESG-oriented activities preferred by some investors at the expense of general shareholder value. The answer may hinge on shareholder engagement or activism, and the impact on firms' share prices.

In this paper, we present a model in which shareholders voice their opinions on an ESG policy via voting. The model considers the interplay between voting and value, using market trades to first establish ownership, followed by a voting stage on the policy proposal. Share prices reflect the firm's free cash flow, and the cost that the firm bears if it adopts a particular ESG policy. We bring this model to the laboratory using induced, rather than homegrown, social preferences, due in part to the lack of reliable field data on preferences for (or even costs of) ESG policies.

In our model, investors have heterogeneous preferences for adopting the costly ESG policy. If the policy is implemented, it then creates a positive utility benefit that

[1] Amel-Zadeh and Serafeim (2018) surveyed senior investment professionals at asset management firms in North America, Asia and Europe and report that 82% of respondents considered ESG information when making investment decisions. The main reasons cited were: 1) ESG information was financially material to investment performance, 2) clients were demanding ESG information, 3) ESG helped with product diversification, and lastly, 4) ESG information was important for ethical investment reasons.

[2] For a critical view, see the July 2022 special report published by *The Economist* on ESG investing, https://www.economist.com/special-report/2022-07-23.

extends to varying degrees to all market participants, regardless of share holdings. This positive utility benefit may arise from mitigating a negative externality of the firm's operations. If the policy is not adopted, then there is no utility benefit but neither is there additional cost borne by shareholders. Similar to Broccardo et al. (2022), the investors in our model derive utility from policy implementation according to an idiosyncratic weight: investors with zero weight receive zero utility from the policy implementation while those with positive weight receive proportional utility increments. The distribution of the policy preference weights matters for adoption of the policy. The model allows for arbitrary distributions, but in the experiment, we consider two specific cases: (i) dispersed preferences, where the weights are uniformly distributed between zero and a maximal weight, and (ii) polarized preferences, where the weights have an extreme bimodal distribution in which each investor's utility includes either zero weight or maximum weight for the ESG policy.

The model has two stages. In the first, shares are traded in a market. In the second stage, the post-trade shareholders vote (one share-one vote) on the ESG policy determined by majority rule. The model's solution concept is subgame perfection: anticipating their potential impact on the voting outcome, forward-looking investors trade shares at prices that reflect the fundamental value (e.g., the firm's free cash flow), and also the cost that the firm bears if it adopts the ESG policy. Specifically, we show that if the ESG policy is approved, the equilibrium share price is the dividend payout less the specified ESG policy cost. If the ESG policy is rejected, then the equilibrium share price is simply equal to the dividend payout. These equilibrium price predictions do not reflect any benefits (or harms) that may accrue from adoption (or non-adoption) of the ESG policy other than the cost of implementing the policy. The equilibrium price and voting outcomes, however, do depend on the distribution of preferences regarding the ESG policy. The model has both approval equilibrium and rejection equilibrium for a wide range of policy costs for some preference distributions, such as the dispersed case, but the equilibrium voting outcome is essentially unique for other preference distributions, such as the polarized case.

In practice, of course, investors may not all be forward looking and they may have only limited or asymmetric information. We view our complete information, rational agent model as an appropriate first step in understanding the interplay between trade, voting and ESG policies: it lays the groundwork for developing more intricate models to examine these interactions and to evaluate the effects of additional complexities.

Further, the simpler model setup facilitates the comprehension by human subjects of their available actions and possible payoffs.

The experiment combines a between-subjects design that varies the distribution of preferences (polarized or dispersed) across sessions, with a within-subjects design that changes the policy cost (low, intermediate, or high) within each session. While the laboratory setup closely follows the theoretical model's voting and trading procedures, it departs from the static nature of the model by allowing fixed groups of subjects to repeatedly make decisions under various parameterizations of the model. Such "stationary repetition" of a static model is standard practice for experiments with induced values that are intended to test the predictive power of equilibrium in models of markets or voting (see, e.g., Smith, 1982). We use neutral language in our instructions and emphasize market level feedback so as to minimize the potential impact of "homegrown" preferences and repeated-game effects.

The experimental results mostly confirm the predictions of our model. As predicted, when the policy cost is high (low), the policy is usually rejected (approved) under both sets of preferences. When the policy cost is intermediate, the policy has a 70% approval rate when preferences are polarized, and a 28% approval rate when preferences are dispersed, a statistically significant difference. Approval is the unique equilibrium outcome in the polarized/intermediate cost case but both approval and rejection are equilibrium outcomes in the dispersed/intermediate cost case, where the voting outcome depends on which investors end up holding shares. In all treatments, the observed share price is generally quite close to the predicted fundamental value when the ESG policy is rejected. However, when the policy is accepted, the observed share price is higher than predicted, and often close to the fundamental value of the firm *without* ESG policies. Hence, the cost of implementing the ESG policy is not fully reflected in share prices. We attribute this departure from theoretical predictions to an endogenous voting premium that our rational agent model does not account for. However, this intriguing behavioral finding suggests that the adoption of costly ESG policies might have a smaller impact on shareholder wealth than one might presume.

In sum, our contribution is both theoretical and empirical. On the theoretical side, we offer a streamlined model of trade and voting in the spirit of Hart and Zingales (2022). Our model allows individual voters to have non-negligible (but not dictatorial) influence on voting outcomes and is amenable to implementation in the laboratory. The empirical contribution is the results from such implementation. The model successfully predicts observed general trends in voting and trading behavior, the main exception being an apparent voting premium in share price.

The next section connects our paper to previous research, especially to the literature on shareholder voting, and on ESG investments. Section 3 presents our model, including some specific examples and the four hypotheses that we test in our experiment. Section 4 describes the experimental procedures and design.

In Section 5, we present the findings of our experiment. The analysis begins by examining aggregate outcomes, including predictions concerning policy adoption and share prices. Subsequently, we delve into the individual choices made by participants, particularly how investors respond to their induced social preferences. In terms of voting, we find that subjects mostly follow the predictions of the model; those who (post trade) would enjoy a net benefit from the ESG proposal reliably vote to approve it while those who would not benefit reliably vote to reject it. We also report a regression analysis on bidding behavior in the first stage that takes advantage of the fact that subjects participate in several repetitions of each treatment. It suggests that the observed voting premium may arise from a struggle to control the policy outcome. Subjects who sold their shares and experience an undesirable outcome tend to increase their bid price for shares in the following period. This behavior is consistent with a well-known behavioral principle sometimes known as "win-stay, lose-shift", wherein individuals make no change to their behavior following a successful outcome but switch their behavior following a failure.

Section 6 summarizes our main findings and offers some ideas for future research. Online Appendix A presents mathematical details for the theoretical model, while online Appendix B provides additional data analysis. The instructions used in the experiment are reported in online Appendix C.

2 Related literature

Our paper is related to several different literatures. The most relevant theory papers are those that, like ours, combine shareholder trading and voting. For instance, Broccardo et al. (2022) and Hart and Zingales (2022) analyze whether stakeholder pressure (from consumers, workers, and shareholders) can result in socially desirable environmental outcomes, e.g., pollution reduction. They explore stakeholders' options of exit (divesting or boycotting) versus voice (voting or engaging), where certain investors and consumers are socially responsible and take into account the well-being of others affected by their decisions. Their model shows that if a majority of agents exhibit positive social responsibility then the shareholders' voice can achieve the desired outcome. However, when the majority does not care about social responsibility, then the voice strategy does not effectively reduce pollution.[3] Differently from our model, they consider multiple firms with production and use a CAPM-like version of competitive equilibrium. Their voice model focuses on limiting cases where individual shareholders bear a miniscule fraction of the policy cost. We allow fully heterogeneous social preferences, include an explicit secondary market format for trading shares, and use game-theoretic equilibrium for investors trading voting shares of a single firm; in our setting, each shareholder bears a nontrivial fraction of the policy cost and potentially can affect the voting outcome.

Levit et al. (2022) study trading and voting in a model where investors differ in their attitudes toward a policy and thus in their valuation of shares, which creates gains from trade. Prior to voting, their investors trade shares, thus endogenously forming the shareholder base. In the case where shareholders have voting rights, expectations about policy acceptance can be self-fulfilling and so give rise to multiple equilibria. In our model, by contrast, shares have a common value for all investors, the differences consist in how much shareholders care about the firm's externality, regardless of share ownership. Multiple equilibria can arise in our model for similar reasons as theirs, despite the difference in how our investors value shares.

Demichelis and Ritzberger (2011) also study a two-stage game with one firm where shareholders first trade and then vote. However, the vote is not about an externality but rather over the firm's production plan. Potential shareholders differ in how much their utility depends on the price they pay as consumers of the firm's output. Abstention is allowed and voting is slightly costly, so insincere voting is dominated by abstention. Their model also has a multiplicity of equilibria, some that support the socially efficient production plan and others that are inefficient.

Gollier and Pouget (2022) similarly study voting on ESG policies. Their investors may care about an externality, and they trade before voting, with rational expectations about the outcome. Their model, like ours, can have multiple equilibria, one in which the firm adopts a socially responsible policy and another in which it does not. However, there are important differences in their setup and predictions relative to ours.

[3] Kaufmann et al. (2024) analyze a model without shareholders. They show that socially responsible consumers may reduce but cannot fully mitigate a negative externality in markets.

Gollier and Pouget (2022) assume a continuum of price- and outcome-taking agents; our agents are finite and aware that they can influence voting outcomes. Their agents' preferences are polarized: some agents are indifferent to the externality while other agents all care to the same degree; our model accommodates arbitrary distributions of agents' preferences regarding the externality. More importantly, their agents' preferences incorporate a proportional-to-shareholding "warm glow" and the agents who care about the externality always vote in favor of the policy; our agents have consequentialist preferences for the externality independent of their shareholdings, and they consider the policy cost that they will bear when they vote on the ESG policy. Gollier and Pouget (2022) assume competitive equilibrium in which agents' demand for shares arises from identical CARA Bernoulli functions, and the share price reflects the CARA parameter and the warm glow as well as the cost of the ESG policy; we assume Nash equilibrium arising from risk-neutral agents' bids in an explicit market institution, and the resulting price reflects the policy cost only when the policy is adopted.

There are also theoretical models in which trading occurs *after* voting. In Meirowitz and Pi (2022), shareholders receive a signal about the asset payoff and then vote on firm policy. After observing the voting outcome, subjects enter the market where shares can be traded. They find that, due to price distortions, the likelihood of sincere voting decreases as the number of shareholders increases. In our model, however, sincere voting always improves payoff whenever it makes any difference. Many voting models have only mixed strategy equilibria; we focus on pure equilibria of our model, which are easier to test empirically. We further differ from these theory papers in empirically testing our model's predictions, which are easier to analyze compared to other voting games that find solely mixed strategies. Another advantage of our model is that it does not rely on risk preferences, which are difficult to measure or control in the laboratory.

Like us, Casella et al. (2012) combine voting theory and laboratory experiments, but their focus is on vote buying, with no role for shareholding per se. Their voters have heterogeneous preferences over a binary social choice. Their model relies on a competitive equilibrium (price-taking), with ex ante but not ex post market clearing. Their main result is that there is always an ex post dictator, who holds a majority of the votes. Our model rules out the possibility of such dictators as we are interested in the behavior of shareholders in widely held companies.

A recent empirical literature shows that shareholder votes can affect stock prices. Fos and Holderness (2023) find that U.S. stock prices typically decline after votes are distributed (ex vote). The magnitude of the decline varies with how controversial the vote is expected to be and how investors are notified of the record date.

There are also a number of studies that focus on how shareholder types may affect voting behavior. Dasgupta et al. (2021) provide a thorough overview of the recent academic literature, both theoretical and empirical, on the role of institutional investors in corporate governance, including voting. They pay close attention to heterogeneity among institutional investors.[4] Levit et al. (2023) develop a model with a minority blockholder and dispersed shareholders. They show that in equilibrium a voting premium emerges from the blockholder's desire for control. Levit et al. (2023) also survey a number of empirical studies that quantify the voting premium.

There is also a substantial literature exploring preferences for ESG investments. Heeb et al. (2023) find that although investors are willing to pay for efforts to reduce CO2 emissions, the magnitude of the impact is less important. These findings remain consistent across different subject populations (experienced private investors, mTurk participants, and students). Bauer et al. (2021) document that about 68% of a Dutch pension fund's participants want the fund to push harder for sustainability of the companies in which it invests. Riedl and Smeets (2017) show that reciprocity plays an important role in determining socially responsible investment, and that investors are willing to pay significantly higher management fees associated with such investments. Using data from private equity funds, Barber et al. (2021) find that investors are willing to forego between 2.5%–3.7% in expected returns by holding funds with impact goals relative to conventional funds. Baker et al. (2022) estimate a model of investor demand for index funds and find that the value that investors place on ESG policies has increased nearly threefold over from 9 basis points (bp) in 2019 to 28 bp in 2022.

Hong and Shore (2023) evaluate 60 empirical papers, exploring shareholder demand for sustainable or ESG investments based on pecuniary and non-pecuniary motivations. With pecuniary motives, shareholders must demand ESG policies because they are profit maximizing in the long-run, though inefficiencies in capital markets, or agency problems, may impede short-run profits. Non-pecuniary motives are that investors care not only about firm profits but also about externalities, and are willing to sacrifice returns; e.g., investors are willing to pay a "greenium" in which they accept lower returns for greater sustainability. They report that a sizeable majority of papers support the non-pecuniary motive for ESG policies. While our framework abstracts

[4]For an overview of early empirical literature on shareholder voting, please see Yermack (2010).

from agency problems, we also conclude that non-pecuniary motivations are behind the voting premium observed in our experiment.

