european corporate governance institute

ecgi

The Political Consequences of Stakeholder Governance

Finance Working Paper N° 1173/2026
November 2026

Doron Levit
University of Washington and ECGI
Markus Parlasca
WU Vienna and VGSF
Paul Voss
HEC Paris

© Doron Levit, Markus Parlasca and Paul Voss 2026. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
https://doi.org/10.67929/ecgi-fin-1173-2026

european corporate governance institute

ECGI Working Paper Series in Finance

The Political Consequences of Stakeholder Governance

Working Paper N° 1173/2026
November 2026

Doron Levit
Markus Parlasca
Paul Voss

© Doron Levit, Markus Parlasca and Paul Voss 2026. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Abstract

We study the political economy of stakeholder governance and corporate lobbying. Lobbying shifts policy toward outcomes that increase the firm's profitability, but it also provokes a backlash from advocacy groups, which push policy back toward the median citizen's preferred outcome. When lobbying is opaque, this backlash generates a ratchet effect: lobbying and backlash escalate, fueling political polarization and instability with little or no policy gain for the firm. To buy quiet, firms adopt stakeholder governance, delegating lobbying to a manager who places less weight on profits and more on citizens' welfare and therefore lobbies less. When the conflict between the firm and society is small, stakeholder governance leads to more socially responsible policies without sacrificing the firm's profits. Paradoxically, when the conflict is large, delegating to a more socially minded manager benefits the firm at society's expense, producing more extreme, socially harmful policies, even when the delegation creates no new agency problems. Our results suggest that stakeholder governance can have unintended political consequences and offer novel implications for corporate governance and lobbying regulation.

Keywords: Stakeholder governance, corporate lobbying, opacity, backlash, delegation, corporate social responsibility, lobbying disclosure

JEL Classification: D72, D82, G34, G38

Doron Levit
University of Washington and ECGI
ORCID 0000-0001-6139-1314

Markus Parlasca
WU Vienna and VGSF
ORCID 0009-0002-1097-2535

Paul Voss
HEC Paris

The Political Consequences of Stakeholder Governance*

Doron Levit† **Markus Parlasca**‡ **Paul Voss**§

November 19, 2026

Abstract

We study the political economy of stakeholder governance and corporate lobbying. Lobbying shifts policy toward outcomes that increase the firm's profitability, but it also provokes a backlash from advocacy groups, which push policy back toward the median citizen's preferred outcome. When lobbying is opaque, this backlash generates a ratchet effect: lobbying and backlash escalate, fueling political polarization and instability with little or no policy gain for the firm. To buy quiet, firms adopt stakeholder governance, delegating lobbying to a manager who places less weight on profits and more on citizens' welfare and therefore lobbies less. When the conflict between the firm and society is small, stakeholder governance leads to more socially responsible policies without sacrificing the firm's profits. Paradoxically, when the conflict is large, delegating to a more socially minded manager benefits the firm at society's expense, producing more extreme, socially harmful policies, even when the delegation creates no new agency problems. Our results suggest that stakeholder governance can have unintended political consequences and offer novel implications for corporate governance and lobbying regulation.

Keywords: Stakeholder governance, corporate lobbying, opacity, backlash, delegation, corporate social responsibility, lobbying disclosure.

JEL classifications: D72, D82, G34, G38.

*We thank John Matsusaka and participants at the 2026 Stigler Center–CEPR Political Economy of Finance Conference for helpful comments.

†University of Washington. Email: dlevit@uw.edu.

‡WU Vienna, VGSF. Email: markus.parlasca@wu.ac.at.

§HEC Paris. Email: voss@hec.fr.

1 Introduction

Firms and citizens often prefer different policies. While firms typically favor policies that increase profits, citizens also weigh their broader societal implications, such as their effects on the environment, wages, safety, and wealth redistribution. Firms therefore have an incentive to exert political pressure to move policy toward their preferred outcome (Stigler, 1971), most prominently through lobbying. Indeed, firms spend billions of dollars annually on lobbyists, campaign contributions, and grassroots organizations (Bombardini and Trebbi, 2020).

Corporate pressure, however, rarely goes unanswered. Organizations representing citizens mobilize to counter this pressure (Becker, 1983). Environmental groups such as Greenpeace campaign to make policymakers account for the environmental effects of policy, often in opposition to corporate interests;[1] trade unions support political parties to advance labor-friendly legislation;[2] and grassroots movements form to oppose externalities created by firms, as with local opposition to AI data centers in 2026.[3] Lobbying thus sets off a political contest whose outcome depends on how hard firms push and how hard citizens push back.[4]

More broadly, concerns that corporate political influence threatens democracy and the functioning of the economy (Zingales, 2017) have fueled calls for stakeholder governance, which directs firms' decision makers to weigh the interests of employees, communities, and society at

[1] A canonical example is Greenpeace's 1995 campaign against Shell's plan to sink the Brent Spar oil platform in the North Atlantic, which led Shell to abandon the plan and shifted government policy on offshore disposal (Jordan, 1998; Huxham and Sumner, 1999). More systematically, Lenox and Eesley (2009) study environmental activist campaigns against US firms and Brulle (2018) documents environmental groups lobbying against fossil-fuel interests on climate policy.

[2] Unions can act as a counterweight to business influence. Feigenbaum et al. (2018) show that unions raise Democratic vote shares through contributions and voter mobilization, and that weakening them moves state policy in a more conservative direction. Becher et al. (2018) find that members of Congress from districts with more union members vote more often for union positions; see Ahlquist (2017) for a review. Most recently, the AFL-CIO has been pressing state legislatures for a worker-focused AI agenda to counter the AI industry's lobbying ("America Isn't Ready for What AI Will Do to Jobs," *The Atlantic*, 2026).

[3] In an October 2026 poll, 76% of Americans said the government is doing too little to regulate AI and 78% said the same of data centers, including majorities of Republicans ("Ahead of the Midterms, Most Americans Want More Regulation of AI and Data Centers," USC, 2026); see also "Nearly two-thirds of Americans oppose data center construction in their community" (Economist/YouGov poll, 2026). Of 47 local data center ballot measures decided in 2026, 46 limited development ("Data center-related ballot measures, 2026," Ballotpedia), and New York adopted the first statewide pause on hyperscale data centers ("The Data-Center Rebellion Is Just Beginning," *The Atlantic*, 2026).

[4] Organizations' leverage in this contest rests largely on their ability to mobilize like-minded voters rather than on financial contributions (Urvoy, 2025). Counterpressure can also come from firms whose interests happen to align with those of citizens, such as clean energy producers opposing fossil-fuel interests (e.g., Kennard, 2020); see also "Lobbyists spent millions to save green energy. Wins were few" (E&E News, 2025).

large alongside those of shareholders. As former Delaware Chief Justice Leo E. Strine Jr. and Nicholas Walter put it:

> [P]ost-*Citizens United*, companies ... will be more able to 'capture' regulators so as to bend the rules of the game in their own favor. As a result, the [stakeholder] argument that corporations should have to consider the best interests of all corporate constituencies ... is strengthened. (Strine and Walter, 2015, p. 390)

These calls can be answered through law, as the proposed Accountable Capitalism Act would do for large U.S. corporations (Warren, 2024), or through voluntary choice, as when firms incorporate as public benefit corporations with a stated public purpose (e.g., Anthropic, PBC, 2026).[5]

In this paper, we study the political consequences of stakeholder governance. We ask whether stakeholder governance leads to more socially desirable public policy. We show that the answer hinges on two factors: the opacity of corporate lobbying and the extent of conflict between the firm and society. When lobbying is transparent, or when the conflict is small, the answer is yes. Perhaps surprisingly, when lobbying is opaque and the conflict is large, stakeholder governance backfires, benefiting the firm and harming citizens although the firm lobbies less. This harm arises even when the firm's decision maker is genuinely more socially responsible, and thus is not a matter of "social washing" or new agency conflicts. Our results therefore call for caution in evaluating stakeholder governance without regard to how information asymmetries shape political competition.

We model the market for political influence as a contest between a firm and an advocacy group. A policymaker whose ideal policy coincides with the median citizen's can be pulled away from it by political pressure. A representative firm exerts costly lobbying effort to generate political pressure which shifts policy toward its own ideal point. A representative advocacy group (e.g., an environmental group, trade union, NGO, or grassroots movement) exerts costly effort to generate counterpressure that pulls policy back toward the median citizen's ideal point. We assume lobbying is opaque, which reflects how lobbying works in practice. Most forms of lobbying are not publicly observable, and disclosure, when it occurs, is often incomplete or

[5]Other mandated forms include codetermination, which gives workers seats on corporate boards. Other voluntary forms include tying executive pay to ESG metrics (Cohen et al., 2023) and CEO pledges to serve all stakeholders (Bebchuk and Tallarita, 2020).

delayed.[6] The advocacy group therefore acts on its rational expectation of lobbying. To capture stakeholder governance, we allow the firm to delegate the lobbying decision to a manager whose preferences may differ from its own. Unlike lobbying, the firm's choice of the manager is publicly observed.

A key ingredient of the model is that the effectiveness of the advocacy group's counterpressure increases with the perceived threat of corporate influence, that is, with expected lobbying. Intuitively, higher expected lobbying motivates the group's members, who realize that absent a response, policy would shift against them. As a result, a given effort by the group generates more counterpressure. This mechanism is well documented. Threats to existing benefits are a central motive for joining and supporting interest groups (Hansen, 1985), and citizens participate more when they perceive that policy may move against them (Campbell, 2003).[7]

As a benchmark, we consider transparent lobbying. When lobbying is transparent, the firm internalizes the counterpressure it provokes and therefore chooses moderate lobbying. At low levels, additional lobbying is cheap and shifts policy effectively, because it triggers only a small increase in counterpressure. At higher levels, additional lobbying becomes more costly and provokes a stronger response from the advocacy group, which dampens its effect on policy. In this benchmark, stakeholder governance always leads to more socially responsible policy: delegating lobbying to a more socially minded manager reduces lobbying and counterpressure, and moves policy toward the median citizen's ideal, at the expense of firm value. Citizens therefore always benefit from stakeholder governance, but the firm never adopts it voluntarily.

Opacity changes these conclusions. Under opacity, the advocacy group's expectation of lobbying is already fixed when the firm lobbies, so the firm no longer internalizes the counterpressure its lobbying provokes and lobbies more than under transparency. The advocacy group

[6] By some estimates, roughly half of all lobbyists in Washington do not report their lobbying at all (Thomas and LaPira, 2017). Similarly, about 40% of firms whose lobbyists visited the IRS are not named in any lobbying disclosure filing (Nagaraj, 2026). Firms can also exert influence through channels that fall entirely outside disclosure requirements, such as charitable giving, whose politically motivated component Bertrand et al. (2020) estimate at about 35% of federal lobbying expenditures, or outright corruption.

[7] Citizens also respond to corporate political activity directly: learning about firms' political donations shifts consumers' behavior toward or against those firms (Panagopoulos et al., 2020). Occupy Wall Street, which emerged in 2011 in reaction to the bank bailouts and the perceived capture of policy by financial interests, is a prominent example of mass mobilization against corporate influence (Calhoun, 2013). More broadly, in the United States, hostility toward the opposing party has become a stronger predictor of turnout than positive attachment to one's own party, suggesting that mobilization is largely reactive to perceived threats (Abramowitz and Webster, 2016).

rationally anticipates this and responds with stronger counterpressure, which in turn strengthens the firm's incentive to lobby. This feedback loop gives rise to a *lobbying ratchet effect*. In equilibrium, both lobbying and counterpressure exceed their levels under transparency, so the heightened lobbying does not result in a commensurate policy gain for the firm, which would prefer to commit to less lobbying. Opacity thus fuels political polarization as a byproduct of the commitment failure it induces: both sides of the political contest exert more pressure than under transparency.[8]

When the conflict of interest between the firm and citizens is small, lobbying remains effective at the margin despite the ratchet effect: an additional unit of lobbying generates more pressure than the counterpressure it provokes. When the conflict is sufficiently large, however, the ratchet effect becomes so strong that lobbying is self-defeating at the margin. That is, the counterpressure provoked by an additional unit of lobbying exceeds the pressure it generates. If the firm could influence the advocacy group's expectations, it would lobby less. But, because those expectations are fixed when the firm lobbies, it is ex post optimal for the firm to lobby intensely even though the backlash dominates.

This commitment problem creates a role for stakeholder governance. Unlike in the transparent case, under opacity the firm voluntarily appoints a manager whose preferences are closer to those of the median citizen, because doing so allows it to commit to less lobbying. Lower expected lobbying induces the advocacy group to scale back its efforts, which reduces counterpressure and lowers the lobbying needed to achieve a given policy. Stakeholder governance thus reverses the ratchet effect: less lobbying leads to less counterpressure, which in turn leads to less lobbying.

Does stakeholder governance make public policy more socially responsible? Not necessarily. A natural conjecture is that a decision maker whose preferences are closer to citizens' delivers a policy closer to their ideal. This conjecture holds under transparency but can fail under opacity. While a more socially minded manager lobbies less, the resulting policy can be *more* extreme, leaving the median citizen worse off even though the manager's preferences are more closely aligned with hers. We call this the *delegation paradox*. Delegation lowers expected

[8]The same feedback loop can also make expectations of lobbying self-fulfilling. If the advocacy group expects intense lobbying, it mobilizes strongly, which in turn makes intense lobbying optimal for the firm; if it expects little lobbying, it mobilizes weakly, and the firm lobbies little. Such multiple equilibria can generate political instability, with episodes of intense lobbying and counterpressure alternating with episodes of political calm.

lobbying, and with it the threat the advocacy group perceives, making the group's effort less effective at generating counterpressure. The group therefore scales back, and when lobbying is self-defeating at the margin, the resulting drop in counterpressure exceeds the drop in corporate pressure. The paradox thus arises when the conflict of interest between the firm and citizens is sufficiently large.

