Strategic innovation and lobbying in response to regulatory uncertainty

Andrew Hultgren

University of Illinois Urbana-Champaign, Department of Agricultural and Consumer Economics*

January 10, 2025

Click here for most recent version.

Abstract

The prospect of environmental regulation can incentivize innovation. However, firms facing potential regulation also have the strategic option to lobby the regulator, increasing or decreasing the expected innovation payoff. The anticipation effects of forward-looking firms make unbiased estimation of their response to regulation difficult, as it is hard to isolate original news about the possibility of future regulation. I overcome this problem by identifying shocks to scientific knowledge from discoveries by independent researchers of previously unknown chemical harms. Scraping the epidemiological literature for scientific publications of chemical harms, I review over 7,000 studies to construct a dataset of initial discoveries. Randomness in the timing and content of these discoveries provides empirical identification. I then ask what is the direct effect of uncertain regulation on firm innovation and lobbying, and how does the option to lobby affect innovation incentives? I show that whether innovation and lobbying are complements or substitutes in firm strategies hinges on whether the firm believes it gains a competitive advantage under future regulation. Empirically, I find that firms slowly increase innovation investments by a total of 1.5% of their market capitalization over the six years following a discovery, but they quickly increase lobbying investments by a total of 17% of pre-period lobbying levels over a 12-18 month period. Further, I find evidence consistent with innovation and lobbying as substitutes in firm strategies. This latter result implies that the option to lobby over regulation may induce an average loss in firm innovation.

*I am indebted to Maximilian Auffhammer, Michael Greenstone, Solomon Hsiang, and James Sallee for their support and feedback on this project. I thank Ashish Arora, Severin Borenstein, Christopher Costello, Thomas Covert, Lucas Davis, Meredith Fowlie, Eyal Frank, Bård Harstad, Koichiro Ito, Jennifer Kao, Larry Karp, Ashley Langer, Jeffrey Perloff, Jacquelyn Pless, David Popp, Gordon Rausser, Joseph Shapiro, Leo Simon, Reed Walker, Brian Wright, and seminar participants at the Energy and Environmental Innovation AEA at ASSA, Berkeley ERE and Climate Lunch, Stanford Political Economics of Environmental Sustainability, UChicago EPIC, and UCSB Occassional Workshop seminars for helpful comments and suggestions. I thank Sarah Dang, Hannah Maeder, Beau Smit, and Xiaoting Sunfor excellent research assistance. All errors and omissions are my own. Corresponding address: ahultgr@illinois.edu.

1 Introduction

Innovation is a critical driver of economic growth and human well-being (Romer, 1990), and environmental regulation can drive innovation, even changing the direction of technical change (Schumpeter, 1942; Porter, 1996; Popp, Newell, and Jaffe, 2010; Acemoglu et al., 2012). Innovators' inability to appropriate the full rents from their discoveries results in the underprovision of innovation (Arrow, 1962). However, a firm's option to lobby over future regulation also affects its own and others' innovation incentives. Faced with potential future regulation, forward-looking firms have a strategic incentive to innovate (should regulation come to pass), but they must trade off this incentive against their strategic option to lobby the regulator. This paper asks what is the direct effect of the possibility of regulation on firm innovation and on firm lobbying, and how does the strategic option to lobby affect the strategic incentive to innovate?

Large literatures have separately studied either firm innovation (Popp, 2006; Aghion et al., 2016; Bloom et al., 2016; Bloom et al., 2019; Akcigit et al., 2022; Dugoua, 2023) or lobbying (Grossman and Helpman, 1994; Kang, 2016; Meng and Rode, 2019; Bertrand et al., 2020) responses to regulatory and tax policy. The contribution of this paper is to recognize that firms jointly optimize over their innovation and lobbying strategies, and to estimate both firm innovation and lobbying responses to an unanticipated shock over the future regulatory state, employing for identification a novel dataset of scientific discoveries by independent researchers. Further, unlike much of the extant literature (reviewed by Popp et al., 2010 and Bombardini et al., 2020), my estimates include forward-looking firms' anticipation effects. I leverage a loophole in U.S. chemicals regulation which allowed tens of thousands of chemicals to enter the U.S. economy without testing for chronic, low-dose exposure health harms. Scraping over 7,000 epidemiological studies, I create a novel dataset of initial scientific discoveries of previously undocumented low-dose harms from widely used industrial chemicals. These discoveries are made by independent epidemiologists, not firm R&D scientists, and represent shocks to firm information sets over their regulatory future – shocks which I use for empirical identification.[1] I pair this dataset with widely-used panel data on firm innovation and lobbying expenditures.[2]

[1] All but one of the scientific discoveries I identify come from the epidemiological research community, not private sector researchers. Results are robust to dropping any discovery.

[2] Firm innovation is often measured using either R&D expenditures or patenting activity (Cohen, 2010; Popp, Newell, and Jaffe, 2010). Here I focus on expenditures data from Compustat as it parallels the structure of firm lobbying data (also firm-level expenditures) and a ready estimate of its returns to the firm exists (Hall, 1999).

Matching the timing of discoveries to the industries they affect, I employ a distributed lag model in first-differences to estimate the dynamic responses of affected firms' innovation and lobbying expenditures. My identifying variation is distinct from that in much of the literature on regulatory-induced innovation or lobbying, where widely anticipated regulatory proceedings may be employed for identification. In this study, identification rests on within-firm variation in exposure to discoveries of previously unknown health harms affecting the firm's products, controlling for sector-wide temporal shocks common to both treated and untreated firms. That is, conditional on sector-wide common information and firm-average private information, I assume a given firm does not know when the next health harm discovery by independent researchers will occur, nor if that discovery will affect its own product offerings versus those of firms operating in other industries.

To fix ideas, consider a well-documented example (Benedick, 1991; Dugoua, 2023). In the 1950s and 60s DuPont was the primary manufacturer of chlorofluorocarbons (CFCs) – a non-toxic set of products in widespread use with no known harms. However, in 1974, two atmospheric chemists posited that CFCs could *theoretically* catalyze the destruction of Earth's ozone layer. DuPont was suddenly faced with a shock to its expectations over the future probability and stringency of CFC regulation. This shock could induce DuPont to innovate, seeking to capture the market for substitute products should regulation come to pass. But DuPont also had the option to lobby the regulator. By lobbying, DuPont could change the probability or stringency of future regulation and thus the expected payoff to innovation. As discussed below, DuPont might lobby for or against regulation, implying innovation and lobbying may be complements or substitutes in the firm's strategies.[3] In this study, I expand beyond this single example to estimate firms' strategic innovation and lobbying behaviors across a wide range of scientific discoveries of previously unknown harms associated with a wide range of industrial chemicals over multiple different industries and decades.

The tradeoffs between innovation and lobbying faced by firms apply to a range of contexts outside of environmental economics. For example, they apply to the field of health economics

A disadvantage of the expenditures measure is it aggregates across a firm's innovation activities. Firm patents, while difficult to value, can provide important insight into the content of firms' innovation activities. In future work, I look to leverage these patent data.

[3]In fact DuPont both innovated and lobbied in response to the regulatory threat, initially lobbying *against* regulation to protect its existing profits and later lobbying *for* regulation once it perceived a competitive advantage in the market for substitutes.

and the regulation of new pharmaceutical drugs found to have previously-unknown side effects, or to finance and the regulation of high-frequency trading over the price volatility it introduces, or to network economics and the regulation of technology firms over questions regarding the spread of misinformation. Further, the empirical approach employed here can readily be extended to other forms of innovation (e.g. patents) and other forms of regulatory influence.

To formalize the problem, I model a simple two-stage game between firms and the regulator in which oligopolistic firms may innovate over production processes and/or lobby over the probability of regulation. The first order conditions yield the testable hypothesis that in response to an increase in the probability of future regulation, firms will increase both their innovation and lobbying investments.[4] However, whether innovation and lobbying enter firm strategies as complements or substitutes hinges on whether the firm believes it gains a competitive advantage under future regulation, either because the firm has a pre-existing advantage in the substitute market or because regulation disproportionately raises competitors' costs (Salop and Scheffman, 1983). Complementarities between innovation and lobbying strategies will vary by firm and by market, and so the average sign must be tested empirically. Finally, because lobbying by any firm affects the probability of regulation for all firms in the market, the model predicts lobbying spillovers to other firms' expected innovation payoffs and subsequent investments.

Taking the model to data, I find that in response to a new discovery of a chemical harm, firms slowly increase their innovation investments by a total of 1.5% relative to their pre-period market capitalization over a six year period, but they quickly increase lobbying investments by a total of 17% of pre-period lobbying levels over a 12-18 month period. Both results are significant at conventional levels, but lobbying response standard errors are large. Testing for heterogeneity in the dynamics of firm responses, I find that both market concentration and pre-existing lobbying capacity may be drivers of heterogeneity in firms' innovation responses, with the latter in particular explaining an important source of heterogeneity in the main innovation estimates (see Section 7.3). It also appears that pre-existing innovation capacity may drive some variation in firm lobbying responses, though these estimates are noisy. Consistent with

[4]For simplicity, the model takes the stringency of potential regulation as given and considers firm lobbying over the probability of regulation. In fact, firms may lobby over both the stringency and probability of regulation. Because these effects operate in the same direction in terms of expectations over firm payoffs, the theoretical predictions are unaffected by this simplification. See Gowrisankaran, Langer, and Zhang (2022) for an example of recent work which models firm technology adoption in response to both the stringency and probability of regulation (but does not model firm lobbying).

theoretical predictions, I find suggestive evidence that within-market lobbying spillovers play an important role in firm innovation responses, and that within-market innovation spillovers may not be as important for firm lobbying responses. The evidence I find for rival firm lobbying spillovers in firm innovation strategies complements and extends the literature on knowledge spillovers in firm innovation decisions (e.g., Bloom, Schankerman, and Van Reenen, 2013; Arora, Belenzon, and Sheer, 2021), identifying a new channel through which firm innovation may be influenced by spillovers from rivals' actions.

Across all tests employed, results are consistent with innovation and lobbying as substitutes in firm strategies on average. Multiple approaches yield results consistent with innovation and lobbying as substitutes, though the estimated effects are noisy in some. Taking the estimates at face value and drawing on a value of innovation from the survey by (Hall, 1999), my estimates imply that the option to lobby may induce in-sample firms on average to reduce innovation with a lost market capitalization value of about $135 billion.[5]

The remainder of the paper proceeds as follows: Section 2 contextualizes this study in the literature, Section 3 formalizes the problem with a simple model, Section 4 describes the construction of the dataset of scientific discoveries of unknown chemical harms and gives details on the innovation and lobbying expenditure data, Section 5 presents a descriptive analysis of the data, Sections 6 and 7 describe the empirical models and results, Section 8 gives a welfare interpretation, and Section 9 concludes.

2 Firm Innovation and Lobbying Under Regulation

Two large literatures separately document the innovation and lobbying responses of firms to the implementation of environmental regulation. But, two issues present themselves. First, the possibility of regulation is typically announced in advance of implementation. Forward-looking firms will respond before regulation is formally developed and implemented, but much of the literature does not account for this anticipation effect. Second, the literature examines firm innovation and lobbying investments over regulation separately. However, faced with the

[5]For context, firms affected by a scientific discovery at any point in my sample represent $2.4 trillion in market capitalization in the year 2020.

possibility of future regulation, firms will jointly optimize over both strategies. I discuss these issues below.

Environmental regulation can serve as a disruptive driver of new innovation (Schumpeter, 1942; Porter, 1996) affecting the direction of technical change (Hicks, 1932; Acemoglu, 1998; Acemoglu et al., 2012). A large literature has examined the effect of environmental regulation on firm innovation (see (Popp, Newell, and Jaffe, 2010) and (Popp, 2019) for recent surveys). Commonly, this literature studies the effect of the introduction of a given regulation on firm innovation, typically measured using patenting data (e.g. Lanjouw and Mody, 1996; Popp, 2006; Calel and Dechezleprêtre, 2016; Brunel, 2019; Dugoua, 2023), though innovation expenditures and new product development have also been studied (e.g. Jaffe and Palmer, 1997; Newell, Jaffe, and Stavins, 1999). Findings of increased patenting activity in response to new environmental regulations appear to be robust – especially when focusing on patents directly related to the production activity being regulated (Popp, Newell, and Jaffe, 2010).

The theoretical literature on the political economy of lobbying goes back to Grossman and Helpman (1994), in which lobby groups pay off the planner to obtain trade protections. Kamenica and Gentzkow (2011) demonstrate that side payments are not necessary for a firm to influence policy and that a well-designed signaling structure can achieve the same end. Harstad and Svensson (2011) study how firms trade off the options to legally lobby versus illegally bribe a regulator in high and low capital stock ("development") contexts. Empirical work in the area has ranged from estimating the effect of corporate lobbying activities on politician behavior such as roll call votes and the probability of legislative enactment (e.g. Ansolabehere, De Figueiredo, and Snyder Jr, 2003; Kang, 2016; Meng and Rode, 2019), to whether lobbyists sell subject matter expertise or access to politicians (i Vidal, Draca, and Fons-Rosen, 2012; Bertrand, Bombardini, and Trebbi, 2014), to charitable giving as an unregulated and subsidized form of lobbying (Bertrand et al., 2020), to returns to corporate lobbying (Kang, 2016), to correlations between lobbying expenditures and legislative consideration of new climate change regulations (Kim, Urpelainen, and Yang, 2016; Brulle, 2018). Empirical findings tend to be consistent with lobbying effects on policymaking being on-net distortionary, though gains from communication cannot be ruled out and are less studied because information transmission is difficult to observe (Bombardini and Trebbi, 2020).

Importantly, none of these works in either literature considers that innovation and lobby-

ing may in fact serve as distinct margins along which profit-maximizing firms simultaneously optimize in response to the prospect of future regulation. In the existing literature, a firm's strategic production function is effectively modeled to include either innovation or lobbying, but not both. However, as I demonstrate and estimate, a firm's payoff to lobbying depends on its innovation efforts, and its payoff to innovation depends on its lobbying efforts.

Further, identification in the extant literature is varied, often studying events (such as the rollout of new regulation) that are announced years in advance and to which forward-looking firms may have already developed a response (e.g. Jaffe and Palmer, 1997; Newell, Jaffe, and Stavins, 1999; Popp, 2006; Johnstone et al., 2012; Aghion et al., 2016; Dugoua, 2023); (Bronars and Lott, 1997; Ansolabehere, De Figueiredo, and Snyder Jr, 2003; De Figueiredo and Silverman, 2006; Bombardini and Trebbi, 2011; Kang, 2016). This anticipation effect implies estimates based on these known events may understate the true causal effects (e.g. Rittenhouse and Zaragoza-Watkins, 2018). By using initial scientific discoveries of previously-unknown chemical harms, I use for estimation unanticipated shocks to the possibility of future regulation and thus seek to avoid bias towards zero in my estimated effects due to firm anticipation.

There is a literature concerned with rent-seeking as a substitute for entrepreneurship (Murphy, Shleifer, and Vishny, 1991, 1993). This theoretical literature is similar in spirit to the types of tradeoffs I consider, and is concerned with appropriative activities that might have a higher payoff than entrepreneurship, limiting economic growth in favor of outright corruption. Correlational work by Lenway, Morck, and Yeung (1996) is the closest empirical application to the innovation and lobbying trade-off that I estimate. There, the authors study U.S. steel mills' stock price changes over four separate announcements of possible trade protection from 1977 to 1984. They find that steel mills that lobbied more on average had lower average expenditures on innovation but higher stock returns to announcements of potential protection, consistent with lobbying as rent-seeking that acts to decrease the returns to innovation in their context.

3 Theoretical Motivation

A simple model motivates the problem, and will show that when faced with an unexpected shock to the probability of future regulation, firms will: 1) increase *both* innovation and lobbying, and 2) will trade off their innovation and lobbying responses in ways that could imply the two are

complements or substitutes in firm strategies. The form of this trade-off is governed by whether the firm perceives itself to have a competitive advantage under the regulated state.

The model Firms, indexed by i, earn rents in the production of a homogeneous good q with convex costs. Production of q also emits what was thought to be a harmless byproduct, with emissions rate e_i.

Then, outside researchers discover that the byproduct emissions are in fact harmful, with marginal damages estimated at τ, and regulators consider Pigouvian taxation of these emissions. Firms now have two margins of adjustment: they can invest in R&D ρ_i to reduce the emissions rate of their production process, and they can make lobbying investments λ_i to influence the common posterior probability of regulation η. Firms endogenously choose to lobby in favor of regulation or against it. Both innovation and lobbying have decreasing marginal returns to spending by each firm.[6]

$$
\begin{aligned}
e_i = e(\rho_i), \qquad & e' < 0, \;\; e'' > 0 \\
\eta = \eta(\lambda_i^f, \lambda_i^a, \lambda_{-i}), \quad & \eta_{\lambda_i^f} > 0, \;\; \eta_{\lambda_i^f \lambda_i^f} < 0 \\
& \eta_{\lambda_i^a} < 0, \;\; \eta_{\lambda_i^a \lambda_i^a} > 0 \\
& \lambda_{-i} \equiv \{\lambda_j^f, \lambda_j^a\} \; \forall \; j \neq i
\end{aligned}
$$

Firms play a two-stage game. In stage 1, firms invest in innovation and lobbying. Then legislators vote on the tax. In stage 2, firms set their Cournot production decisions q_i based on whether the tax has passed. The firm's second stage emissions rate e_i is fixed by its first stage innovation investment. The game is solved via backward induction; firms first solve stage 2 for each case of tax or no-tax, then use this information to solve their first stage ρ_i and λ_i allocations.

The stage 2 solution for tax $t \in \{0, \tau\}$ is:

[6]In particular, this excludes decreasing marginal returns only to aggregate lobbying expenditures. If only aggregate lobbying for or against an issue mattered, then only the firms with the most to gain and lose on a given issue would lobby. All others would free ride, which is rejected by the data (e.g. (Meng and Rode, 2019)).

$$q_i^* = \underset{q_i}{\text{argmax}} \ \pi_i\Big(q_i, t, e(\rho_i), e(\rho_{-i})\Big)$$

$$\implies \pi_i^* = \pi_i^*\Big(t, e(\rho_i), e(\rho_{-i})\Big)$$

where $e(\rho_{-i}) \equiv \{e(\rho_j)\} \ \forall \ j \neq i$.[7]

In the case of taxation at $t = \tau$, denote the firm's payoff as $\pi_i^*\big(\tau, e(\rho_i), e(\rho_{-i})\big)$. This payoff depends on firm i's previous period innovation investment, which controls the firm's exposure to the emissions tax. It also depends on competitor firms' $(-i)$ previous period innovation investments, which determine the post-taxation market structure. In the case of no taxation, emissions do not enter the profit function of any firm since taxation is zero, so denote the firm's payoff as $\pi_i^*(0)$. Then, the firm's first stage problem is:

$$\max_{\lambda_i, \rho_i} \ \eta\Big(\lambda_i^f, \lambda_i^a, \lambda_{-i}\Big) \times \pi_i^*\Big(\tau, e(\rho_i), e(\rho_{-i})\Big) \ + \ \Big[1 - \eta(\lambda_i^f, \lambda_i^a, \lambda_{-i})\Big] \times \pi_i^*(0) - \rho_i - \sum_{f,a} \lambda_i^{(\cdot)}$$

That is, risk-neutral firms choose innovation and lobbying investments in order to maximize the expected value of profits.