Some recent experimental work has studied the impact of subjects' homegrown preferences (as opposed to the induced preferences that we employ) on valuing assets with externalities. Guenster et al. (2022) find a higher willingness-to-pay for socially responsible assets than for conventional assets. On average, the price premium for socially responsible assets is about 8%. Humphrey et al. (2021) show that subjects are less willing to invest in a risky asset that creates a negative externality. Bonnefon et al. (2022) report on an experiment eliciting willingness to pay for fictional companies that either give some fraction of profits to charities or take money away from charities or create no such externalities. They find that subjects' bids in excess of the shareholder dividend incorporates the charitable externality in a linear fashion, i.e., excess bids are higher (lower) for companies donating (taking) to (from) charities, regardless of whether or not those transfers require that the investor owns shares in the company. These results suggest that investors' bids are driven by value-alignment rather than whether their investment choices actually impact outcomes. By contrast, in our setting, ownership conveys voting rights, and votes determine whether an ESG policy is implemented or not. Further, we do not consider the case of negative externalities. However, similar to all of these papers we find that when ESG policies are implemented, investors are willing to pay in excess of the fundamental value for such assets.

Finally, some experiments study the effect of vote trading on welfare. Casella et al. (2012) find welfare losses when vote trading is allowed while Tsakas et al. (2021) show welfare gains when the payoff function includes vote shares. Dittmann et al. (2014) design an experiment with dual-class shares. They conclude that a voting premium appears because subjects value participation for purely instrumental reasons. Similarly, in our experiment, subjects can trade shares, and therefore voting privileges to approve or reject a costly policy that generates a positive external benefit. We also find some evidence for a voting premium in that market prices are greater than equilibrium predictions when approval of the policy is likely.

3 A Model of Trading and Voting

We capture key aspects of ESG policy-making in a two-stage game of complete information. A firm has a fundamental value of v per share. After trade, shareholders vote on whether to adopt an ESG policy with a known cost of $\tau \in (0, v)$ per share and a per-capita social benefit (or harm reduction) to all agents of $h \geq \tau$. The idea is that the firm has a specific opportunity to benefit the community at large. For example, a monopolist firm might consider reducing the pollution that it generates even though doing increases its production cost and reduces its profits. Alternatively, a firm with limited pricing power might alter its supply chain, increasing its input costs but reducing its indirect use of forced labor.

Following Broccardo et al. (2022) and Hart and Zingales (2022), we focus on the impact of idiosyncratic utility bonuses $\lambda_i h$ that investors might gain, regardless of their shareholdings, if the ESG policy is implemented. As Hart and Zingales (2022) stress, such utility bonuses are included in shareholder *welfare* but are excluded from the traditional measure of shareholder *value* (or, for the sake of clarity, what we will call shareholder *wealth*), because these bonuses are not reflected in the fundamental value of an asset share.

More formally, investors $i = 1, ..., n \geq 2$ are each characterized by a share endowment $e_i \geq 0$, normalized so that $\sum_{i=1}^{n} e_i = n$, and by a (social) preference parameter $\lambda_i \geq 0$.[5] In the first stage of the game, investors trade shares in a Bid-only Call Market. That is, all investors simultaneously submit bid prices $b_i \geq 0$ and maximum trade quantities $q_i \in [0, q_M]$. As detailed below, the market then finds a clearing price P and net trade vector $\mathbf{a} = (a_1, ..., a_n)$ with $\sum a_i = 0$ and $|a_i| \leq q_i$, resulting in final share allocation $\mathbf{x} = (x_1, ..., x_n)$ where $x_i = a_i + e_i$.

In the second stage, shareholders vote on a (ESG) policy proposal. We assume one-share-one-vote, so a total of $\sum x_i = \sum e_i = n$ votes are cast; investor i casts all x_i of their votes regarding the proposal either Y (Yea) or N (Nay). We assume majority rule, so the proposal is rejected (denoted $\xi = 0$) if there are less than $n/2$ votes Y, in which case each share has value $v > 0$. If there are at least $n/2$ votes Y, then the proposal is accepted and the ESG policy is implemented. In that case, denoted $\xi = 1$, each share has value $v - \tau > 0$, and each investor i (shareholder or not) receives their idiosyncratic utility increment $\lambda_i h$.

Thus if the proposal is rejected ($\xi = 0$), then investor i's payoff is $x_i v - a_i P$, where the second term represents payments for purchased shares when $a_i > 0$ or receipts from sales when $a_i < 0$. If the proposal is accepted ($\xi = 1$), then investor i gains $\lambda_i h$ but

[5] We ignore the possibility of spiteful social preferences, $\lambda_i < 0$, because such preferences of limited interest in our context.

loses $x_i\tau$. Therefore the payoff function is

$$u_i = e_i v + a_i(v - P) + [\lambda_i h - (e_i + a_i)\tau]\xi. \tag{1}$$

3.1 Equilibrium

We now characterize equilibrium behavior using subgame perfect Nash equilibrium with the refinement that, in the second stage, investors follow the obvious threshold strategy (SPNE*). That is, SPNE* requires that investors vote Y if and only if implementing the proposal increases their own payoff, i.e., iff the bracketed term in equation (1) is nonnegative, i.e.,

$$\text{Vote Y} \iff \lambda_i \geq \lambda^*(x_i) = \frac{(e_i + a_i)\tau}{h}. \tag{2}$$

It is easy to see that this behavior is weakly dominant: a deviation either has no impact on the outcome or else it lowers the deviator's payoff.

We now detail the Bid-only Call Market (BOCM) mechanism that we use in the first stage of our game.[6] Recall that each investor i submits a bid price $b_i > 0$ and quantity $q_i \in [0, q_M]$, where $q_M > 0$ is the maximum individual order size that the market will process. Using the indicator function — $I[z] = 1$ if expression z holds and otherwise $I[z] = 0$ — we can write the revealed demand as

$$D(p) = \sum_{i=1}^{n} q_i I[b_i \geq p]. \tag{3}$$

Thus D is a decreasing left-continuous step function; in the usual (q, p) graph, the steps have widths q_i at heights $p = b_i$ and are ordered in decreasing b_i.

The supply of asset shares is constant: $S(p) = \sum_{i=1}^{n} e_i = n$. The market clearing price P is a solution to demand = supply, or to $D(p) - n = 0$. If there is an interval (between adjacent bid prices) of solutions, then P is the midpoint. For all investors i with $b_i > P$ the market completely fills the buy order at price P, so $a_i = q_i$ and the investor pays Pa_i. Likewise, each investor with $b_i < P$ sells the specified quantity: $a_i = -q_i$ at price P and so receives payment Pa_i. By construction, the market clears

[6] We use the BOCM because it is tractable theoretically, and even more importantly because it is easy for laboratory subjects to understand. More complicated formats such as the continuous double auction are notoriously difficult to model game theoretically, and take much more time to implement in a laboratory experiment. Our BOCM format therefore is a separate methodological contribution of this paper.

at price P: share purchases $\sum a_i I[a_i > 0]$ equal share sales $-\sum a_i I[a_i < 0]$.[7] Since supply = number of investors, it can be seen that P is the median q-weighted bid price.

Note that any given strategy profile $(\mathbf{b}, \mathbf{q})$ produces unique outcomes $\mathbf{a}$ and P at the first stage, and that those outcomes, together with the endowment and preference parameters $(\mathbf{e}, \lambda)$, produce a unique second stage outcome ξ via equation (2).

Definition. The strategy profile $(\mathbf{b}, \mathbf{q})$ with outcome vector $(\mathbf{a}, P, \xi)$ is a SPNE* if, for each investor i, replacing (b_i, q_i) by any other bid (b'_i, q'_i) produces outcomes $(\mathbf{a}', P', \xi')$ that do not increase investor i's payoff as specified in equation (1).

Remark 1. SPNE* considers only pure strategy equilibria. It is a relatively strong equilibrium concept which requires that investors anticipate the impact their bids will have on the second stage voting outcome, and rules out equilibria in which investors vote against their preferred policy outcome (which otherwise would be possible when they don't expect their vote to matter). As noted in the Introduction and further argued in Section 6, using such a strong equilibrium concept is appropriate for a first investigation of voting by an endogenously determined electorate, but future studies may find it useful to consider weaker equilibrium concepts.

Definition. There are two types of SPNE* : *approval equilibrium*, in which the SPNE* voting outcome is $\xi = 1$, and *rejection equilibrium*, in which it is $\xi = 0$.

Definition. The parameter vector $(n, v, h, q_M, \mathbf{e}, \lambda)$ is *admissible* if $n \geq 4$ is an integer; $v, h > 0$; $q_M = \frac{n}{4}$; $e_i \in [0, \frac{n}{4})$ with $\sum_{i=1}^{n} e_i = n$; and $0 \leq \lambda_1 \leq ... \leq \lambda_n$ with $\lambda_n, \lambda_{n-1} > 0$.

Remark 2. The substantive restrictions for admissibility are intended to rule out monopoly, i.e., an ex-post dictator; see Section 6 for further comment.

Definition. An allocation $\mathbf{x} = (x_1, ..., x_n)$ is *feasible* if there is some SPNE* strategy profile for which the BOCM generates trade vector $\mathbf{a}$ such that $\mathbf{x} = \mathbf{a} + \mathbf{e}$.

Proposition 1. *Given any admissible parameter vector* $(n, v, h, q_M, \mathbf{e}, \lambda)$*,*
(i) $\exists\ \hat{\tau} \in [0, v)$ *s.t. an approval equilibrium exists iff* $\tau \in [0, \hat{\tau}]$ *,*
(ii) $\exists\ \check{\tau} \in [0, v)$ *s.t. a rejection equilibrium exists iff* $\tau \in (\check{\tau}, v]$*,*
(iii) $\check{\tau} \leq \hat{\tau}$*.*

[7] It is possible that $b_j = P$ for $m > 0$ investors, i.e., that one or more investors reveal indifference to buying or selling at what turns out to be the market clearing price. For such investors, the BOCM sets $a_j = -\frac{1}{m} \sum a_i I[b_i \neq P]$, i.e., they equally split the residual supply or demand. Thus the market clears even when there are bids exactly at the market price and the purchases and sales at other bid prices do not balance.

Appendix A provides proofs of the Propositions. In the proof of Proposition 1, the upper bound $\hat{\tau}$ on approval equilibrium is obtained by finding, for a given final allocation $\mathbf{x}$, the $\tau_A(\mathbf{x})$ where (2) binds for the pivotal voter, and taking the maximum τ_A over feasible final allocations. Finding the lower bound $\check{\tau}$ on rejection equilibria is analogous, and the constructions ensure that (iii) holds. Example 2 below shows that there are parameter vectors for which $\check{\tau} = \hat{\tau}$, and Examples 1 and 3 show that there are parameter vectors for which the two types of equilibrium coexist over a non-trivial interval of costs $[\check{\tau}, \hat{\tau}]$.

Turning now to equilibrium pricing, it turns out that difficulties can arise for acceptance equilibria if one or more investors are exactly indifferent as to whether the project is approved and their combined vote is pivotal. Appendix A defines a *regular* equilibrium as one where this knife-edge case does not occur.

Proposition 2. *For any given admissible parameter vector, let $(P, \mathbf{a}, \xi)$ be the outcome of a regular SPNE*. Then $P = v - \tau\xi$. In particular,*
(i) $P = v - \tau$ in any regular approval equilibrium, and
(ii) $P = v$ in any rejection equilibrium.

Proposition 2 asserts that in SPNE* there is no room for a voting premium — in the market, a share is worth no more (and no less) than the net liquidating dividend $v - \tau\xi$. The proof in Appendix A exploits the fact that other prices provide arbitrage opportunities. When $P > v - \tau\xi$, sellers will want to deviate from their present bid profile to sell more, and buyers will want to deviate to buy less. The opposite deviations are profitable when $P < v - \tau\xi$, and so $P \neq v - \tau\xi$ seems inconsistent with SPNE*. Care must be taken, however, to ensure that those deviations remain profitable when they might alter the voting outcome ξ.

Corollary 1. *Regular SPNE* exists for every admissible parameter vector and every $\tau \in [0, v]$. At every $\tau \in [0, v]$ there is either a unique rejection equilibrium price or a unique regular approval equilibrium price (or both, for $\tau \in (\check{\tau}, \hat{\tau})$).*

The first part of the Corollary follows easily from Proposition 1, noting that regularity can be ensured by slight adjustments to the bid profile. The second part of the Corollary follows directly from Proposition 2, which shows that the equilibrium price is unique given the equilibrium voting outcome ξ. Of course, the bid profiles that produce these equilibrium outcomes are not unique, nor are the trade outcomes

a. However, the next result shows that the trade outcomes do not affect equilibrium payoffs.

Corollary 2. *For any given cost $\tau \in (0, v]$ and admissible parameter vector, investors' payoffs u_i in equilibrium depend on vote outcome ξ but not on trade outcome* $\mathbf{a}$.

Rewrite equation (1) as

$$u_i = e_i(v - \tau\xi) + a_i[v - \tau\xi - P] + \lambda_i h\xi. \tag{4}$$

Proposition 2 ensures that the square-bracketed expression in (4) is zero in equilibrium, so each payoff u_i is independent of a_i and the Corollary follows.