Our results highlight how stakeholder governance affects both the efficiency and the division of surplus in the market for political influence. Under transparency, firms never adopt stakeholder governance voluntarily and regulation that mandates it would benefit citizens at the firm's expense. Under opacity, firms may adopt it voluntarily, with very different consequences depending on the extent of conflict. When the conflict is small, stakeholder governance is Pareto improving: citizens obtain less extreme policy, and the firm's savings on lobbying outweigh its policy loss. Thus, contrary to the transparent benchmark, citizens' gains need not come at the firm's expense. However, when the conflict of interest is sufficiently large, the firm optimally appoints a more socially minded manager who delivers more extreme policy and leaves citizens strictly worse off, i.e., the firm optimally appoints a manager in the region of the delegation paradox. The very feature that makes stakeholder governance privately attractive, its ability to quiet the opposition, is what makes it socially costly.

These welfare effects have broader implications for the regulation of political influence. First, the effect of *mandated* stakeholder governance on policy is non-monotone in how strongly the mandate moderates the firm's decision maker. When lobbying is not yet self-defeating at the margin, any such mandate moves policy toward citizens. When lobbying is already self-defeating at the margin, however, a mild mandate moves policy away from citizens, and only a sufficiently stringent mandate moves it closer to them. The required stringency increases with the conflict of interest between the firm and citizens. Stakeholder mandates are therefore most likely to backfire precisely where they appear most needed: in firms whose political objectives lie furthest from citizens'. Moreover, whenever a mild mandate backfires, the firm strictly gains from it, so the absence of corporate resistance to a stakeholder mandate is not evidence that it serves citizens.

Second, the same logic extends beyond governance to other forces that restrain lobbying without making it observable. When lobbying is self-defeating at the margin, a reform that raises the cost of lobbying reduces lobbying but moves policy toward the firm's ideal one. The

insight also sheds light on why firms in some industries coordinate their lobbying while firms in others lobby individually. When several firms lobby for the same policy, free-riding among them acts like a higher cost of lobbying, which coordination removes. When lobbying is self-defeating at the margin, firms are therefore better off not coordinating: free-riding restrains lobbying and reduces counterpressure by more than it reduces the industry's influence, yielding a more industry-friendly policy at a lower cost.

Beyond mandates, our analysis also offers a novel rationale for the *voluntary* adoption of stakeholder governance by purely profit-maximizing firms, one that arises only when lobbying is opaque. It predicts that stakeholder orientation is most pronounced where backlash is strong relative to the influence lobbying buys, as in politically contested environments. It also predicts that firms with very different political objectives appoint observationally similar managers, so the governance a firm adopts reveals little about its political objectives. For policy, the key implication is that voluntary adoption is no substitute for public intervention: it benefits citizens when the conflict is small but harms them when the conflict is large.

Finally, our analysis clarifies how stakeholder governance interacts with the most direct form of public intervention: lobbying disclosure. For the firm, transparency and voluntary stakeholder governance are substitutes: both eliminate the ratchet effect and yield the same policy. Consequently, mandating transparency adds nothing for citizens when firms already adjust their governance optimally. When they do not, it delivers exactly what optimal delegation would: it reduces both lobbying and counterpressure, and it benefits citizens when the conflict of interest is small but harms them when it is large. From citizens' perspective, however, transparency and stakeholder mandates are complements. Under transparency, a more socially minded decision maker always moves policy toward citizens, so stakeholder mandates work as intended, albeit at the expense of firm value. Under opacity, the same mandates can backfire. Disclosure requirements and stakeholder mandates should therefore be evaluated jointly.

Related Literature. Our paper contributes to the literature on corporate lobbying by introducing an informational friction, the unobservability of lobbying. As in Grossman and Helpman (1994), lobbying shifts the policymaker's ideal point. Whereas lobbyists in their model commit to policy-contingent contributions, our firm cannot commit and its lobbying is unobservable. This friction also differs from the one in the literature on informational lobbying, in which in-

terest groups hold private information about a policy-relevant state (see, e.g., Potters and van Winden, 1992; Austen-Smith, 1993; Bennedsen and Feldmann, 2002; Grossman and Helpman, 2001), and our focus on pressure that shifts the policymaker's choice distinguishes our approach from models of vote buying (Snyder, 1991; Dekel et al., 2008). Our model is a contest between competing pressure groups in the spirit of Becker (1983) (see Konrad, 2009, for a survey). Dixit (1987) shows that the value of precommitment in a contest depends on the slope of the rival's best response. In our setting, supermodularity of the counterpressure technology makes expected lobbying and backlash strategic complements. We show that the resulting escalation can be interpreted as political polarization, and we study corporate governance as a response to the underlying commitment problem. More broadly, we contribute to the literature on firms as shaping policy rather than only responding to it (Stigler, 1971; Laffont and Tirole, 1993; Dal Bo, 2006; Zingales, 2017).[9] Political backlash has been studied against globalization (Pastor and Veronesi, 2021) and against actions taken by shareholders (Doettling et al., 2025). We study backlash against expected corporate lobbying.

Our paper also contributes to the literature on corporate social responsibility and stakeholder capitalism. Much of this literature studies how the social preferences of shareholders and investors shape firms' objectives, cost of capital, and real decisions (Hart and Zingales, 2017, 2022; Broccardo et al., 2022; Pástor et al., 2021; Oehmke and Opp, 2025), while Bebchuk and Tallarita (2020) question whether stakeholder governance delivers on its promise. We study its political consequences instead. We provide a positive theory of why profit-maximizing firms may voluntarily adopt stakeholder-oriented governance, and we show that, whether adopted voluntarily or mandated by regulation, it can harm citizens. Our delegation paradox relates to work showing that pro-social preferences can backfire. Building on Heinkel et al. (2001), several papers show that responsible investing can have adverse consequences through a capital-allocation channel or through the responsible investor's own commitment problem (Davies and Van Wesep, 2018; Bisceglia et al., 2022; Gupta et al., 2025; Edmans et al., 2026). Outside the investing context, Geelen et al. (2022) show that an agent's stronger pro-social preferences can lead a manager to centralize a task, resulting in a less pro-social outcome, and Allen et al. (2023) show that pro-social preferences and financial market tools can clash. In our model, by

[9]For a review of the literature on politics and finance, see Akey et al. (2025); for a review of the empirical literature on lobbying, see Bombardini and Trebbi (2020); and for evidence that political connections can lead to rent-seeking, see Faccio and Zingales (2021).

contrast, the pro-social manager does not create a commitment problem but solves one, and the paradox arises from her effect on the opposition.

Finally, we apply the logic of time inconsistency in economic policy to corporate political activity. The lobbying ratchet effect is the analogue of the inflation bias that arises when a central bank cannot commit (Kydland and Prescott, 1977; Barro and Gordon, 1983).[10] Unlike in Cukierman and Meltzer (1986), where a central bank may prefer opacity about its objectives, opacity in our model is a feature of the lobbying environment that the firm would remove if it could. That delegation can substitute for commitment echoes Rogoff (1985), who proposes appointing a conservative central banker to fight inflation bias, and the literature on strategic delegation in oligopoly, in which owners appoint managers with non-profit-maximizing objectives to commit to a more aggressive output strategy (Vickers, 1985; Fershtman and Judd, 1987; Sklivas, 1987). In our model, the firm instead delegates to a *less* aggressive manager to commit to restraint, and, unlike in Rogoff (1985), delegation fully restores commitment.

2 Model Setup

Consider an economy populated by a continuum of citizens of measure one, a representative firm (the corporate sector), an advocacy organization representing citizens, and a policymaker who chooses a policy $x \in \mathbb{R}$. We normalize $x = 0$ to be the median citizen's ideal policy and the status quo.

The policymaker. The policymaker's ideal policy is also $x = 0$, which can be read as reflecting either electoral accountability to the median citizen or the personal costs she incurs by departing from the status quo. Organized interests, however, can pull her away from this ideal by exerting political pressure through campaign contributions, public-relations campaigns, endorsements, or turnout assistance. Denoting corporate pressure by f and countervailing pressure by Φ, the policy maker chooses x to maximize

$$U \equiv -\frac{1}{2}x^2 + \big(f - \Phi\big)x. \tag{1}$$

[10] The term "ratchet effect" is also used in the literature on dynamic incentive contracts (Freixas et al., 1985; Laffont and Tirole, 1988), where an agent's good performance today raises the standard to which she is held tomorrow. Our usage instead refers to the escalation of lobbying and backlash under rational expectations.

Her policy choice therefore depends on the balance of organized pressure, i.e., $x = f - \Phi$.

The firm. The firm's profits depend on the policy x: they are maximized at $b_F > 0$ and decline in the squared distance between x and b_F. The assumption $b_F > 0$ captures the tendency of corporate interests to prefer policies that depart from the median citizen's, such as lower corporate taxes, weaker labor regulation, lighter environmental oversight, and more lenient competition policy.

To influence the policy outcome, the firm can lobby. As noted above, lobbying takes many forms and, in extreme cases, may even shade into corruption through the exchange of favors or payments for regulatory benefits. Formally, the firm chooses a lobbying intensity $l \geq 0$ at a convex cost $c\,l^2$, where $c > 0$.[11] Lobbying intensity l generates political pressure $f(l)$ on the policymaker, where f is twice continuously differentiable with $f(0) = 0$, $f'(l) > 0$, and $f''(l) \leq 0$ for all $l \geq 0$. Lobbying thus has nonincreasing marginal returns: each additional unit of effort yields no more influence than the one before, reflecting the idea that the firm targets the most receptive access points in government first. The firm's objective function is $\Pi(b_F)$, where

$$\Pi(b) \equiv -(b - x)^2 - c\,l^2. \tag{2}$$

The firm can either lobby directly or delegate the lobbying lobbying decision to a manager with ideal point b_M, which may differ from b_F. The manager then chooses l to maximize $\Pi(b_M)$. A leading example is the appointment of a socially responsible manager, who places intrinsic value on public-goods and therefore at least partially internalizes the citizens' preferences when choosing lobbying intensity. The firm's own payoff remains $\Pi(b_F)$, evaluated at the lobbying intensity the manager chooses.

Throughout, we assume there exists $\bar{l} > 0$ such that

$$c\,\bar{l} > b_F\,f'(\bar{l}). \tag{3}$$

As shown below, this condition guarantees existence of an interior equilibrium. It states that at a sufficiently high lobbying intensity $\bar{l}$, the marginal cost of lobbying exceeds the maximum

[11] The parameter c captures both direct lobbying costs and indirect ones, such as effects on the firm's cost of capital (Grotteria, 2024).

marginal benefit any manager $b \in (0, b_F]$ could obtain. Consequently, no such manager will choose to lobby beyond $\bar{l}$ in equilibrium. This condition reflects real-world constraints such as finite resources or reputational risks.

The informational friction. Lobbying activity is not directly observed by the citizens, who instead have an expectation l_e of it. Much of the influence firms exert over policy operates through channels that are opaque, for example, campaign contributions are funneled through intermediaries and lobbyists meet privately with regulators. Even where disclosure laws exist, as in the United States under the Lobbying Disclosure Act of 1995, they cover only a narrow subset of influence activities and are widely regarded as incomplete.[12] More generally, our analysis requires only that disclosure be imperfect or delayed in practice, so that citizens must form expectations about lobbying from limited or lagged information. This information asymmetry is the key friction in the model and drives the backlash mechanism we analyze. In Section 4.1 we study transparent lobbying as a benchmark.

The advocacy organization. Citizens can act through an advocacy organization (e.g., a trade union or an environmental or consumer group) that represents those opposed to the firm and shares their ideal point. Members bear the cost of mobilizing, while the policy their campaign obtains benefits all citizens. The organization runs a campaign of intensity $a \geq 0$ at a convex cost γa^2, with $\gamma > 0$.[13]

The decision to campaign is taken before the firm's lobbying is observed, so the organization acts on its expectation of lobbying, l_e. Because the model has no unobserved heterogeneity, this expectation is degenerate and, on the equilibrium path, correct. We therefore write everything directly as a function of l_e rather than as an expectation. Campaign effort a translates into counter-pressure $\Phi(a, l_e)$, which we assume is twice continuously differentiable, with $\Phi(0, l_e) = 0$

[12] Nagaraj (2026) estimates the extent of undisclosed corporate lobbying directly: intersecting 4.5 trillion smartphone location pings with building footprints, he finds that roughly 40% of firms initiating in-person contact with the IRS are not named in any LDA filing.