The first order conditions are given by:[8]

$$\eta'\big(\lambda_i^f, \lambda_i^a, \lambda_{-i}\big)\Bigg[\underbrace{\pi_i^*\big(\tau, e(\rho_i), e(\rho_{-i})\big) - \pi_i^*(0)}_{\substack{\text{Profit wedge between regulated} \\ \text{and unregulated states}}}\Bigg] = 1 \tag{1}$$

$$\eta\big(\lambda_i^f, \lambda_i^a, \lambda_{-i}\big)\frac{\partial \pi_i^*\big(\tau, e(\rho_i), e(\rho_{-i})\big)}{\partial e_i}\frac{\partial e_i}{\partial \rho_i} = 1 \tag{2}$$

Equations 1 and 2 define the firm's lobbying and innovation best response strategies to the threat of regulation, given the lobbying and innovation actions of its competitors.[9] Equation 1

[7]Note that q_i^* represents the equilibrium of the Cournot game in which firms best-respond to their competitors' quantity decisions, conditional on the value of the tax and – in the case of nonzero taxation – on all firms' previous innovation investments.

[8]As discussed in the text, firm i will either lobby for regulation or against it, but never both. If lobbying for regulation, η' refers to the derivative of η with respect to λ_i^f. If lobbying against regulation, η' refers to the derivative of η with respect to λ_i^a.

[9]To fix an example, consider two identical firms i and j in Cournot competition. The equilibrium is solved via backward induction. Absent a tax in stage 2, firms best respond to the other's quantity move and in equilibrium they split the market at $q_i^* = q_j^*$, each earning profits $\pi(0)$. Under the tax $t = \tau$ in stage 2, firms again best respond to the other's quantity move, but now they take the previous innovation investments (ρ_i, ρ_j) as given.

states that firms will lobby until the marginal payoff from lobbying (the change in the probability of regulation due to lobbying times the profit wedge between the regulated and unregulated states) equals the dollar value of the marginal investment (by definition, a value of 1). Importantly, note that if the profit wedge $\pi_i^*(\tau, e_i, e_{-i}) - \pi_i^*(0) < 0$ (i.e. the firm is worse off under regulation) then it must be that $\eta' < 0$, that is the firm lobbies against regulation. However, if competitors have relatively high baseline emissions rates such that mitigation is costly, it may be that the firm gains a competitive advantage under regulation and $\pi_i^*(\tau, e_i, e_{-i}) - \pi_i^*(0) > 0$. In this case, $\eta' > 0$ and the firm lobbies *for* regulation, because even though its own costs increase under regulation, those of its competitors increase more (consistent with, e.g., Salop and Scheffman, 1983). In either case, because the profit wedge between the regulated and unregulated states depends on the innovation activities of firms i and $-i$, optimal lobbying is determined jointly with optimal innovation.

Equation 2 states that firms will invest in innovation until the expected marginal payoff (the probability of being in the regulated state times the marginal benefit of emissions-rate reductions) equals the dollar value of the marginal investment. This is true whether the firm expects to be positively or negatively affected by regulation. In a symmetric fashion to the previous first order condition, here optimal innovation is determined jointly with optimal lobbying. That is, the probability weight η in the marginal innovation payoff depends on the lobbying activities of firms i and $-i$.

Now, we wish to understand whether innovation and lobbying are complements or substitutes in this model. The cross-partial derivative on expected profits is informative.

$$\frac{\partial^2 E[\pi_i^*]}{\partial \lambda_i^{(\cdot)} \partial \rho_i} = \eta'\big(\lambda_i^f, \lambda_i^a, \lambda_{-i}\big) \frac{\partial \pi_i^*\big(\tau, e(\rho_i), e(\rho_{-i})\big)}{\partial e_i} \frac{\partial e_i}{\partial \rho_i} \tag{3}$$

Note that the sign of Equation 3 is controlled by the sign of η' (the signs of the other terms are fixed). The sign of η' is controlled by the sign of the profit wedge $\pi_i^*(\tau, e_i, e_{-i}) - \pi_i^*(0)$, which determines whether the firm lobbies for or against regulation. So, if the firm loses under regulation then innovation and lobbying are substitutes in firm strategies and Equation 3 states

For each firm, the taxed marginal emission $\tau \times e(\rho_i)$ and $\tau \times e(\rho_j)$ represents a firm-specific marginal cost shifter in the stage 2 Cournot game. Thus, firm i's best-response function in stage 2 is $q_i = q_i(q_j|\tau, \rho_i, \rho_j)$. In stage 1, both firms lobby against regulation (as profits would be strictly lower) and assign probability weights to the stage 2 payoff functions according to their lobbying investments. This becomes a single-stage game in innovation and lobbying with a pure strategy Nash equilibrium (Debreu, 1952; Glicksberg, 1952; Fan, 1952).

that the returns to innovation fall the higher the lobbying investment (and vice versa). The intuition for this result is straightforward: innovation investments pay off only in the state of the world where the tax is enacted, whereas lobbying pays off by shifting probability mass away from the tax state. Symmetrically, innovation pays off by reducing the profit wedge between the no-tax and tax states, which reduces the payoff to shifting probability mass into the no-tax state. So, the payoff to innovation decreases the more spending there is on lobbying, and the payoff to lobbying decreases the more spending on innovation.

However, if the firm gains a competitive advantage under regulation, then as discussed above $\pi_i^*(\tau, e_i) - \pi_i^*(0) > 0 \implies \eta' > 0$. In this case innovation and lobbying are complements in firm strategies and the returns to innovation increase the higher the lobbying investment (and vice versa). The intuition for this result is equally straightforward: innovation investments pay off only in the state of the world where the tax is enacted, and now lobbying pays off by shifting probability mass *into* the tax state. Symmetrically, innovation now pays off by *increasing* the profit wedge between the no-tax and tax states, which increases the payoff to shifting probability mass into the tax state. So, in this case, the payoff to innovation increases the more spending there is on lobbying, and the payoff to lobbying increases the more spending on innovation.

Thus, whether innovation and lobbying are complements or substitutes in firm strategies is theoretically ambiguous and depends on the particular market being regulated. The average sign of Equation 3 must be tested in the data.

Additionally, recall that η is influenced by lobbying from other firms as well. Note that we can apply the logic of Equation 3 to an increase in some element of λ_{-i} in an identical manner to an increase in λ_i^f or λ_i^a. Thus, by the same logic as that outlined above, a change in λ_{-i} will affect the equilibrium value of ρ_i; that is, lobbying by one firm has spillover effects on the innovation activities of other firms.[10] However, unless technology spillovers occur on the timescale of firms' strategic responses, it is not necessarily the case that the reverse is true. That is, innovation by firm i does not necessarily have spillover effects on the lobbying of a firm $-i$, because innovation returns may be private on the timescale of firms' strategic responses to the regulatory threat. I empirically test both of these spillover predictions in Section 7 and

[10]Institutionally, public reporting on lobbying activities is required every six months, providing a mechanism for information transmission over lobbying even between firms that are not already coordinating.

find suggestive support for both of them. Finally, Appendix F derives the comparative statics of the multi-firm strategic game with respect to a change in market concentration, which I test empirically in Section 7.

This simple model describes three key tests regarding how firms will respond to externality-related information shocks over the possibility of future regulation. Equations 1 and 2 predict firms will increase both lobbying and innovation in response to an information shock over the increased probability of future regulation. These are two testable hypotheses that can be taken to the data, and the same predictions hold whether the firm stands to gain or lose from regulation. Additionally, Equation 3 states that innovation and lobbying may be either complements or substitutes in firm strategies depending on whether firms perceive competitive advantage under regulation. The average sign of this effect needs to be determined from the data.

4 Constructing a dataset of scientific discoveries

As discussed previously, regulatory events are often announced years in advance of their implementation, potentially leading forward-looking firms to respond. These anticipation effects may bias existing empirical estimates employing event study or difference-in-difference models to regulatory proceedings and rollout. To address this concern, I construct a dataset of exogenous shocks to the probability of future regulation. Specifically, I construct a dataset of initial scientific discoveries of previously unknown harms from commonly used industrial chemicals. These discoveries result from the work of academic and public-sector epidemiological researchers, not private sector researchers.[11] Construction of this dataset required scraping the internet for scientific publications of potential low-dose, chronic-exposure human health harms from chemicals and reviewing over 7,000 individual publications to identify 108 initial discoveries of previously unknown human health harms from widely used chemicals. I discuss the data set construction in detail here.

[11]With one exception, as noted previously. Results are robust to dropping this, or any other, discovery.

4.1 Institutional context for introducing new chemicals in the U.S.

One might ask: why were the low-dose harms whose discoveries I document not already known at the time a given chemical was introduced into the economy? Under Section 5 of the 1976 Toxic Substances Control Act (TSCA), the EPA was to regulate chemicals if it *"...finds that there is a reasonable basis to conclude that... a chemical substance or mixture... presents or will present an unreasonable risk of injury to health or the environment..."* (15 USC 2605(a)).[12] This may sound liek a strong regulatory mandate, however several shortcomings in the TSCA new chemical review process rendered it ineffective (Schmidt, 2016; Gerlach, 2016). First, the 62,000 chemicals already in industrial use in 1976 were exempted from testing and were simply grandfathered in under the act. Second, EPA had limited ability to mandate that firms generate toxicity data for new chemicals. EPA could only require toxicity data if a chemical potentially presented an "unreasonable health risk" based on existing data. But EPA could not require data in the first place, meaning there was often insufficient existing data to prove the potential for an unreasonable health risk. Firms, of course, have no private incentive to produce this data. Third, if EPA did not make a ruling on a new chemical within a 90 day period following notification, it entered the economy by default without further regulatory requirements.[13]

By 2011, EPA had issued regulations under TSCA covering only nine of the more than 85,000 chemicals in circulation in the U.S. economy (Schmidt, 2016). Low-dose harms from new or grandfathered (pre-1976) chemicals were not required to be studied, and the U.S. has been called the "Wild West" of chemical regulation (Gerlach, 2016). Given this regulatory context, it has been left to independent scientists to identify these low-dose health effects, and so I draw from that literature for my dataset construction.

4.2 Scraping and evaluating scientific publications

[12] https://www.govinfo.gov/content/pkg/USCODE-2009-title15/html/USCODE-2009-title15-chap53-subchapI-sec2605.htm

[13] In 2016, forty years after its initial passage, Congress amended TSCA to address the grandfathering and data requirement issues discussed above (see Denison (2017) for a detailed overview). The discoveries I study here relate to chemicals introduced into the economy before the 2016 TSCA amendments. And, while improved screening capacity by EPA should improve detection of health harms, there is no reason to believe such detection will be perfect.

Scraping scientific publications I begin by scraping the internet for scientific publications of low-dose, chronic human health effects from exposure to commonly used industrial chemicals. As an independent list of chemicals to scrape, I draw on the chemicals listed on the Toxics Release Inventory (TRI; EPA, 2021) The search phrase I use requires that the chemical name be present in each matching publication. The search phrase gives additional weight to exact matches and near matches of the following set of terms: *endocrine, estrogenic, carcinogen, cancer, chronic, toxic, human, health.* The last six terms are self-explanatory. The first two terms "endocrine" and "estrogenic" describe chemicals that mimic human hormones and their interactions with the human hormonal ("endocrine") system, which can lead to various cancers (e.g. breast cancer) and other health effects. I analyze TRI chemicals in reverse order from their date of addition to the list, and have reviewed results for 78 chemicals.

Evaluating scraped publications For each chemical, the top 100 hits were returned and compiled, giving a total of 7,800 scientific publications that have been manually reviewed. These 100 findings for each chemical were reviewed in chronological order – from oldest to newest – and the *first* finding was reported of either: 1) a low-dose, chronic exposure health effect in laboratory animal testing, or 2) a human health effect associated with routine chemical use following manufacturer guidelines. For laboratory testing, I define "low dose" as under 1,000 ppm (about 65 mg/kg in rats and 140 mg/kg in mice) and "chronic" as a minimum of four weeks of laboratory exposure,[14] though many studies last for one to two years. Different types of health effects could be reported for the same chemical in different years. This process resulted in 97 scientific discoveries of low-dose, chronic exposure health effects. An additional eleven discoveries compiled from early work was also included (see Appendix E), for a total of 108 initial scientific discoveries of previously unknown health effects from commonly used industrial chemicals.

Matching discoveries to firms Because I do not observe which firms use or produce each chemical, I cannot match chemicals directly to firms in the innovation or lobbying data. However, I can match chemicals to industries. The chemicals database PubChem[15] contains an

[14]For reproductive studies in rats and mice only one week of exposure is permitted, since total gestation is only about 21 days.

[15]https://pubchem.ncbi.nlm.nih.gov

array of properties of a wide range of known chemicals, with a separate entry dedicated to each chemical. PubChem also has a section titled "Use and Manufacturing" which lists the industries in which a particular chemical is used. I draw on this field to match each chemical to the industry identifiers available in the innovation data (four-digit SIC codes[16]) and in the lobbying data ("catcode" industry codes). There are 312 four-digit SIC codes and 459 lobbying catcodes, so the two sets of industry definitions are of similar granularity.

4.3 Data sources for firm innovation and lobbying investments

Data on firm innovation and lobbying outcomes used in this study have been used in a wide range of previous studies (see Arora, Belenzon, and Sheer, 2021; Bombardini and Trebbi, 2020, for relevant reviews). Details on these data are described below.

Innovation data Firm innovation data were obtained from the Compustat Fundamentals dataset, accessed through Wharton Research Data Services. This dataset contains firm-level annual observations of key financial data from 1950–2021.[17] The reported financial data are wide-ranging and include innovation expenditures, market capitalization, net income, and sales as well as firm SIC sector codes. The data cover all publicly-traded firms in the U.S. and are compiled from required SEC 10-K filings. Mergers are tracked and the identifier for the parent company persists.

Firm innovation investments exhibit both many zeros and a long right tail, so I normalize a firm's innovation spending by its pre-period market capitalization, following Chan, Lakonishok, and Sougiannis (2001).[18] To avoid capturing changes in firm market capitalization when estimating treatment effects, I normalize treated firms by their pre-period average firm market capitalization over the two years ending before the first scientific discovery affecting them. Untreated firms do not have a clearly defined "baseline" period and so are normalized by their panel-average market capitalization. See Section 6 for additional discussion. I winsorize out-

[16]Following, e.g., Bloom, Schankerman, and Van Reenen (2013).

[17]Quarterly data are also available, starting in 1989. Most firms report in the annual dataset, but some firms only report in the quarterly file. Data for these firms were annualized and included in this analysis.

[18]Normalizing innovation expenditures by sales is another alternative in the literature, but criticized by Jaffe and Palmer (1997) when used across industries with varying levels of market power. Jaffe and Palmer (1997) study industry aggregate innovation and so are able to take logs; the many zeros in the firm-level panel prevent me from doing this.

lier values of firm innovation (in first differences) at the top and bottom 0.5 percentiles of the distribution.

Lobbying data Firm lobbying data were obtained from the Center for Responsive Politics (CRP). Lobbying data are required to be reported semi-annually to the Senate Office of Public Records for any lobbying activities in excess of $200. The data specify the lobbyist, the client paying for the lobbying, the subject matter being lobbied, and a general identifier of the body lobbied (e.g. "House", "Senate", or a specific Federal agency), from January 1998–June 2021. Because the lobbying panel is much shorter than the innovation panel, fewer scientific discoveries enter the lobbying regressions than the innovation regressions below.

Like the innovation data, firm lobbying data exhibit both many zeros and a long right tail. However, unlike the innovation data, the lobbying data only contain information on firm lobbying expenditures. No other firm covariates (e.g. market capitalization, net income, or sales) are present to account for scale effects. This leads me to normalize a firm's lobbying spending to the only measure available: the firm's average pre-period level of lobbying. I avoid changes in the denominator when estimating treatment effects by normalizing treated firms by their pre-period average lobbying spending over the three half-year periods ending before the first scientific discovery affecting them. Untreated firms do not have a clearly defined "baseline" period and so are normalized by their panel-average lobbying expenditures. Figure 2e indicates that this normalization is stable over time. Consistent with the approach in the innovation data, I winsorize outlier values of firm lobbying (in first differences) at the top and bottom 0.5 percentiles of the distribution.

5 Descriptives: Scientific discoveries and firm outcomes

The process of scraping scientific publications of low-dose harms to human health, reviewing the over 7,000 publications to identify first-time discoveries of previously unknown chemical harms, and matching the chemical behind each discovery to its associated industries in the innovation and lobbying data results in a novel dataset of 108 discoveries affecting 60 industries, with discoveries occurring in 44 distinct years and affecting over 300 firms in the balanced panel. Here I present an overview of these data.

Cross sectional variation in innovation and lobbying Table 1 displays firm annual financial, innovation, and lobbying data (columns 1–4) averaged by SIC division. Averages across all SIC divisions are reported in the bottom row. Two additional columns (5–6) report SIC-division average innovation and lobbying expenditures both normalized by SIC-division average market capitalization. Rows are ordered from the highest to the lowest fraction of market capitalization committed to innovation expenditures (that is, rows are decreasing in column 5).

Several patterns emerge in the data. First, firm average innovation and lobbying both vary substantially by SIC division, but are not highly correlated with each other. Some divisions have large annual innovation and lobbying investments (e.g. Manufacturing), some have high levels of innovation but relatively low levels of lobbying (e.g. Services), some have low levels of innovation but high levels of lobbying (e.g. Finance, Insurance, & Real Estate), and some are low in both measures (e.g. Construction). Second, there is much more variation in division average innovation (two orders of magnitude) than there is in division average lobbying (a factor of about 3x), suggesting more variation across SIC divisions in returns to innovation than in returns to lobbying. Third, divisions with higher market capitalizations do appear to innovate more, but the correlation is not as clear in the sales data. This descriptively suggests innovators may chase, or seek to protect, rents. Such patterns are harder to discern in firm lobbying investments, though the more limited variation in division average lobbying may be a factor. Finally, innovation and lobbying intensities (here crudely reported as ratios of SIC division averages, columns 5–6) appear to be negatively correlated. Innovation intensities are generally higher for low lobbying intensity divisions, and vice versa.