Remark 3. Outside of equilibrium, the square-bracketed expression is negative (resp. positive) when P is above (resp. below) its equilibrium value. Thus we have the natural result that prices above equilibrium are good for sellers ($a_i < 0$) and prices below equilibrium are good for buyers ($a_i > 0$).

Remark 4. Intuition might suggest that free riding could pose difficulties for approval outcomes: everyone with $\lambda_i > 0$ benefits from approval, but it might seem that they can free ride by not buying shares, which bear the costs, hoping that other investors favoring approval will hold enough shares. However, an implication of Proposition 2 and its corollaries is that the endowed shareholders, not the final shareholders, will bear all the costs because, in approval equilibrium, trade takes place only at a share price that fully reflects the anticipated costs. Of course, the free rider problem resurfaces when $P > v - \tau$ outside of regular approval equilibrium.

Example 1: Homogeneous preferences. To improve intuition and to set the stage for our experiment, consider the following examples with $n = 11$ investors, fundamental share value $v = 1$, social benefit $h = 1$ and uniform endowment $e_i = 1, \quad i = 1, ..., 11$. The first example shows that the cost interval $[\check{\tau}, \hat{\tau}]$ allowing both types of equilibrium can be fairly wide when preference parameters are tightly bunched.

Indeed, suppose that preferences are completely homogeneous, $\lambda_i = \lambda_0 > 0$ for all traders. To find the upper bound $\hat{\tau}$ on approval equilibria, consider a SPNE* bid profile that produces no trade ($\mathbf{a} = \mathbf{0}$) and $P = v - \tau = 1 - \tau$ in the first stage. (Appendix A verifies the existence of such profiles.) Since $x_i = 1$, equation (2) tells us that each investor i will vote Y as long as $\tau \leq \frac{\lambda_0 h}{x_i} = \lambda_0$. Hence $\hat{\tau} \geq \lambda_0$. In any approval equilibrium we must have at least one investor holding at least one share who votes Y, but that will not happen for $\tau > \lambda_0$. Thus $\hat{\tau} \leq \lambda_0$, and therefore $\hat{\tau} = \lambda_0$.

To find the lower bound $\check{\tau}$ on rejection equilibria, consider a SPNE* bid profile that results in $P = v = 1$ and $x_1 = x_2 = 3$. Then investors 1 and 2 will vote N and force $\xi = 0$ as long as $\tau > \frac{\lambda_0 h}{x_1} = \frac{\lambda_0}{3}$. Thus $\check{\tau} \leq \frac{\lambda_0}{3}$, so both kinds of equilibria exist over the range $\tau \in (\frac{\lambda_0}{3}, \lambda_0]$ when preferences are homogeneous at λ_0.

Example 2: Polarized preferences. A second example illustrates the case of an essentially unique equilibrium outcome ($\check{\tau} = \hat{\tau}$) and will be used in our laboratory experiment. As before, $n = 11$ and $v = h = e_i = 1$. Here and below we simplify the analysis by imposing the bidding constraint $q_i = 1$.[8] Consider the case of extremely polarized preferences: $\lambda_i = 0$ for $i = 1, ..., 6$ and $\lambda_i = 1$ for $i = 7, ..., 11$. Any shares held by investors 1-6 will be voted N for any positive τ, so approval equilibria are only possible when at least one of the other investors acquires an additional share. A specific such example detailed in Appendix A has outcomes $P = 1 - \tau$ and $\mathbf{a} = \mathbf{0}$ except that $a_7 = a_8 = a_9 = 1,\ a_1 = a_2 = a_3 = -1$. Investors 7-9 will vote their two shares Y (and a fortiori investors 10-11 will also vote Y, ensuring $\xi = 1$) as long as $\tau \leq \frac{\lambda_7 h}{x_7} = \frac{1}{2}$. Appendix A verifies that indeed $\hat{\tau} = 0.5$. If $\tau > 0.5$ then no investor will find it profitable to buy any units, so $\check{\tau} = 0.5$ as well.

Example 3: Dispersed preferences. Our final example is the same as the previous example except that $\lambda_i = \frac{i-1}{11}$. Note that the mean parameter is $\bar{\lambda} = \frac{1}{n} \sum \lambda_i = \frac{5}{11}$ in both examples, but here preferences are distributed uniformly rather than polarized. In the experiment this case is referred to as the "dispersed preferences" treatment.

To find the upper bound $\hat{\tau}$ for approval equilibrium, note that, with no trade from the uniform endowment $e_i = 1$, the pivotal voter is $i = 6$ with $\lambda_6 = \frac{5}{11}$. She will vote Y iff $\tau \leq \tau_6(\mathbf{e}) = \frac{\lambda_6 h}{x_6} = \lambda_6 = \frac{5}{11}$, so $\hat{\tau} \geq \frac{5}{11}$. The only other candidates for achieving the upper bound are allocations $\mathbf{x}$ with $x_{11} = 2$. This investor will then vote Y only if $\tau \leq \tau_{11}(\mathbf{x}) = \frac{\lambda_{11} h}{x_{11}} = \frac{10/11}{2} = \frac{5}{11}$. We conclude that $\hat{\tau} = \frac{5}{11}$.

The lower bound $\check{\tau}$ for rejection equilibrium comes from trades that induce negative assortativity, where higher λ investors sell shares to lower λ investors. Suppose that $P = v$, and $a_1 = a_2 = a_3 = a_4 = 1 = -a_5 = -a_6 = -a_7 = -a_8$. Investor 4 (and a fortiori 1, 2 and 3) will vote N if $\tau > \tau_4(\mathbf{x}) = \frac{\lambda_4 h}{x_4} = \frac{3/11}{2} = \frac{3}{22}$. Investors 5-8 have no shares and therefore no vote. Investors 9-11 will vote their endowed shares Y, but the proposal will fail on an 8-3 vote. We conclude (see Appendix A for details omitted

[8] Since the number of bidders is odd, this simplification implies that P is unilaterally set by the 6th highest bid $b_{(6)}$. Absent ties at $b_{(6)}$, that investor neither buys nor sells, and so can not favorably manipulate price. Nor can any other investor.

here) that $\check{\tau} = \frac{3}{22}$ in this example. Figure 1 below shows the ranges of ESG policy costs τ that support approval and rejection equilibrium for Examples 2 and 3.

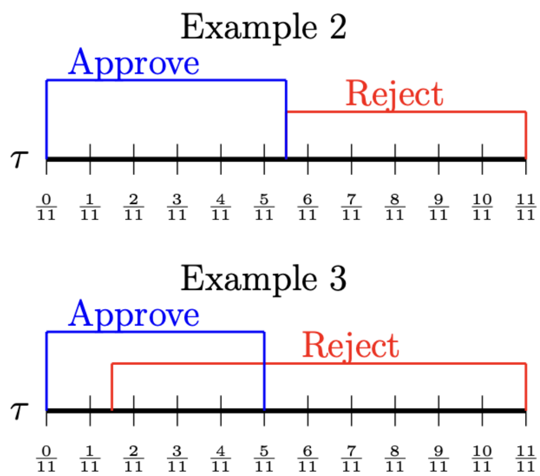


Figure 1: Intervals of ESG policy cost τ that support each type of equilibrium for polarized (Example 2) and dispersed (Example 3) preferences.

3.2 Efficiency analysis

The parameter h is described as the per-capita benefit of the ESG proposal to the entire population. If there are $m >> n$ people in the world, then the total benefit mh may far exceed the cost $n\tau$. However, the model has nothing to say about what benefits accrue to the $m - n$ non-investors versus the n investors in the game; the benefits to non-investors might be vast or negligible or anywhere in between.

Here we focus on a narrower notion of efficiency, the payoff sum over the n investors. There are two observations of interest. First, efficiency in this sense doesn't depend on market outcomes $(P, \mathbf{a})$; it depends only on voting outcome ξ. To see this, note that the first term in (1) always sums to nv, the second term always sums to 0, and the third term sums to $[h \sum \lambda_i - n\tau]\xi$. Therefore all approval equilibria (and non-equilibria with $\xi = 1$) are equally efficient. By the same token, each rejection equilibrium (and non-equilibrium) has the same efficiency as any other rejection outcome.

Second, approval outcomes are strictly better than rejection outcomes iff the bracketed expression above is positive. Note that the game is unchanged if we multiply each λ_i by a positive constant c and divide h by the same constant. Thus we can set $h = 1$

with no loss of generality, and write that approval is efficient iff

$$\tau \leq \bar{\lambda} = \frac{1}{n} \sum_{i=1}^{n} \lambda_i; \tag{5}$$

that is, if and only if the per capita cost for investors is less than the per capita utility increment. In Examples 2 and 3, and so in our experiment, $\bar{\lambda} = \frac{5}{11}$. Thus approval outcomes are efficient in our experiment iff $\tau \leq \frac{5}{11}$.

3.3 Testable Predictions

We now formulate hypotheses for polarized preferences (Example 2) and for dispersed preferences (Example 3), which we shall test in the laboratory using ESG costs $\tau \in \{0.2, 0.35, 0.6\}$. Hence, the first hypothesis focuses on those specific values.

Hypothesis 1. *With dispersed preferences, the ESG policy can be rejected or accepted for $\tau \in \{0.2, 0.35\}$; for $\tau = 0.6$, the policy is always rejected. With polarized preferences, the policy is always accepted for $\tau \in \{0.2, 0.35\}$ and always rejected for $\tau = 0.6$.*

This hypothesis comes from Proposition 1 and is illustrated in Figure 1. For $\tau \in \{0.2, 0.35\}$, only approval equilibria exist with polarized preferences, whereas both approval and rejection equilibria occur with dispersed preferences, as these costs fall within the interval $(\check{\tau}, \hat{\tau}] = (3/22, 5/11]$ associated with multiple equilibria. Our model does not address equilibrium selection explicitly, but reasonable refinements would suggest that rejection is less common towards the lower end of this cost range.[9] When $\tau = 0.6$ the policy cost is above $\hat{\tau}$ for both sets of preferences, and so the policy is always rejected in equilibrium.

Hypothesis 2. *When the policy is rejected (resp. accepted), the price per share is equal to v (resp. $v - \tau$).*

Hypothesis 2 comes directly from Proposition 2, which shows how the equilibrium share price P depends on the equilibrium voting outcome ξ. Even outside of equilibrium, if all investors confidently anticipate $\xi = 1$ then any deviation from $P = v - \tau$

[9] Our intuition is that there are fewer bid profiles supporting rejection equilibrium and that type of equilibrium is more delicate (i.e., overturned by smaller perturbations of the bid profile) as we decrease the policy cost towards $\check{\tau} = 3/22$.

creates arbitrage opportunities. Likewise, if all investors expect $\xi = 0$ then arbitrage should ensure that $P = v$.

Hypothesis 3. *Investors with $\lambda_i \geq \lambda^*$ (resp. $\lambda_i < \lambda^*$) as defined in equation (2) will vote to approve (resp. reject) the policy.*

Hypothesis 3 is our assumption that investors follow their weakly dominant strategy in the second stage of the game. It is possible that human subjects will deviate from this threshold behavior due to, e.g., noise or other-regarding preferences.

Hypothesis 4. *Subjects who experience their less preferred policy outcome in a given period will tend to increase their bid prices in the next period of an experiment, especially when they hold no shares and the vote is close.*

Hypothesis 4 does not follow from our static model, which assumes equilibrium behavior, but it nevertheless can shed light on what sort of equilibrium is relevant. In a generalized Competitive Equilibrium (CE) investors are price-takers and ξ-takers. By contrast, in a SPNE* investors are aware that their bids can affect prices and that their votes can affect the outcome ξ. Hypothesis 4 tests the notion that investors are ξ-takers, and asserts to the contrary that investors might adjust their prices in response to voting outcomes. Doing so can be a best response when an investor may be able to acquire the pivotal vote.

4 Laboratory procedures

The experiment was programmed using the oTree software (Chen et al., 2016). Subjects were undergraduate students pursuing a variety of majors at the University of California, Irvine with no prior experience with our study. In each session, subjects were first given written instructions that were read aloud, and then asked to complete a quiz to check their understanding; copies of these materials can be found in the Online Appendix, section C. Subjects received feedback as to which quiz questions they answered correctly or incorrectly and in the latter case, they were instructed about the correct answer. The experimenter (one of the authors) answered any remaining questions privately. Subjects then moved on to the main task, which involved the play of 48 two-stage periods of the asset market/voting game. Finally, participants filled out an exit survey and were paid. A summary of the number of sessions and subjects is provided in Table 1.

Table 1: Session information

Preferences	#subjects	#sessions
Polarized	55	5
Dispersed	55	5

Several model parameters are common across all sessions and treatments. Following the examples in the previous section, we always have $n = 11$ subjects per session. We express price and value parameters in ECUs or points, scaled by 100 (e.g., $v = h = 100$) to allow for more noticeable variation. As in the Examples, each subject i is endowed with $e_i = 1$ share per period and was restricted to buying or selling at most $q_i = 1$ indivisible share per period. Subjects were also endowed with a zero-interest loan of 200 ECUs to be repaid at the end of each period; this loan amount is more than sufficient to purchase a unit of the asset at equilibrium prices.