[13] Convexity can be microfounded as follows: if members' private costs of acting are drawn from a distribution K, and a campaign of intensity a mobilizes every member with cost at most a, the organization bears total cost $\int_0^a s\, dK(s)$, which reduces to γa^2 when K is uniform.

and $\Phi_a > 0$. The organization chooses $a \in [0, \bar{a}\,]$ to maximize

$$W \equiv -x^2 - \gamma a^2. \tag{4}$$

Throughout, subscripts on Φ denote partial derivatives evaluated at (a, l_e): Φ_a and $\Phi_{a\,a}$ are the marginal product of campaign intensity and its slope, while Φ_l and $\Phi_{a\,l}$ are the level and slope effects of the anticipated lobbying threat.[14] Arguments are suppressed except where clarity requires them. Our key assumption is that the counter-pressure technology is supermodular,

$$\Phi_{a\,l} > 0. \tag{5}$$

A leading example, which we will use for illustrative purposes, is $\Phi(a, l_e) = \sigma\, l_e\, a$. Supermodularity means that counter-pressure is more effective at the margin when anticipated lobbying is larger. Intuitively, large expected lobbying may motivate the advocacy group's members as they realize that policy would shift unfavorably absent a counter response. Thus, a given activity of the advocacy group becomes more effective. In the same spirit, large expected lobbying may give voters a reason to organize, making the politician more responsive to organized interests (Olson, 1965). As noted in the introduction, the evidence is consistent with this mechanism, as corporate lobbying and perceived policy threats trigger counter-mobilization among opposing citizens, raising turnout and grassroots opposition activity among those adversely affected. Since $\Phi(0, l_e) = 0$, (5) also implies $\Phi_l > 0$: a larger threat raises the opposition's influence at any given effort, not merely the return to increasing it.

We place no restriction on the sign of $\Phi_{a\,a}$, but we limit how convex Φ can be in a. Specifically, we assume[15]

$$\Theta > 0 \quad \text{for all } a \in [0, \bar{a}], \qquad \text{where } \Theta \equiv 1 - a\frac{\Phi_{a\,a}}{\Phi_a}. \tag{6}$$

Timing. The game proceeds in four stages, summarized in Figure 1. In the first stage, the firm publicly appoints a manager with ideal point b_M, chosen to maximize its payoff $\Pi(b_F)$ in (2); not delegating corresponds to $b_M = b_F$. In the second stage, the appointed manager

[14] A subscript l on Φ denotes differentiation with respect to its second argument, the anticipated lobbying intensity l_e, not the realized lobbying. The distinction matters off the equilibrium path.

[15] $\Theta > 1$ when Φ is concave in a, $\Theta = 1$ when Φ is linear in a, and $\Theta < 1$ when Φ is convex in a.

chooses the lobbying intensity l. In the third stage, the advocacy organization moves. It observes b_M but not l, and so forms an expectation l_e of the lobbying it faces and mobilizes at campaign intensity $a(l_e)$.[16] This is why delegation matters: b_M is the one instrument through which the firm can move what the organization expects. In the fourth stage, lobbying l and counter-pressure $\Phi(a(l_e), l_e)$ are realized, and the policymaker implements

$$x(l, l_e) \equiv f(l) - \Phi\big(a(l_e), l_e\big). \tag{7}$$

Equilibrium. We study rational-expectations equilibria. Given b_M, lobbying l maximizes $\Pi(b_M)$ taking the organization's expectation l_e as given; campaign intensity $a(l_e)$ maximizes W given that expectation; and expectations are correct, $l_e = l$. Anticipating this, the firm chooses b_M to maximize $\Pi(b_F)$. We abuse notation and write

$$x(l) \equiv x(l, l) \tag{8}$$

for policy along the rational-expectations locus.

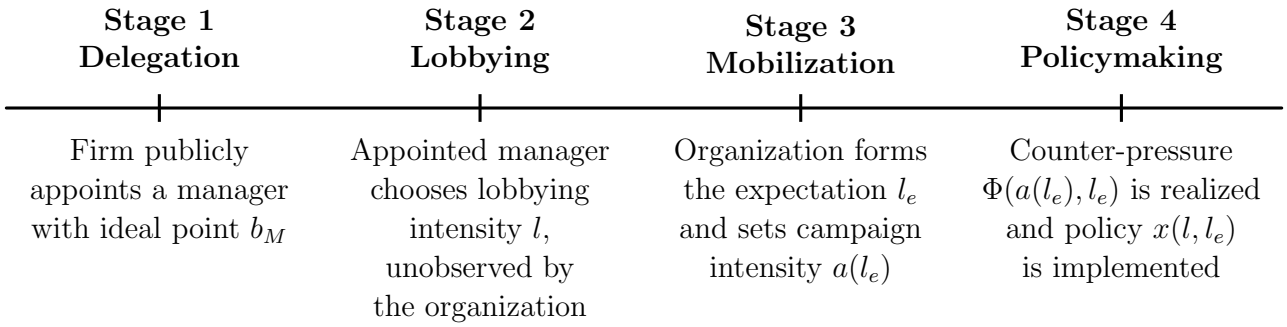


Figure 1: Sequence of events

[16] Because the organization never observes l, the second and third stages effectively form a simultaneous-move game played after the delegation decision, and the order in which they are presented is immaterial. We let the organization move second because doing so makes its inference problem transparent. It also reflects that the status quo already coincides with the organization's ideal point, so the organization is purely reactive. It mobilizes only against an anticipated attempt to move policy away from it.

3 Political Backlash

We begin with the organization's problem and characterize its optimal campaign effort, $\alpha(l_e)$. All omitted proofs are in Appendix A.

Lemma 1 *The organization's problem* (4) *has a unique solution* $\alpha(l_e)$*, characterized by*

$$\Phi_a(\alpha, l_e)\, x(l_e) = \gamma\alpha. \tag{9}$$

Moreover, the organization campaigns if and only if it anticipates lobbying: $\alpha(l_e) > 0$ *for every* $l_e > 0$*, and* $\alpha(0) = 0$*.*

The first-order condition (9) equates the organization's marginal benefit of campaigning with its marginal cost, γa. The marginal benefit is the additional counter-pressure a campaign generates, Φ_a, multiplied by the policy distortion it helps undo, x. The organization therefore campaigns harder when its effort is more effective or when the policy it faces is farther from its ideal. Because campaigning is costly, the organization never fully offsets the firm's pressure: at an interior optimum, $\gamma a > 0$ and $\Phi_a > 0$, so (9) implies $x > 0$.

Notably, α need not increase in l_e, because a larger lobbying threat has two opposing effects. First, it makes campaigning more effective at the margin, raising Φ_a, since $\Phi_{al} > 0$. Second, holding campaign effort fixed, it changes the policy gap x at rate $f' - \Phi_l$. If $f' \geq \Phi_l$, the gap weakly widens, and both effects push α up. If $f' < \Phi_l$, the gap narrows, and when this second effect dominates, the organization scales back its campaign. Counter-pressure Φ nevertheless always rises with anticipated lobbying, as we show below, even when α does not.

Substituting α into (7) reduces the model to a single relation between lobbying, its anticipation, and policy. The next lemma formalizes this reduction and establishes the properties we use throughout.

Lemma 2 *The equilibrium policy is additively separable*

$$x\big(l, l_e\big) \;=\; f(l) - h(l_e), \tag{10}$$

where

$$h(l_e) \equiv \Phi\big(\alpha(l_e), l_e\big) \tag{11}$$

is the continuously differentiable backlash function, satisfying $h(0) = 0$ *and* $h'(l_e) > 0$*. Moreover,* $x(0) = 0$*,* $x(l) > 0$ *for all* $l > 0$*, and*

$$x'(l) < 0 \quad \Longleftrightarrow \quad a\,\Phi_{a\,l} \;>\; \Theta\big(f' - \Phi_l\big), \tag{12}$$

evaluated at $\alpha(l)$*.*

According to Lemma 2, realized lobbying enters policy only through f, and anticipated lobbying only through h. We call h the *backlash function* because more anticipated lobbying always raises counter-pressure. Two channels explain $h' > 0$, as the decomposition

$$h' = \Phi_a\,\alpha' + \Phi_l \tag{13}$$

shows: the organization may campaign harder (the first term, positive whenever α is increasing), and any given campaign is more potent against a larger threat (the second term, positive because $\Phi_l > 0$). The first channel can reverse, but the second cannot, and it dominates whenever the first turns negative. Since $x(l) > 0$ for all $l > 0$, counter-mobilization restrains corporate influence but never reverses it.

Lemma 2 ensures $h' > 0$ but not $h' < f'$. Since $x' = f' - h'$, backlash can be strong enough that more lobbying moves policy away from the firm. Condition (12) states when this happens. The left-hand side measures how much harder the organization pushes back when the firm pushes harder, which supermodularity makes positive. The right-hand side is the direct effect of lobbying net of Φ_l, scaled by Θ, which measures how quickly the return to additional campaigning falls. Supermodularity is therefore necessary: without it, the left-hand side is zero and the right-hand side is positive, so (12) cannot hold. Backlash is more likely to dominate when lobbying is already high, so that f' is small; when the organization is heavily mobilized and its campaign responds strongly to the threat, so that $a\,\Phi_{a\,l}$ is large; or when Θ is small, so that the organization escalates readily.

By contrast, backlash cannot dominate at low lobbying, $x'(0) > 0$. With no anticipated lobbying, the organization does not campaign, $\alpha(0) = 0$, and since $\Phi(0, l_e) = 0$, a small threat raises counter-pressure only by triggering a campaign. Accordingly, the left-hand side of (12) is zero, while the right-hand side is $\Theta f'(0) > 0$. The firm's lobbying is therefore productive at

low intensities and can become self-defeating only at higher ones, so x may be hump-shaped.

For use below, let $l_{\max}$ denote the smallest maximizer of x on $[0, \bar{l}\,]$. At times we assume that x is single-peaked, rising up to $l_{\max} \in (0, \bar{l}\,)$ and falling beyond it.[17]

Leading example. To build intuition for the backlash function, consider

$$f(l) = \sqrt{l}, \qquad \Phi(a, l_e) = \sigma\, a\, l_e, \tag{14}$$

with $\sigma > 0$. All maintained assumptions hold: $f(0) = 0$, $f' > 0 > f''$; $\Phi(0, l_e) = 0$, $\Phi_a = \sigma l_e > 0$ and $\Phi_{a\,l} = \sigma > 0$; and since $\Phi_{a\,a} = 0$ we have $\Theta \equiv 1$. The example has a closed-form solution:

$$\alpha(l_e) = \frac{\sigma\, l_e^{3/2}}{\gamma + \sigma^2 l_e^2}, \qquad h(l_e) = \frac{\sigma^2 l_e^{5/2}}{\gamma + \sigma^2 l_e^2}, \qquad x(l) = \frac{\gamma\sqrt{l}}{\gamma + \sigma^2 l^2}, \tag{15}$$

where policy x is single-peaked with an interior peak at

$$l_{max} = \frac{1}{\sigma}\sqrt{\frac{\gamma}{3}}. \tag{16}$$

Panel (a) of Figure 2 below plots the three objects of Lemma 2. Gross influence $f(l)$ rises steeply at first and then flattens. Backlash $h(l)$ does the opposite: the first units of lobbying provoke almost no counter-mobilization, but backlash accelerates as lobbying grows. Policy $x(l)$ is the vertical gap between the two curves. It widens until $l_{\max}$, where $f' = h'$, and narrows beyond it. As lobbying grows, the campaign neutralizes ever more of the influence the firm buys, and $x(l) \to 0$.

4 Firm Lobbying

Having characterized how the organization responds to anticipated lobbying, we now turn to the firm's lobbying choice; delegation is analyzed in Section 5.

We consider two regimes: a benchmark in which lobbying is transparent, and our main case in which it is not. The next result shows that, in both regimes, the firm stops short of its ideal

[17] By Lemma 2, x is single-peaked whenever $a\Phi_{a\,l} - \Theta\big(f' - \Phi_l\big)$, evaluated at $\alpha(l)$, crosses zero once on $(0, \bar{l}\,)$, from below.

policy.

Lemma 3 *Whether or not lobbying is transparent, equilibrium policy falls short of the firm's ideal:* $x^* < b_F$.

Intuitively, the firm's policy loss is flat near its ideal, so the last unit of lobbying needed to reach b_F yields only a second-order gain while its cost is first-order. The firm therefore always stops short of b_F.

4.1 Transparent lobbying

When lobbying is transparent, the organization observes the firm's lobbying before deciding how much to campaign. It therefore responds to realized rather than anticipated lobbying, so that $l_e = l$. When choosing l, the firm thus fully internalizes the backlash its lobbying provokes. Substituting the policy constraint into (2), the firm chooses $l \geq 0$ to maximize

$$\Pi_T \equiv -(b_F - x(l))^2 - cl^2. \tag{17}$$

Let l_T denote the firm's optimal lobbying intensity under transparency, and $x_T = x(l_T)$ the corresponding equilibrium policy.

Proposition 1 *Under transparency, the firm's optimal lobbying is interior,* $l_T \in \left(0, \min\{l_{\max}, \bar{l}\}\right)$, *and satisfies*

$$l_T = \frac{1}{c} x'(l_T) \left(b_F - x(l_T)\right). \tag{18}$$

Moreover, l_T *and* x_T *are locally strictly decreasing in* c *and strictly increasing in* b_F.

Proposition 1 implies that $l_T < l_{\max}$: the firm never lobbies past the level that maximizes x, because any policy reached beyond $l_{\max}$ can also be reached with less lobbying, and thus at lower cost. It follows that $x'(l_T) > 0$. At the optimum, an additional unit of lobbying still moves policy toward the firm's ideal point, so the firm never lobbies to the point where lobbying becomes self-defeating at the margin. Neither conclusion depends on the shape of x, both follow from the fact that the firm internalizes the backlash.[18] Panel (b) of Figure 2 shows the transparent optimum in the leading example.