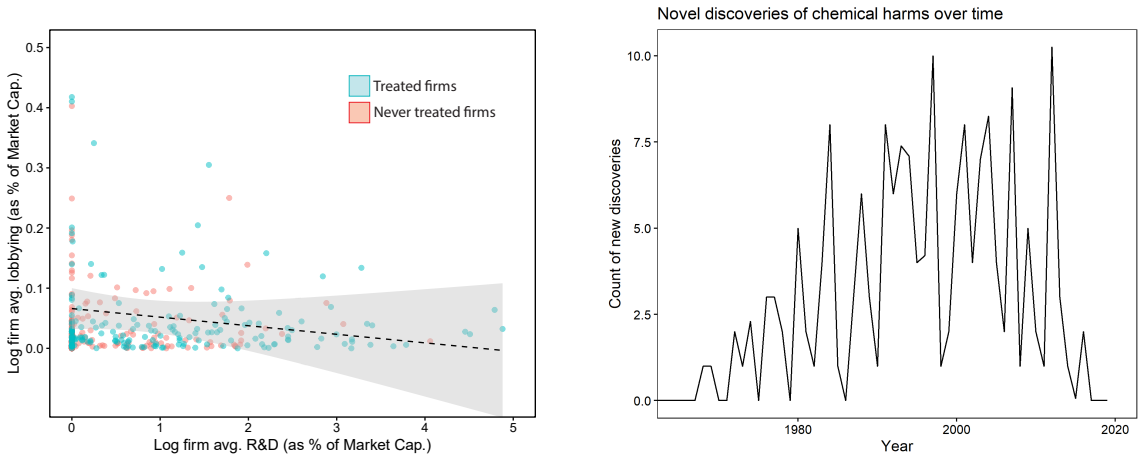
Table 1 reports SIC-division average innovation and lobbying normalized by SIC-division average market capitalization. This has the obvious problem of omitting within-division, firm-level correlations in innovation or lobbying investments and their market capitalizations. For a subset of firms that can be matched across the separate Compustat and CRP datasets (see Section 7.3.2), I normalize firm innovation and lobbying against firm market capitalization and report 4-digit SIC code averages in Figure 1(a). I separate 4-digit SIC industries that are ever exposed to a scientific discovery in my data ("Treated") from those that are not ("Untreated"). Two noteworthy features appear in the data. First, treated and untreated industries are evenly represented across levels of normalized innovation and lobbying investments. Second, the neg-

ative correlation between normalized innovation and normalized lobbying observed in Table 1 still appears in this firm-matched subset of the data.

Finally, the time series of discoveries of chemical harms (Figure 1b) displays clear temporal structure, with the rate of discoveries generally increasing over time. Economic growth, regulatory regimes, and market structures have all also varied over this period of expanding knowledge of chemical harms. Thus, addressing temporal confounds in estimation will be critically important and is discussed further below. However, it is also noteworthy that discoveries of chemical harms are not heavily clustered in any one period, but rather exhibit substantial interannual variation – a useful feature for empirical estimation.

Table 1: **Descriptive statistics of firm financial and lobbying data.** Average sales, market capitalization, R&D expenditures, and lobbying expenditures by SIC division displayed in columns (1)–(4). SIC-division-average R&D and lobbying, both normalized by SIC-division-average market capitalization, displayed in columns (5)–(6). The average scale of lobbying in the economy is about two orders of magnitude smaller than the average scale of innovation (columns (3) and (4), bottom row). SIC divisions are ordered by normalized R&D expenditures (column (5)), highlighting the negative correlation between division-wide scale-normalized innovation and lobbying. The "Public Administration" division was excluded because it consists of government entities and other uncategorized firms.

	(1)	(2)	(3)	(4)	(5)	(6)
SIC Div.	Sales ($M)	Mkt Cap ($M)	R&D ($M)	Lobby ($k)	R&D/MktCap (—)	Lobby/MktCap ($\times 10^{-3}$)
Manufacturing	2,358	2,785	79	189	0.028	0.068
Services	936	2,256	43	95	0.019	0.042
Ag, Forestry, and Fishing	943	540	4.1	113	0.008	0.210
Retail Trade	3,882	2,628	13	255	0.005	0.097
Transportation & Public Utilities	3,112	4,251	13	160	0.003	0.038
Mining	1,816	3,159	9.2	216	0.003	0.068
Wholesale Trade	2,800	824	1.7	94	0.002	0.114
Construction	1,246	663	0.48	80	0.001	0.120
Finance, Insurance, & Real Estate	1,846	1,945	0.41	203	0.0002	0.104
All Firms	2,228	2,591	39	113	0.015	0.044



(a) (b)

Figure 1: (a) Cross sectional variation in firms' innovation and lobbying investments, for a subset of firms that can be matched across the Compustat and CRP datasets (see Section 7.3.2). Innovation and lobbying normalizations are calculated at the firm level, then averaged to 4-digit SIC code industries. An industry is labeled as "treated" (blue) if it is affected by at least one discovery of a previously-unknown chemical harm at some point, and it is labeled as "untreated" (red) otherwise. (b) Time series of the count of discoveries of previously-unknown chemical harms in each year. Discoveries of harms are by independent epidemiological researchers, not R&D scientists employed at the sample firms. While the rate of new discoveries of chemical harms increases over time, discoveries are not themselves highly clustered in any one period and exhibit substantial interannual variation over all periods in the data.

Time series variation in innovation and lobbying Turning to the time series of firm investments, both innovation and lobbying are increasing in real terms for both treated and untreated firms (Panels (a) and (d) of Figure 2). Additionally, levels of innovation and lobbying are both higher on average for treated firms than for untreated firms. This points to the classic concern that, at least in outcome levels, selection into treatment is present both within the cross section as well as over time. Because the scientific discoveries of previously-unknown chemical harms that I compile only affect treated firms (by definition) and are increasing in frequency over time (Figure 1b), both the cross sectional and temporal patterns exhibited here would contribute to an upward bias in the estimated effect of a scientific discovery on both innovation and lobbying investments by firms. I address these concerns in the discussion which follows and in Section 6.

To account for scale effects, I normalize innovation investments by the firm's market capitalization (Figure 2b) and lobbying investments by the firm's average lobbying expenditure[19] (Figure 2e). These normalizations largely eliminate average cross sectional differences between treated and untreated firms. Temporal trends are reduced, but still present, and are not always parallel across treated and untreated firms (see e.g. the end of both panels).

Because innovation expenditures often exhibit a unit root (which I also cannot reject, see Section 6), I use first differences in normalized outcomes in my empirical specification. First differencing eliminates cross-sectional differences over treated and untreated firms on average, and eliminates average time trends as well (Panels (c) and (f) of Figure 2). Of course, it may be the case (for example) that some firms are trending upward in their innovation or lobbying investments while others are trending downward, and these groups of firms are differentially affected by scientific discoveries over time. These differential trends can still be present even if cross-sectional and temporal differences in aggregate are not. In my empirical specification I rely on within-firm and within-SIC-division-year variation in the timing of initial scientific discoveries of chemical harms. Thus I compare a treated firm to itself over time (first differences) and its own trend over time (firm fixed effect in first differences), and to other firms in the same sector of the economy which were not treated in that year. Identification rests on the assumption that, conditional on a firm's own average level of and trend in exposure to discoveries of chemical

[19] As noted previously, the lobbying data does not contain any firm covariate information that can be used as a measure of firm scale.

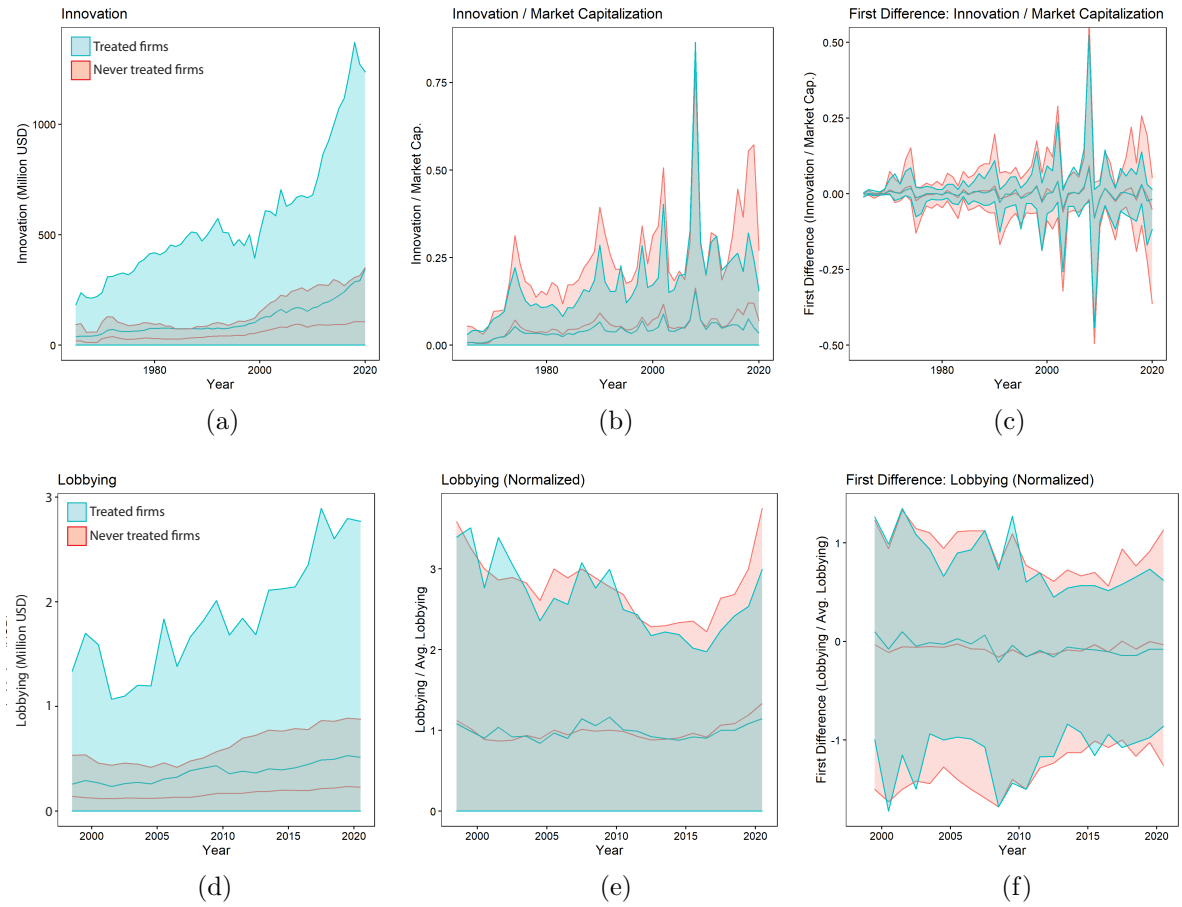


Figure 2: Time series variation in innovation (a-c) and lobbying (d-f), across treated firms (blue) and untreated firms (red). Innovation and lobbying expenditures are shown in real terms. Levels of firm innovation and lobbying investments exhibit average differences in the cross section and over time (a and d). Normalizing for scale effects (b and e) largely eliminates average cross-sectional differences, but temporal trends persist and are not always parallel even across aggregate treated and untreated categories. First differencing the normalized outcomes largely eliminates temporal trends (c and f). Shaded areas reflect the 5^{th} and 95^{th} quantiles of each year's distribution over firms.

harms, and on common sector-wide shocks and trends in discoveries, a given firm does not know if the *next* discovery of a chemical harm will affect its own operations versus those of other firms in the same sector of the economy.

One additional detail in the data is worth pointing out. The first-differenced, normalized innovation data (Figure 2c) exhibit a large positive spike for both treated and untreated firms in 2008 and a subsequent negative spike in 2009. This appears to be due to the Great Recession, which began in late 2008. Firm innovation budgets were largely set or perhaps largely spent, but their market values collapsed, leading to very high levels of innovation normalized by market capitalization in that year and a large first difference with the previous year. In the following year innovation expenditures fell and firm market values partially recovered, leading to low levels of normalized innovation and a large (negative) first difference. Interannual shocks such as these are controlled for by SIC-division-by-year fixed effects in my empirical specification.

6 Empirical specification

6.1 Estimating the innovation and lobbying responses

The first order conditions (Eq. 1 and 2) imply that firms will increase both their innovation and lobbying expenditures in response to a regulatory threat. I first test these two predictions. Begin by considering a simple model of firm innovation or lobbying investments, which may be dynamic and which scale with some measure of firm size:

$$Y_{it} = \rho Y_{it-1} + SCALE_{it} \times e_{it}$$

where Y_{it} represents the level of innovation or lobbying investment by firm i in year t, $SCALE_{it}$ is a measure of firm size, and e_{it} is a disturbance term. Such a model might arise, for example, if firm managers take last year's innovation budget as an input into planning the subsequent year's budget, but respond to aggregate shocks (such as changes in macroeconomic or regulatory conditions) in a way that scales with the size of the firm.[20]

[20]Estimation in logged scale-adjusted outcomes is typical, at least in the innovation literature where it is based in a CES production function (e.g. Bloom, Griffith, and Van Reenen, 2002). However, both the lobbying and R&D expenditures data contain many zeros at the firm level, requiring in logs that these firms be dropped or imposing other adjustments to the data. I avoid such issues by estimating in scale-adjusted levels instead.

The disturbance term e_{it} can be further decomposed into a firm-specific regulatory shock r_{it} (e.g., an information shock over the possibility of future regulation), a firm-specific trend in innovation or lobbying investments μ_i,[21] a set of sector-by-year shocks ϕ_{st} reflecting both stochasticity in the economy as well as sectoral regulatory trends and shocks, and a noise term ϵ_{it}. This yields the estimating equation

$$Y_{it} = \rho Y_{it-1} + SCALE_{it} \times \Big(\beta r_{it} + \mu_i + \phi_{st} + \epsilon_{it} \Big).$$

In the literature on firm innovation, it has been difficult to reject a unit root in the data generating process (e.g., Okunade and Murthy, 2002; Apergis, Economidou, and Filippidis, 2008), and I cannot reject a unit root in this analysis for either firm innovation or lobbying expenditures.[22] This leads me to first-difference innovation and lobbying expenditures, to eliminate potential confounding from a unit root in the undifferenced data (Greene, 2003). The objective is to isolate the firm response to the regulatory shock r_{it} for estimation, so I normalize the differenced outcome by firm scale, giving

$$\frac{\Delta Y_{it}}{SCALE_{it}} = \beta r_{it} + \phi_{st} + \epsilon_{it}.$$

Finally, three additional points remain. First, any measure of firm scale may itself be affected by an information shock over the firm's future regulatory exposure, changing the outcome variable (through its denominator) even if actual investments are held constant. Thus, I normalize by $\overline{SCALE}_i$, the average of $SCALE_{it}$ over the periods directly preceding the first regulatory shock faced by each firm.[23] Second, firm responses r_{it} may evolve over multiple periods. Such a dynamic response might occur if the perceived probability of regulation evolves

[21] If y_i^{∞} is scale-adjusted steady state investment, then conditional on sectoral and regulatory shocks, $\mu_i = (1-\rho)y_i^{\infty}$ would be consistent with a firm at steady state, with constant investment in each period perhaps replacing depreciated knowledge or lobby network stocks. μ_i different from $(1-\rho)y_i^{\infty}$ would imply time-trending behavior in firm investments, consistent with approaching a steady state from a point of disequilibrium, perhaps due to stochastic shocks to the firm's business or regulatory environment.

[22] An augmented Dickey-Fuller test with drift and a time trend fails to reject a unit root in R&D expenditures for 88% of firms at $p < 0.05$, and still fails to reject a unit root at $p < 0.10$ for 81% of firms. The failure to reject a unit root further increases if one or two lags in the dependent variable are added. The failure to reject a unit root is even higher for firms' lobbying expenditures (e.g. 91% of firms fail to reject a unit root at $p < 0.05$), and this holds regardless of lag specification.

[23] As discussed in Section 4.3, I scale innovation investments by the firm's two-year pre-period average market capitalization, which I choose instead of sales to account for differences in pricing power across firms and industry sectors (Jaffe and Palmer, 1997). I scale lobbying investments by the 18-month pre-period average lobbying, which is the only information about the firm available in that dataset.

after the initial shock (as new evidence comes to light or due to strategic actions by the players), if new innovation or lobbying programs require costly adjustments like developing influence networks or product offerings, or if firms benefit from an option value of waiting (particularly for new innovation investments). I employ a distributed lag estimator to capture the dynamics of firm responses. This estimator flexibly allows for either transitory or permanent effects of information shocks on firm innovation or lobbying investments. Third, firms may be differentially exposed to a given information shock. I capture treatment intensity in the R&D data using Compustat data on firm net sales across SIC 4-digit industries. Treatment intensity in firm innovation responses is measured as the fraction of a firm's average net sales occurring in industries affected by a given information shock. That is,

$$INTENSITY_{it} = \sum_{SIC4} \left(\overline{SALES}_{SIC4,i} \times TREAT_{SIC4,t}\right) / \sum_{SIC4} \overline{SALES}_{SIC4,i}.$$

This information is not available in the lobbying panel of data, so treatment intensity is set to 1 for all shocks in the lobbying response. The final empirical specification for firm i in industry sector s and period t is given as

$$\Delta y_{it} = \sum_{\ell \in b} \beta_\ell LAG_{\ell it} + \phi_{st} + \epsilon_{it}. \tag{4}$$

Recall that Eq. 4 represents an empirical test of the theoretical predictions (Eq. 1 and 2) that firms will increase both their innovation and lobbying expenditures in response to a regulatory threat. Here, y_{it} either represents firm annual innovation or biannual lobbying investments (normalized as described in Section 4.3), b represents the bandwidth (set of leads and lags) over which lagged effects β_ℓ are estimated, and s indexes SIC 2-digit industries in the innovation data and sector codes in the lobbying data. $LAG_{\ell it}$ is defined as $\sum_{yr \in \tau_i} \mathbb{1}\{t - \ell = yr\} \times INTENSITY_{it}$, where τ_i denotes the set of treatment years yr for firm i, $\mathbb{1}\{t - \ell = yr\}$ is an indicator for whether period t is ℓ lags from a treatment year in the set τ_i, and $INTENSITY_{it}$ is treatment intensity as described above.[24] Firm fixed effects in levels of the outcome are implicitly captured by first-differencing the outcome variable, purging the data of

[24]Typically, $\sum_{yr \in \tau_i} \mathbb{1}\{t - \ell = yr\}$ is either 0 or 1. Either the firm was exposed to a discovery of a low-dose harm ℓ lags ago, or it wasn't. However, this term can be greater than 1 if a given firm is exposed to more than one discovery in the same year.

any constant differences between firms. The term ϕ_{st} absorbs any remaining sector-wide shocks to innovation or lobbying expenditures, such as changes to innovation tax incentives, changes in the overall regulatory environment, or the cyclical nature of the economy, any of which may have differential effects across sectors of the economy. The coefficients of interest, β_ℓ, are estimated off of differences between the trajectories of within-firm deviations in innovation and lobbying investments, comparing the within-firm trajectories of treated firms to those of similar, untreated firms before and after an information shock. Standard errors are cluster-robust by firm, allowing for arbitrary correlations in innovation and lobbying activities at the firm level over time.