Following longstanding tradition in experimental economics, preferences were *induced* to enable the control necessary to test model predictions.[10] Specifically, subjects were randomly *assigned* values for λ, and the way in which these values were assigned differed between two preference distribution treatments, polarized and dispersed. These treatments were implemented in a *between subjects* manner: each session was conducted under a single preference distribution. In the five sessions of the polarized treatment, the first 6 investors, $i = 1, 2, \ldots, 6$, were assigned $\lambda = 0$ and the remaining 5 investors, $i = 7, 8, \ldots, 11$, have $\lambda = 1$, as in Example 2. In the dispersed preference treatment, as in Example 3, $\lambda_i = \frac{i-1}{11}$ for investor i. We shuffled the subjects' assigned λs mid-session, after 24 periods, so that low-λ subjects had a chance to make decisions as high-λ investors and vice versa. Specifically, in the polarized treatment, following period 24, the 5 subjects with $\lambda = 1$ swap with 5 of the subjects with $\lambda = 0$; the other $\lambda = 0$ subject sees no change in her preference parameter. In the dispersed treatment, mid-session, the subject assigned to λ_i is switched to $\lambda_{12-i} = \frac{11-i}{11}$.

The other treatment variable is the policy cost per share, $\tau \in \{20, 35, 60\}$ which in any given period was the same for all 11 subjects. This policy cost varied across 8-period blocks within a session. Specifically, in both the polarized and the dispersed

[10] We used neutral language and avoided any mention of ESG; subjects were only told that the "policy will also generate a personalized benefit to each participant...[which] ranges from 0 to 100 points" and the possible policy benefit weights/values (which differed according to whether the treatment was polarized or dispersed). While it would also be interesting to consider subjects' "homegrown preferences" toward ESG policies, we would not know in advance how such preferences map to the theoretical preference distributions that we study.

treatments, during the first 24 periods, subjects faced each of the 3 policy costs once for 8 consecutive periods (in a random order) all under a constant personal value for λ. They faced the same 3 costs again in the remaining 3 blocks randomly under a (typically) *different* constant personal value for λ. This block structure allows play to settle down and, with both an early and a late block for each cost, provides a consistency check and a strong test of Hypotheses 1-3.

Each session lasted just under 120 minutes. One of the 48 periods played was randomly selected at the end for payment. Points in excess of the 200 point loan earned cash at the fixed conversion rate of $20 per 100 points. Each subject also received a show-up payment of $10. In all, we conducted 10 sessions, 5 each of the polarized and dispersed treatments. Average subject earnings were $31.86.

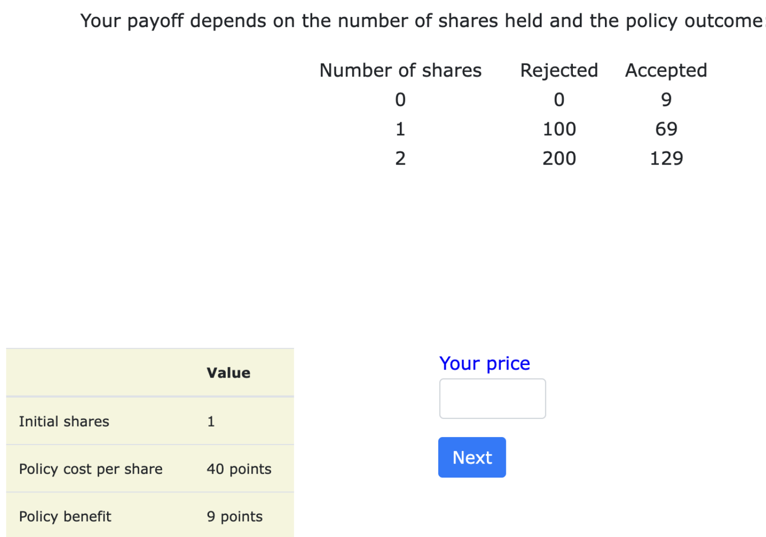


Figure 2: User interface for bidding stage

Note: Subjects submit bids in the box labeled *Your price*. The top panel displays payoff information, which depends on the number of shares held after trading, whether the policy is approved, and λ_i. We also provide information on the initial shares, policy cost per share (τ) and policy benefit ($\lambda_i h$).

Figure 2 presents the user interface (UI) that subjects used to submit their bid price to a Bid-Only Call Market in the first (trading) stage of each period. As noted in Section 3, the market clearing price P is the median bid (since all bids are for a single unit); subjects with bids above P are buyers of a single share, and those with bids below P are sellers. Absent tied bids at P, a subject bidding exactly P just keeps her endowment.[11] The top of Figure 2 displays components of the subject's payoff,

[11] Ties at the median bid value are broken randomly so that the expressions in footnote 7 hold in

equation (1), for relevant contingencies (final allocation and voting outcome). The numbers to the left of the price entry box remind subjects of their share endowment ($e_i = 1$) as well as the common-to-all policy cost per share (τ) and their personal policy benefit ($\lambda_i h$) if the policy is accepted. Consistent with the complete information assumption of the model, the instructions tell subjects the economy-wide distribution of such policy benefit values (polarized or dispersed).

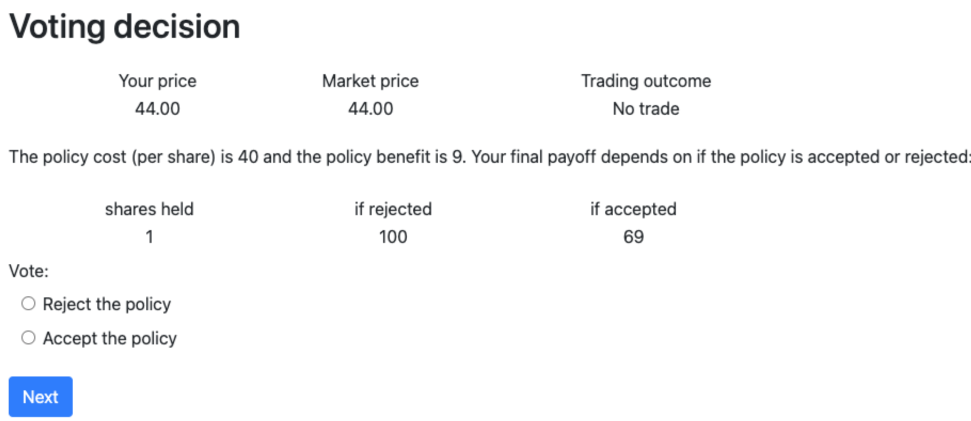


Figure 3: User interface for voting stage

Note: The subject casts her vote by clicking on *Reject the policy* or *Accept the policy*. The top panel shows the bid submitted, the market clearing price, and the trading outcome. Below the market information, the interface displays payoff relevant information, depending on the policy state.

Following the trading stage, subjects enter the voting stage as illustrated in Figure 3. The screen for this stage shows the trading outcome, including the market price and their own trade (if any). It also shows their contingent payoff for policy rejection or approval given their realized share holdings. Subjects holding 1 or 2 shares are then asked to vote their share(s) by clicking on either the Reject or Accept button as shown in Figure 3. Under majority rule, ties are not possible since there are 11 shares total. Subjects holding zero shares saw the same screen with the vote buttons omitted and were prompted to click a 'next' button.

The subjects finish the period with a final screen (omitted for brevity) that reminds them of the trading outcome and shows the number of votes to approve and to reject the policy, the resulting policy outcome, and their own payoff for those outcomes, as in equation (1).

expectation. E.g., if everyone submits the same bid value b_0, then $P = b_0$ and there is no trade. Subjects were informed of this rationing rule in the FAQ of the instructions.

5 Results

5.1 Aggregate results

Recall that Hypotheses 1 and 2 use equilibrium (SPNE*) results to predict observed aggregate outcomes: approval or rejection, and share prices. Since equilibrium is intended to represent outcomes after behavior has settled down, we focus on the late periods, defined as the last 4 periods within a block of 8 periods. Since there are two blocks for each cost, we have 8 periods per τ in each session.[12]

Figure 4 shows mean policy approval rates in the Polarized (dark bars) and the Dispersed treatments (light bars). The data are further disaggregated according to the other treatment variable, the policy cost, where $\tau \in \{20, 35, 60\}$. Figure 4 also reports p-values from a non-parametric Wilcoxon test for differences in acceptance rates between preference treatments for a specific value of τ. Table 2 presents a summary of predictions regarding approval rates and prices and also reports means (medians) for approval rates (prices) per treatment condition.

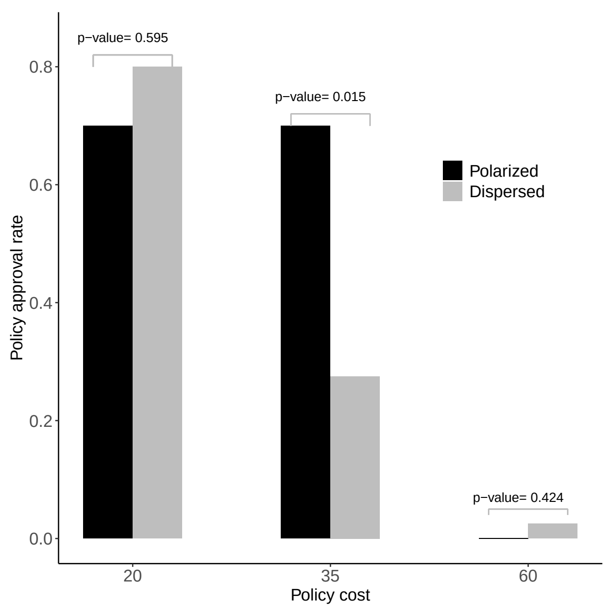


Figure 4: Approval rate under polarized and dispersed preferences
Note: The data are from the last four periods of each block; for each treatment combination there are five relevant sessions with two blocks per session. p-values are from a Wilcoxon-test using each session as an observation.

We summarize the aggregate findings regarding policy approval rates as follows.

[12] Extending the sample to all periods does not qualitatively change our results.

Table 2: Summary of predictions and results

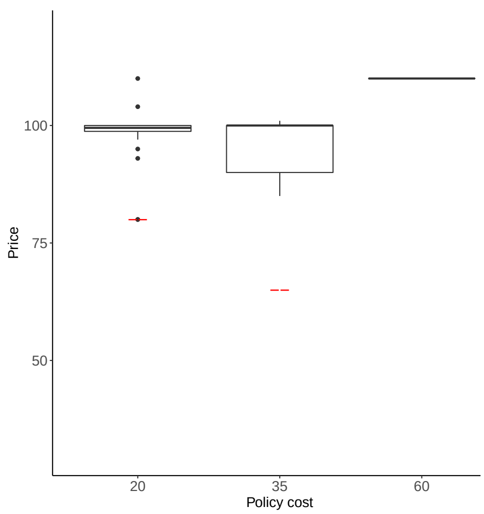
	Predictions		Results	
Treatment	Approve (A) / Reject (R)	Price	Approval rate (%)	Price (P)
Polarized preferences				
$\tau = 20$	A	80	70	95
$\tau = 35$	A	65	70	90
$\tau = 60$	R	100	0	97
Dispersed preferences				
$\tau = 20$	{A, R}	{80, 100}	80	100
$\tau = 35$	{A, R}	{65, 100}	28	98
$\tau = 60$	R	100	3	100

Note: Data are from 5 Polarized and 5 Dispersed sessions, and comprise the last four periods of each of two blocks per session with the given policy cost.

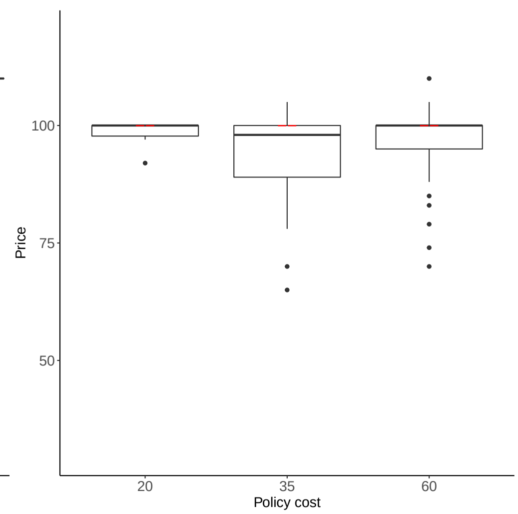
Result 1. *Consistent with Hypothesis 1, when the policy cost per share is high ($\tau = 60$), the policy is almost always rejected under both sets of preferences. When the cost per share is intermediate ($\tau = 35$), we find that approval rates are significantly greater in the polarized treatment where they average 70% as compared with the dispersed treatment where they average 28% (p-value=0.015). When the cost per share is low ($\tau = 20$), we find a similar, high rate of approval across both preference treatments.*

Of course, the equilibrium predictions are not perfect. Under polarized preferences, we see in Table 2 that the actual approval rate is indeed 0% for $\tau = 60$, but for $\tau \in \{20, 35\}$ the observed 70% approval rate falls well short of the predicted 100% rate. Recall that under polarized preferences, the default (i.e., no-trade) outcome at all three cost levels is rejection, since the $\lambda = 0$ investors are endowed with the majority of shares (6/11). Later we will investigate whether the approval rate shortfall is due to some $\lambda = 0$ investors who fight to control the policy outcome and are reluctant to sell their shares.