[18] Assumption (3) serves only to keep the upper bound $\bar{l}$ from binding.

The comparative statics are intuitive. A higher lobbying cost c makes each unit of influence more expensive and induces the firm to scale back its effort. A more extreme ideal point b_F widens the gap between the implemented policy and the firm's preferred outcome, raising the marginal value of closing it.

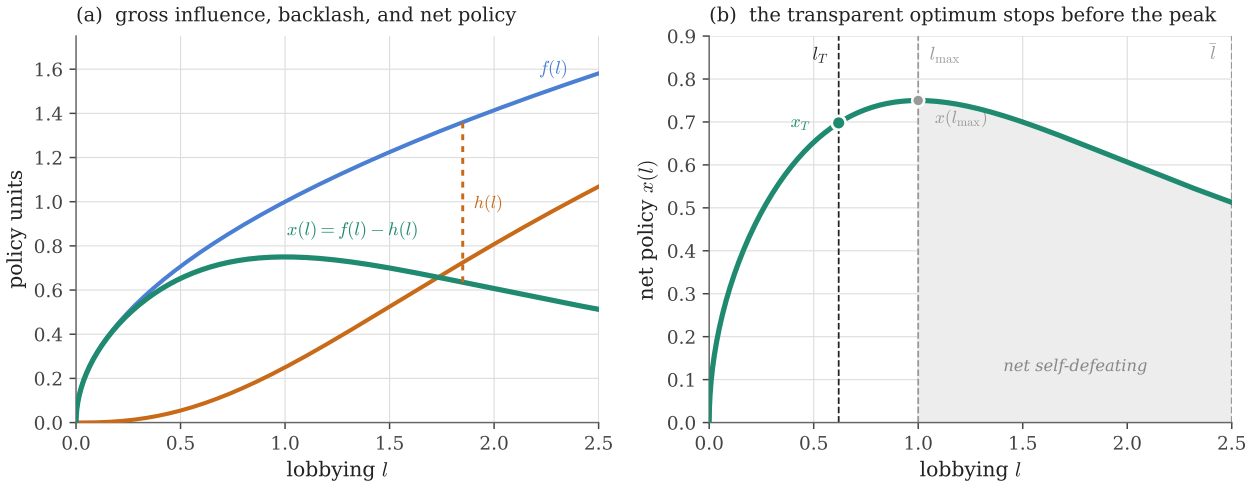


Figure 2: Transparent lobbying ($\gamma = 3$, $\sigma = 1$, $c = 0.15$, $b_F = 1$, $\bar{l} = 2.5$). Panel (a) decomposes gross influence into $x(l) = f(l) - h(l)$. Panel (b) locates the transparent optimum: $l_T = 0.62 < l_{max} = 1$ and $x_T = 0.70 < x(l_{max}) = 0.75$.

4.2 Opaque lobbying

We now turn to our main case, in which the organization cannot observe the firm's lobbying before deciding how much to campaign. Taking the organization's expectation l_e as given, the firm chooses $l \geq 0$ to maximize

$$\Pi_O \equiv -(b_F - x(l, l_e))^2 - cl^2, \tag{19}$$

where, by Lemma 2, $x(l, l_e) = f(l) - h(l_e)$. Since $f'' \leq 0$ and $c > 0$, the objective function is strictly concave in l, ensuring a unique solution. We denote this best response by $\ell(l_e)$, which is characterized by the solution to the first-order condition:

$$l = \frac{1}{c} f'(l)(b_F - x(l, l_e)). \tag{20}$$

Since $h'(l_e) > 0$, we have $\ell'(l_e) > 0$: when the organization expects greater lobbying, the resulting backlash $h(l_e)$ is stronger. This pushes the implemented policy further below the firm's ideal point b_F, thereby increasing the marginal benefit of lobbying and inducing the firm to respond with higher effort.

In equilibrium, expectations will be consistent with the firm's actual lobbying choice, so that $l^* = \ell(l^*)$ and the net policy effect is $x^* = f(l^*) - h(l^*) = x(l^*)$.

Stability. We impose a standard stability criterion, which reduces (though does not eliminate) equilibrium multiplicity. An equilibrium is stable if, when the organization's beliefs rise (fall) slightly above (below) the equilibrium level, the firm's best response rises (falls) by less than the change in beliefs. Graphically, an interior equilibrium is stable if $\ell(\cdot)$ crosses the 45-degree line from above at l^*, so that a small perturbation in beliefs triggers a corrective rather than amplifying response.

Stability condition: An equilibrium l^* is stable if for all $\epsilon > 0$ sufficiently small: either $l^* = 0$ or $\ell(l^* - \epsilon) > l^* - \epsilon$; and either $l^* = \bar{l}$ or $\ell(l^* + \epsilon) < l^* + \epsilon$.

For brevity, from here on we refer to a stable equilibrium simply as an equilibrium, and denote equilibrium lobbying and policy by l_F^* and x_F^*, respectively.

Proposition 2 *Under opacity, a rational expectations equilibrium always exists. In any equilibrium, lobbying is interior, $l_F^* \in (0, \bar{l})$, and satisfies*

$$l = \frac{1}{c} f'(l) \left(b_F - x(l)\right), \tag{21}$$

so that the equilibrium policy is

$$x_F^* = b_F - c\frac{l_F^*}{f'(l_F^*)}. \tag{22}$$

There exists $\bar{c} > 0$ such that the equilibrium is unique whenever $c > \bar{c}$. Moreover, l_F^* is locally strictly decreasing in c and strictly increasing in b_F, while x_F^* is strictly decreasing in c and strictly increasing in b_F if and only if $x'(l^*) > 0$.

To understand Proposition 2, recall that higher expected lobbying provokes a stronger backlash ($h'(l_e) > 0$), which the firm responds by lobbying more. If this backlash is strong

enough, the firm's best response rises faster than expectations themselves, so that $\ell'(l_e) > 1$. Expectations and lobbying then feed on each other: each increment in expected lobbying generates backlash that calls for yet more lobbying. Assumption (3) ensures that the marginal benefit of lobbying eventually declines fast enough for the escalation to stop, which guarantees that a rational expectations equilibrium exists.

More generally, the model can admit multiple stable equilibria, reflecting a coordination problem in the political system that opacity creates. On the one hand, if the organization expects aggressive lobbying, it responds with strong counter efforts. This, in turn, induces the firm to lobby heavily to counter the organization, validating both the high lobbying expectations and the heightened backlash. On the other hand, if the organization expects limited lobbying, the anticipated backlash is muted, the marginal return to political influence is low, and the firm restrains itself, validating the moderate expectations. Multiplicity thus reflects a strategic complementarity: under opacity, expectations of lobbying can be self-fulfilling.

The comparative statics in Proposition 2 contrast sharply with those in Proposition 1. As under transparency, equilibrium lobbying l^* decreases in c and increases in b_F. The policy outcome, however, need not follow. Recall that under transparency, optimal lobbying always satisfies $x'(l_T) > 0$: at the margin, additional lobbying moves policy toward the firm's ideal point. Policy therefore responds in the intuitive way, decreasing in the lobbying cost c and increasing in the firm's ideal point b_F.

Under opacity, by contrast, the feedback between expectations and lobbying described above can produce an equilibrium in which $x'(l^*) < 0$. The firm then over-lobbies, trapped by the organization's expectation of aggressive lobbying. In this case, a higher lobbying cost acts as a commitment device: it lowers the firm's lobbying in a way the organization anticipates, which dampens the backlash. Because $x'(l^*) < 0$, the backlash falls by more than the firm's direct influence, so policy moves toward b_F. Similarly, a more extreme firm (higher b_F) lobbies harder, but the additional backlash more than offsets its additional influence, and policy moves away from the firm's ideal point. This last comparative statics is the key to the delegation paradox we derive in Section 5.

Lobbying ratchet effect. With the equilibrium characterized, we can now define and analyze the lobbying ratchet effect.

Definition 1 *The rational expectations equilibrium exhibits a lobbying ratchet effect if and only if* $l_F^* > l_T$.

That is, the firm lobbies more when lobbying is opaque than it would if it were transparent. Because l_T is unique, this comparison is well-defined.[19] The next result characterizes when the ratchet effect arises.

Lemma 4 *A rational expectations equilibrium that exhibits the lobbying ratchet effect always exists. If* $\Pi_T(l)$ *is concave, then every rational expectations equilibrium exhibits the lobbying ratchet effect.*

The ratchet effect arises because, under opacity, the firm cannot commit to a level of lobbying and instead takes the organization's expectation l_e as given when choosing l. It therefore ignores the backlash that additional lobbying provokes (its marginal return is $f'(l)$ rather than $x'(l)$) and lobbies more than it would under transparency. In equilibrium, the organization anticipates this over-lobbying. Higher expectations trigger a stronger backlash, which in turn leads the firm to lobby more. The resulting equilibrium involves strictly more lobbying than the transparent benchmark, where the firm fully internalizes the backlash it generates.[20] Since h is increasing, the counter-pressure the firm faces is higher as well, $h(l^*) > h(l_T)$, even though the organization's campaign effort itself need not rise. Both sides of the political contest thus push harder than under transparency, and opacity generates political polarization as a byproduct of the commitment failure it induces.

Leading example (continued). Panel (a) of Figure 3 draws the best response $\ell(l_e)$ against the 45° line. The equilibrium l_F^* is the crossing. The transparent optimum l_T, at which the firm internalizes the backlash, lies far to its left, and the vertical distance $\ell(l_T) - l_T > 0$ is the lobbying ratchet effect (Lemma 4): at the level of lobbying the firm would commit to, it strictly wants to deviate upward. Panel (b) illustrates the comparative statics of Proposition 2. Since l_F^* rises with b_F, x_F^* is single-peaked in b_F. Its maximum is attained exactly when $l_F^* = l_{max}$. This is the leading example's version of the statement that under opacity a more extreme firm

[19] Although equation (18) may have multiple solutions, the global maximizer of $\Pi_T(l)$ is unique.

[20] With multiple equilibria, the ratchet effect is guaranteed only in the largest equilibrium. Another equilibrium may be sustained by low expectations and feature less lobbying than under transparency, $l^* < l_T$.

may obtain a *less* favorable policy: beyond the maximum, the additional lobbying an extreme firm undertakes provokes more backlash than it overcomes, and the firm is worse off for being more aggressive.

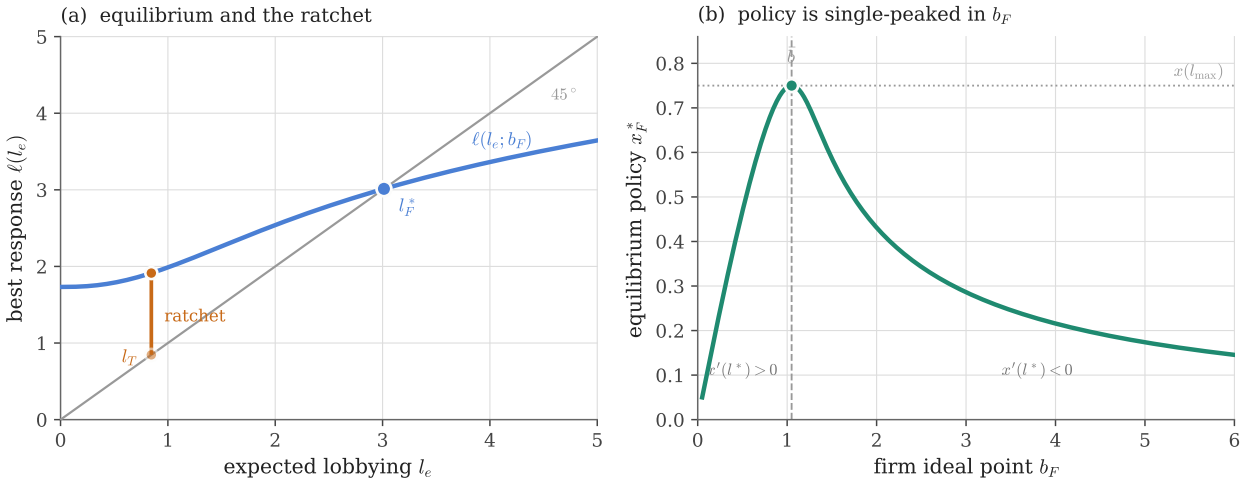


Figure 3: Equilibrium under opacity in the leading example ($\gamma = 3$, $\sigma = 1$, $c = 0.15$, so $l_{max} = 1$, $x(l_{max}) = 0.75$, $\bar{b} = 1.05$, and $\bar{c} = 0.048 < c$ so the equilibrium is unique). Panel (a): the best response $\ell(l_e)$ for $b_F = 2$ and the 45° line. The equilibrium is $l_F^* = 3.01$; the transparent optimum is $l_T = 0.85$, and at l_T the firm's best response is 1.91; it strictly wants to escalate, which is the lobbying ratchet. Panel (b): the equilibrium policy x_F^* as a function of the firm's ideal point.

5 Delegation and Stakeholder Governance

So far, we have assumed that the firm makes its own lobbying decisions to maximize shareholder value, $\Pi(b_F)$. In practice, lobbying decisions are made by managers whose objectives may differ from the firm's. Firms can shape those objectives through ESG-linked compensation, sustainability-oriented board appointments, and explicit stakeholder mandates. Regulations, such as labor co-determination laws, fiduciary duty reforms, and mandatory ESG disclosure, can also shift managers' objectives toward those of a broader set of stakeholders. We capture these channels in reduced form through the first stage of the game, in which a manager with ideal point b_M is appointed, either by the firm or through regulation, and then makes the lobbying decision.