Because the model is estimated in first differences, the cumulative effect Ω_k of a shock on an outcome at some time k periods after the shock occurred is given by

$$\Omega_k = \sum_{\ell=0}^{k} \beta_\ell \ . \tag{5}$$

The theoretical predictions of Eqs. 1 and 2 are that $\Omega_k > 0$ for both the innovation and lobbying estimated responses. Results in Section 7 are plotted in terms of these cumulative effects Ω_k, with uncertainties that capture the full intertemporal covariance structure of the estimated dynamic response.

6.2 Testing the sign of the relation between innovation and lobbying

As described in Section 3, whether innovation and lobbying enter firms' strategic responses to uncertain regulation as complements or as substitutes depends on whether firms perceive themselves to be winners or losers under regulation. Firms might perceive themselves to be winners if competitive products are made relatively more costly by regulation or if they believe they have a competitive advantage in innovating into a new market for substitute products created by regulation. Firms might perceive themselves to be losers under future regulation if they lack a competitive advantage in innovation or if their payoff in the market for substitutes under regulation is small relative to their incumbent market payoff. Perceived winners under regulation will lobby for regulation, enhancing their innovation payoffs and further increasing the privately optimal level of innovation; i.e. innovation and lobbying are complements for these firms. The reverse holds for perceived losers under regulation. Since the theoretical prediction

is ambiguous (see Eq. 3), I construct an empirical test to estimate which effect dominates in the data. That is, I test the average sign of $\frac{\partial^2 E[\pi_i^*]}{\partial \lambda_i^{(\cdot)} \partial \rho_i}$ as given in Eq. 3.

To test whether innovation and lobbying are complements or substitutes in firm strategies, one might think to interact lobbying investments with the estimated innovation response and interact innovation investments with the estimated lobbying response. Positive interaction effects, for example, might indicate the two are complements – innovation is increasing in lobbying. However, post-period innovation and lobbying are jointly-determined endogenous outcomes of firm optimization, and associated interaction estimates are subject to simultaneity bias. Consider a situation where the rents at stake are large: firms have a larger first-order payoff to both innovation and lobbying, regardless of whether the two are complements or substitutes. As such, even if the two are substitutes, firms will both innovate more and lobby more relative to other firms (facing different shocks) operating in markets with lower rents.

To address this simultaneity bias, I take averages of pre-period innovation and lobbying investments as measures of pre-existing innovation and lobbying capacities, avoiding endogeneity with post-period responses. I then use these pre-period averages as interaction terms in the distributed lag estimators. Pre-period average lobbying is interacted with each lead and lag of the innovation response, which tests whether firms with larger pre-existing lobbying networks tend to innovate more or less in response to an information shock relative to those firms with smaller pre-existing lobbying networks. A positive association would suggest lobbying and innovation are complements in firm strategies, while a negative association would suggest the opposite. I further construct an analogous (but separate) test of innovation and lobbying complementarities by examining interactions in firm lobbying responses over their pre-period innovation investments. This is a symmetric test of the cross-partial derivative given in Eq. 3. These estimators are shown in Equations 6 and 7.

Importantly, recall the theoretical prediction that lobbying spillovers will be important for firm innovation responses (because lobbying by one firm affects the probability of regulation for all firms). This, combined with the empirical fact that industry trade groups coordinate lobbying activities among their members, leads me to account for spillovers by default in Equations 6 and 7. In particular, a firm's measure of average pre-period lobbying and innovation is not only its own investments, but the sum of all such investments by firms in the same industry (4-digit SIC code or its equivalent in the lobbying data). I empirically test for these spillover

effects in Section 7.3.2, finding consistent with theoretical predictions that lobbying spillovers to other firms' innovation responses are indeed important while innovation spillovers to lobbying responses appear not to be.

$$\Delta y_{it}^{R\&D} = \sum_{\ell \in b} \beta_{\ell}^{R\&D} LAG_{\ell it} + \sum_{\ell \in b} \gamma_{\ell}^{R\&D} LAG_{\ell it} \times \text{High Pre-Lobby}_I \; + \; \phi_{st} + \epsilon_{it} \tag{6}$$

$$\Delta y_{it}^{Lobby} = \sum_{\ell \in b} \beta_{\ell}^{Lobby} LAG_{\ell it} + \sum_{\ell \in b} \gamma_{\ell}^{Lobby} LAG_{\ell it} \times \text{High Pre-R\&D}_I \; + \; \phi_{st} + \epsilon_{it} \tag{7}$$

The notation here is identical to that of Equation 4, with the addition of the set of interactions between pre-period lobbying or innovation and each of the leads and lags ℓ. Pre-Lobby$_I$ and Pre-R&D$_I$ are indicators for pre-period lobbying and innovation being above the median value for treated firms. Pre-period expenditures are defined as the average expenditure across firms in each of the most highly resolved industry designations I available in the data.[25] The pre-period is the three periods which precede the first discovery an industry is exposed to. While this specification addresses the simultaneity issue discussed above, it should be noted that the pre-period interaction terms take one value per industry and thus represent a cross-sectional source of variation for estimating heterogeneity in firm responses.

The specifications in Equations 6 and 7 represent two separate tests of whether innovation and lobbying are complements or substitutes in firm strategies. The cumulative effects for "high" firms (e.g. $\sum_{\ell} \left(\beta_{\ell}^{R\&D} + \gamma_{\ell}^{R\&D} \times \text{High Pre-Lobby}_I\right)$) versus those for "low" firms (e.g. $\sum_{\ell} \left(\beta_{\ell}^{R\&D}\right)$) are the estimates of interest. Cumulative effects increasing in the interaction value would be consistent with innovation and lobbying as complements in firm strategies (that is, the option to lobby increases firm innovation), while cumulative effects decreasing in the interaction value would be consistent with innovation and lobbying as substitutes in firm strategies (the option to lobby decreases firm innovation). However, because these two regressions are separate estimates of the effects of interest across different outcomes, they need not agree on sign and may empirically support opposing conclusions.

[25] I indexes four-digit SIC codes in the innovation data, and their equivalent industries ("Catcodes") in the lobbying data.

7 Results

7.1 Testing first-order predictions

A simple model demonstrates that firms faced with the unexpected possibility of future regulation will increase both innovation and lobbying investments in response, and will do so until the expected marginal payoff to each investment equals the cost of investment funds to the firm (Section 3). Here, I test the prediction that firms will increase both innovation and lobbying investments in response to a shock over the prospect of future regulation.

7.1.1 The innovation response

The distributed lag model employed estimates the overall effect of a shock over the prospect of future regulation as well as the dynamic evolution of firm responses (3). Each point in Figure 3 represents the cumulative average treatment effect of a scientific discovery of a previously unknown externality on the innovation investments of affected firms ℓ years after the discovery. Here, the length of a single time period is one year. Time $\ell = 0$ is the year of publication of the discovery (by independent epidemiological researchers, not firm R&D scientists). Cumulative effects are estimated relative to innovation in the first year before the discovery, $\ell = -1$.

Firms slowly increase innovation in response to the scientific discovery of a previously unknown externality. On average across treated firms in my sample, innovation investments increase by about 1.5% of a firm's market capitalization, and this response is realized over a period of about six years following a shock. This six-year cumulative effect reflects an average annual increase in innovation investments of about 0.25% of firm market capitalization per year.

Lead effects in Figure 3 test whether firms increase innovation in the three years before a discovery occurs, offering a partial falsification test of the identifying assumption in the model. The parallel trends identifying assumption of the model is an assumption over the post-period counterfactual which can never be observed; specifically the assumption that, were it the case that treated firms were not actually treated, their cumulative path of innovation would follow that of the control firms. The lead effects of Figure 3 test this assumption in the pre-period rather than the post-period. Positive leads would indicate that pre-period innovation investments of treated firms are trending downward relative to those of control firms, raising the concern that post-period treatment effects could understate the true effect. Symmetrically, negative leads

would indicate that pre-period innovation investments of treated firms are trending upward relative to those of control firms, raising the concern that post-period treatment effects could overstate the true effect or even simply reflect a differential time trend when in fact the true effect is zero. The lead effects estimated here are reassuringly flat, indicating that, at least in the pre-period, control firms provide a good counterfactual to the innovation behavior of treated firms. Further, as demonstrated in Appendix A, pre-trends are flat even when estimating 10 years of leads and lags (though post-period estimates are noisier).

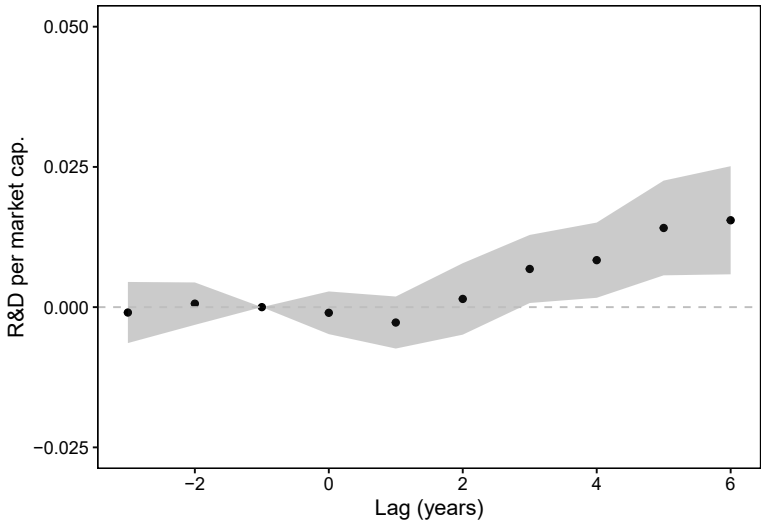


Figure 3: The cumulative effect of the discovery of a previously unknown potential externality on the innovation investments of affected firms. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect. Time 0 is the year of publication of a discovery, and the estimator includes six subsequent years of lagged effects. The estimator also includes three years of lead effects which serve as a partial test for differential trends. Grey bands represent 95% confidence intervals.

7.1.2 The lobbying response

Firms' average lobbying response to the discovery of a previously unknown externality (Figure 4) is estimated in a similar fashion to the innovation response above. Because the lobbying data are reported every half year, each point in Figure 4 represents a six-month period (versus the annual periods shown in Figure 3).

In contrast to the slower, six-year innovation response, firms rapidly increase their lobbying over a 12-18 month period in response to the scientific discovery of a previously unknown externality. The cumulative firm response is to increase lobbying investments by about 17%

of pre-period levels on average. Pre-period lead effects are generally flat, again providing a partial test of the parallel trends identification assumption and indicating that, at least in the pre-period, control firms provide a good counterfactual to the innovation behavior of treated firms. There is a non-significant increase in lobbying in the first lead (lag $\ell = -0.5$) in Figure 4, which may be due to noise in the data or may indicate limited leakage of information in the months prior to a discovery's publication (recall that lobbying is reported in six-month, rather than annual, intervals). Leads before this first lead are flat.

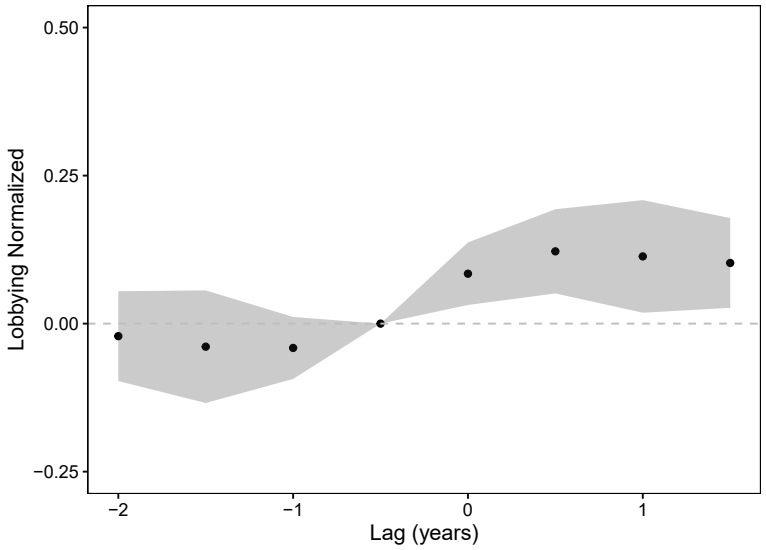


Figure 4: The cumulative effect of the discovery of a previously unknown potential externality on the lobbying investments of affected firms. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect. Time 0 is the half-year of publication of a discovery, and the estimator includes three subsequent periods (half-years) of lagged effects. The estimator also includes four periods of lead effects which serve as a partial test for differential trends. Grey bands represent 95% confidence intervals.

7.1.3 Testing pre-period parallel trends

As discussed, estimating lead effects in a distributed lag model provides a partial test of the parallel-trends assumption (in the pre-period, rather than the untestable assumption over the post-period). The main innovation and lobbying specifications include three years and two years of leads, respectively. An extended test shows results for 10 periods of leads and lags (Figure A.1). Estimated effects show more noise for two reasons. First, estimating more period treatment effects leaves less residual variation for the estimation of any given treatment effect,

leading to larger standard errors. Second, estimating effects over 10 periods requires that treated firms have more pre- and post-period data in order to enter the balanced panel, and as such fewer total firms are available for treatment effects estimation over this longer time horizon. These caveats aside, the point-estimates of overall cumulative effects are consistent with those estimated in the main models, as is the temporal structure of the responses. Importantly, leads are relatively flat over the longer pre-period for both models. These estimates, though considerably more noisy than my preferred specifications, demonstrate that the parallel trends assumption — which of course cannot be tested in the post-period — does hold for many periods before treatment. They also demonstrate that the point-estimates of the treatment effects stabilize within the shorter lag periods estimated in my preferred specifications (Figures 3 and 4), and they stabilize at similar levels.

7.1.4 Additional robustness tests

The main estimates of Figures 3 and 4 are robust to multiple variants of the main specification. These include: alternative fixed effects specifications, alternative bandwidths over which leads and lags are estimated, and dropping individual discoveries from the estimating data set (indicating that estimated average effects are not driven by any one single discovery). All of these robustness tests are available in Appendices B and C.

7.2 Market concentration as a factor in firm strategies

A large literature exists studying the role of market power in overall innovation throughout the economy (e.g., Arrow, 1962; Blundell, Griffith, and Van Reenen, 1999; Aghion et al., 2005; Cohen, 2010). Here I ask how market power, as measured using the Herfindahl-Hirschman Index (HHI), might mediate firms' strategic innovation and lobbying investments. As described briefly in Section 3 and fully derived in Appendix F, the sign of the effect of adding more firms to the strategic game I model is ambiguous. The effect on firm innovation strategies will be controlled by the differential form of rivalry presented by a new entrant across regulated states. For example, if the firm wins under regulation and has used regulation to raise barriers to entry in the form of new entrant marginal costs (e.g., Salop and Scheffman, 1983), then it may stand to lose less from new entry in the regulated state versus the unregulated state, and so the marginal

payoff to innovation increases with regulation. However, if the firm wins under regulation but faces new entrants who are disruptive innovators (e.g., Blundell, Griffith, and Van Reenen, 1999) then it may stand to lose more from new entry in the regulated state and so the marginal payoff to innovation decreases with regulation. The effect on firm lobbying strategies is likewise ambiguous, depending on how the regulator weights concentrated lobbying efforts from a few firms versus diffuse lobbying efforts from many firms.

As is standard, I calculate the HHI each year at the 3-digit SIC code level. I then run the following set of regressions to examine heterogeneity over market concentration in firms' innovation and lobbying responses.

$$\Delta y_{it}^{R\&D} = \sum_{\ell \in b} \beta_{\ell}^{R\&D} LAG_{\ell it} + \sum_{\ell \in b} \xi_{\ell}^{R\&D} LAG_{\ell it} \times \text{HHI}_{s3,t} \; + \; \phi_i + \phi_{st} + \epsilon_{it}$$

$$\Delta y_{it}^{Lobby} = \sum_{\ell \in b} \beta_{\ell}^{Lobby} LAG_{\ell it} + \sum_{\ell \in b} \xi_{\ell}^{Lobby} LAG_{\ell it} \times \text{HHI}_{s3,t} \; + \; \phi_i + \phi_{st} + \epsilon_{it}$$

Estimated heterogeneity in the innovation response (Figure 5) is noisy (high- and low-HHI cumulative responses are not significantly different at p-values of 0.12 in lag 5 and 0.22 in lag 6), though consistent in sign for all periods following the information shock. Results in Figure 5 suggest that firms operating in low-concentration markets (with relatively less market power) tend to innovate more in response to the regulatory threat than do firms in high concentration markets with relatively more market power. This is consistent with theories and empirical findings of "creative destruction" associated with the innovative threat coming from a competitive fringe (Arrow, 1962; Blundell, Griffith, and Van Reenen, 1999; Cohen, 2010), though estimates are noisy.

Little meaningful heterogeneity is uncovered in the lobbying response (Figure D.2), with both low- and high-HHI industries exhibiting similar patterns the their lobbying investments following an information shock. This includes both overlap in 95% confidence intervals as well as the magnitude of the differences between (and evening ordering of) the point estimates.

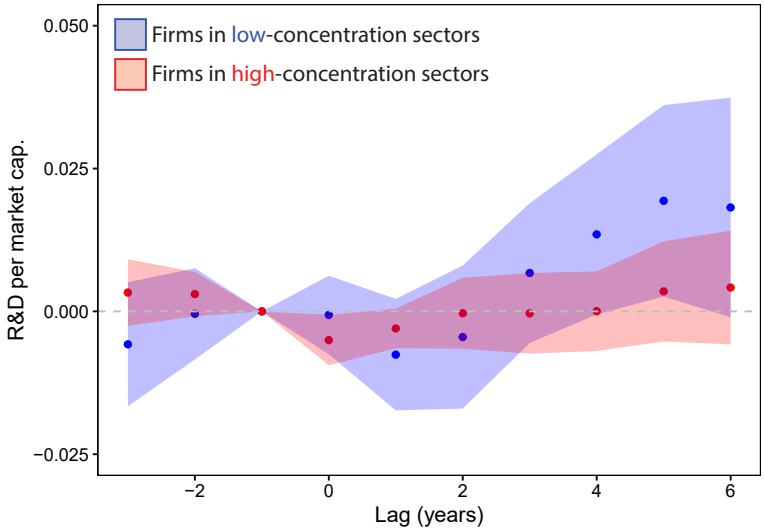


Figure 5: Heterogeneity over market concentration (HHI) in the cumulative effect of the discovery of a previously unknown potential externality on the innovation and lobbying investments of affected firms. The 90^{th} percentile of HHI in the data is 0.4, measured on the [0, 1] interval. Heterogeneity in the innovation response is noisy, though consistent in sign for all periods following the information shock. Results here suggest that firms operating in a low-concentration market (with relatively less market power) tend to innovate more in response to the regulatory threat than do firms in a high concentration market with relatively more market power. The corresponding estimate for the lobbying response is shown in Figure D.2. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

7.3 Evidence of firm substitution between innovation and lobbying

Here, I test whether innovation and lobbying are complements or substitutes in firms' strategic responses to the threat of future regulation. This test provides suggestive evidence regarding whether the option to lobby induces more or less innovation from treated firms.