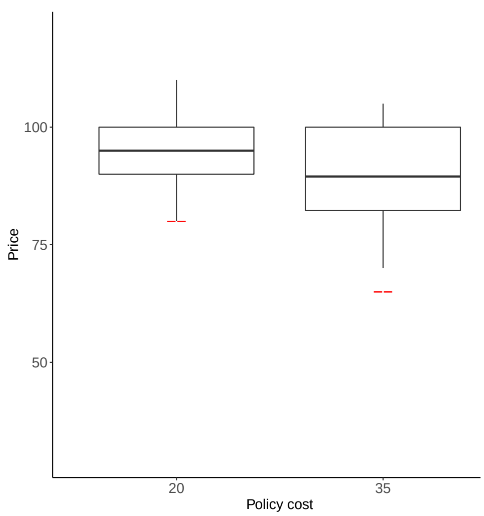
Under dispersed preferences, a low policy cost ($\tau = 20$) facilitates the highest observed approval rate (80%) in Table 2, even though rejection is also an equilibrium outcome here. As noted earlier in the paragraph following Hypothesis 1, an intermediate policy cost ($\tau = 35$) increases the difficulty of coordinating on approval equilibrium, and indeed the approval rate falls to 28%. At a higher cost ($\tau = 60$) there is no approval equilibrium and the observed approval rate (3%) is close to the predicted 0%. Thus the observed approval rates are roughly in line with the Hypothesis 1 predictions.



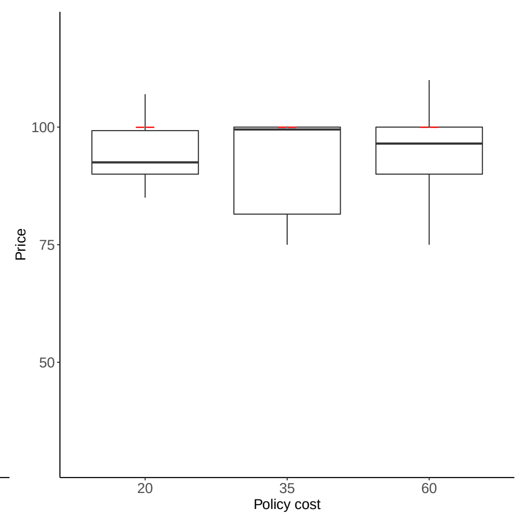
(a) Dispersed, Approve



(b) Dispersed, Reject



(c) Polarized, Approve



(d) Polarized, Reject

Figure 5: Prices under dispersed and polarized preferences by policy outcome

Note: Red dashes (– –) indicate the predicted price; in the dispersed case, the prediction depends on whether the policy is rejected or approved. Data are from the last 4 periods in each block. For $\tau = 60$, there is only 1 observation in panel (a) and none in panel (c) because, consistent with predictions, Approval outcomes are rare or nonexistent.

We now turn to the market clearing share prices P observed in the first stage Bid-only Call Market. Figure 5 shows Box plots of P for each treatment, disaggregated by whether the policy was accepted or rejected. Together with the price data in Table 2, the Figure supports the following:

Result 2. *Consistent with Hypothesis 2 when the policy is rejected, share prices P are very close to the predicted fundamental value $v = 100$. However, inconsistent with Hypothesis 2 when the policy is accepted, the price per share is well above its fundamental value, $v-\tau$. Overall, we cannot reject that prices are equal across treatment conditions.*

When the policy is costly, $\tau = 60$, Hypothesis 1 states that the policy will be rejected and Hypothesis 2 states that under rejection, the price is equal to 100, the fundamental value per share. Figure 5 panels b, d provide support for this price prediction. Also, with only one exception, the policy is indeed always rejected.

When $\tau = 35$ and preferences are dispersed, Hypothesis 1 allows for both acceptance and rejection equilibria. Figure 5 panels a, b show that most market share prices are close to 100 in this case, regardless of whether or not the policy is accepted. Here we find a low rate of acceptance (28%). In the polarized treatment with $\tau = 35$, acceptance is the unique equilibrium prediction and the median share price is a bit below 90 when the policy is accepted. Still, it is significantly higher than the predicted fundamental value of 65, and close to the predicted value of 100 when the policy is rejected. This price difference across policy outcomes is not statistically significant according to a conservative standard non-parametric test applied to median prices per session.

In the low cost $\tau = 20$ treatment, share prices are also above the fundamental value when the policy is usually accepted. When preferences are polarized and acceptance is the unique equilibrium prediction, the acceptance rate is 70%, and the median price is 95 when the policy is accepted. In the dispersed treatment, where rejection and acceptance equilibria coexist, the acceptance rate is 80% and the median price when the policy is accepted is 100. In both treatments, these prices are greater than the predicted price of 80 under acceptance. Further, there is again no significant difference in prices between policy outcomes when $\tau = 20$.

5.2 Individual results

To better understand the anomalous price behavior when the proposal is accepted, and also because it is interesting in its own right, we now examine individual voting behavior. Guided by Hypothesis 3, observed individual votes in Table 3 are separated according to whether or not the subject's induced policy preference parameter λ_i is strictly below the threshold $\lambda^*(x_i)$ in equation (2) given post trade holdings $x_i = 1$ or 2 that period. Table 3 provides support for the following:

Result 3. *Consistent with Hypothesis 3 and SPNE*, investors generally vote according to the optimal threshold strategies. Shareholders with values for λ_i greater than λ^* as defined in equation (2), vote to approve the policy 94% of the time on average when $\tau \in \{20, 35, 60\}$. Shareholders whose λ_i values are below the optimal threshold, generally vote to reject the policy (14% approval on average).*

Table 3: Approval rates and threshold conditions

	cost = 20		cost = 35		cost = 60	
	Approve (%)	Obs.	Approve (%)	Obs.	Approve (%)	Obs.
			Dispersed preferences			
$\lambda < \lambda^*$	0.11	97	0.12	146	0.16	225
$\lambda \geq \lambda^*$	0.95	152	0.87	101	0.95	22
			Polarized preferences			
$\lambda < \lambda^*$	0.19	104	0.15	122	0.12	221
$\lambda \geq \lambda^*$	0.94	143	0.94	124	1.00	22

Note: Data consist of all choices in all periods for subjects who hold one or two shares post trade; columns labeled Approve(%) report the fraction of subjects who voted Y for the given threshold condition; columns labeled Obs. report the number of such observations; the weakly dominant strategy is to vote Y iff $\lambda \geq \lambda^*$.

Table 3 shows that shareholders with preferences greater than or equal to the threshold value, $\lambda \geq \lambda^*$, vote as predicted at least 87% of the time for each cost level in the dispersed treatment, and at least 95% of the time in for each cost level in the polarized treatment. That is, high λ shareholders generally vote to approve the policy while low λ shareholders — those with $\lambda < \lambda^*$ — also generally vote as predicted, here to reject the ESG policy. They vote to reject at least 84% of the time across cost treatments with dispersed preferences, and at least 81% of the time with polarized preferences. We perform a t-test using each shareholder as an observation, and check whether subjects cast their votes according to the threshold condition. For both treatments, polarized and dispersed, the p-value is < 0.001 when comparing investors' decisions against an alternative hypothesis that the outcome is generated by chance.

Although deviations from threshold behavior are not especially frequent, voting to approve the policy when one should reject it is clearly more common than the opposite deviation. Perhaps the first post-hoc explanation that comes to mind is that some subjects have homegrown preferences for efficiency. Recall from equation (5) that acceptance outcomes in our experiment are socially efficient iff $\tau h \leq \frac{5}{11}100$. A preference for efficiency thus might explain the more frequent deviations towards approval that we observe in the $\tau = \{20, 35\}$ treatments but not in the $\tau = 60$ treatments.

We next address Hypothesis 4, which predicts how investors will react to policy outcomes that lower their payoff. Such adverse outcomes arise in the two distinct cases considered in the two columns of Table 4. Column (1) considers the case where the investor has a strong induced preference for policy acceptance (λ_i is at or above the optimal threshold even when the investor holds 2 shares) but the policy is rejected in a given period. We analyze an investor's bid adjustment ΔBid, defined as their bid in the current period minus their bid in the previous (adverse) period.

What should an investor do when she desires a policy that was just defeated? If she held no shares and so did not vote, then she may want to increase her bid next time to improve her chances of acquiring shares in order to vote to approve the policy, especially if the recent vote was close. If she held 2 shares (which she presumably voted Y) and the policy was nevertheless rejected, then she may become discouraged and reduce her bid next time.

To test such stories, we define the dummy variables Held None (=1 if the subject held no shares in the previous period) and Pivotal (=1 if there were 6 or 7 votes N). Table 4 reports OLS regressions of ΔBid on those dummy variables and their interaction. Observations for subjects who only held one share in the previous period are excluded because they are not common (about 12% of the total observations) and because we have no clear story of how they might behave. Observations for $\tau = 60$ are also excluded because an individual has very little scope to alter those voting outcomes (virtually always rejection). We pool data from the dispersed and polarized treatments because we see no reason why, controlling for voting outcomes, the policy preference distribution would affect bid adjustment.[13]

[13] As a robustness check we ran separate regressions for the polarized and dispersed session data. Those regressions, omitted for brevity, indeed produce similar coefficient estimates.

Table 4: ΔBid (OLS regression)

	(1) Policy rejected $\lambda \geq \lambda^*(2)$	(2) Policy accepted $\lambda < \lambda^*(2)$
Constant	-4.76	-27.00
	[0.484]	[0.000]
Held None	22.54	33.85
	[0.009]	[0.000]
Pivotal	-6.65	13.31
	[0.397]	[0.019]
Held None × Pivotal	3.87	-12.93
	[0.699]	[0.053]
Obs.	240	503
R^2	0.14	0.13

Notes: The dependent variable is the current period bid less the previous period bid. The constant captures the behavior of subjects who held two shares in the previous period. The regression includes dummy variables for holding zero shares and proximity to being pivotal, which occurs when 6 or 7 votes are cast for the undesired policy outcome. Fixed effects at the subject level are included. p-values are presented in brackets.

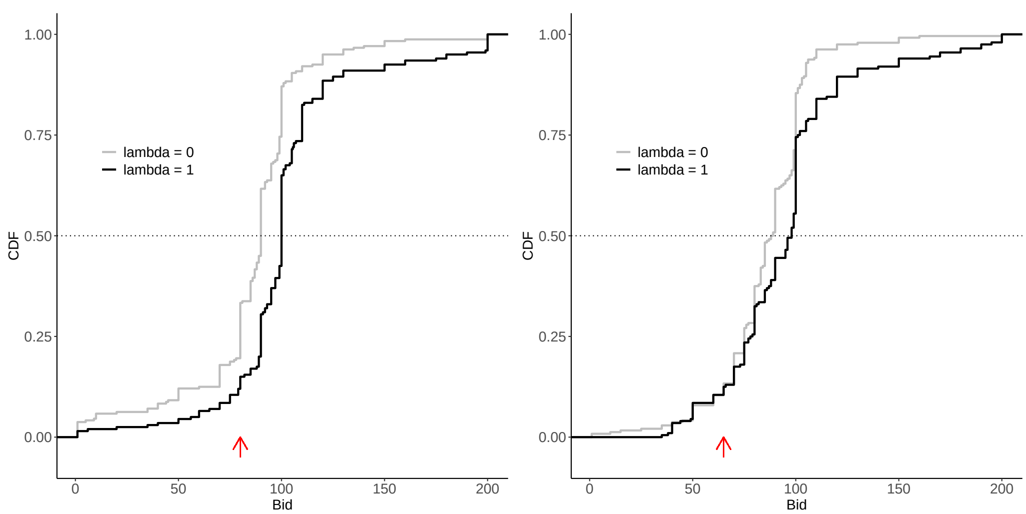
The coefficient estimates reported in column (1) are consistent with our stories. The Constant coefficient represents the bid adjustment of an investor holding 2 shares in the previous (adverse) period; it is indeed negative but nowhere near significant. The highly significant Held None coefficient confirms the story that a subject holding zero shares will tend to increase her bid the following period. The coefficients for the Pivotal dummy variable and its interaction with Held None are not significant here.

Column (2) reports estimates for the other adverse case, where an investor wants to see the policy rejected (at least when she holds two shares) but the policy is accepted in a given period. If she does hold 2 shares, then she might become discouraged about her influence on the vote outcome and become less interested in holding shares, and so reduce her bid next time. This story is supported by the large negative and highly significant Constant coefficient. As one might expect, that discouragement effect is weaker when the policy barely passes, as indicated by the significantly positive coefficient on the Pivotal dummy variable. On the other hand, disappointed subjects who held no shares increase their bids very substantially when the policy is approved, as indicated by the Held None coefficient estimate of almost 34. A close vote has little effect for such subjects: the coefficient on the Held None × Pivotal interaction variable almost completely offsets the Pivotal coefficient.

Result 4. *Consistent with Hypothesis 4, subjects who hold no shares and see an adverse policy outcome will tend to increase their bid prices in the following period. A close vote reinforces that tendency when the policy is accepted, but otherwise has little impact.*

Evidently investors do not take the policy outcome as a given: those whose additional votes might favorably alter their payoff tend to increase their bid prices significantly. This suggests that the price anomaly may be due in part to a fight for control over policy-making decisions. When the policy is likely to be approved, low-lambda subjects (who would prefer rejection) may be unwilling to sell their shares at the equilibrium price of $v - \tau$ and instead may raise their bids and push the price up in the hope of acquiring enough votes to change the policy outcome.

To follow up on this thought, let us focus on the distributions of individual bids in the polarized treatment with $\tau \in \{20, 35\}$, where acceptance of the policy is the unique equilibrium prediction and therefore likely. We focus on this treatment because policy acceptance is ambiguous in the dispersed treatment and acceptance almost never occurs when $\tau = 60$ in either treatment.



(a) cost = 20 **(b)** cost = 35

Figure 6: CDFs of bid prices under polarized preferences

Note: The empirical cumulative distribution function (CDF) is constructed using investors' bid prices in the last 4 periods of a block. The red arrows denote the equilibrium predicted prices, 80 when $\tau = 20$ and 65 when $\tau = 35$.