5.1 Delegation under transparency

Consider first the transparent benchmark. Because the organization responds to realized lobbying, the manager's identity cannot shape its expectations. The firm's own choice $l_T(b_F)$ is therefore its optimal lobbying level. Since l_T is strictly increasing in the decision-maker's ideal point (Proposition 1), any manager with $b_M \neq b_F$ chooses a different level, which strictly lowers firm value. This yields the following corollary.

Corollary 1 *Under transparency, the firm strictly prefers keeping control, that is,* $\Pi_T(b_M) < \Pi_T(b_F)$ *for all* $b_M \neq b_F$.

Moreover, under transparency the decision-maker internalizes the backlash its lobbying provokes. Its first-order condition, $(b - x)\, x' = c\, l$, implies $x' > 0$ whenever $l_T(b) > 0$. Policy $x(l_T(b))$ is therefore increasing in b. This confirms a natural intuition: a manager whose ideal point is closer to the citizens' faces a smaller conflict of interest with them, so she lobbies less and delivers a policy closer to their ideal point. To compare regulations, we call a regulation *efficient* if it implements its policy outcome x at the lowest cost to firm value.

Corollary 2 *Under transparency, any efficient regulation that moves policy toward the citizens' ideal point sets* $b_M < b_F$ *and lowers firm value.*

Under transparency, citizens can therefore gain only at the expense of firm value. As we show next, both this trade-off and the intuition behind it fail when lobbying is opaque.

5.2 Delegation under opacity

Under opacity, the organization observes b_M but not l, so the manager's identity shapes what the organization expects. Writing the manager's best response as $\ell(l_e; b_M)$, equilibrium lobbying satisfies $l_M^* = \ell(l_M^*; b_M)$ and induces the policy $x_M^* \equiv x(l_M^*)$; setting $b_M = b_F$ recovers l_F^* and x_F^*. Throughout this section, we assume

$$c > \bar{c} \qquad \text{and} \qquad x(l) \text{ is single-peaked}, \tag{23}$$

and do not restate (23) in each result. By Proposition 2, $c > \bar{c}$ ensures that the equilibrium is unique for every b_M, so the consequences of each delegation decision are unambiguous.

Delegation can benefit the firm only by mitigating the lobbying ratchet effect. A manager with $b_M > b_F$ does the opposite. He values policy movement more than the firm does, so he lobbies even more ($l_M^* > l_F^*$). This provokes more backlash and worsens the ratchet effect; the firm never benefits from delegating to a manager with extreme preferences.[21]

Lemma 5 (No over-delegation) $\Pi_O(b_M) < \Pi_O(b_F)$ *for every* $b_M > b_F$.

We therefore focus on managers who are more moderate than the firm, with $b_M < b_F$. We first characterize the equilibrium when lobbying is delegated to a manager with an exogenous ideal point b_M, which isolates the policy and welfare implications of preference misalignment between the firm and its manager. We then endogenize the firm's choice of b_M and show that its optimal delegation decision can harm the median citizen.

A key object of the analysis is

$$\Lambda(l) \;\equiv\; x(l) + c\,\frac{l}{f'(l)}. \tag{24}$$

By Proposition 2, l is an equilibrium under a manager with ideal point b if and only if

$$\Lambda(l) = b. \tag{25}$$

The map Λ therefore *inverts* the equilibrium correspondence: for each lobbying intensity, it gives the unique manager who would choose it in equilibrium. Appointing a manager is thus equivalent to choosing equilibrium lobbying directly, with $b_M = \Lambda(l)$. The uniqueness condition $c > \bar{c}$ implies that Λ is strictly increasing, so $l^*(\cdot) = \Lambda^{-1}$ is continuous and strictly increasing. Moreover, since $\Lambda(0) = 0$ and, under Assumption 3, $\Lambda(\bar{l}\,) > b_F$, every manager $b \in (0, b_F]$ induces an interior equilibrium. A more moderate manager therefore lobbies strictly less; delegation always de-escalates.

Since delegation always de-escalates, one might expect the intuition from the transparent case to carry over, so that a less extreme manager, $b_M < b_F$, also moderates policy, $x_M^* < x_F^*$. We show that the opposite can hold. Delegating to a less extreme manager can produce a

[21] Lobbying past the peak of x does not help either, since any policy reached there was already attainable more cheaply before the peak.

more extreme policy, leaving the median citizen strictly worse off even though the manager's preferences are closer to hers. We call this the *delegation paradox*.

Proposition 3 (Delegation paradox) *Let $\bar{b} \equiv \Lambda(l_{\max})$ denote the ideal point of the manager who would lobby exactly $l_{\max}$ in equilibrium.*

(i) If $b_F > \bar{b}$, there exists a threshold $\underline{b}(b_F) \in (0, \bar{b})$, decreasing in b_F, such that delegating to any manager with $b_M \in \left(\underline{b}(b_F), b_F\right)$ raises policy, $x_M^* > x_F^*$. Such delegation strictly benefits the firm and strictly harms the median citizen.

(ii) If $b_F \leq \bar{b}$, delegating to any manager with $b_M \in (0, b_F)$ lowers policy, $x_M^* < x_F^*$, and strictly benefits the median citizen.

The paradox arises because de-escalation in lobbying need not mean moderation in policy. The policy $x(l)$ is non-monotonic. On its downward-sloping part, the backlash that lobbying provokes more than offsets the influence it buys at the margin, so lobbying less *raises* policy.

The upper cutoff $\bar{b}$ does not depend on b_F. Because Λ is increasing, $b_F > \bar{b}$ is equivalent to $l_F^* > l_{max}$. The paradox therefore requires that the ratchet effect has already pushed the firm past the point at which lobbying becomes self-defeating at the margin. In other words, the delegation paradox is a large-conflict phenomenon. It is not enough that lobbying provokes a backlash; at the margin, the backlash must already outweigh the influence the firm is buying. When this condition fails, part (ii) shows that the natural intuition holds.

The lower cutoff $\underline{b}(b_F)$ limits how moderate the manager can be for the paradox to hold. A manager only slightly less extreme than the firm cuts lobbying back toward the peak, which raises policy. Once lobbying falls below the peak, however, further moderation lowers policy again. The cutoff $\underline{b}(b_F)$ is the manager whose policy exactly matches the firm's from the other side of the peak, and any manager more moderate than this pushes policy below x_F^*. The cutoff decreases in b_F. A more extreme firm lobbies further past the peak, so backlash erodes more of its policy, and the matching point on the rising side lies lower. The range of managers for whom the paradox arises therefore widens as the firm becomes more extreme, which again shows that the delegation paradox is a large-conflict phenomenon.

From the firm's perspective, however, the paradox is good news. The firm strictly gains from such delegation, and the gain involves no trade-off. Under delegation the firm obtains a

policy closer to its ideal point ($x_F^* < x_M^* < b_M < b_F$) *and* pays a lower lobbying bill ($l_M^* < l_F^*$). Appointing a more moderate manager acts as a commitment device that restrains the firm's self-defeating lobbying. The firm is not buying influence; it is buying quiet, and quiet is cheaper than influence precisely when the opposition is responsive. Within this region, the firm gains at the median citizen's expense. This has a counterintuitive implication for governance reforms that push firms toward more moderate or socially responsible management. When $b_F \leq \bar{b}$, such reforms unambiguously benefit citizens. When $b_F > \bar{b}$, they can harm them, and they are most likely to backfire precisely where they appear most needed: in firms whose political objectives lie furthest from citizens.

Leading Example (continued). Figure 4 illustrates Proposition 3 at $b_F = 2$. Without delegation, the firm lobbies $l_F^* = 3.01$, far past the peak $l_{\max} = 1$, and obtains policy $x_F^* = 0.43$. Appointing a manager with $b_M = 1.23$ cuts lobbying to 1.44 and *raises* policy to 0.71 (Panel a). More generally, policy peaks at $\bar{b} = 1.05$, and the median citizen is strictly worse off under any manager in $(\underline{b}, b_F)$, where $\underline{b} = 0.46$ (Panel b). The firm strictly gains throughout this region, and its optimal manager, $b^* = 0.98$, lies inside it (Panel c).

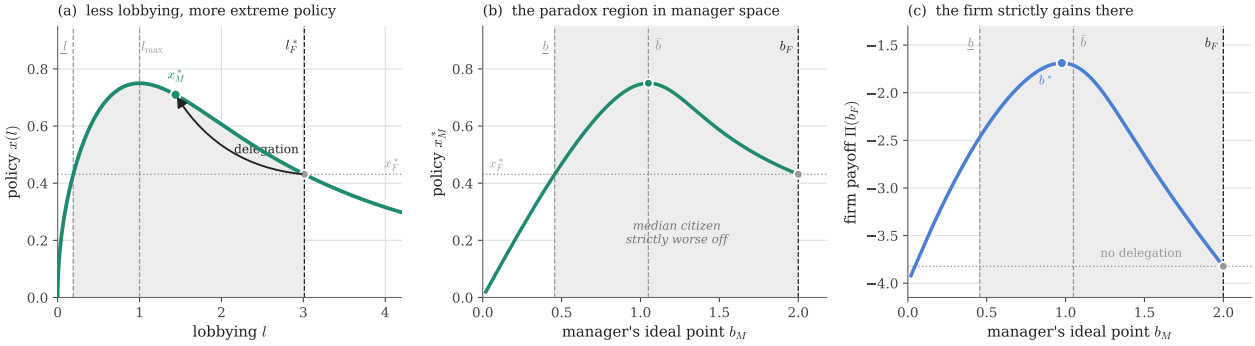


Figure 4: The delegation paradox in the leading example ($b_F = 2$). Panel (a): the policy locus $x(l)$, with the firm's equilibrium l_F^* and the delegated equilibrium l_M^* for $b_M = 1.23$. The shaded interval $(\underline{l}, l_F^*)$ contains the lobbying levels that yield a higher policy than l_F^*. Panel (b): equilibrium policy as a function of the manager's ideal point. The shaded region $(\underline{b}, b_F)$ is where the median citizen is strictly worse off than without delegation. Panel (c): the firm's payoff as a function of the manager's ideal point, with its optimum b^*.

Proposition 3 shows that delegation can move policy in either direction, depending on the

manager. This raises the question of which manager the firm would actually appoint.

Proposition 4 (Optimal delegation) *The firm's optimal delegation is unique and given by*

$$b^* \;=\; \Lambda(l_T), \tag{26}$$

it satisfies $b^* < b_F$*, and it implements* $l^*(b^*) = l_T < \min\{l_{\max}, l_F^*\}$*. The extent of moderation is*

$$b_F - b^* \;=\; \big(b_F - x_T\big)\,\frac{h'(l_T)}{f'(l_T)}. \tag{27}$$

Moreover, there exists a unique threshold $\hat{b} > \bar{b}$ *such that citizens are strictly worse off under optimal delegation,* $x_T > x_F^*$*, if and only if* $b_F > \hat{b}$*.*

The first part of Proposition 4 says that delegation restores commitment. Because the manager's ideal point pins down the equilibrium lobbying one for one, choosing b_M is choosing l. The firm's problem therefore collapses to the transparent problem, whose solution is l_T, and the firm attains its full-commitment payoff. In this sense, governance and transparency are substitutes for the firm, that is, lobbying transparency and delegation to a moderate manager are worth exactly the same to it.

One might expect the firm simply to appoint a manager whose ideal point equals the policy it wants, x_T. Such a manager, however, would lobby too little. At l_T, policy would already coincide with her ideal point, so the marginal unit of lobbying would bring her no benefit while still being costly. She would therefore lobby less than l_T, and policy would fall short of x_T. To induce l_T, the firm must give the manager a stake in policy beyond x_T, but a smaller stake than its own. The manager, like the firm, takes l_e as given, so she perceives the gross return f' rather than the net return $x' = f' - h'$. With the firm's stake, she would therefore lobby too aggressively. The firm instead gives her a stake, $b^* - x_T$, that is smaller than its own, $b_F - x_T$, by exactly the fraction of influence that backlash destroys, $h'(l_T)/f'(l_T)$. This is Equation (27). The firm's choice of a stakeholder mandate therefore reflects not only its own objective but also the strength of the opposition it expects to provoke.

Two limiting cases illustrate this. With no backlash, $h' \equiv 0$ and $b^* = b_F$. Opacity then creates no commitment problem, and delegation is pointless, just as under transparency (Corollary 1). As $h' \to f'$, the firm must appoint a manager whose ideal point nearly coincides with

the policy it seeks to implement. Stakeholder governance (i.e., moderation) should therefore be most pronounced where backlash is strong relative to the influence lobbying buys, as in politically contested environments where opposition is easily mobilized, and not merely where the firm's own preferences are extreme.

Another interesting implication is that b^* is bounded, even though the gap $b_F - b^*$ is not. Since $l_T < l_{\max}$ and Λ is increasing, $b^* = \Lambda(l_T) < \Lambda(l_{\max}) = \bar{b}$. As b_F grows, l_T approaches $l_{\max}$ and b^* approaches $\bar{b}$. Firms with very different political objectives thus appoint observationally similar managers. What varies across these firms is not so much the governance they adopt as how far it sits from their own interest.