7.3.1 A test for innovation and lobbying as complements or substitutes in firm strategies

As discussed in Sections 3 and 6.2, whether innovation and lobbying enter firm strategies as complements or substitutes depends on whether or not the firm perceives a larger payoff under the regulated or unregulated states. Thus it is theoretically ambiguous whether or not the option to lobby increases or decreases firm innovation incentives (and vice versa), and the average effects must be measured in data. These results are shown in Figures 6 and 7.

On average for treated firms in my estimating sample, I find that firms in sectors with lower pre-period lobbying tend to innovate more in response to the threat of future regulation (Figure 6), which is consistent with innovation and lobbying as substitutes in firm strategies. These heterogeneous responses are significantly different from one another through lag 3 (p-value 0.028) and marginally different at lag 6 (p-value 0.106), and describe an important source of heterogeneity in the main innovation response (Figure 3). Furthermore, the interaction effect estimated here is highly flexible: each lead and lag indicator in my estimator has a separate interaction with pre-period lobbying, and nothing constrains the cumulative effects to be consistently above or below each other in every period. Yet, for every post-period in the cumulative effects estimate, I find that firms in industries with below-median pre-period lobbying exhibit a larger innovation response, and firms in industries with above-median pre-period lobbying exhibit a diminished innovation response.

The symmetry of the cross-partial derivative over firm FOCs (Eq. 3) implies an additional test. If innovation and lobbying are indeed substitutes in firm strategies, then it will not only be the case that lobbying investments reduce the payoff to innovation (Figure 6), but it will also be the case that innovation investments reduce the payoff to lobbying (Equation 3). This is because innovation reduces the profit wedge induced by the threat of regulation, making lobbying less beneficial to the firm.

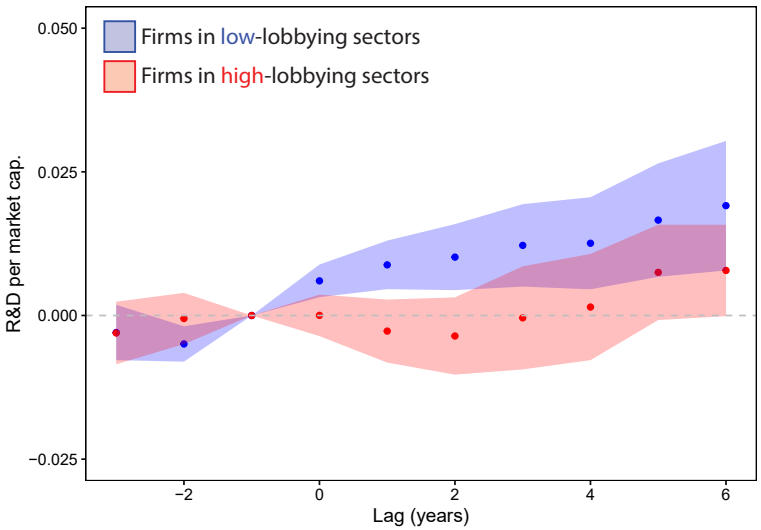


Figure 6: Heterogeneity over pre-period lobbying in the cumulative effect of the discovery of a previously unknown potential externality on the innovation investments of affected firms. These estimates show that firms that had been lobbying more in the pre-period tend to innovate less in the post-period, an important source of heterogeneity in the main innovation response (Figure 3). This is consistent with innovation and lobbying as substitutes in firm strategies: the option to lobby may induce firms to innovate less than they otherwise would have. Note that pre-period lobbying is a cross-sectional source of variation, and thus the heterogeneity in effects shown here should be considered as suggestive in nature. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

I implement this test as Equation 7, interacting an indicator for above-median industry average pre-period innovation with the firm's lobbying response. Estimates here are quite noisy (Figure 7). The point estimates are consistent with innovation and lobbying as substitutes, where firms in industries with higher pre-period innovation have lower point estimates in the lobbying response to the threat of future regulation, though results here are considerably more noisy.[26] As before, the interaction effect estimated here is highly flexible: each lead and lag indicator in my estimator has a separate interaction with an indicator over pre-period innovation and nothing constrains the cumulative effects to be consistently above or below each other in every period (and indeed there is one period where the point estimates flip).

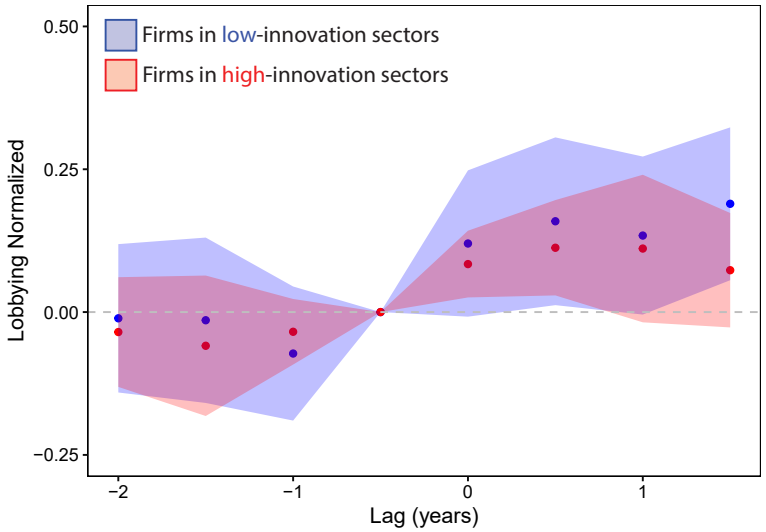


Figure 7: Heterogeneity over pre-period innovation in the cumulative effect of the discovery of a previously unknown potential externality on the lobbying investments of affected firms. It may be that firms that had been innovating more in the pre-period tend to lobby less in the post-period, though estimates are too noisy to draw clear conclusions. Pre-period innovation is a cross-sectional source of variation, and thus the heterogeneity in effects shown here should be considered as suggestive in nature. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

Turning back to heterogeneity in firm innovation responses, recall that market concentration (HHI) was also found to be a driver of heterogeneity in the innovation response. As such, we might be concerned that one interaction term is simply serving to estimate an omitted variable effect for the other. Below, I again test for heterogeneity in innovation responses over

[26]It turns out that some of this noise is due to the finding that – consistent with theoretical predictions – innovation spillovers into firm lobbying responses may not be very important. Thus the aggregate measure of industry-wide preperiod average innovation I use here likely introduces noise in the more relevant firm-specific innovation capacity. See Section 7.3.2 and Figure D.1 for detailed results.

pre-period lobbying, but this time simultaneously controlling for interactions between lag indicators and HHI. Thus the heterogeneous effects estimated here are conditional on heterogeneity in responses over market concentration (Figure 8). This estimator pushes the data harder, including main lagged effects, interacted lagged effects over a pre-period lobbying indicator, and interacted lagged effects over HHI. While the estimated effects are a little more noisy, the results are otherwise very similar to those of Figure 6, suggesting that market concentration is not an important omitted variable in that result.

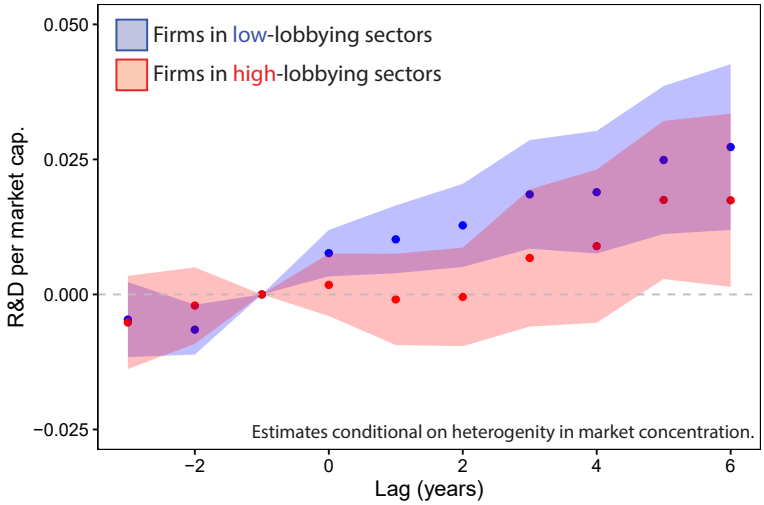


Figure 8: Heterogeneity in the innovation response over a pre-period lobbying indicator (as Figure 6), while simultaneously controlling for heterogeneity in the innovation response over market concentration (HHI). Estimates are consistent with those of Figure 6, indicating that the estimated effects of lobbying capacity on firm innovation responses is not due to HHI as an omitted variable. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

7.3.2 The role of spillovers

As discussed in Section 3, lobbying by firm j should theoretically affect innovation payoffs to firm i in a fashion directly analogous to how firm i's lobbying affects its own incentive to innovate. This is because lobbying by any firm shifts the common probability of future regulation experienced by all firms in the market, along with the fact that lobbying activities are publicly reported every six months (and in practice often coordinated by industry trade groups). However, unless technology spillovers occur on the timescale of firms' strategic lobbying responses, it is not necessarily the case that the reverse is true. That is, innovation by firm j

may not have spillover effects on the lobbying of firm i, because innovation outcomes may be private on the timescale of firms' strategic responses to the regulatory threat. Here I empirically test both of these predictions.

The left panel of Figure 9 replicates Figure 6, where the red response includes an indicator for above median pre-period average lobbying within an industry. Here, lobbying by all firms within an industry affects the innovation response of any given firm. In the right panel of Figure 9, the pre-period lobbying interaction is defined using a firm's own lobbying only, not the lobbying of other firms in the industry.[27] Here, a firm's own lobbying has almost no measurable effect on its own innovation response, suggesting the spillover effects from industry-wide lobbying empirically dominate the effects of a firm's own lobbying on its own innovation.

One important caveat must be noted here, however. A firm's own lobbying cannot fully be matched back to itself, due to the presence of industry trade groups in the lobbying data. Part of a firm's lobbying can occur through industry trade groups as intermediaries: firms donate to trade groups, who then lobby on behalf of the firm and/or industry. These trade groups are not required to report the sources of their funding. So, while I fully observe lobbying at the industry level (since trade groups are linked to their industries), I do not fully observe it at the firm level as this lobbying-via-trade-group cannot be linked back to individual firms. Because trade groups often serve as a coordinating mechanism for industry lobbying efforts, this data issue may not affect the interpretation of Figure 9. However, if trade groups lobby explicitly on behalf of their individual members rather than the industry as a whole, then there exists a portion of individual firm lobbying that I cannot observe in the data and could partially drive the null interaction observed in the right-hand panel of Figure 9.

Overall, consistent with the theoretical prediction, evidence here suggests that lobbying spillovers are an important determinant of firms' innovation strategies. This is consistent with a setting in which lobbying by any one firm shifts expectations over future regulation for all firms, generating spillovers to all affected firms' innovation investments.[28]

[27]Firm names are not standardized in the lobbying data, so matching firms across the innovation and lobbying datasets was not straightforward. To overcome this issue, I automatically queried Google for each firm name in each dataset. I then allowed the degree to which Google returned the same set of web addresses for any given pair of firm names to determine the degree to which those two firms were a match across the two datasets. Across the datasets, 2,894 individual firms were successfully matched (out of 20,688 firms in the innovation data and 41,080 firms in the lobbying data). Of the 824 treated firms in the innovation data, 282 were matched in the lobbying data. Of the 902 treated firms in the lobbying data, 193 were matched in the innovation data.

[28]Unfortunately, the data do not permit a robust test of innovation spillovers to firm lobbying investments.

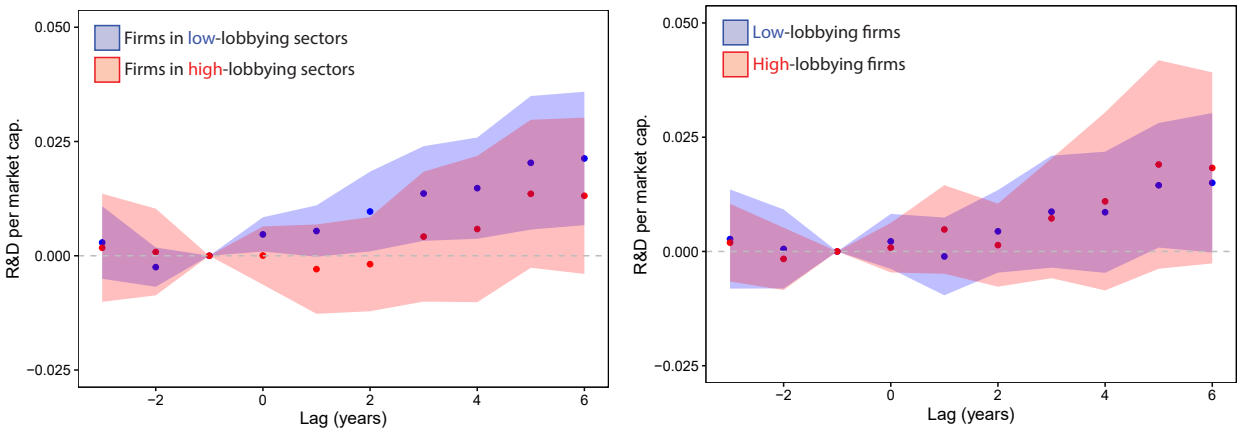


Figure 9: Lobbying spillovers to rival firm innovation may be important. Left: The result from Figure 6, where the pre-period lobbying interaction term is defined as the average lobbying within a 4-digit SIC code. Right: Same as left, except pre-period lobbying is defined using a firm's own lobbying only, for the set of firms that could be individually matched across the innovation and lobbying datasets. Heterogeneity in firm innovation responses is apparent in industry-wide lobbying but not firm-specific lobbying, suggesting that industry-wide lobbying spillovers are important drivers of firms' individual innovation responses. In both cases, pre-period lobbying is a cross-sectional interaction term, thus the heterogeneity in effects shown here should be considered suggestive rather than causal. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

8 Welfare implications of the option to lobby

Considering the welfare implications of firms' option to lobby requires caution for two key reasons. First, lobbying over regulation may itself have positive or negative welfare implications (Bombardini and Trebbi, 2020). If the regulatory instrument considered requires information about firm private costs to determine optimal regulation (e.g. a performance standard or command and control regulation), then it is entirely possible that lobbying by firms communicates information necessary to the regulator for the optimal design of regulation. On the other hand, if the regulatory instrument considered does not rely on firm private information (e.g. a tax[29]), then lobbying by firms may be nothing more than an effort to appropriate rents. However even in the case of Pigouvian taxation, the welfare implications of lobbying in a second-best policy context may not be obvious. For example, the regulator may be concerned about leakage when regulating a trade-exposed sector, in which case firm information over supply in the regulated

In an analysis similar to that of Figure 9, I examine within-industry innovation spillovers on firms' lobbying responses. Unfortunately, in the more limited firm-matched dataset, stable firm lobbying results with meaningful heterogeneity are difficult to recover. Results are shown in Appendix D.

[29]The Pigouvian tax is set at the level of marginal damages, which is not something that firms generally have private information about.

and unregulated economies as well as demand for output may be necessary to estimate leakage effects (Baylis, Fullerton, and Karney, 2014). To be conservative, I do not take a stand here on the direct social welfare value of firms' lobbying expenditures in my sample.

Second, the heterogeneity in firm innovation responses to the threat of regulation that I estimate is over cross sectional variation in sector average lobbying (Equation 6). While the information shocks firms face may be plausibly exogenous, cross sectional variation in sectoral average lobbying is not. Thus, it may be the case in Figure 7 that firms with stronger innovation responses simply happen to be in industry sectors where lobbying is less prevalent, and vice versa. While my findings are consistent with innovation and lobbying as substitutes in firm strategies, they are not conclusive evidence.

With these two caveats in mind, I conduct a back-of-the-envelope evaluation of the welfare implications of my findings. In doing so, I set aside any direct welfare implications of lobbying on optimal regulatory design and I take my estimate of heterogeneity in firms' innovation responses over pre-period sectoral average lobbying (Figure 6) at face value. Average pre-period lobbying across all treated firms in my sample is just over $230,000. I estimate an average reduction in firms' six-year cumulative innovation responses from 2.2% to 0.6% of firm market capitalization for firms in industries with below median versus above median pre-period lobbying. The total pre-period market capitalization of treated firms in my sample is $2.4 trillion, which implies a reduction in total innovation investments of $38 billion. Taking a value of the private returns to innovation from (Hall 1999) of 5x implies the value of forgone innovation is $190 billion. This estimate is conservative, in that it excludes the social value of positive innovation spillovers that cannot be appropriated by the firm.

9 Conclusion

This paper introduces the idea that firms with the option to lobby over future regulation affect their own and others' innovation incentives. Faced with a shock to the probability of future regulation, forward-looking firms have a strategic incentive to innovate (should regulation come to pass), but they also have the strategic option to lobby the regulator. I show that firms have a first-order incentive to increase both innovation and lobbying investments in response to a shock to the probability of future regulation – predictions which I test empirically. However, because

lobbying changes the probability of future regulation for all firms in the market, lobbying by a focal firm affects that firm's payoff to its own innovation as well as the innovation payoffs of other potentially regulated firms. This suggests lobbying will have spillover effects on the innovation responses of firms operating in the same market. A firm's own innovation also affects its own lobbying incentives, but innovation spillovers to other firms' lobbying investments may not be important if information transmission regarding firm innovation outcomes is slow relative to the timescale of firm lobbying responses. In addition to the first order predictions, both of these spillover predictions are empirically tested in the data.

This paper also introduces a new dataset of scientific discoveries by independent researchers over human health harms, which I employ for empirical identification of firm strategic responses to a regulatory threat. I scrape the scientific literature for publications of harms from commonly used industrial chemicals, and review over 7,000 publications to construct a novel dataset of scientific discoveries of previously-unknown harms from commonly used industrial chemicals. I then ask how firms invest in innovation and lobbying in response to an information shock over their regulatory future, and does the option to lobby induce more or less innovation from firms on average.

Consistent with theoretical predictions, I find that firms slowly increase innovation investments by a total of 1.5% of their market capitalization over a six year period, but they quickly increase lobbying investments by a total of 17% of pre-period lobbying levels over a 12-18 month period. Innovation results are marginally significant at conventional levels and lobbying results are significant, but standard errors for both are large. Testing for cross-sectional heterogeneity in firms' dynamic responses, I find evidence consistent with innovation and lobbying as substitutes in firm strategies. In particular, pre-existing lobbying capacity is an important identifier of heterogeneity in firm innovation responses, with "low-lobbying" firms exhibiting a significantly larger innovation response than "high-lobbying" firms. Further, evidence suggests that within-market lobbying spillovers play an important role in firm innovation responses, while the reverse (within-market innovation spillovers to firm lobbying) may be limited. Both of these findings are consistent with theoretical predictions regarding firm spillovers.