Figure 6 presents CDFs of investors' bids in the $\tau = 20$ and $\tau = 35$ treatments of the polarized case. We see that bids are higher (in the sense of first order stochastic dominance) for high lambda ($\lambda = 1$) subjects than for low lambda ($\lambda = 0$) subjects. We also observe that the low lambda subjects' median bid price is around 90 when $\tau \in \{20, 35\}$, which is much higher than the predicted bids of 65 and 80, respectively. A Wilcoxon test, using the median bid price at the subject level for individuals with $\lambda = 0$, confirms that bid prices for $\tau = 20$ are similar to those for $\tau = 35$. This result suggests that low-λ investors, who are natural sellers, are reluctant to lose control and/or are trying to extract higher prices from the high-λ investors who benefit from implementation of the policy. A further possibility is that low-λ voters simply under-appreciate their potential losses from bidding up prices in the low τ case where the policy is likely to be adopted.[14]

For completeness, we also analyze median bid prices for investors who benefit from the policy approval ($\lambda = 1$). The left panel of Figure 6 shows that when $\tau = 20$, the median bid for those with $\lambda = 1$ is 100 and is higher than the median bid for those with $\lambda = 0$ (p-value< 0.001 for a Wilcoxon test). For $\tau = 35$, the median bid prices are similar between the high and low lambda types (Wilcoxon p-value is 0.049).

Thus, we find that when the policy is likely to be approved (i.e., the polarized preference case with low policy costs), the natural sellers (those with $\lambda = 0$) submit higher than predicted bid prices that only partially reflect the reduction in the fundamental value due to the policy cost τ. In response, the natural buyers (those with $\lambda = 1$) also submit higher bids. This overbidding by both lambda types is the proximate cause of the market prices exceeding approval equilibrium predictions.[15]

On a tangentially related topic, details of theoretical proofs suggest negative assortativity of trades a and preferences λ for rejection equilibria, at least near the lower bound $\check{\tau}$, and positive assortativity for acceptance equilibria, at least near the upper bound $\hat{\tau}$. To investigate, we calculated the Spearman rank correlation for observed trades a and induced λ values. For the polarized case, the coefficients are 0.366 (-0.140) when the policy was accepted (rejected), while for the dispersed case, the coefficients are 0.199 (0.053) when the policy was accepted (rejected). These results indicate a small positive correlation when the policy is accepted, and no consistent correlation when the policy is rejected.

Another way to examine assortativity is to compare share holdings of high lambda

[14]This is reminiscent of Dal Bó et al. (2018)'s finding that voters' bad policy choices may be driven by their mistaken assessments of the equilibrium outcomes resulting from those policy choices.

[15]Table B.1 in the Appendix B shows that such overpricing of the asset favors subjects who sold their shares, yet the difference from baseline profit is not especially large.

subjects to those of low lambda investors. The relevant CDFs appear in Appendix B, Figure B.1. The median combined share holdings of the 5 high lambda investors (who can hold up to 10 shares in total) is 6 when the cost is 35 and 7 when the cost is 20. This evidence suggests that some of these high λ investors do not acquire a second share, which may indicate a type of free-riding behavior, since (at observed prices, above the approval equilibrium level) holding shares reduces their payoff. We also present the analysis for dispersed preferences. In that case, the median total shareholdings by the top 5 lambda investors is 4 when the cost is 35 and 7 when the cost is 20.

6 Discussion and Conclusion

Our paper makes two contributions. First, on the theoretical side, it develops a tractable model to help sort out issues regarding shareholder voting on a costly policy that, if approved, would mitigate negative externalities of the firm's operations and thus generate positive benefits to shareholders and non-shareholders alike. Voting rights in the model are endogenous: shares can be bought and sold before the vote takes place. Investors have heterogeneous preferences; typically some shareholders will favor the policy while others oppose it. Free riding can arise because share purchase increases a shareholder's policy cost burden, but does not increase their policy benefit.[16]

The model has approval equilibria for all policy costs below an upper bound $\hat{\tau}$, and rejection equilibria for all policy costs above a lower bound $\check{\tau}$. In many cases, including our dispersed preference example, there is a wide interval $(\check{\tau}, \hat{\tau}]$ of policy costs with both types of equilibria. In other cases, including our polarized preference example, the equilibrium type is unique. The price of shares in the secondary market (modelled as a single bid call market) is unique in each type of equilibrium, and equal to (cost-adjusted) fundamental value.

Our other contribution is empirical. As investors' preferences for and costs of ESG policies are difficult to measure, we induce chosen values in a laboratory experiment where we evaluate the main theoretical predictions. While the model is static, we examine how subjects adapt to repeated play and to the behavior of other subjects.

[16] As noted in Remark 4 of Section 3, the free rider problem disappears in approval equilibrium: it is the endowed shareholders who bear the costs of the policy, not those who acquire shares (cheaply) to vote it in. However, given the higher share prices observed in the experiment, the free rider problem recurs and those who acquire shares to vote Y typically incur negative trading profits when the ESG policy passes.

We find that approval and rejection of ESG policies in the laboratory largely follows equilibrium predictions. In the case of dispersed preferences, observed approval rates are decreasing in the policy cost, including the cost range that permits both types of equilibria. We also find that observed prices are in line with equilibrium predictions when the policy is rejected. However, when the policy is approved, prices are generally above the equilibrium prediction and not far from the fundamental value of the firm had it not adopted the policy. Overall, we conclude that ESG policy adoption in our laboratory experiment is less costly to incumbent shareholders than predicted. Our interpretation is that there is a voting premium not recognized in our model, arising from a struggle to control the voting outcome. To the extent that similar forces are at work in the wider world, this observation might help explain the recent shift towards greater consideration of ESG policies in the valuation of firms.

It may be worth mentioning that our model allows interpretations other than ESG policy adoption. A much different *anti*-social responsibility interpretation is that a policy proposal (e.g., to change auditors) may allow investors heterogeneous opportunities λ_i for personal benefit (e.g., sourcing from cronies, buying corporate jets, or siphoning funds) at a cost $n\tau$ to shareholders.

Future work building on our contributions could take several directions. On the theoretical side, our 2-period model could be extended in various ways. Perhaps the simplest would be to relax some of the restrictions we impose on admissible parameters in order to rule out monopolistic behavior. (Recall that the restrictions $e_i, a_i < n/4$ ensure that no investor i can acquire a majority of shares and thus unilaterally determine the voting outcome.) Another extension is to incorporate a third stage (after trading and voting) where multiple firms compete in the market. This could endogenize τ, the cost of an ESG policy, which the current model treats as an exogenous parameter. Theorists might also want to consider weaker notions of equilibrium than subgame perfection, which may help elucidate the observed voting premium. Although we don't expect it to greatly alter our results, one may wish to consider more standard trading formats such as call markets that allow separate bids and asks. More ambitiously, one could investigate a continuous double auction market format that allows trading divisible or multiple shares.

On the empirical side, it would be of interest to shed more light on the price anomaly we document. Future experiments can consider adding a belief elicitation stage, which we omit given the complexity of the tasks, or increase the number of repetitions to check whether coordination can be enhanced to a particular outcome. Another possible focus for future experiments is on "homegrown" preferences for ESG. We tightly controlled the distribution of preferences by assigning utility bonuses λ to individual subjects, using neutral language with no mention of ESG. Future work could build on our setup (as well as recent work mentioned in the literature review) to try to measure subjects' actual preferences for ESG policies and assess their impact on laboratory outcomes.

Beyond the laboratory, our findings might help inform new studies of archival data. For instance, we observe that subjects fight for voting control and adjust their bids in response to voting outcomes. Can this win-stay, lose-shift behavior also be observed in field data?

Appendix for Online Publication Only

A Proofs

For convenience, we reproduce here equations (1) and (2) from Section (3) of the text.

$$u_i = e_i v + a_i(v - P) + [\lambda_i h - (e_i + a_i)\tau]\xi. \tag{A.1}$$

$$\text{Vote Y} \iff \lambda_i \geq \lambda^*(x_i) = \frac{(e_i + a_i)\tau}{h}. \tag{A.2}$$

Proposition 1. *Given any admissible parameter vector* $(n, v, h, q_M, \mathbf{e}, \lambda)$,
(i) $\exists\ \hat{\tau} \in [0, v)$ *s.t. an approval equilibrium exists iff* $\tau \in [0, \hat{\tau}]$ *,*
(ii) $\exists\ \check{\tau} \in [0, v)$ *s.t. a rejection equilibrium exists iff* $\tau \in (\check{\tau}, v]$*,*
(iii) $\check{\tau} \leq \hat{\tau}$*.*

Proof. Take any allocation $\mathbf{x} = \mathbf{e} + \mathbf{a}$. Rearranging (A.2), we see that each investor i who can vote (i.e., with $x_i > 0$) will vote Y iff

$$\tau \leq \tau_i(\mathbf{x}) \equiv \frac{\lambda_i h}{x_i}. \tag{A.3}$$

Let $A(\tau, \mathbf{x}) = \sum x_i I[\tau \leq \tau_i(\mathbf{x})]$. Since A is left-continuous and decreases in τ from n at $\tau = 0$ to 0 at $\tau > \max_i \tau_i(\mathbf{x})$, there is a largest solution, call it $\tau_A(\mathbf{x})$, to $A(\tau, \mathbf{x}) \geq n/2$. Thus at least $n/2$ shares vote Y (and so $\xi = 1$) $\iff$ $\tau \in [0, \tau_A(\mathbf{x})]$.

Let $\hat{\tau} \equiv \sup\{\tau_A(\mathbf{x}) : \mathbf{x} \text{ is feasible}\}$. Of course, we can replace sup by max when the set of feasible final allocations is finite, as it is in Examples 2 and 3. Lemma 1 below guarantees that we are taking the sup (or max) over a non-empty set, so $\hat{\tau}$ is well defined. By construction, $\hat{\tau}$ satisfies part (i).

For part (ii), note that the $\iff$ characterization above tells us that $\xi = 0$ for a given allocation $\mathbf{x}$ whenever $\tau > \tau_A(\mathbf{x})$. Let $\hat{\tau} \equiv \inf\{\tau_A(\mathbf{x}) : \mathbf{x} \text{ is feasible}\}$. By construction, for any $\tau > \hat{\tau}$ there is some feasible allocation $\mathbf{x}$ such that voting according to (A.3) guarantees $\xi = 0$, i.e., rejection, while no such $\mathbf{x}$ exists when $\tau \leq \hat{\tau}$. Part (ii) follows. Part (iii) follows from the fact that the infimum of any set is smaller than its supremum. □

Lemma 1. For any admissible parameter vector, there is an acceptance equilibrium bid profile for some $\tau > 0$. Its equilibrium outcome includes a trade vector $\mathbf{a}$, with feasible final allocation $\mathbf{x} = \mathbf{a} + \mathbf{e}$, so $\{\tau_A(\mathbf{x}) : \mathbf{x} \text{ is feasible}\} \neq \emptyset$.

Proof. From the definition of admissibility, we see that $0 < \lambda_{n-1} \leq \lambda_n$. Pick some $\tau \in (0, \frac{\lambda_{n-1} h}{n})$. Then, even with their largest possible final allocation, investors n and $n-1$ prefer acceptance to rejection and will vote all their shares Y. With this in mind, it is routine to verify that there is no profitable deviation from the following bid profile: $b_n = b_{n-1} = v$ (or a bit less, but more than $v-\tau$), $q_n = q_M - e_n = \frac{n}{4} - e_n$, $q_{n-1} = \frac{n}{4} - e_{n-1}$, while for $i = 1, ..., n-2$ we set $b_i = v - \tau$ and $q_i = q_M = \frac{n}{4}$. That profile produces the outcome $P = v - \tau,\ a_n = q_n,\ a_{n-1} = q_{n-1},\ a_i = -\frac{0.5n - e_n - e_{n-1}}{n-2}$ for $i = 1, ..., n-2$, and feasible final allocation $\mathbf{x} = \mathbf{a} + \mathbf{e}$ with $x_n + x_{n-1} = n/2$. Thus, even without support from others, investors n and $n-1$ ensure that $\xi = 1$ according to equation (A.2). □

Remark A1. By convention, we say that an investor will vote Y when indifferent between acceptance and rejection of the policy, i.e., when (A.2) or (A.3) holds with equality. An an approval outcome would seem fragile if such indifferent investors are pivotal in maintaining it. Since investor i's indifference is broken by the slightest change in x_i and such a change can arise from a tiny change in i's bid or even in the bid of another investor, it might seem that this knife-edge case is of little practical interest. However, it arises naturally at the upper bound $\hat{\tau}$ of ESG costs that allow approval equilibrium. At any higher cost Proposition 1 guarantees a rejection equilibrium, and the convention on ties guarantees that all rejection equilibria are maintained by investors who (hold a majority of shares and) *strictly* prefer N. The main theoretical reason for mentioning knife edge approval equilibrium is to provide a minor but necessary qualification for the pricing result below, that $P = v - \tau$ in approval equilibrium. To make such matters more precise, we begin with the following:

Definition. Let $B(\mathbf{x}) = \{i : \ \lambda_i = \lambda^*(x_i)\}$ be the set of investors that are exactly indifferent to ξ.