The last part of Proposition 4 sharpens the delegation paradox. Proposition 3 shows that, when $b_F > \bar{b}$, the paradox arises for managers in $(\underline{b}, b_F)$. Proposition 4 shows that whether it arises for the manager the firm actually appoints depends on a single threshold. As b_F grows, the firm's commitment lobbying approaches the peak, so x_T rises toward $x(l_{max})$. Under opacity, by contrast, a firm with $b_F > \bar{b}$ is pushed further past the peak as b_F grows, so x_F^* falls. Since $x_T < x_F^*$ for all $b_F \leq \bar{b}$, the two cross exactly once, at $\hat{b}$. For $b_F < \hat{b}$, the firm's choice is Pareto improving. Citizens gain from a less extreme policy, and the firm gains because its savings in lobbying outweigh the policy loss. For $b_F > \hat{b}$, the firm appoints a manager who leaves citizens strictly worse off. In this region, stakeholder governance is valuable to the firm precisely because it ends the wasteful lobbying that backlash more than neutralizes, and ending it harms citizens. The mechanism that makes delegation privately attractive is the same one that makes it socially costly.

Leading Example (continued). In the leading example, the commitment and opaque policies cross at $\hat{b} = 1.20$ (Figure 5, Panel a), so the firm's optimal delegation harms the median citizen whenever $b_F > 1.20$. At that point the firm's chosen manager enters the paradox region, which opens at $\bar{b} = 1.05$ (Panel b). Consistent with the bound, b^* is nearly flat. It moves only from 0.87 to 1.02 as b_F runs from 1.1 to 6, while x_T approaches $x(l_{\max}) = 0.75$.

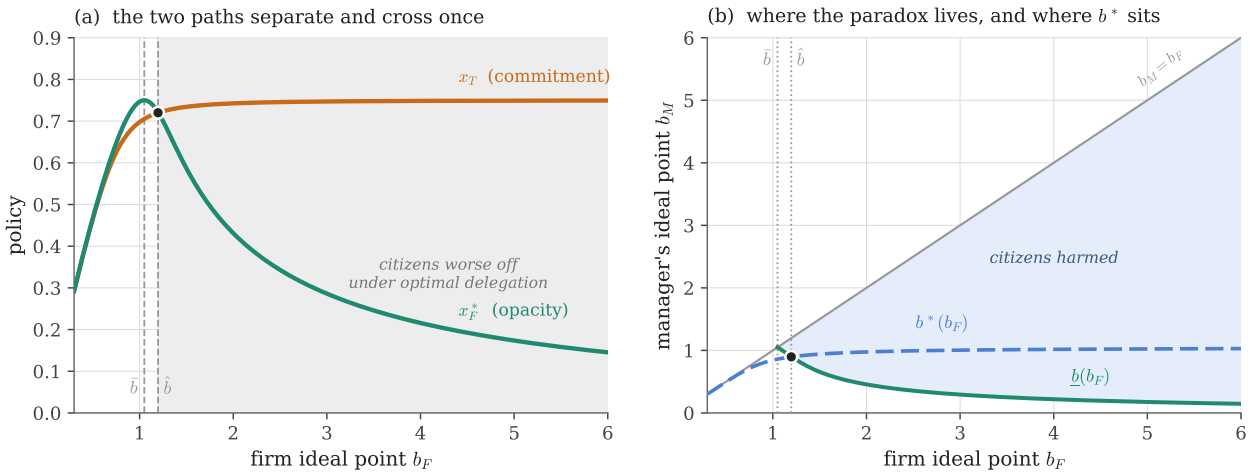


Figure 5: Optimal delegation in the leading example. Panel (a): the commitment policy x_T and the opaque policy x_F^* as functions of b_F. The shaded region is where the firm's optimal delegation strictly harms the median citizen. Panel (b): the paradox region in (b_F, b_M) space, bounded below by $\underline{b}(b_F)$ (green curve), and the firm's optimal manager $b^*(b_F)$ (dashed blue line). The marked point is $b_F = \hat{b}$.

Implications for lobbying disclosure. Proposition 4 also speaks to mandatory lobbying disclosure. Under transparency, the firm has no reason to delegate (Corollary 1). It keeps $b_M = b_F$ and lobbies l_T. Under opacity, optimal delegation implements the same l_T. The two regimes therefore deliver identical lobbying, counter-pressure $h(l_T)$, policy x_T, and firm value.

Two implications follow. First, if the firm already delegates optimally, mandating transparency has no effect on policy or on the median citizen. It merely replaces delegation as the firm's commitment device. Second, if the firm's governance does not adjust, so that under opacity lobbying is chosen according to b_F, mandating transparency moves the equilibrium from (l_F^*, x_F^*) to (l_T, x_T). Since $l_T < l_F^*$ and h is increasing, disclosure lowers both lobbying and the counter-pressure it provokes, and it raises firm value. Its effect on the median citizen, however, is exactly that of optimal delegation. By Proposition 4, policy moves toward citizens if $b_F < \hat{b}$ but away from them if $b_F > \hat{b}$. In the latter case, disclosure backfires for the same reason the delegation paradox arises. Lower lobbying makes the organization's campaign less effective, and the resulting fall in counter-pressure, $h(l_F^*) - h(l_T)$, exceeds the fall in the firm's influence, $f(l_F^*) - f(l_T)$.

What transparency does accomplish is to restore the monotone link between the decision maker's preferences and policy (Corollary 2). This is why, from citizens' perspective, disclosure requirements and stakeholder mandates are complements.

6 Conclusion

The debate on stakeholder governance has mainly focused on how it affects the firm's choices in product and labor markets. Our analysis shows that it also affects the firm's strategy in the market for political influence. When corporate lobbying is opaque, stakeholder governance changes not only how much a firm lobbies but also how strongly citizens push back, and the second effect can dominate. A more pro-social manager therefore need not result in more pro-social policy. Evaluating stakeholder governance thus requires accounting for its political consequences, and these are most likely to be adverse where firms' interests diverge most from citizens'.

This has implications for voluntary and mandatory stakeholder governance. Voluntary adoption can be a profit-maximizing response to opacity rather than a concession to stakeholders. Its spread is therefore not evidence that it benefits citizens, and the governance a firm adopts reveals little about its political objectives. Mandated stakeholder governance can backfire in exactly the firms it targets, and the firm gains when it does, so corporate acceptance of a mandate does not show that it serves citizens. The same caution applies to other reforms that curb lobbying without making it observable. Finally, disclosure and stakeholder mandates should not be designed in isolation: transparency is what makes stakeholder mandates reliably move policy toward citizens. Policies that aim to align corporations with society should therefore consider not only how firms are governed, but also how visible their political influence is.

References

Abramowitz, A. I. and S. Webster (2016, March). The rise of negative partisanship and the nationalization of u.s. elections in the 21st century. *Electoral Studies 41*, 12–22.

Ahlquist, J. S. (2017). Labor unions, political representation, and economic inequality. *Annual Review of Political Science 20*, 409–432.

Akey, P., N. Gupta, and S. Lewellen (2025). Politics and finance. *ECGI Working Paper*.

Allen, F., A. Barbalau, and F. Zeni (2023). Reducing carbon using regulatory and financial market tools. *Working Paper*.

Anthropic, PBC (2026). Form S-1 registration statement. U.S. Securities and Exchange Commission. Filed October 26, 2026.

Austen-Smith, D. (1993). Information and influence: Lobbying for agendas and votes. *American Journal of Political Science 37*(3), 799–833.

Barro, R. J. and D. B. Gordon (1983). A positive theory of monetary policy in a natural rate model. *Journal of Political Economy 91*(4), 589–610.

Bebchuk, L. A. and R. Tallarita (2020). The illusory promise of stakeholder governance. *Cornell Law Review 106*, 91–178.

Becher, M., D. Stegmueller, and K. Käppner (2018). Local union organization and law making in the US congress. *Journal of Politics 80*(2), 539–554.

Becker, G. S. (1983). A theory of competition among pressure groups for political influence. *Quarterly Journal of Economics 98*(3), 371–400.

Bennedsen, M. and S. E. Feldmann (2002). Lobbying legislatures. *Journal of Political Economy 110*(4), 919–946.

Bertrand, M., M. Bombardini, R. Fisman, and F. Trebbi (2020). Tax-exempt lobbying: Corporate philanthropy as a tool for political influence. *American Economic Review 110*(7), 2065–2102.

Bisceglia, M., A. Piccolo, and J. Schneemeier (2022). Externalities of responsible investments. Working paper, SSRN 4183855.

Bombardini, M. and F. Trebbi (2020). Empirical models of lobbying. *Annual Review of Economics 12*, 391–413.

Broccardo, E., O. Hart, and L. Zingales (2022). Exit versus voice. *Journal of Political Economy 130*(12), 3101–3145.

Brulle, R. J. (2018). The climate lobby: A sectoral analysis of lobbying spending on climate change in the USA, 2000 to 2016. *Climatic Change 149*(3–4), 289–303.

Calhoun, C. (2013). Occupy Wall Street in perspective. *British Journal of Sociology 64*(1), 26–38.

Campbell, A. L. (2003). Participatory reactions to policy threats: Senior citizens and the defense of Social Security and Medicare. *Political Behavior 25*(1), 29–49.

Cohen, S., I. Kadach, G. Ormazabal, and S. Reichelstein (2023). Executive compensation tied to ESG performance: International evidence. *Journal of Accounting Research 61*(3), 805–853.

Cukierman, A. and A. H. Meltzer (1986). A theory of ambiguity, credibility, and inflation under discretion and asymmetric information. *Econometrica 54*(5), 1099–1128.

Dal Bo, E. (2006). Regulatory capture: A review. *Oxford Review of Economic Policy 22*(2), 203–225.

Davies, S. W. and E. D. Van Wesep (2018). The unintended consequences of divestment. *Journal of Financial Economics 128*(3), 558–575.

Dekel, E., M. O. Jackson, and A. Wolinsky (2008). Vote buying: General elections. *Journal of Political Economy 116*(2), 351–380.

Dixit, A. (1987). Strategic behavior in contests. *American Economic Review 77*(5), 891–898.

Doettling, R., D. Levit, N. Malenko, and M. Rola-Janicka (2025). Voting on public goods: Citizens vs shareholders. *Working Paper*.

Edmans, A., D. Levit, and J. Schneemeier (2026). Socially responsible divestment. ECGI Finance Working Paper No. 823/2022.

Faccio, M. and L. Zingales (2021, April). Political determinants of competition in the mobile telecommunications industry. *Review of Financial Studies 35*(4), 1983–2018.

Feigenbaum, J., A. Hertel-Fernandez, and V. Williamson (2018). From the bargaining table to the ballot box: Political effects of right to work laws. Working Paper 24259, National Bureau of Economic Research.

Fershtman, C. and K. L. Judd (1987). Equilibrium incentives in oligopoly. *American Economic Review 77*(5), 927–940.

Freixas, X., R. Guesnerie, and J. Tirole (1985). Planning under incomplete information and the ratchet effect. *Review of Economic Studies 52*(2), 173–191.

Geelen, T., J. Hajda, and J. Starmans (2022). Sustainable organizations. *Working Paper*.

Grossman, G. M. and E. Helpman (1994). Protection for sale. *American Economic Review 84*(4), 833–850.

Grossman, G. M. and E. Helpman (2001). *Special Interest Politics*. Cambridge, MA: MIT Press.

Grotteria, M. (2024, March). Follow the money. *Review of Economic Studies 91*(2), 1122–1161.

Gupta, D., A. Kopytov, and J. Starmans (2025). The pace of change: Socially responsible investing in private markets. *Review of Financial Studies*. Forthcoming. https://doi.org/10.1093/rfs/hhaf083.

Hansen, J. M. (1985). The political economy of group membership. *American Political Science Review 79*(1), 79–96.

Hart, O. and L. Zingales (2017). Companies should maximize shareholder welfare not market value. *Journal of Law, Finance, and Accounting 2*(2), 247–275.

Hart, O. and L. Zingales (2022). The new corporate governance. *University of Chicago Business Law Review 1*(1), 195–216.

Heinkel, R., A. Kraus, and J. Zechner (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis 36*(4), 431–449.

Huxham, M. and D. Sumner (1999). Emotion, science and rationality: The case of the Brent Spar. *Environmental Values 8*(3), 349–368.

Jordan, G. (1998). Indirect causes and effects in policy change: The Brent Spar case. *Public Administration*.

Kennard, A. (2020). The enemy of my enemy: When firms support climate change regulation. *International Organization 74*(2), 187–221.

Konrad, K. A. (2009). *Strategy and Dynamics in Contests*. Oxford: Oxford University Press.

Kydland, F. E. and E. C. Prescott (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy 85*(3), 473–491.

Laffont, J.-J. and J. Tirole (1988). The dynamics of incentive contracts. *Econometrica 56*(5), 1153–1175.

Laffont, J.-J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation*. MIT Press.

Lenox, M. J. and C. E. Eesley (2009). Private environmental activism and the selection and response of firm targets. *Journal of Economics & Management Strategy 18*(1), 45–73.

Nagaraj, E. (2026, June). Shadow lobbying during tax audits. Working paper, University of Florida.

Oehmke, M. and M. M. Opp (2025). A theory of socially responsible investment. *Review of Economic Studies 92*(2), 1193–1225.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

Panagopoulos, C., D. P. Green, J. Krasno, M. Schwam-Baird, and K. Endres (2020). Partisan consumerism: Experimental tests of consumer reactions to corporate political activity. *Journal of Politics 82*(3), 996–1007.