Overall, two separate tests – reflecting the same cross-partial derivative taken over the innovation and lobbying first order conditions – yield results consistent with substitution between innovation and lobbying strategies. Further, both tests are highly flexible, yet the sign of the

interaction is consistent in nearly all periods following a discovery. However, the caveats noted above apply. Taking the estimates at face value and drawing on a value of innovation from the literature, my estimates imply that the option to lobby may induce in-sample firms on average to reduce innovation with a forgone private value of about $135 billion. This estimate is conservative in that it does not account for positive social gains from innovation spillovers that firms cannot appropriate. Whether there is a net welfare loss associated with this possibility of lost innovation depends on whether lobbying by firms contains information of use to the regulator in optimally designing regulation. The materiality of lobbying informational content (and thus the welfare implications of the innovation tradeoff) may depend on the type of regulation under consideration, and particularly the role of firm private information in the design of optimal regulation.

This analysis does face limitations which present avenues for future work. While I can test the predicted sign of the first order conditions over firm innovation and lobbying, the data do not permit me to test whether firms are correctly optimizing on the margin. Internal frictions (e.g. managerial competition) may prevent firms from fully internalizing their own innovation and lobbying spillovers, and I look forward to exploring this in future work. Further, because I must pool across industries for statistical power, I do not observe the market structure of each affected industry. This means my estimates should be interpreted as the average treatment effect (ATE) of a scientific discovery of a chemical harm, where the industry affected is drawn at random from the industries in my sample. Effectively, this ATE can be interpreted as the expected effect of some future scientific discovery of a low-dose chemical harm, where we do not know which industry that particular discovery will affect. Finally, my response heterogeneity estimates are consistent with innovation and lobbying as substitutes in firm strategies on average, however the heterogeneity estimated is cross-sectional in nature and some of the estimated interactions are noisy, and thus these estimates should be held with caution.

Multiple other extensions to this work are possible; I highlight a few of them here. First, I test for aggregate innovation and lobbying investments in response to the threat of future regulation, but insight regarding how firms innovate and lobby could be important and may be facilitated by systematic (machine learned) textual analysis of firm patents and long-form lobbying reports. Second, the reduced-form heterogeneity in firm responses that I estimate is an average effect over a wide range of industries and market structures. Particularly if more data

on discoveries is collected, a separate effort focusing on industries with a well-known structure to demand and competition may yield insight into the role of competition in firms' strategic responses, beyond the simple HHI measure I estimate here. Third, the model in this paper is a two-stage game which abstracts away from firm dynamics. A dynamic model incorporating incomplete information over competitors' research portfolios and knowledge stocks could provide a more detailed characterization of the temporal evolution of a given firms' perceived competitive advantage. Fourth, the rapid lobbying response relative to the slower innovation response could suggest that firms employ lobbying as a means to resolve regulatory uncertainty before investing in innovation. Exploring lobbying as a modulator of firms' option value to the timing of innovation investments (Dixit and Pindyck, 1994) represents an interesting avenue for future work. Finally, I construct my dataset of initial discoveries of chemical harms by manually reviewing scraped scientific publications. This dataset could be used to train a machine learning approach to more fully automate the construction of an expanded dataset of the health-harm discovery process, which could itself be an asset to the larger research community.

References

Acemoglu, Daron. 1998. "Why do new technologies complement skills? Directed technical change and wage inequality." *The Quarterly Journal of Economics* 113 (4):1055–1089.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. "The environment and directed technical change." *American Economic Review* 102 (1):131–66.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and innovation: An inverted-U relationship." *The Quarterly Journal of Economics* 120 (2):701–728.

Aghion, Philippe, Antoine Dechezleprêtre, David Hemous, Ralf Martin, and John Van Reenen. 2016. "Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry." *Journal of Political Economy* 124 (1):1–51.

Akcigit, Ufuk, John Grigsby, Tom Nicholas, and Stefanie Stantcheva. 2022. "Taxation and innovation in the twentieth century." *The Quarterly Journal of Economics* 137 (1):329–385.

Ansolabehere, Stephen, John M De Figueiredo, and James M Snyder Jr. 2003. "Why is there so little money in US politics?" *Journal of Economic Perspectives* 17 (1):105–130.

Apergis, Nicholas, Claire Economidou, and Ioannis Filippidis. 2008. "Innovation, technology transfer and labor productivity linkages: evidence from a panel of manufacturing industries." *Review of World Economics* 144 (3):491–508.

Arora, Ashish, Sharon Belenzon, and Lia Sheer. 2021. "Knowledge spillovers and corporate investment in scientific research." *American Economic Review* 111 (3):871–898.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention (pp. 609-626)." *The Rate and Direction of Inventive Activity: Economic and Social Factors (ed. R. Nelson)* .

Baylis, Kathy, Don Fullerton, and Daniel H Karney. 2014. "Negative leakage." *Journal of the Association of Environmental and Resource Economists* 1 (1/2):51–73.

Benedick, Richard Elliot. 1991. *Ozone diplomacy: new directions in safeguarding the planet*. Harvard University Press.

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

Bertrand, Marianne, Matilde Bombardini, and Francesco Trebbi. 2014. "Is it whom you know or what you know? An empirical assessment of the lobbying process." *American Economic Review* 104 (12):3885–3920.

Bloom, Nicholas, Mirko Draca, and John Van Reenen. 2016. "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity." *The Review of Economic Studies* 83 (1):87–117.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying technology spillovers and product market rivalry." *Econometrica* 81 (4):1347–1393.

Bloom, Nicholas, John Van Reenen, and Heidi Williams. 2019. "A toolkit of policies to promote innovation." *Journal of Economic Perspectives* 33 (3):163–184.

Bloom, Nick, Rachel Griffith, and John Van Reenen. 2002. "Do R&D tax credits work? Evidence from a panel of countries 1979–1997." *Journal of Public Economics* 85 (1):1–31.

Blundell, Richard, Rachel Griffith, and John Van Reenen. 1999. "Market share, market value and innovation in a panel of British manufacturing firms." *The Review of Economic Studies* 66 (3):529–554.

Bombardini, Matilde and Francesco Trebbi. 2011. "Votes or money? Theory and evidence from the US Congress." *Journal of Public Economics* 95 (7-8):587–611.

———. 2020. "Empirical models of lobbying." *Annual Review of Economics* 12:391–413.

Bräuning, Falk, Jose L Fillat, and Gustavo Joaquim. 2023. "Firms' Cash Holdings and Monetary Policy Transmission." *Federal Reserve Bank of Boston Research Paper Series Current Policy Perspectives Paper* (97115).

Bronars, Stephen G and John R Lott, Jr. 1997. "Do campaign donations alter how a politician votes? Or, do donors support candidates who value the same things that they do?" *The Journal of Law and Economics* 40 (2):317–350.

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

Brunel, Claire. 2019. "Green innovation and green Imports: Links between environmental policies, innovation, and production." *Journal of Environmental Management* 248:109290.

Calel, Raphael and Antoine Dechezleprêtre. 2016. "Environmental policy and directed technological change: evidence from the European carbon market." *Review of Economics and Statistics* 98 (1):173–191.

Chan, Louis KC, Josef Lakonishok, and Theodore Sougiannis. 2001. "The stock market valuation of research and development expenditures." *The Journal of Finance* 56 (6):2431–2456.

Cohen, Wesley M. 2010. "Fifty years of empirical studies of innovative activity and performance." *Handbook of the Economics of Innovation* 1:129–213.

De Figueiredo, John M and Brian S Silverman. 2006. "Academic earmarks and the returns to lobbying." *The Journal of Law and Economics* 49 (2):597–625.

Debreu, Gerard. 1952. "A social equilibrium existence theorem." *Proceedings of the National Academy of Sciences* 38 (10):886–893.

Denison, Richard. 2017. "A primer on the new Toxic Substances Control Act (TSCA) and what led to it." *Whitepaper: Environmental Defense Fund* .

Dixit, Avinash K and Robert S Pindyck. 1994. *Investment under uncertainty*. Princeton University Press.

Dugoua, Eugenie. 2023. "Induced innovation and international environmental agreements: evidence from the ozone regime." *Review of Economics and Statistics* :1–45.

EPA, US. 2021. "TRI-Listed Chemicals." https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals. [Accessed October 29, 2021].

Fan, Ky. 1952. "Fixed-point and minimax theorems in locally convex topological linear spaces." *Proceedings of the National Academy of Sciences of the United States of America* 38 (2):121.

Gerlach, C. 2016. "New Toxic Substances Control Act: An End to the Wild West for Chemical Safety." https://sitn.hms.harvard.edu/flash/2016/new-toxic-substances-control-act-end-wild-west-chemical-safety/. [Accessed October 26, 2021].

Glicksberg, Irving L. 1952. "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points." *Proceedings of the American Mathematical Society* 3 (1):170–174.

Gowrisankaran, Gautam, Ashley Langer, and Wendan Zhang. 2022. "Policy uncertainty in the market for coal electricity: The case of air toxics standards." *Working Paper. National Bureau of Economic Research* .

Greene, William H. 2003. *Econometric Analysis*. Pearson Education India.

Grossman, Gene M and Elhanan Helpman. 1994. "Protection for Sale." *The American Economic Review* :833–850.

Hall, Bronwyn H. 1999. "Innovation and market value."

Harstad, Bård and Jakob Svensson. 2011. "Bribes, lobbying, and development." *American Political Science Review* 105 (1):46–63.

Hicks, John. 1932. *The Theory of Wages*. Springer.

i Vidal, Jordi Blanes, Mirko Draca, and Christian Fons-Rosen. 2012. "Revolving door lobbyists." *The American Economic Review* 102 (7):3731.

Jaffe, Adam B and Karen Palmer. 1997. "Environmental regulation and innovation: a panel data study." *Review of Economics and Statistics* 79 (4):610–619.

Jaffe, Adam B and Manuel Trajtenberg. 2002. *Patents, citations, and innovations: A window on the knowledge economy*. MIT press.

Johnstone, Nick, Ivan Haščič, Julie Poirier, Marion Hemar, and Christian Michel. 2012. "Environmental policy stringency and technological innovation: evidence from survey data and patent counts." *Applied Economics* 44 (17):2157–2170.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101 (6):2590–2615.

Kang, Karam. 2016. "Policy influence and private returns from lobbying in the energy sector." *The Review of Economic Studies* 83 (1):269–305.

Kim, Sung Eun, Johannes Urpelainen, and Joonseok Yang. 2016. "Electric utilities and American climate policy: lobbying by expected winners and losers." *Journal of Public Policy* 36 (2):251–275.

Lanjouw, Jean Olson and Ashoka Mody. 1996. "Innovation and the international diffusion of environmentally responsive technology." *Research Policy* 25 (4):549–571.

Lenway, Stefanie, Randall Morck, and Bernard Yeung. 1996. "Rent seeking, protectionism and innovation in the American steel industry." *The Economic Journal* 106 (435):410–421.

Meng, Kyle C and Ashwin Rode. 2019. "The social cost of lobbying over climate policy." *Nature Climate Change* 9 (6):472–476.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny. 1991. "The allocation of talent: Implications for growth." *The Quarterly Journal of Economics* 106 (2):503–530.

———. 1993. "Why is rent-seeking so costly to growth?" *The American Economic Review* 83 (2):409–414.

Newell, Richard G, Adam B Jaffe, and Robert N Stavins. 1999. "The induced innovation hypothesis and energy-saving technological change." *The Quarterly Journal of Economics* 114 (3):941–975.

Okunade, Albert A and Vasudeva NR Murthy. 2002. "Technology as a "major driver" of health care costs: a cointegration analysis of the Newhouse conjecture." *Journal of Health Economics* 21 (1):147–159.

Popp, David. 2006. "International innovation and diffusion of air pollution control technologies: the effects of NOX and SO2 regulation in the US, Japan, and Germany." *Journal of Environmental Economics and Management* 51 (1):46–71.

———. 2019. "Environmental policy and innovation: a decade of research." *Working Paper. National Bureau of Economic Research* .

Popp, David, Richard G Newell, and Adam B Jaffe. 2010. "Energy, the environment, and technological change." *Handbook of the Economics of Innovation* 2:873–937.

Porter, Michael. 1996. "America's green strategy." *Business and the Environment: A Reader* 33:1072.

Rittenhouse, Katherine and Matthew Zaragoza-Watkins. 2018. "Anticipation and environmental regulation." *Journal of Environmental Economics and Management* 89:255–277.

Romer, Paul M. 1990. "Endogenous technological change." *Journal of Political Economy* 98 (5, Part 2):S71–S102.

Salop, Steven C and David T Scheffman. 1983. "Raising rivals' costs." *The American Economic Review* 73 (2):267–271.

Schmidt, Charles W. 2016. "TSCA 2.0: A new era in chemical risk management." *Environmental Health Perspectives* 124 (10).

Schumpeter, Joseph. 1942. "Creative destruction. Capitalism." *Socialism and Democracy* 825 (82-85).

Appendix

A Testing pre-period parallel trends

The main innovation and lobbying specifications include three years and two and a half years of leads, respectively. Here I show results for 10 periods of leads and lags for the innovation response and 8 leads and 10 lags for the lobbying response (Figure A.1). Note that lobbying is observed every six months so 10 periods is five years, and the lobbying panel (which starts in 1998) is much shorter than is the innovation panel (which starts in 1950). Estimated effects show more noise for two reasons. First, estimating more period treatment effects leaves less residual variation for the estimation of any given treatment effect, leading to larger standard errors. Second, estimating effects over 10 leads and lags requires that treated firms have more pre- and post-period data in order to enter the balanced panel, and as such fewer total firms are available for treatment effects estimation over this longer time horizon. These caveats aside, the point-estimates of overall cumulative effects are remarkably consistent with those estimated in the main models, as is the temporal structure of the responses.

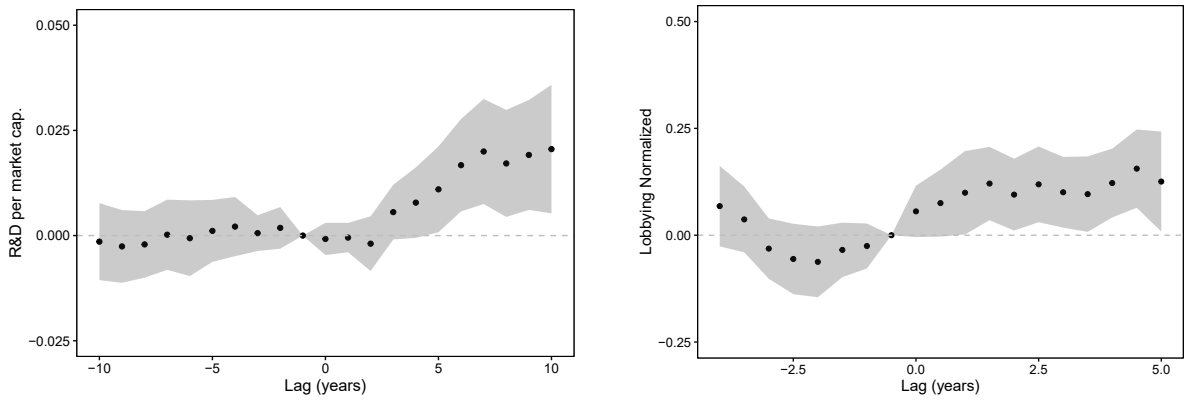
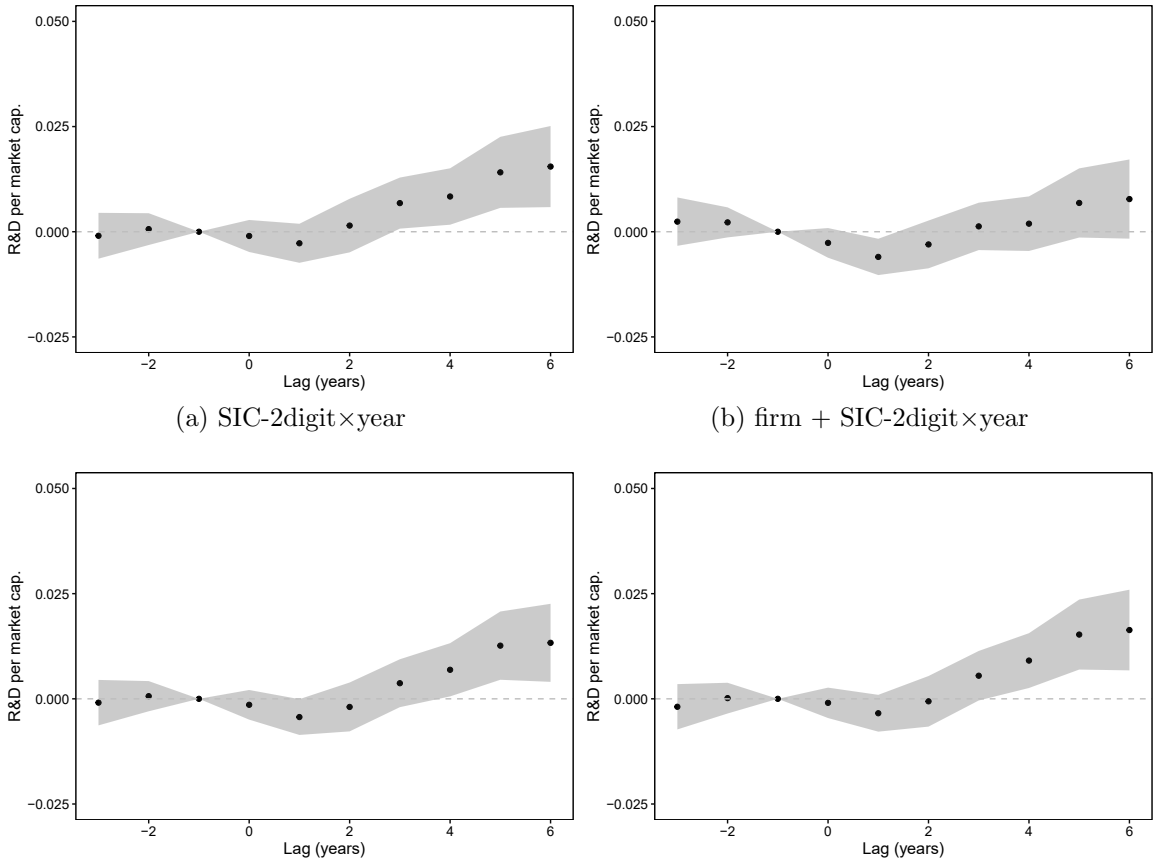


Figure A.1: The cumulative effect of the discovery of a previously unknown potential externality on the innovation (right) and lobbying (left) investments of affected firms. These estimates include an extended set of leads and lags in firm responses, increasing the noise in the estimated effects considerably. However, additional leads serve as an even stronger partial test for differential trends, showing that treated and control firms exhibit parallel trends for many periods before treatment occurs. The additional lags show that the point-estimates of the treatment effects stabilize within the shorter lag periods estimated in the preferred specifications (Figures 3 and 4), and they stabilize at similar values. Estimates here include 10 years of leads and lags for the innovation response and 4 years of leads and 5 years of lags for the lobbying response (8 and 10 half-year periods). 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

B Robustness of the innovation response

This section demonstrates robustness of the innovation response to alternative fixed effects specifications (Figure B.1), to alternative lag specifications (Figure B.2), and to dropping individual discoveries (Figure B.3, only the most influential dropped discoveries shown).