Remark A2. According to equation (2), investors $i \in B(\mathbf{x})$ will vote Y. However, for almost all allocations $\mathbf{x} \in S \equiv \{\mathbf{y} \in \Re^n : y_i \geq 0 \ \wedge \ \sum_{i=1}^n y_i = n\}$, the set B is empty, i.e., it has cardinality $|B(\mathbf{x})| = 0$ except on a set of (relative) Lebesgue measure zero, i.e., generically there are no such investors.

Definition. An outcome $(\mathbf{a}, P, \xi)$ with final allocation $\mathbf{x} = \mathbf{e} + \mathbf{a}$ is *fragile* if (i) $|B(\mathbf{x})| \geq 1$ and (ii) $X_+ + \sum_{i \in B(\mathbf{x})} x_i \geq n/2 > X_+$, where X_+ is the sum of shares held by investors for which $\lambda_i > \lambda^*(x_i)$, i.e., those who strictly prefer $\xi = 1$. A SPNE* is *regular* if its outcome is not fragile.

Proposition 2. *For any given admissible parameter vector, let* $(P, \mathbf{a}, \xi)$ *be the outcome*

of a regular SPNE. Then $P = v - \tau\xi$. In particular,*
(i) $P = v - \tau$ in any regular approval equilibrium, and
(ii) $P = v$ in any rejection equilibrium.

Proof. Let $(b_i, q_i)_{i=1}^n$ be a SPNE* bid profile that supports $\xi = 0$ (rejection). Suppose, for the sake of contradiction, that $P = v + c$ for some $c > 0$. Then there must be some trader j with $b_j \geq P$ and $x_j > 0$ who votes N; otherwise the BOCM could not have produced that price (and outcome ξ). By (A.3), j will continue to vote N at $x'_j = x_j - \epsilon$ for sufficiently small $\epsilon > 0$. The following deviation is therefore profitable:

$$b'_j = v + 0.5c, \; q'_j = a_j - \epsilon. \tag{A.4}$$

To see this, note that the second term in (1) is negative for j at the original profile, but is less so (or is possibly positive) at the deviation, since the resulting $P' \in [v+0.5c, v+c]$ reduces the loss at the old x_j and the resulting a'_j strictly reduces the loss further by $P'\epsilon$. Finally, note that the deviation does not affect rejection: $\xi' = \xi = 0$ because the deviation to q'_j constrains a'_j so that j continues to vote N, while any counterparties (whose a'_i's offset j's change in net trade) are, if anything, more inclined to vote N. Hence the deviation only affects the second term in (1) and so is indeed profitable, a contradiction. We conclude that $P > v$ is not possible in a rejection equilibrium. A similar argument rules out $P < v$: one finds a profitable deviation for a seller that will increase P and/or decrease the amount sold, without altering $\xi = 0$.

To establish the result for acceptance equilibrium, suppose that we have a SPNE* bid profile that supports $\xi = 1$. If $P < v - \tau$, then (applying the logic of the previous paragraph with the appropriate changes in sign, etc.) we can find a beneficial deviation for some particular seller: it will decrease her loss (a negative second term in (1)) but not push the price above $v - \tau$ nor alter ξ. If $P > v - \tau$ then there again are profitable deviations as long as the outcome is not fragile. We provide more details for this case since it is the least intuitive and turns out to be the most empirically problematic.

Let $(b_i, q_i)_{i=1}^n$ be a regular SPNE* bid profile that supports $\xi = 1$ (acceptance), with price P and trade vector $\mathbf{a}$ resulting in final allocation $\mathbf{x} = \mathbf{e} + \mathbf{a}$. Suppose for sake of contradiction that $P > v - \tau$. Let $j = \arg\max\{x_i : \; \lambda_i < \lambda^*(x_i)\}$ be the investor that votes the most shares N. (If no investor with positive shares votes N, then anyone can unilaterally profitably deviate without overturning $\xi = 1$, so the arbitrage argument is straightforward.) Consider the unilateral deviation $b'_j = b_j - \epsilon$, with either $q'_j = a_j - \epsilon$

if $a_j > 0$ or else, in the (implausible) case that j is not a buyer, set $q'_j = |a_j| + \epsilon$. Then $x'_j < x_j$ in the resulting final allocation, so the second term in (1) increases. If $|B(\mathbf{x})| = 0$, no investor j will change her vote for $\epsilon > 0$ sufficiently small, and the sum X_- of N votes will not increase no matter which investors k absorb the impact of j's decrease in final holdings. Hence $\xi' = \xi = 1$ and the deviation is indeed profitable. If $|B(\mathbf{x})| > 0$ then the regularity of the equilibrium ensures that the indifferent agents $k \in B(\mathbf{x})$ can't overturn $\xi = 1$. Again the reallocation away from N-voter j will not change the vote of any of the non-indifferent voters for $\epsilon > 0$ sufficiently small, nor increase the sum of their N votes. Hence again $\xi' = 1$ and the deviator increases her payoff. This contradicts the supposition that $P > v - \tau$ is compatible with regular acceptance SPNE*. □

Remark A3. The qualification about regular acceptance equilibrium requires modification to apply to trade in indivisible shares. In the parametric setup of the experiment, a price $P > v - \tau$ is undermined by arbitrage when someone selling an extra (indivisible) share should believe that $\xi = 1$ will still prevail. Two alternative conditions suffice. Either (a) $\xi = 1$ is supported by more than $\frac{n+1}{2}$ votes Y,[17] or (b) none of the (exactly) $\frac{n+1}{2}$ votes Y would be altered by acquiring another share according to equation (2).[18]

Remark A4. The fear that $\xi = 1$ may flip to $\xi = 0$ if an investor tries to arbitrage is not, in our view, a plausible reason for the overpricing $P > v - \tau$ often observed in approval outcomes of our experiment. Condition (b) above always holds in the polarized $\tau = 20, 35$ cases. Condition (a) usually holds in the other cases with frequent approval outcomes. Moreover, when approval outcomes have predominated in previous rounds, it is implausible that investors will believe that selling (or not buying) a single share is so likely to cause the voting outcome to flip to $\xi = 0$ that they would be reluctant to sell at a price P much closer to v than to $v - \tau$.

[17] Recall that the experiment features an odd number of indivisible shares. If the Y vote is unanimous, it is not hard to see that profitable arbitrage will break the supposed equilibrium. Otherwise, let any N voter i deviate by selling her share (or not buying an additional share). This will decrease the N vote by 1 (or possibly by 2 if she bought another share and was close to indifference). At worst, the buyer j of the share could be pushed from 1 vote Y to 2 votes N so the net effect is at worst decreasing the Y vote by 1. Hence condition (a) suffices to ensure that $\xi = 1$ still prevails and that i indeed has a profitable deviation.

[18] Here the bare majority of Y votes are held by investors who either are already maxed out with $x_j = 2$ and thus will not change their vote, or by investors with sufficiently large λ_j that an extra share (x_j going from 1 to 2) will not change their vote. Thus condition (b) ensures that the number of Y votes does not decrease and so is also sufficient to ensure profitable arbitrage of $P > v - \tau$.

Example 1 details. We need to exhibit a SPNE* bid profile that produces outcomes $\xi = 1, \mathbf{a} = 0$ and $P = v - \tau$ for some $\tau > 0$. Let $\tau = 0.5\lambda_0$, and consider the bid profile $b_i = v - \tau = 1 - 0.5\lambda_0$ and $q_i = q_M$ for every investor i. Given these identical bids, it is easy to see that the BOCM produces outcomes $\mathbf{a} = 0$ and $P = v - \tau$. Since here $\tau < \tau_i(\mathbf{x}) = \lambda_0$, all investors will vote Y and the second stage outcome indeed will be $\xi = 1$. Note that unilateral deviations will affect neither P nor ξ and so will not increase payoff. Hence the given profile is indeed a SPNE* of the desired form.

We also need a SPNE* bid profile that results in $P = v = 1$ and $x_1 = x_2 = 3$ when $\tau < \frac{\lambda_0}{3}$. Consider $b_1 = b_2 > 1 = b_3 = ... = b_{11}$ with $q_1 = ... = q_{11} = 2$. Clearly this profile yields the desired rejection outcome. Given the constraint $q_i' \leq q_M = \frac{11}{4}$, no unilateral bid deviation can affect price. Any unilateral deviation that could affect ξ according to (2) would reduce the deviator's payoff at that price. Deviations that affect $\mathbf{a}$ but not P or ξ have no effect on payoff. Hence the profile is indeed SPNE*.

Example 2 details. Let $\tau = 0.5$ and consider the profile $b_7 = b_8 = b_9 > P = v - \tau = 0.5 = b_{10} = b_{11} = b_4 = b_5 = b_6 > b_1 = b_2 = b_3$. (Recall that we are working with the constraint that $q_i = 1$, so strategy profiles here are just bid price vectors.) Then $P = 0.5$ and $a_7 = a_8 = a_9 = 1,\ a_1 = a_2 = a_3 = -1$, while $a_i = 0$ for the other 5 investors. According to equation (2), investors 7-9 each will vote their two shares Y (barely) as will investors 10, 11 their single shares. Thus $\xi = 1$. It is easy to see that no unilateral deviation will affect P or ξ. Since $P = v - \tau$, changes in a_i have no effect on payoff in equation (1). Hence this profile is an acceptance equilibrium for $\tau = 0.5$. Approval equilibria are not possible for $\tau > 0.5$, since no one is willing to hold a second share. Thus, all equilibria are rejection equilibria for $\tau > 0.5$, i.e., $\hat{\tau} = \check{\tau} = 0.5$.

Example 3 details. The details for establishing that $\hat{\tau} = \frac{5}{11}$ are routine. To verify that $\check{\tau} = \frac{3}{22}$, first note that the bid profile $b_1 = b_2 = b_3 = b_4 > 1 = b_9 = b_{10} = b_{11} > b_5 = b_6 = b_7 = b_8$ generates the outcome noted in the text. Given any unilateral deviation, the price $P' = P = v = 1$ remains unchanged, and such a deviation can cause at most a 2 vote change. However, the 8-3 vote for rejection prevents such a change from altering the voting outcome, so $\xi' = \xi = 0$. As usual, this implies that payoff u_i' in (1) is also unchanged, and thus we indeed have a rejection equilibrium. Consequently $\check{\tau} \leq \frac{3}{22}$.

There are only a finite number of other possible allocations that could yield $\check{\tau}$ strictly less than $\frac{3}{22.}$. Here we will check only the most plausible sort of allocation, with $x_1 = x_2 = x_3 = 2$, which arises from a profile such as $b_1 = b_2 = b_3 > 1 = b_7 =$

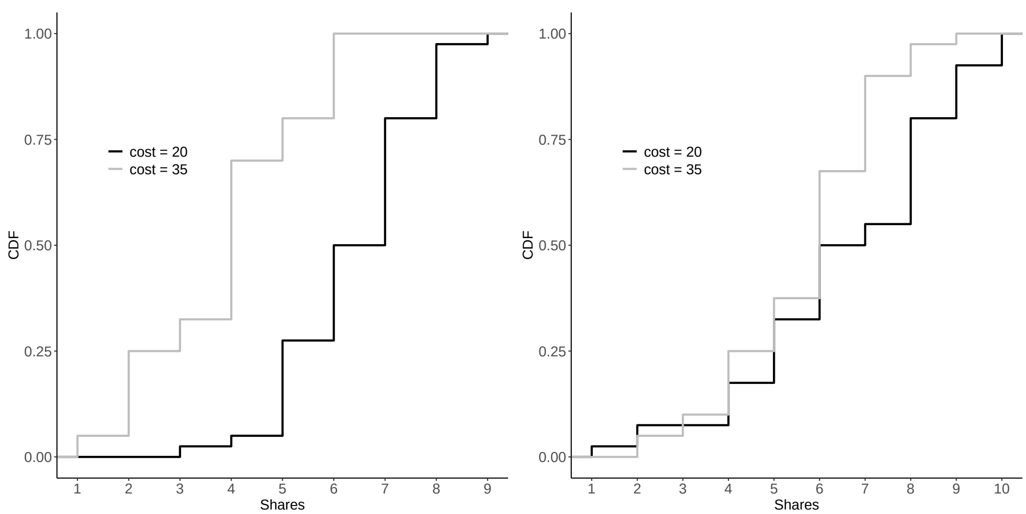
$b_8 = b_9 = b_{10} = b_{11} > b_4 = b_5 = b_6$. The allocation provides at least 6 Nay votes when $\tau > \tau_3(\mathbf{x}) = \frac{\lambda_3 h}{x_3} = \frac{2/11}{2} = \frac{1}{11}$, and so it might seem that $\check{\tau} = \frac{1}{11.}$ However, the supporting profile is not a rejection equilibrium because there is a profitable deviation: investor 3 can bid $b_3' < 1$, selling her share at price $P = 1$. Two other shares will end up in the hands of traders 4-11, who will vote Y and overturn the rejection for any $\tau < \frac{3}{22}$. Investor 3 then earns profits $u_3' = u_3 + \tau + \lambda_3 > u_3$, so the deviation is strictly profitable and breaks the proposed equilibrium.

B Additional data analysis

Table B.1: (mean) Observed Profit - Benchmark profit

	Approve		Reject	
Cost	Shareholder	Seller	Shareholder	Seller
	Dispersed preferences			
20	-11.69	20.69	-5.38	-7.66
35	-14.73	31.49	1.67	-6.94
60	–	–	2.80	-3.59
	Polarized preferences			
20	-11.76	14.68	–	–
35	-20.12	24.79	–	–
60	–	–	3.74	-4.62

Note: The benchmark profit replaces an investor's actual bid by her bid in the previous round. For dispersed preferences with cost < 60, it also assumes sincere voting by other actual shareholders. For other cases, the benchmark profit is the equilibrium profit prediction.