Pástor, Ľ., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. *Journal of Financial Economics 142*(2), 550–571.

Pastor, L. and P. Veronesi (2021). Inequality aversion, populism, and the backlash against globalization. *Journal of Finance 76*(6), 2857–2906.

Potters, J. and F. van Winden (1992). Lobbying and asymmetric information. *Public Choice 74*(3), 269–292.

Rogoff, K. (1985, November). The optimal degree of commitment to an intermediate monetary target. *The Quarterly Journal of Economics 100*(4), 1169–1189.

Sklivas, S. D. (1987). The strategic choice of managerial incentives. *RAND Journal of Economics 18*(3), 452–458.

Snyder, J. M. (1991). On buying legislatures. *Economics & Politics 3*(2), 93–109.

Stigler, G. J. (1971). The theory of economic regulation. *The Bell Journal of Economics and Management Science 2*(1), 3–21.

Strine, Jr., L. E. and N. Walter (2015). Conservative collision course? The tension between conservative corporate law theory and *Citizens United*. *Cornell Law Review 100*(2), 335–390.

Thomas, H. F. and T. M. LaPira (2017). How many lobbyists are in Washington? Shadow lobbying and the gray market for policy advocacy. *Interest Groups & Advocacy 6*(3), 199–214.

Urvoy, C. (2025). Organized voters: Elections and public funding of nonprofits. *American Economic Review 115*(1), 183–219.

Vickers, J. (1985). Delegation and the theory of the firm. *The Economic Journal 95*, 138–147.

Warren, E. (2024). Accountable capitalism act. S. 5493, 118th Congress, 2nd Session. Introduced in the U.S. Senate, December 11, 2024.

Zingales, L. (2017). Towards a political theory of the firm. *Journal of Economic Perspectives 31*(3), 113–130.

A Appendix

Proof of Lemma 1. Write $x = f(l_e) - \Phi(a, l_e)$ for policy at $l = l_e$. Since $\frac{\partial x}{\partial a} = -\Phi_a$, the derivative of (4) is

$$W_a = 2(\Phi_a\, x - \gamma a). \tag{28}$$

Any critical point with $a > 0$ therefore satisfies (9), which requires $x > 0$; this proves $x(l) > 0$ for all $l > 0$. Differentiating (28) and using (9) to substitute $x = \gamma a/\Phi_a$,

$$W_{aa} = -2(\gamma\Theta + \Phi_a^2). \tag{29}$$

By Assumption (6), $W_{aa} < 0$ at every critical point, so every critical point is a strict local maximum and there is exactly one; W is therefore strictly quasiconcave on the region where (9) can hold, and α is unique. For $l_e > 0$, $W_a(0) = 2\Phi_a(0, l_e) f(l_e) > 0$, so $\alpha > 0$; for $l_e = 0$, $x \le 0$ for all a and $W_a < 0$ for $a > 0$, so $\alpha(0) = 0$. ■

Proof of Lemma 2. Separability (10) is immediate from (7), since l enters only through $f(l)$ and l_e only through $\Phi(a^*(l_e), l_e)$. Invoking the Implicit Function Theorem on (9) with respect to l_e and substituting $x = \frac{\gamma a}{\Phi_a}$ and $\frac{\partial x}{\partial l_e} = f' - \Phi_l$, we get

$$\frac{da^*}{dl_e} = \frac{\frac{\gamma a\, \Phi_{a\,l}}{\Phi_a} + \Phi_a\big(f' - \Phi_l\big)}{\gamma\Theta + \Phi_a^2}. \tag{30}$$

Since $h = \Phi(a^*(l_e), l_e)$,

$$\begin{aligned} h' &= \Phi_a \frac{da^*}{dl_e} + \Phi_l \\ &= \Phi_a \frac{\dfrac{\gamma a\, \Phi_{a\,l}}{\Phi_a} + \Phi_a\big(f' - \Phi_l\big)}{\gamma\Theta + \Phi_a^2} + \Phi_l \\ &= \frac{\gamma a\, \Phi_{a\,l} + \Phi_a^2 f' + \gamma\Theta\, \Phi_l}{\gamma\Theta + \Phi_a^2}, \end{aligned} \tag{31}$$

where the last line collects the two occurrences of Φ_l. Every term in the numerator of (31) is strictly positive under assumptions (6), (5), and $\Phi_l > 0$, hence $h' > 0$.

Finally, $x'(l) = f'(l) - h'(l)$ and, using (31),

$$\begin{aligned} x' < 0 &\iff f'\big(\gamma\Theta + \Phi_a^2\big) < \gamma a\,\Phi_{a\,l} + \Phi_a^2 f' + \gamma\Theta\,\Phi_l \\ &\iff \gamma\Theta\, f' < \gamma a\,\Phi_{a\,l} + \gamma\Theta\,\Phi_l \\ &\iff a\,\Phi_{a\,l} > \Theta\big(f' - \Phi_l\big), \end{aligned} \tag{32}$$

which is (12). The terms in $\Phi_a^2 f'$ cancel between the two sides, and dividing by $\gamma > 0$ removes the mobilization cost. At $l = 0$ we have $a^* = 0$ and, since $\Phi(0, l_e) = 0$ for all l_e, also $\Phi_l(0,0) = 0$; (12) then reads $0 > \Theta(0)f'(0)$, which fails, so $x'(0) > 0$, as required. ■

Proof of Proposition 1. We first prove $x_T > 0$. From Lemma 1, $a^*(0) = 0$. Thus, if on the contrary $x_T < 0$, then it must be $a^*(l^T) > 0$, and hence, $l_T > 0$. But if so, the firm is strictly better off choosing $l = 0$, in this case, $x\,(0) = 0$ so the firm obtains a better policy and a lower lobbing cost, creating a contradiction. Lemma 2 implies $x'(0) > 0$, and hence $x^T = 0$ is also suboptimal. Therefore, it must be $x_T > 0$ and $l_T > 0$.

We next show that $l_T \leq l_{\max}$. The firm's payoff depends on l only through the policy $x(l)$ and the cost $c\,l^2$, so among all lobbying levels that deliver the same policy it must choose the cheapest. Suppose, then, that $l_T > l_{\max}$. By definition of $l_{\max}$ we have $x(l_T) \leq x(l_{\max})$, while Lemma 2 gives $x(0) = 0 < x(l_T)$. Since x is continuous, the intermediate value theorem yields some $l' \in (0, l_{\max}]$ with $x(l') = x(l_T)$. Such an l' implements exactly the same policy at strictly lower cost, because $l' \leq l_{\max} < l_T$, contradicting the optimality of l_T.

Thus, the firm's problem is reduced to choosing $l \in [0, l_{max}]$ to maximize Π_T. Since Π_T is continuous, a solution exists. For interior solution, the FOC is

$$(b_F - x(l))x'(l) - cl = 0 \iff l = \frac{1}{c}x'(l)\,(b_F - x(l))\,. \tag{33}$$

We argue $x'(l_T) > 0$. Since $l^T > 0$ and $x^T < b_F$ (Lemma 3), the FOC implies $x'(l_T) > 0$. Since $x'(l_{max}) = 0$, this also implies that $l_T \neq l_{max}$, and hence $l_T < l_{max}$.

We prove $l_T < \bar{l}$. By Lemma 3, the firm never chooses l such that $b_F \leq x(l)$. Thus, consider $l \geq \bar{l}$ satisfying $b_F > x(l)$. Then,

$$cl \geq c\bar{l} > f'(\bar{l})b_F \geq f'(l)b_F \geq f'(l)(b_F - x(l)) \geq \big(f'(l) - h'(l)\big)(b_F - x(l)), \tag{34}$$

where the inequalities follow from $l \geq \bar{l}$, Assumption (3), $f'' \leq 0$, $b_F > x(l)$, and $h'(l) > 0$, respectively. Since $x'(l) = f'(l) - h'(l)$, this chain of inequalities implies $cl > x'(l)(b_F - x(l))$, which establishes that $\Pi_T'(l) < 0$. Thus, any $l \geq \bar{l}$ is strictly suboptimal, guaranteeing that $l_T < \bar{l}$.

Consider the comparative statics with respect to c and b_f. Since $x(l)$ is invariant to b and c, differentiating (33), $\frac{\partial^2 \Pi_T}{\partial l \partial b} = 2x'(l_T)$ and $\frac{\partial^2 \Pi_T}{\partial l \partial c} = -2l_T$, so at an interior optimum with $\Pi''_T < 0$ the implicit function theorem gives

$$\frac{\partial l_T}{\partial b} = -\frac{2x'(l_T)}{\Pi''_T} > 0, \qquad \frac{\partial l_T}{\partial c} = \frac{2l_T}{\Pi''_T} < 0,$$

using $x'(l_T) > 0$ and $l_T > 0$. Since $x_T = x(l_T)$, the same sign of $x'(l_T)$ transmits both effects to policy. ■

Proof of Proposition 2. Recall $x(l, l_e) = f(l) - h(l_e)$. The firm's problem is choosing $l \in [0, \bar{l}]$ to maximize Π. Since Π is continuous, a solution exists. For interior solution, the FOC is

$$(b_F - x(l, l_e))f'(l) - cl = 0 \iff l = \ell(l_e) \equiv \frac{1}{c} f'(l) \left(b_F - x(l, l_e)\right).$$

At $l_e = 0$, the FOC reads as $(b_F - f(l))f'(l) = 0$, and since $(b_F - f(0))f'(0) > 0$, we have $\ell(0) > 0$. By Lemma 3, the firm never lobbies such that $b_F \leq x$, and hence $\ell(l_e) > 0$ for all $l_e > 0$. The SOC is

$$(b_F - f(l) + h(l_e))f''(l) - (f'(l))^2 - c < 0.$$

Substituting in $l = \ell(l_e)$, the SOC becomes

$$cl\frac{f''(l)}{f'(l)} - (f'(l))^2 - c < 0,$$

which always hold by the assumption $f'' \leq 0$. Moreover, applying the Implicit Function Theorem on the FOC, and using the SOC,

$$\begin{aligned} \frac{\partial \ell(l_e)}{\partial l_e} &= \frac{f'(l)h'(l_e)}{c - f''(l)\big(b_F - f(l) + h(l_e)\big) + \big(f'(l)\big)^2} \\ &= \frac{f'(l)h'(l_e)}{c - cl\frac{f''(l)}{f'(l)} + \big(f'(l)\big)^2} > 0. \end{aligned} \tag{35}$$

where the second equality follows from substituting in the FOC.

If a rational expectations equilibrium exists, then it must solve

$$l^* = \ell(l^*) \iff l^* = \frac{1}{c} f'(l^*) \left(b_F - x(l^*)\right). \tag{36}$$

Notice $l^* > 0$ and that the equilibrium policy is $x^* = b_f - c\frac{l^*}{f'(l^*)}$.

Next, we prove that a stable equilibrium exists. A stable equilibrium exists if and only if the best-response function $\ell(l)$ intersects the 45-degree line from above. Recall that $\ell(0) > 0$. Furthermore, we show that $\ell(l) < l$ for all $l \geq \bar{l}$ such that $x(l) \leq b_F$. Indeed,

$$cl \geq c\bar{l} > f'(\bar{l})b_F \geq f'(l)b_F \geq f'(l)(b_F - x(l)), \tag{37}$$

where the inequalities follow from $l \geq \bar{l}$, Assumption (3), $f'' \leq 0$, and $b_F - x(l) \geq 0$, respectively. This chain of inequalities implies $cl > f'(l)(b_F - x(l))$, which establishes that $\ell(l) < l$. Consequently, no $l \geq \bar{l}$ can be an equilibrium. Since $\ell(0) > 0$ and $\ell(\bar{l}) < \bar{l}$, the Intermediate Value Theorem guarantees the existence of a fixed point $l^* \in (0, \bar{l})$ such that $\ell(l^*) = l^*$, where $\ell(l)$ intersects the 45-degree line from above. Notice that the smallest solution of $\ell(l) = l$, is necessarily stable, and that would be an equilibrium.

Next, notice that $\lim_{c\to\infty} \ell'(l) = 0 < 1$ for every $l \in [0, \bar{l}]$. Therefore, for c sufficiently large, $\ell(l) = l$ has a strictly positive unique solution.

Finally, consider the comparative statics. Applying the Implicit Function Theorem on (35):

$$\begin{aligned}
\frac{\partial \ell(l^*)}{\partial c} &= \frac{-\ell}{-cl\frac{f''(\ell)}{f'(\ell)} + (f'(\ell))^2 + c} < 0 \\
\frac{\partial \ell(l^*)}{\partial b_F} &= \frac{f'(\ell)}{-cl\frac{f''(\ell)}{f'(\ell)} + (f'(\ell))^2 + c} > 0.
\end{aligned}$$

In any stable interior equilibrium, $\ell(l)$ crosses the the 45 degrees line from above, and hence $\ell'(l^*) < 1$. Thus, by the implicit function theorem

$$\frac{\partial l^*}{\partial c} = -\frac{\frac{\partial \ell(l^*)}{\partial c}}{\ell'(l^*) - 1} < 0 \quad \text{and} \quad \frac{\partial l^*}{\partial b_F} = -\frac{\frac{\partial \ell(l^*)}{\partial b_F}}{\ell'(l^*) - 1} > 0,$$

as required.