(a) SIC-2digit×year

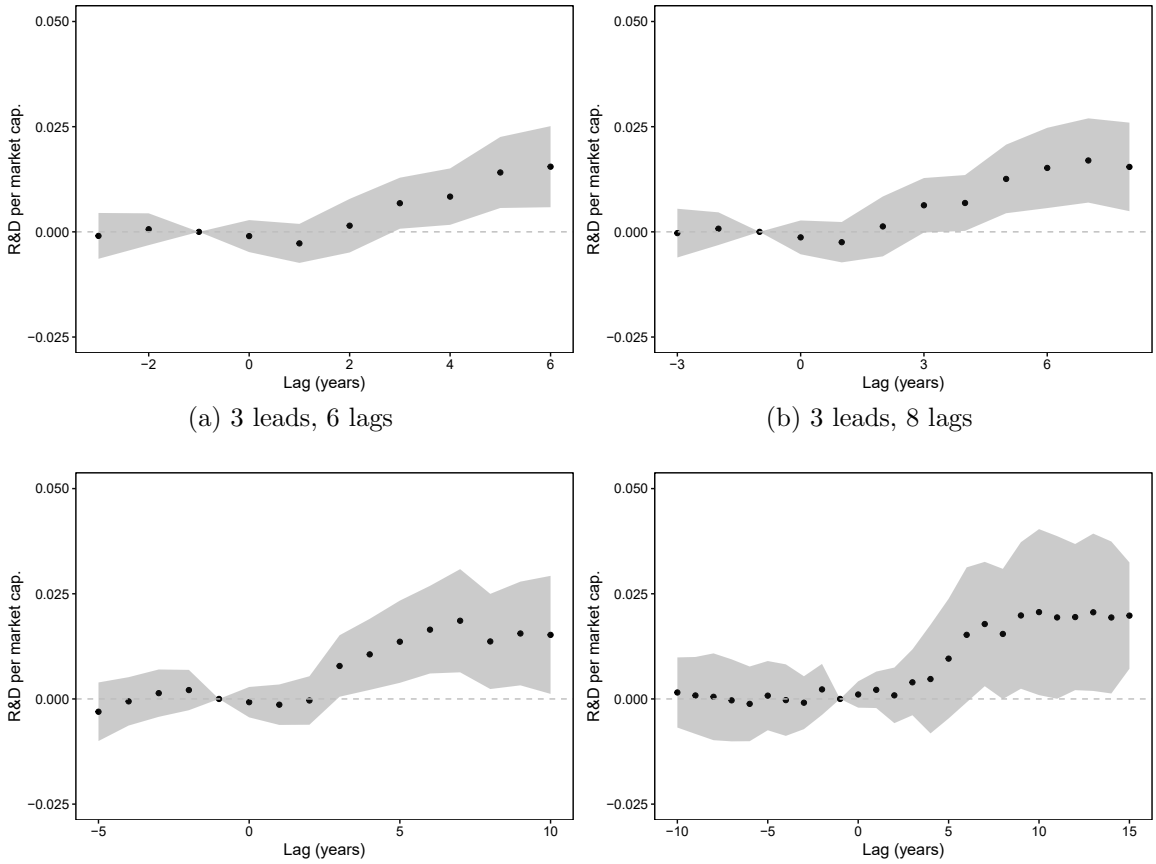
(b) firm + SIC-2digit×year

(c) SIC-2digit + year

(d) year

Figure B.1: Robustness of the innovation response to alternative fixed effects specifications.

Panel (a) is the preferred specification presented in the main text, included here for comparison. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.



(a) 3 leads, 6 lags

(b) 3 leads, 8 lags

(c) 5 leads, 10 lags

(d) 10 leads, 15 lags

Figure B.2: Robustness of the innovation response to alternative lag specifications.

Panel (a) is the preferred specification presented in the main text, included here for comparison. All estimates are on a balanced panel, so the number of firms entering each estimate varies as the number of estimated leads and lags varies. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

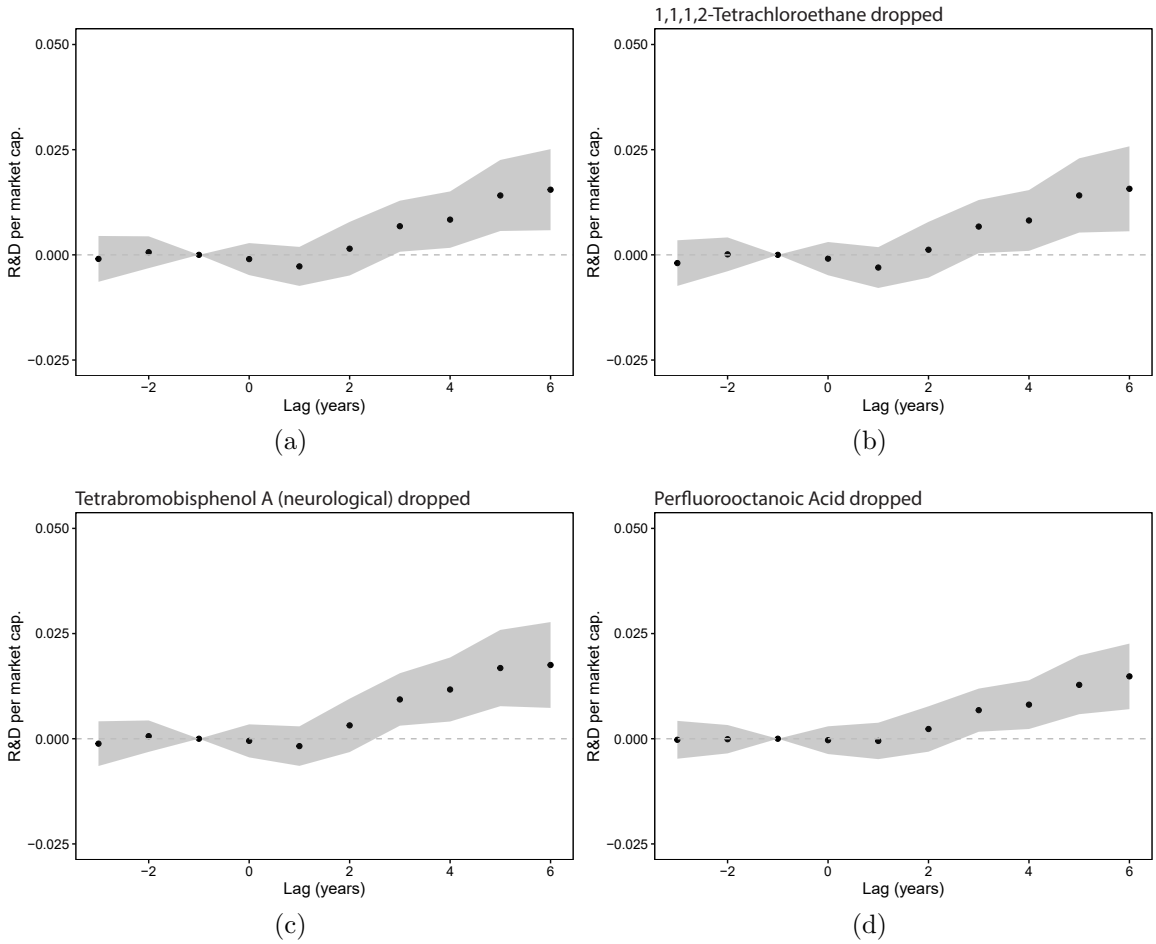
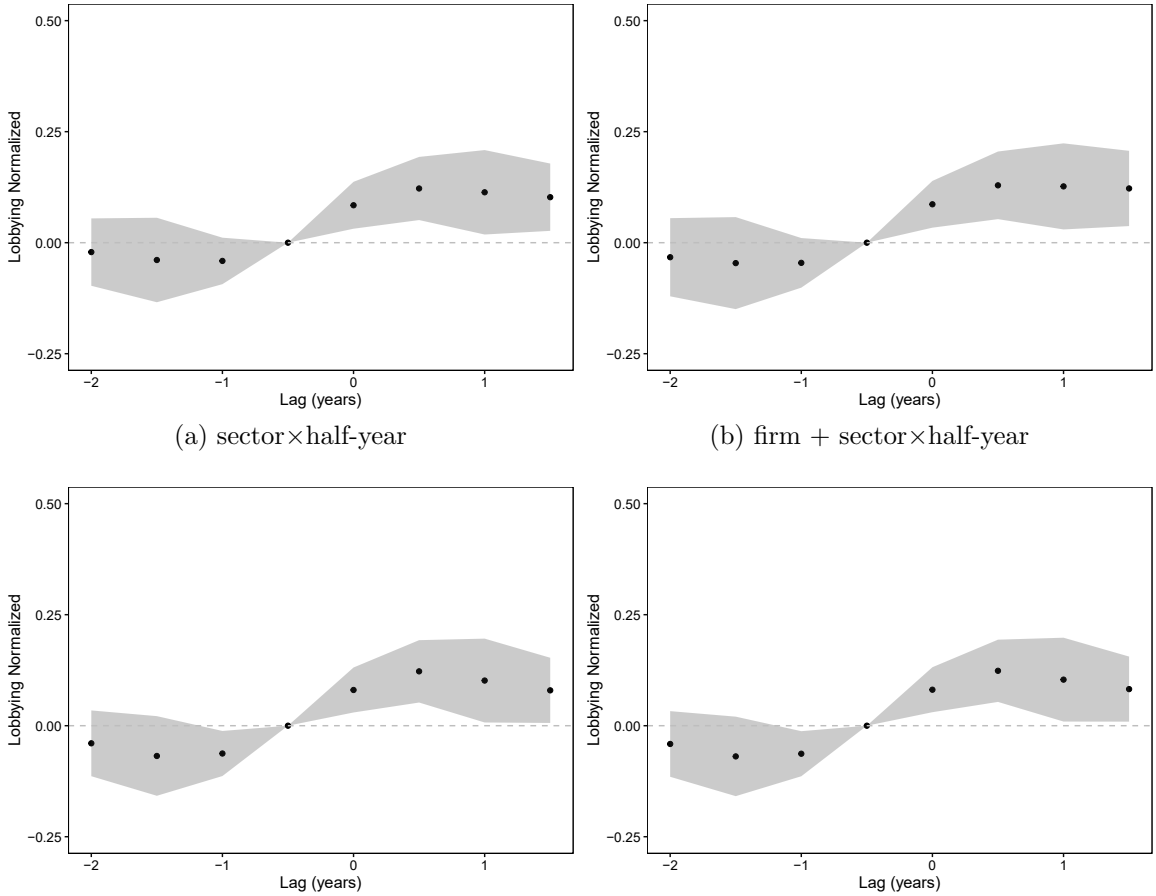


Figure B.3: Robustness of the innovation response to dropping individual discoveries.

Panel (a) is the preferred specification presented in the main text, included here for comparison. Panel (b) is representative of dropping a random discovery. Panels (c) and (d) are the two discoveries that *most* affect the estimates when they are dropped. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

C Robustness of the lobbying response

This section demonstrates robustness of the lobbying response to alternative fixed effects specifications (Figure C.1), to alternative lag specifications (Figure C.2), and to dropping individual discoveries (Figure C.3, only the most influential dropped discoveries shown).



(a) sector×half-year

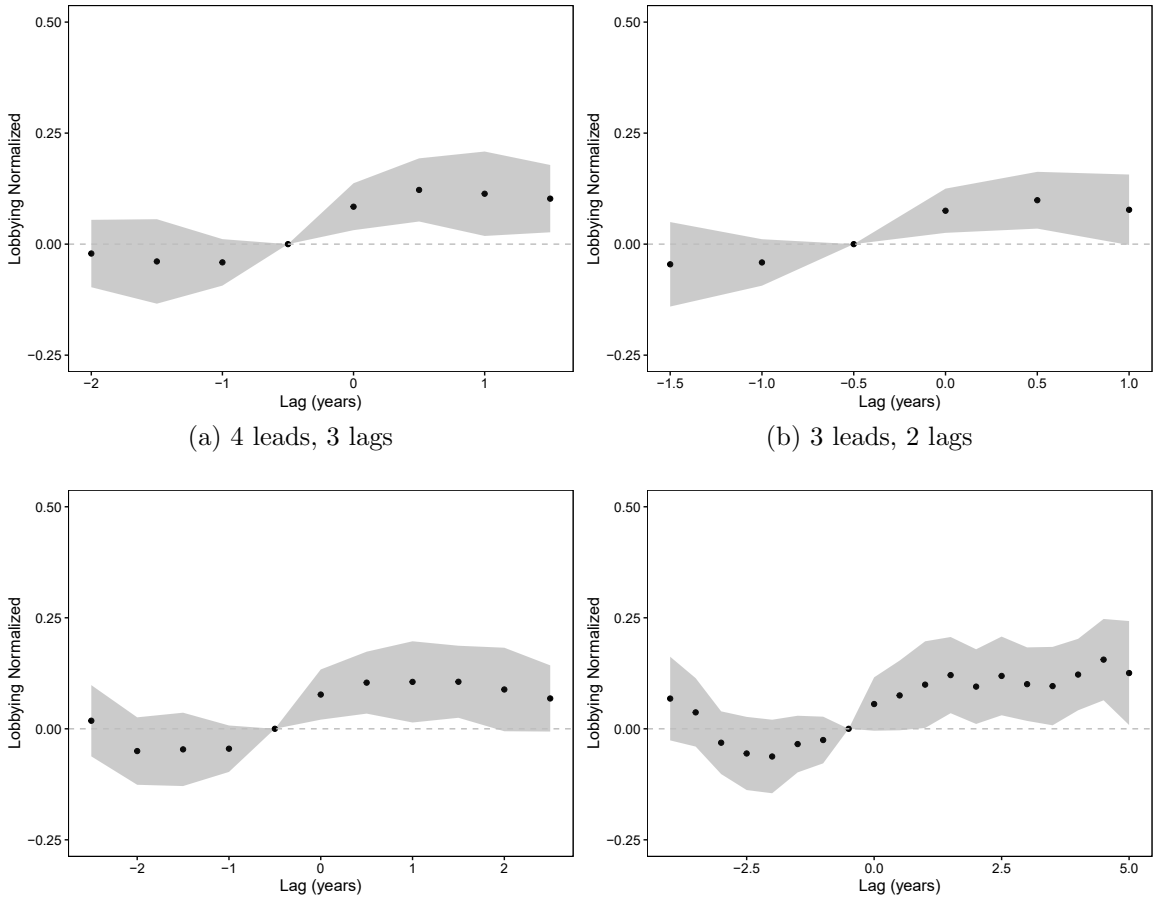
(b) firm + sector×half-year

(c) sector + half-year

(d) half-year

Figure C.1: Robustness of the lobbying response to alternative fixed effects specifications.

Panel (a) is the preferred specification presented in the main text, included here for comparison. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.



(a) 4 leads, 3 lags

(b) 3 leads, 2 lags

(c) 5 leads, 5 lags

(d) 8 leads, 10 lags

Figure C.2: Robustness of the lobbying response to alternative lag specifications.

Panel (a) is the preferred specification presented in the main text, included here for comparison. All estimates are on a balanced panel, so the number of firms entering each estimate varies as the number of estimated leads and lags varies. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

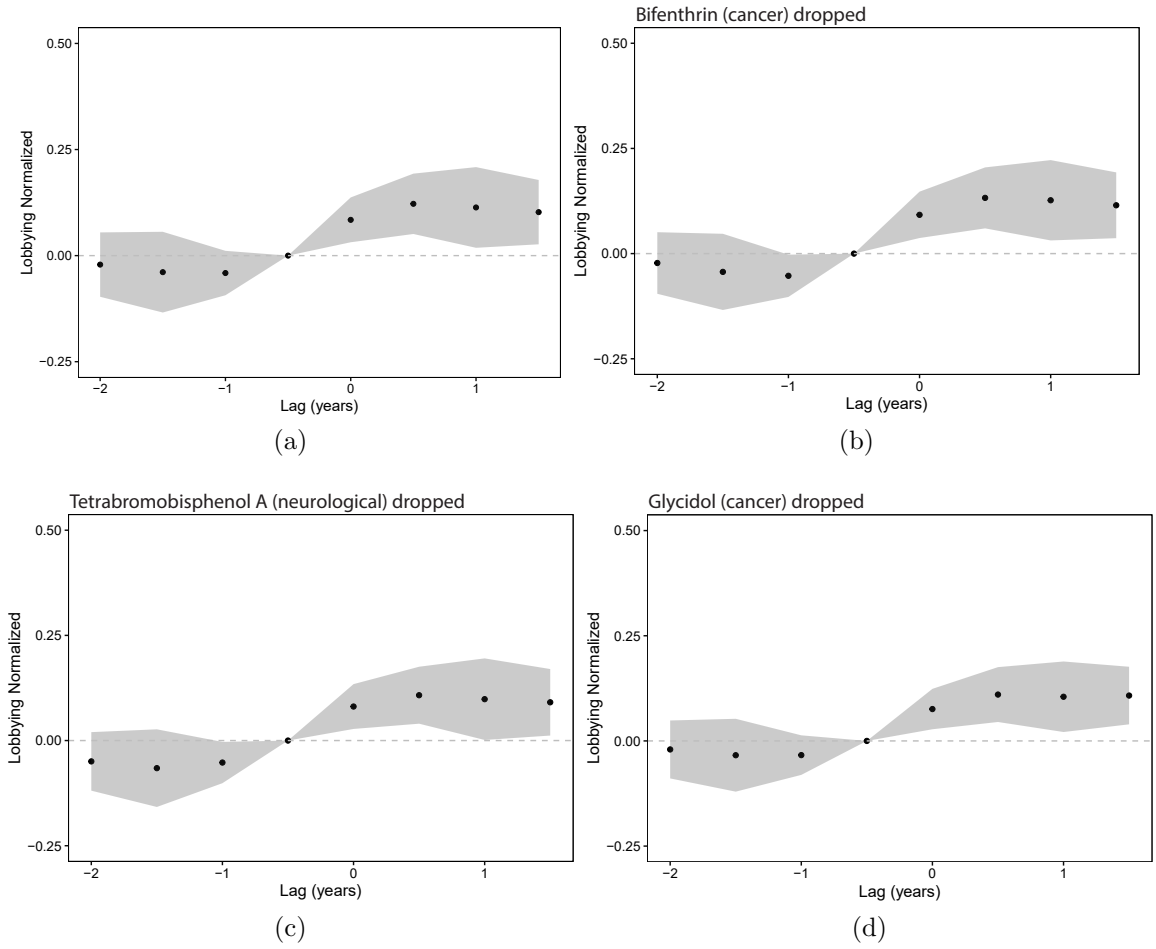


Figure C.3: Robustness of the lobbying response to dropping individual discoveries.