(a) Dispersed preferences **(b)** Polarized preferences

Figure B.1: CDF of share holdings for the top 5 λ-value subjects

Note: The CDF is constructed using the share final holdings per subject for those with the 5 highest value of λ. The maximum number of shares they can hold is 10.

C Experimental Instructions

Here we report the instructions for the Dispersed treatment. Note that λ_i from the model is represented here by w_i. The instructions for the Polarized treatment are similar, except for the values of w_i reported.

Instructions

Welcome to this experiment in decision-making. Each participant is guaranteed $10 for showing up and completing today's session. You can earn additional money based on the decisions that you make. Your total earnings will be paid to you in US dollars following the completion of today's session according to your chosen payment method. Kindly silence all electronic devices and do not talk with other participants during today's session. If you have any questions, or need assistance of any kind, please contact the experimenter.

General Information

This study consists of 48 separate rounds. Each round consists of 1) a trading stage which is followed by 2) a voting stage. In the trading stage of each round, you can trade shares of an asset. You start the trading stage with one share of the asset and with a cash loan. You can buy a second asset share using your cash, or sell the share for cash (or neither). Following completion of the trading stage you enter the voting stage. All participants holding one or more asset shares will vote on a policy. The number of votes you have is equal to the number of shares you hold at the end of the trading stage. The policy will be adopted if it wins a majority of votes. In that case, shareholders will pay a cost of c points per asset share held, but each market participant will gain a personalized benefit from adoption of the policy (even participants who vote against the policy or who do not hold any asset shares). The personalized benefit is greater for some participants than for others, as explained below. If the policy does not win a majority of votes then it is not adopted and so it incurs no costs and brings no benefits.

After playing all rounds, your final payment will be computed using a randomly selected round. The total points you earn in the selected round will be converted into your real money earnings at the rate of 100 points= $20.

Market Stage Details

Each market consists of 11 participants, who have the option of buying or selling asset shares. At the beginning of the market stage of each round you and all other participants will receive a loan of 200 points and 1 share of the asset. Notice that since there are 11 participants, there are a total of 11 shares of the asset in each round.

At the end of each market stage, you will receive 100 points for each asset share that you own after trading has taken place. If you bought an asset, your point total will be reduced by the market price that you paid for that asset, but if you sold your asset share, your point total will be increased by the market price at which you sold the asset. All participants repay their loan of 200 points at the end of the round.

Trading rules and market price To trade in the market, you need to submit a price for which you are indifferent between buying or selling one share of the asset. Your submitted price and the market price will determine whether you actually buy or sell a share of the asset. If the market price turns out to be greater than the price you submitted, then you will sell your single share of the asset (you want to sell when market prices are higher than your indifference price). If the market price turns out to be lower than the price you submitted, then you will buy another share of the asset (you want to buy when your indifference price is greater than the market price).

The market price, P, at which all trades take place is the **median** of all submitted indifference prices. The median price is the price for which 50% of all submitted prices lie above it and 50 of all submitted prices lie below it. If your indifference price lies above the market (median) price P, then you are a buyer and if it lies below the market (median) price then you are a seller. Note that you will end the trading stage of each round holding either 0, 1 or 2 shares of the asset.

Voting Stage Details

After trading is completed, participants holding shares will vote to approve or reject the policy. Each asset share provides one vote, so a participant who holds one share has one vote, a participant who has two shares has two votes, and a participant who sold their share does not vote.

If you are eligible to vote, you will see a screen like the one in Figure C.1 below.

After voting by the shareholders is completed, the outcome will be displayed to all participants.

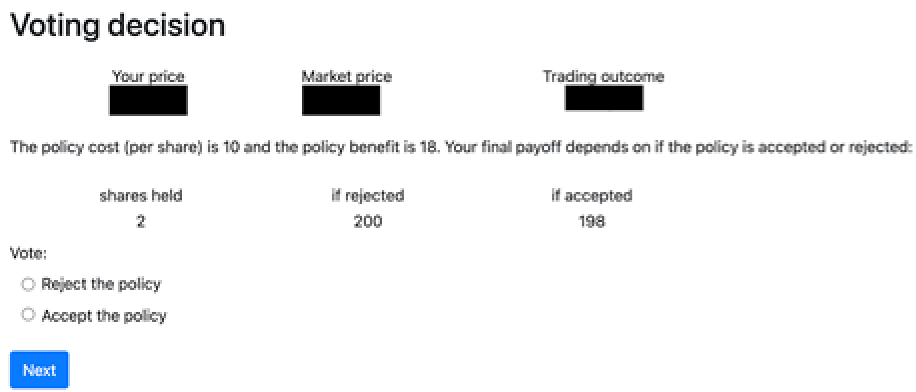


Figure C.1: Voting example (market information is blacked out). The policy cost per share is common for all participants in a given round but the policy benefit will vary across participants.

Calculating the benefit and cost

If the policy is approved, it will cost each shareholder **c** points per share held. The cost **c** will be announced each round and is the same for all shareholders. The policy will also generate a personalized benefit to each participant. That benefit ranges from 0 to 100 points and depends on each investor i's value weighting of the policy, $\mathbf{w_i}$, and the overall policy benefit 100, but it does not depend on how you voted or how many shares you hold. Thus, if the policy is approved, then you will receive the personalized benefit $\mathbf{100} * \mathbf{w_i}$ where your value weight $\mathbf{w_i}$ could be 0, 1/11,...,10/11. A participant with $w_i = 0$ does not receive any benefit from the policy being implemented. You will be told your personal policy benefit each round $(100 * w_i)$. Be sure to check your policy benefit and the policy cost, because $\mathbf{w_i}$ and **c** can change from one round to the next. To summarize:

- If the policy is rejected, a shareholder receives 100 points per share held.
- If the policy is accepted, a shareholder receives 100 points per share held minus the policy cost c per share plus the personalized benefit, as shown in Table C.1 below.

Recall that only those with shares after the trading stage can vote, and one share gives the right to one vote. Since shareholders can hold up to 2 shares, this means that one shareholder can have at most 2 votes. The policy is passed according to simple majority rule: the policy is approved if it receives at least 6 votes to accept,

Table C.1: Payoff calculation

Number of shares held	Policy is rejected	Policy is accepted	Profit from Trade
0	0	100*w_i	+P
1	100	$100 - c + 100 * w_i$	0
2	200	$200 - 2 * c + 100 * w_i$	−P

Note: Participants keep the cash earned from selling shares, and must repay the cash loaned (200).

and rejected otherwise.

Your final payoff each round is determined according to Table 1 which shows the profit according to the policy status (rejected or accepted) and your profit from the market trading stage. Note that if you sold an asset at the market price P, you held 0 shares and the profit is +P after repaying the loan of 200. If you do not sell or buy an asset, then the profit from market trading is zero; and if you bought a share, you hold 2 assets, and the profit from market trading is −P. Thus, your final payoff in points for a round depends on the number of shares you end up holding (first column in Table 1) and consists of the profit according to the policy status (second or third column in Table 1) plus your profit from market activity (fourth column in Table 1).

Recall that the computer program will randomly select one round from all 48 rounds played to be paid out to you. Your payment will equal your points for the chosen round converted into dollars at the rate of 100 points=$20, or 1 point=20 cents.

Are there any questions before we move on to the quiz?

FAQ

Q1. What happens if my submitted price turns out to be exactly equal to the market price P?

A1. If nobody else submitted exactly the same price as you, then in this case you will neither buy nor sell; you just keep the one share of the asset that you had before trading began. If two or more traders submit the same price, and if that price turns out to be the market price P, then some of them may turn out to be buyers or sellers. That should be OK for everyone, since the submitted price is the one that makes you indifferent between buying or selling (or neither).

Quiz

The correct answer per question is highlighted.

1. If you submit a price of 70 and the market clearing price is 60 you will
   (a) **Buy a share**
   (b) Sell a share
   (c) Not transact
2. You hold 2 shares. Your weight w_i is 0.7. 100 is the benefit amount from the policy and c is the cost of the policy. What is your net benefit if the policy is approved?
   (a) 2*100 - 0.7*c
   (b) 0.7*100 - c
   (c) **0.7*100 - 2*c**
3. If you sell an asset then
   (a) Your cash holdings will decrease by the market price
   (b) **Your cash holdings will increase by the market price**
   (c) Your cash holdings will increase by your submitted price
4. If a majority vote to approve the proposal then what is the net payoff for those who hold one share?
   (a) **100 - c + $\mathbf{w_i}$*100**
   (b) 2*(100 - c)
   (c) 100
5. If a majority vote to reject the proposal then what is the share payout for those that hold two shares?
   (a) $200 + w_i * 100$,

(b) 2*(100 - c)

(c) **200**

References

Amel-Zadeh, Amir and George Serafeim, "Why and how investors use ESG information: Evidence from a global survey," *Financial Analysts Journal*, 2018, *74* (3), 87–103.

Baker, Malcolm, Mark L Egan, and Suproteem K Sarkar, "How do investors value ESG?," 2022. National Bureau of Economic Research Working Paper No. 30708.

Barber, Brad M, Adair Morse, and Ayako Yasuda, "Impact investing," *Journal of Financial Economics*, 2021, *139* (1), 162–185.

Bauer, Rob, Tobias Ruof, and Paul Smeets, "Get real! Individuals prefer more sustainable investments," *Review of Financial Studies*, 2021, *34* (8), 3976–4043.

Bonnefon, Jean-François, Augustin Landier, Parinitha R Sastry, and David Thesmar, "The moral preferences of investors: Experimental evidence," 2022. National Bureau of Economic Research Working Paper No. 29647.

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales, "Exit vs. Voice," *Journal of Political Economy*, 2022, *130* (12), 3101–3145.

Casella, Alessandra, Aniol Llorente-Saguer, and Thomas R Palfrey, "Competitive equilibrium in markets for votes," *Journal of Political Economy*, 2012, *120* (4), 593–658.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Dal Bó, Ernesto, Pedro Dal Bó, and Erik Eyster, "The demand for bad policy when voters underappreciate equilibrium effects," *Review of Economic Studies*, 2018, *85* (2), 964–998.

Dasgupta, Amil, Vyacheslav Fos, Zacharias Sautner et al., "Institutional investors and corporate governance," *Foundations and Trends® in Finance*, 2021, *12* (4), 276–394.

Demichelis, Stefano and Klaus Ritzberger, "A general equilibrium analysis of corporate control and the stock market," *Economic Theory*, 2011, *46*, 221–254.

Dittmann, Ingolf, Dorothea Kübler, Ernst Maug, and Lydia Mechtenberg, "Why votes have value: Instrumental voting with overconfidence and overestimation of others' errors," *Games and Economic Behavior*, 2014, *84*, 17–38.

Fos, Vyacheslav and Clifford G Holderness, "The distribution of voting rights to shareholders," *Journal of Financial and Quantitative Analysis*, 2023, *58* (5), 1878–1910.

Friedman, Milton, "A Friedman doctrine– The Social Responsibility of Business Is to Increase Its Profits," *New York Times*, 13 Sep 1970. Available at: https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html.

Gollier, Christian and Sébastien Pouget, "Investment strategies and corporate behaviour with socially responsible investors: A theory of active ownership," *Economica*, 2022, *89* (356), 997–1023.

Guenster, Nadja, Daniel Brodback, Sébastien Pouget, and Ruichen Wang, "The valuation of corporate social responsibility: A willingness-to-pay experiment," 2022. Available at SSRN.

Hart, Oliver D and Luigi Zingales, "The New Corporate Governance," 2022. National Bureau of Economic Research Working Paper No. 29975.

Heeb, Florian, Julian F Kölbel, Falko Paetzold, and Stefan Zeisberger, "Do investors care about impact?," *Review of Financial Studies*, 2023, *36* (5), 1737–1787.

Hong, Harrison and Edward Shore, "Corporate social responsibility," *Annual Review of Financial Economics*, 2023, *15* (1), 327–350.

Humphrey, Jacquelyn, Shimon Kogan, Jacob Sagi, and Laura Starks, "The asymmetry in responsible investing preferences," 2021. National Bureau of Economic Research Working Paper No. 29288.

Kaufmann, Marc, Peter Andre, and Botond Kőszegi, "Understanding markets with socially responsible consumers," *The Quarterly Journal of Economics*, 2024, pp. 1–47.

Levit, Doron, Nadya Malenko, and Ernst Maug, "Trading and shareholder democracy," *Journal of Finance*, 2022.

_ , _ , and _ , "The voting premium," 2023. National Bureau of Economic Research Working Paper no. 31892.

Meirowitz, Adam and Shaoting Pi, "Voting and trading: The shareholder's dilemma," *Journal of Financial Economics*, 2022, *146* (3), 1073–1096.

Riedl, Arno and Paul Smeets, "Why do investors hold socially responsible mutual funds?," *Journal of Finance*, 2017, *72* (6), 2505–2550.

Smith, Vernon L, "Microeconomic systems as an experimental science," *American Economic Review*, 1982, *72* (5), 923–955.

Tsakas, Nikolas, Dimitrios Xefteris, and Nicholas Ziros, "Vote trading in power-sharing systems: A laboratory investigation," *Economic Journal*, 2021, *131* (636), 1849–1882.

Yermack, David, "Shareholder voting and corporate governance," *Annual Review of Financial Economics*, 2010, *2* (1), 103–125.