Next, differentiating the equilibrium relationship $x^* = b_F - \frac{cl^*}{f'(l^*)}$ directly yields the explicit derivatives:

$$\frac{\partial x^*}{\partial c} = -\frac{l^*}{f'(l^*)} - c\frac{\frac{\partial l^*}{\partial c} f'(l^*) - l^* f''(l^*)\frac{\partial l^*}{\partial c}}{(f'(l^*))^2}, \tag{38}$$

$$\frac{\partial x^*}{\partial b_F} = 1 - c\frac{\frac{\partial l^*}{\partial b_F} f'(l^*) - l^* f''(l^*)\frac{\partial l^*}{\partial b_F}}{(f'(l^*))^2}. \tag{39}$$

To sign these derivatives concisely, define $\Delta \equiv c\left(1 - l^*\frac{f''(l^*)}{f'(l^*)}\right) > 0$. Factoring $\frac{\partial l^*}{\partial c}$ and $\frac{\partial l^*}{\partial b_F}$ in

(38) and (39) simplifies the conditions to:

$$\begin{aligned}\frac{\partial x^*}{\partial c} < 0 &\iff \frac{\partial l^*}{\partial c} > -\frac{l^*}{\Delta} \iff \frac{-\ell(l^*)}{\Delta + f'(l^*)\left(f'(l^*) - h'(l^*)\right)} > -\frac{l^*}{\Delta} \\ &\iff f'(l^*) - h'(l^*) > 0 \iff x'(l^*) > 0. \end{aligned} \tag{40}$$

$$\begin{aligned}\frac{\partial x^*}{\partial b_F} > 0 &\iff \frac{\partial l^*}{\partial b_F} < \frac{f'(l^*)}{\Delta} \iff \frac{f'(l^*)}{\Delta + f'(l^*)\left(f'(l^*) - h'(l^*)\right)} < \frac{f'(l^*)}{\Delta} \\ &\iff f'(l^*) - h'(l^*) > 0 \iff x'(l^*) > 0. \end{aligned} \tag{41}$$

Thus, both $\frac{\partial x^*}{\partial c} < 0$ and $\frac{\partial x^*}{\partial b_F} > 0$ hold if and only if $x'(l^*) > 0$. ■

Proof of Lemma 4. Recall that l_T solves

$$l = \frac{1}{c}\big(f'(l) - h'(l)\big)\big(b_F - x(l)\big),$$

while $\ell(l)$ solves

$$l = \frac{1}{c} f'(l)\big(b_F - x(l)\big).$$

Since $h'(l) > 0$, the right-hand side of the former is strictly smaller than that of the latter for all $l \geq 0$. In particular, this implies $\ell(l_T) > l_T$. Because $\ell(\bar{l}) < \bar{l}$ by Assumption (3), the Intermediate Value Theorem guarantees the existence of a stable fixed point $l^* \in (l_T, \bar{l})$ such that $\ell(l^*) = l^*$ and $\ell'(l^*) < 1$.

Furthermore, if $\Pi_T(l)$ is strictly concave, then

$$l < \frac{1}{c}\big(f'(l) - h'(l)\big)\big(b_F - x(l)\big)$$

for all $l \in (0, l_T)$, which implies $l < \ell(l)$ on this interval. Consequently, even the smallest rational expectations equilibrium must satisfy $l^* > l_T$. ■

Proof of Proposition 3. Fix $b_F > \bar{b}$. By Proposition 2, if l_b^* is the smallest rational-expectations equilibrium under a manager with ideal point $b \in (0, b_F]$, then $\Lambda(l_b^*) = b$. Moreover, l_b^* is strictly increasing in b.

Let l_F^* be any equilibrium under b_F, so $\Lambda(l_F^*) = b_F$. If $l < l_{max}$ then $x'(l) > 0$, and since $f'' \leq 0$, we have $\Lambda'(l) > 0$. Therefore, if $l_F^* \leq l_{max}$ then $\Lambda(l_F^*) \leq \Lambda(l_{max}) = \bar{b} < b_F$, a contradiction. Hence every equilibrium under b_F, satisfies $l_F^* > l_{max}$.

Since $0 < x_F^* < x(l_{max})$, x is continuous and strictly increasing on $[0, l_{max}]$ with $x(0) = 0$, there is a unique $\underline{l} \in (0, l_{max})$ with $x(\underline{l}) = x_F^*$. Set $\underline{b} \equiv \Lambda(\underline{l})$. Because Λ is strictly increasing on $[0, l_{max}]$ with $\Lambda(0) = 0$, we have $0 < \underline{b} < \Lambda(l_{max}) = \bar{b} < b_F$, and $\underline{l}$ is the smallest solution of

$\Lambda(l) = \underline{b}$, i.e. $l^*(\underline{b}) = \underline{l}$. Notice that, since $\underline{l}$ increases in x_F^*, which decreases in b_F in this range (Proposition 2), $\underline{b}$ is decreasing in b_F.

Since l_b^* is strictly increasing in b, for every $b_M \in (\underline{b}, b_F)$ we have

$$\underline{l} \;<\; l_M^* \;<\; l_F^*. \tag{42}$$

Since $\underline{l} < l_{max} < l_F^*$, exactly one of two cases obtains. If $l_M^* \in (\underline{l}, l_{max}]$, then $x_M^* > x(\underline{l}\,) = x_F^*$ because x is strictly increasing on $[0, l_{max}]$. If $l_M^* \in (l_{max}, l_F^*)$, then $x_M^* > x(l_F^*) = x_F^*$ because x is strictly decreasing on $(l_{max}, \bar{l}\,]$. In either case $x_M^* > x_F^*$, as required.

The median citizen is therefore strictly worse off. The firm is strictly better off since the policy is closer to its ideal point

$$x_F^* < x_M^* = b_M - \frac{c\, l_M^*}{f'(l_M^*)} \;<\; b_M \;<\; b_F, \tag{43}$$

and by (42), $l_M^* < l_F^*$, so its lobbying cost strictly falls as well. Therefore, $\Pi(b_F)$ is strictly higher under delegation to b_M.

Finally, suppose $b_F \le \bar{b} = \Lambda(l_{\max})$, and recall Λ is strictly increasing on $[0, l_{\max}]$. As $\Lambda(0) = 0 < b_F \le \Lambda(l_{\max})$, the intermediate value theorem places a solution of $\Lambda(l) = b_F$ in $(0, l_{\max}]$, so the smallest one satisfies $l_F^* \le l_{\max}$. For $b_M < b_F$, strict monotonicity of $l^*(\cdot)$ gives $l_M^* < l_F^* \le l_{\max}$. Both points therefore lie on $[0, l_{\max}]$, where x is strictly increasing, so $x_M^* < x_F^*$. ■

Proof of Lemma 5. A manager with ideal point b induces the unique rational-expectations lobbying level $l^*(b)$, with $\frac{\partial l^*}{\partial b} > 0$, so the firm's payoff from appointing him is $\Pi(b) = -\big(b_F - x(l^*(b))\big)^2 - c\, l^*(b)^2$. Suppressing the argument $l^*(b)$ of x, f, h and their derivatives,

$$\tfrac{1}{2}\,\Pi'(b) = \Big[(b_F - x)\, x' - c\, l^*\Big]\, \frac{\partial l^*}{\partial b}.$$

The firm values a marginal unit of lobbying at $(b_F - x)x'$ but pays its full marginal cost $c\, l^*$, which by the appointed manager's first-order condition equals the manager's own valuation, $c\, l^* = (b - x) f'$. Substituting this, and in the first line also $x' = f' - h'$, gives two equivalent forms of the bracket:

$$\begin{aligned} (b_F - x)\, x' - c\, l^* &= (x - b_F)\, h' - (b - b_F)\, f' \\ &= -(x - b_F)\, x' - (b - x)\, f'. \end{aligned} \tag{44}$$

Since $\frac{\partial l^*}{\partial b} > 0$, the sign of $\Pi'(b)$ is that of the bracket. Fix $b > b_F$.

Case 1: $l^*(b) \le l_{\max}$, that is, $b \le \bar{b}$. At every point of $(b_F, \bar{b}]$ the induced lobbying is at most $l_{\max}$, and $x' \ge 0$ because x is single-peaked. If $x \le b_F$, the first line of (44) is strictly negative since $h' > 0$ and $f' > 0$. If instead $x > b_F$, the second line is strictly negative: its first term is weakly negative because $x' \ge 0$, and its second is strictly negative because $x < b$ by Lemma 3. Hence $\Pi' < 0$ on $(b_F, \bar{b}]$, so Π is strictly decreasing on $[b_F, \bar{b}]$ and $\Pi(b) < \Pi(b_F)$.

Case 2: $l^*(b) > l_{\max}$. Recall $l^*(b) > l^*(b_F)$ and $x(l^*(b_F)) < b_F$ by Lemma 3. If $x(l^*(b)) \le x(l^*(b_F))$, manager b delivers a policy weakly farther from b_F than the firm's own choice, at strictly higher lobbying cost, so $\Pi(b) < \Pi(b_F)$. If instead $x(l^*(b)) > x(l^*(b_F))$, then, since x is nonincreasing beyond $l_{\max}$, we must have $l^*(b_F) < l_{\max}$ and $x(l^*(b_F)) < x(l^*(b)) \le x(l_{\max})$. By the intermediate value theorem, some $l' \in \big(l^*(b_F), l_{\max}\big]$ satisfies $x(l') = x(l^*(b))$. Since l^* is continuous and strictly increasing, some manager $b' \in (b_F, \bar{b}]$ has $l^*(b') = l'$. He delivers the same policy as b with strictly less lobbying, so by Case 1, $\Pi(b) < \Pi(b') < \Pi(b_F)$.

Therefore $\Pi(b) < \Pi(b_F)$ for every $b > b_F$. ■

Proof of Proposition 4. A manager b produces the equilibrium $l^*(b)$ and hence the policy $x\big(l^*(b)\big)$ and the lobbying cost $c\,l^*(b)^2$. The firm's payoff from appointing b is therefore $\Pi_T(l^*(b))$, where $\Pi_T(\cdot)$ is the objective of the transparent benchmark. Hence the firm can do no better than the transparent outcome.

Next, let $b^* \equiv \Lambda(l_T)$. By construction l_T is a rational-expectations equilibrium under b^*. Since l_T is the unique maximizer of Π_T, and $l^*(b)$ is strictly increasing in b, b^* is the unique b with $\Pi_T(l^*(b)) = \Pi_T(l_T)$.

To show $b^* < b_F$, substitute the transparent first-order condition $c\,l_T = \big(b_F - x_T\big)x'(l_T)$ into (26):

$$b^* = x_T + \big(b_F - x_T\big)\frac{x'(l_T)}{f'(l_T)} = x_T + \big(b_F - x_T\big)\left(1 - \frac{h'(l_T)}{f'(l_T)}\right), \tag{45}$$

using $x' = f' - h'$. Since $h' > 0$ by Lemma 2 and $b_F > x_T$ by Lemma 3, the bracket lies strictly in $(0, 1)$, so $x_T < b^* < b_F$.

Since $b^* < b_F$ and $l^*(\cdot)$ is strictly increasing, $l_T = l^*(b^*) < l^*(b_F) = l_F^*$. That $l_T < l_{max}$ follows from Proposition 1.

Finally, suppose $x(l)$ is single-peaked. For $b_F > \bar{b}$ we have $l_F^* > l_{max}$ and hence $x'(l_F^*) < 0$. By Proposition 1, l_T is strictly increasing in b_F and $x'(l_T) > 0$, so $x_T = x(l_T)$ is strictly increasing in b_F. By Proposition 2, x_F^* is strictly increasing in b_F if and only if $x'(l^*) > 0$; since $x'(l_F^*) < 0$, x_F^* is strictly decreasing in b_F. The difference $x_T - x_F^*$ is therefore strictly increasing in b_F. At $b_F = \bar{b}$ we have $l_F^* = l_{max}$, so $x_F^* = x(l_{max}) > x_T$ because $l_T < l_{max}$ strictly; the difference starts negative, and single crossing follows. The threshold $\hat{b}$ lies inside the admissible range whenever x_T exceeds $x(\bar{l}\,)$ at the largest b_F consistent with Assumption (3).

Assumption (3) restricts b_F to $(0, B)$ with $B \equiv c\bar{l}/f'(\bar{l})$. Let l_B solve $\Lambda(l) = B$. Then, a sufficient condition for $\hat{b} < B$ is

$$x(l_B) \;<\; x\big(l_T(\bar{b})\big), \tag{46}$$

i.e. the opaque policy at the top of the admissible range falls short of the commitment policy at its bottom. ■

european corporate governance institute

About ECGI

ECGI is an international non-profit membership organisation that provides a platform for debate and dialogue among academics, policymakers, and business leaders, with a focus on major corporate governance, ESG, and stewardship issues.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

european corporate governance institute

ECGI Working Paper Series in Finance

Editorial Board

Editor	**Nadya Malenko,** Professor of Finance, Boston College
Consulting Editors	**Renée Adams,** Professor of Finance, University of Oxford
	Franklin Allen, Professor of Finance and Economics, Imperial College London
	Julian Franks, Professor of Finance, London Business School
	Mireia Giné, Professor of Finance, IESE Business School
	Marco Pagano, Professor of Economics, Facoltà di Economia Università di Napoli Federico II
Series Manager	**Asif Malik,** ECGI Working Paper Series Manager

https://www.ecgi.global/publications/working-papers