Panel (a) is the preferred specification presented in the main text, included here for comparison. Panel (b) is representative of dropping a random discovery. Panels (c) and (d) are the two discoveries that *most* affect the estimates when they are dropped. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

D Additional heterogeneity in the lobbying response

This appendix presents additional tests for heterogeneity in firm lobbying responses to a regulatory threat. These results are generally too noisy to lend them a high level of confidence, and so should be interpreted with caution.

Tests for innovation spillovers to firm lobbying

This section presents tests for spillover effects of firm innovation on other firms' lobbying investments, and is analogous to the lobbying spillover tests on firm innovation in Section 7.3.2. However, results here were too unstable for inclusion in the main text.

The theoretical motivation for lobbying spillovers to firm innovation activities – as examined in Section 7.3.2 – is straightforward: lobbying by one firm affects the probability of regulation for all firms, and thus innovation payoffs in response to a regulatory threat. However, it is also possible that innovation by one firm could affect the lobbying payoffs of another firm if there are innovation knowledge spillovers on the timescale of firm responses to a regulatory threat. Such knowledge spillovers could occur, for example, through joint research partnerships between firms, research consortia, or public-private research partnerships. Such knowledge spillovers might alter firms' payoff wedge between the regulated and unregulated states, thus altering their lobbying incentives.

Here I test for the presence of such knowledge spillovers in firm lobbying responses to a regulatory threat, using a limited dataset of firm-specific matches across the Compustat and Center for Responsive Politics datasets (see Section 7.3.2). The left and right panels of Figure D.1 are directly analogous to the spillover test in Figure 9. That is, the left panel displays an estimate of main specification for firm lobbying heterogeneity over pre-period sectoral innovation (Figure 7), but using the more limited firm-matched dataset. In this limited dataset, firms in low-innovation sectors (blue) appear to lobby more in response to a regulatory threat than do firms in high-innovation sectors (red), but results estimates show more noise than those in the full dataset. In the right panel of Figure D.1, heterogeneity is estimated over a firm's own pre-period innovation (rather than sectoral innovation in the left panel). Here, noise in the estimates makes it impossible to distinguish any meaningful heterogeneity in firm responses. Thus, unfortunately, the data do not yield meaningful insights regarding the potential role of innovation spillovers on firm lobbying responses to a regulatory threat.

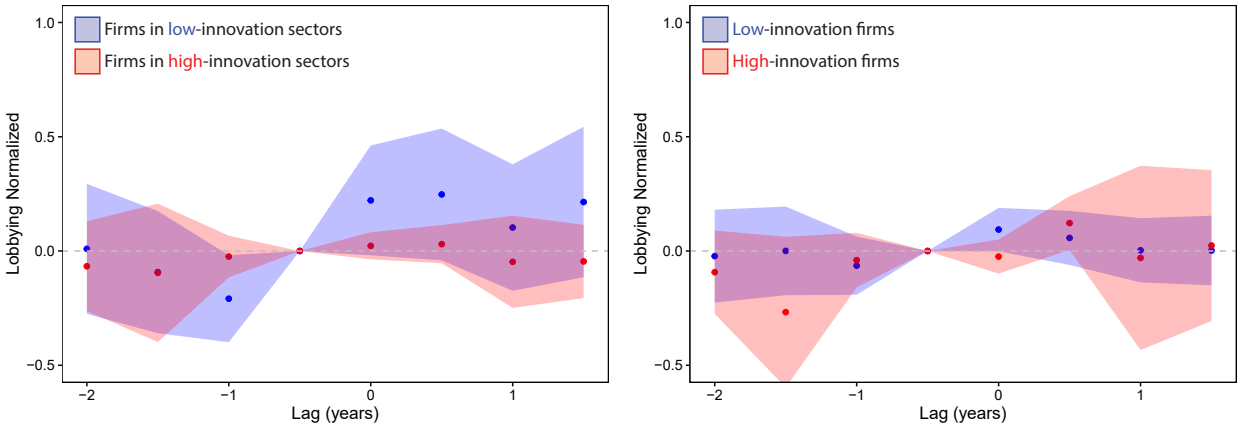


Figure D.1: Innovation spillovers to rival firm lobbying are noisy. Left: The main specification for firm lobbying heterogeneity over pre-period sectoral innovation (as displayed in Figure 7 with modified y-axis), but using a more limited dataset composed of individual firm matches across the Compustat and Center for Responsive Politics datasets. See Section 7.3.2 for the matching procedure. Firms in low-innovation sectors (blue) appear to lobby more in response to a regulatory threat than do firms in high-innovation sectors (red), but estimates in this limited dataset are subject to more noise. Right: Identical to left panel, except heterogeneity is estimated over a firm's own pre-period innovation (rather than sectoral average innovation). In this limited dataset, results are too noisy to distinguish any meaningful heterogeneity. In both panels, pre-period innovation is a cross-sectional interaction term, thus the heterogeneity in effects shown here should be considered suggestive. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

Heterogeneity over market concentration in the lobbying response

This section displays heterogeneity over market concentration (industry sector HHI) in firm lobbying responses to a regulatory threat. This estimate is analogous to the innovation response estimate of Figure 5, but with firm lobbying as the outcome. As can be seen, there is little meaningful heterogeneity to be estimated, and the result is reported here for completeness.

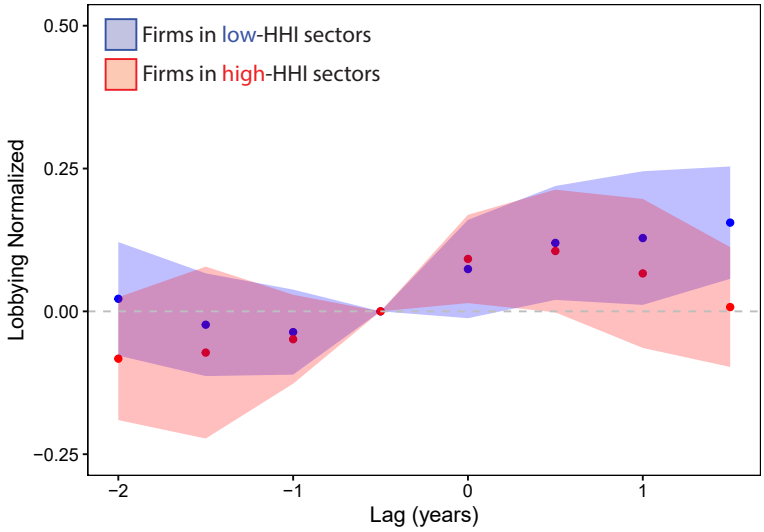


Figure D.2: Heterogeneity over market concentration (HHI) in the cumulative effect of the discovery of a previously unknown potential externality on the lobbying investments of affected firms.

The 10^{th} and 90^{th} percentiles of HHI in the data are 0.05 and 0.3, measured on the [0, 1] interval. Little meaningful heterogeneity in the lobbying response over HHI is resolved. 95% confidence intervals are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

E Non-scraped discoveries

Early work involved a small number of discoveries chemical harms that were manually identified, before the process for this analysis of scraping the literature was developed. Manual identification of discoveries involved reading literature surveys and media reports to identify the earliest publications associated with previously unknown, low-dose, chronic health effects. Those discoveries are documented here, as are plots showing robustness of the innovation and lobbying responses to dropping all of these "non-scraped" discoveries.

Chemical	Discovery date	Source	Description
Chlorofluorocarbons (CFCs)	6/28/1974	Molina and Rowland (1974)	Aerosol propellant, coolant, electronics cleaning
Bisphenol A (BPA)	6/1/1993	Krishnan et al (1993)	Accidental discovery; widely use plasticizer
Parabens	11/1/1998	Routledge et al (1998)	Preservative used in cosmetics, toiletries, and some foods
GenX	6/30/2015	Rae et al (2015)	Used in production of Teflon, other fluorinated polymers
Perfluorooctane sulfonate (PFOS)	9/2/1997	Anderson and Mulvana (1997)	Fabric stain repellent, cleaning products, firefighting, metal plating and semiconductor cleaning
Perfluorooctanoic acid (PFOA)	9/1/1993	Gilliland and Mandel (1993)	Similar uses as PFOS; voluntarily phased out (announced in May 2000)
Methyl tert-butyl ether (MTBE)	9/30/1994	Moolenaar et al (1994)	Blending agent in gasoline; solvent in plastics and pharaceuticals
Fracking – Water contamination	12/31/2011	EPA (2011)	EPA draft study investigating water quality complaints
Fracking – Earthquakes	8/31/2012	BCOGC (2012)	British Columbia Oil and Gas Commission investigation
Glyphosate – Weed resistance	7/10/1996	Pratley et al (1996)	Conference report reporting weed resistance
Glyphosate – Human exposure	11/2/2004	Benedetti et al (2004)	Low dose chronic exposure

Table E.1: Discoveries that were manually identified.

This table lists discoveries that were manually identified, before the process for this analysis of scraping the literature was developed. Manual identification of discoveries involved reading literature surveys and media reports to identify the earliest publications associated with previously unknown, low-dose, chronic health effects.

F Innovation, lobbying and market competition

The model presented in the main text describes the innovation and lobbying strategic game faced by firms under regulatory uncertainty. This section expands that discussion to consider a change in the competitive structure of the market. As described in the main text, the comparative statics of the innovation and lobbying best-response functions with respect to a change in competitive structure cannot in general be signed. This appendix details why that is the case. In the following, I break down the analysis into its components to more easily describe the signs of the various terms and set up the application of the Implicit Function Theorem.

Consider a firm that enters the market at the start of the two stage game, entering with arbitrary emissions technology, constant marginal cost weakly lower than that of all N incumbents, and atomistic capacity. This results in an inward shift of each incumbent firm's residual demand curve by an infinitesimal amount. In the non-regulated state, the inward shift occurs at the entrant's marginal cost; in the regulated state it occurs at the entrant's marginal cost plus emissions rate times the level of the tax τ. Thus, we can construct a vector of infinitesimally small downward willingness-to-pay shifts $\mathbf{d}$ around the pre-entrant second stage equilibria.

Now consider the Jacobian of firm i's first order conditions over the vector of residual demand shocks $\mathbf{d}$. Denote the left-hand side of the first order conditions (Eqs. 1 and 2) for firm i as $[f_{1i}, f_{2i}] = \mathbf{f}_i$. Note that Eqs. 1 and 2 give $\lambda_i^* = \lambda_i^*(\rho_i; \boldsymbol{\rho}_{-i}, \boldsymbol{\lambda}_{-i}, \tau)$ and $\rho_i^* = \rho_i^*(\lambda_i; \boldsymbol{\rho}_{-i}, \boldsymbol{\lambda}_{-i}, \tau)$. The partials of $\mathbf{f}_i$ with respect to the j^{th} element of the market vector of residual demand shocks d_j make up the elements of the $2 \times N$ Jacobian matrix $J_{\mathbf{f}_i,\mathbf{d}}$. These elements of $J_{\mathbf{f}_i,\mathbf{d}}$ are given as[30]

$$\frac{\partial f_{1i}}{\partial d_j} = \Delta\pi_i \nabla_{\boldsymbol{\lambda}} \eta' (\nabla_{d_j} \boldsymbol{\lambda}^*)^T + \eta' \frac{\partial \Delta\pi_i}{\partial d_j} \tag{F.1}$$

$$\frac{\partial f_{2i}}{\partial d_j} = \frac{\partial \pi(\tau, \rho_i, \rho_{-i})}{\partial e_i} \frac{\partial e_i}{\partial \rho_i} \nabla_{\boldsymbol{\lambda}} \eta (\nabla_{d_j} \boldsymbol{\lambda}^*)^T, \tag{F.2}$$

where in this model a marginal change in the vector of residual demands $\mathbf{d}$ does not affect the emissions margin in firm profits nor the innovation margin in firm emissions (Eq. F.2), and $\Delta\pi_i$ refers to the profit wedge between the regulated and unregulated states.

The first term of Eq. F.1 reflects how changes in firms' residual demands lead to changes in lobbying strategies in the market, affecting the marginal effectiveness of lobbying by the focal firm i in shifting the probability of regulation (weighted by the focal firm's profit wedge). If the marginal effectiveness of lobbying by the focal firm is decreasing in its own lobbying efforts as well as those of its competitors (i.e. η' and $\partial^2\eta/\partial\lambda_i\partial\lambda_k$ have opposite signs $\forall\ i, k$), then the terms $\Delta\pi_i$ and $\nabla_{\boldsymbol{\lambda}}\eta'$ will have opposite signs.[31] However, the signs of the terms in $\nabla_{d_j}\boldsymbol{\lambda}^*$ are

[30] Note that $\nabla_{\boldsymbol{\lambda}} \frac{\partial \eta}{\partial \lambda_i} (\nabla_{d_j} \boldsymbol{\lambda})^T = \begin{bmatrix} \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_1} & \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_2} & \dots & \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_N} \end{bmatrix} \times \begin{bmatrix} \frac{\partial \lambda_1^*}{\partial d_j} & \frac{\partial \lambda_2^*}{\partial d_j} & \dots & \frac{\partial \lambda_N^*}{\partial d_j} \end{bmatrix}^T$
$= \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_1} \frac{\partial \lambda_1^*}{\partial d_j} + \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_2} \frac{\partial \lambda_2^*}{\partial d_j} + \dots + \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_N} \frac{\partial \lambda_N^*}{\partial d_j} = \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_1} \frac{\partial \lambda_1^*}{\partial \Delta\pi_1} \frac{\partial \Delta\pi_1}{\partial d_j} + \dots + \frac{\partial^2 \eta}{\partial \lambda_i \partial \lambda_N} \frac{\partial \lambda_N^*}{\partial \Delta\pi_N} \frac{\partial \Delta\pi_N}{\partial d_j}$, and parallel patterns hold for the other similar terms in these equations.

[31] Recall that the sign of η' equals the sign of $\Delta\pi_i$; firms lobby for regulation if they stand to gain from it, and against if they stand to lose.

ambiguous, as they depend on the terms $\partial\lambda_k^*/\partial\Delta\pi_k$ and $\partial\Delta\pi_k/\partial d_k$. The sign of the former is positive (lobbying is increasing in a firm's own payoff wedge), but the latter cannot be signed without knowledge of the structure of the particular market facing regulation. In particular, the effect of a change in a firm's residual demand on its payoff wedge between the regulated an unregulated states will depend on the shape of the demand curve around the regulated and unregulated equilibria as well as the emissions technology of the entering firm.[32]

The second term in Eq. F.1 reflect how changes in firms' residual demands lead to changes in the regulatory payoff wedge, weighted by the marginal effectiveness of lobbying, and the overall sign is again ambiguous. As previously discussed, η' will take the opposite sign of $\Delta\pi_i$. However, the sign of $\partial\Delta\pi_i/\partial d_j$ depends (as it did before) on the structure of the market facing regulation and whether the focal firm faces a residual demand curve that is more or less downward sloped in the vicinity of the regulated vs. unregulated equilibrium quantity choice.

Finally, Eq. F.2 reflects the marginal change in the probability of regulation as equilibrium lobbying efforts shift, weighted by the marginal payoff to innovation. The profitability payoff to innovation is positive, but the overall sign of Eq. F.2 is ambiguous. The signs of the terms in $\nabla_{d_j}\boldsymbol{\lambda}^*$ are as discussed above, and the signs of the terms in $\nabla_{\boldsymbol{\lambda}}\eta$ depend on whether individual firms lobby for or against regulation.

The Jacobian over f_1 and f_2 with respect to firm i's choice parameters is given below (following the convention in the main text, primes refer to derivatives with respect to the focal firm's own choice parameter).

$$J_{\mathbf{f}_i,[\lambda_i,\rho_i]} = \begin{bmatrix} \eta''\Delta\pi_i & \eta'\pi'(\tau,\rho_i,\rho_{-i})e' \\ \eta'\pi'(\tau,\rho_i,\rho_{-i})e' & \eta\pi'(\tau,\rho_i,\rho_{-i})e'' \end{bmatrix}$$

The signs of the terms in the Jacobian are relatively straightforward. To sign the first term, note that η' will take the same sign as $\Delta\pi_i$, so that η'' takes the opposite sign of $\Delta\pi_i$ and the overall term is negative. The off-diagonals are controlled by the sign of η' (or equivalently, the sign of $\Delta\pi_i$). The remaining diagonal element is straightforwardly negative in sign. Additionally, note that an interior optimum to the firm's problem requires that the product of the diagonals be greater than the product of the off-diagonals.

Finally, the full comparative statics of the problem with respect to a vector of residual demand shocks $\mathbf{d}$ is given below. As discussed above, while the signs of $J_{\mathbf{f}_i,[\lambda_i,\rho_i]}$ are all unambiguous, the signs of the strategic reactions in the best-response functions to demand shocks to other firms in the market (contained in $J_{\mathbf{f}_i,\mathbf{d}}$) are not. Thus the comparative statics of lobbying and innovation investments are ambiguous.

$$\begin{bmatrix} \frac{\partial\lambda_i^*}{\partial d_1} & \cdots & \frac{\partial\lambda_i^*}{\partial d_j} & \cdots & \frac{\partial\lambda_i^*}{\partial d_N} \\ \frac{\partial\rho_i^*}{\partial d_1} & \cdots & \frac{\partial\rho_i^*}{\partial d_j} & \cdots & \frac{\partial\rho_i^*}{\partial d_N} \end{bmatrix} = -J^{-1}_{\mathbf{f}_i,[\lambda_i,\rho_i]} \begin{bmatrix} \frac{\partial f_{1i}}{\partial d_1} & \cdots & \frac{\partial f_{1i}}{\partial d_j} & \cdots & \frac{\partial f_{1i}}{\partial d_N} \\ \frac{\partial f_{2i}}{\partial d_1} & \cdots & \frac{\partial f_{2i}}{\partial d_j} & \cdots & \frac{\partial f_{2i}}{\partial d_N} \end{bmatrix}$$

[32]Consider the Cournot competition first order condition a firm faces in the regulated state: $p_i(q;q_{-i},e_{-i}) + q_i p_i' = c_i' + \tau e_i$. Note that residual demand facing firm i depends on the emissions rates of firms $-i$, and note that p_i' need not be constant in q.