

STRATEGIC INNOVATION AND LOBBYING IN RESPONSE TO REGULATORY UNCERTAINTY

Andrew Hultgren

University of Chicago, Department of Economics and EPIC*

January 9, 2022

[Click here for most recent version.](#)

Abstract

The prospect of environmental regulation can drive innovation. However, firms facing potential regulation also have the strategic option to lobby the regulator, increasing or decreasing the expected payoff to innovation. 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 initial discoveries of previously unknown chemical harms. To do so, I scrape the epidemiological literature for scientific publications of chemical harms and review over 7,000 studies to construct a dataset of initial discoveries. I then ask what is the direct effect of uncertain regulation on firm innovation and on firm lobbying, and additionally how does the strategic option to lobby affect the strategic incentive to innovate? 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. I find that firms slowly increase innovation investments by a total of 2% of their market capitalization over the four years following a discovery, but they quickly increase lobbying investments by a total of 20% 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, though the estimated effect is noisy. This latter estimate, while not statistically significant, implies that the option to lobby may induce a loss in innovation by firms.

*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, Fiona Burlig, Thomas Covert, Lucas Davis, Meredith Fowle, Eyal Frank, Koichiro Ito, Larry Karp, Jeff Perloff, Gordon Rauser, Joseph Shapiro, Leo Simon, Reed Walker, Shaoda Wang, Brian Wright, my ARE 2015 cohort, and seminar participants at the Berkeley ERE, Berkeley Climate Lunch, UChicago EPIC, and UCSB Occasional Workshop seminars for helpful comments and suggestions. I thank Sarah Dang, Hannah Maeder, Xiaoting Sun, and Joshua Tang for excellent research assistance. All errors and omissions are, of course, my own. Corresponding address: hultgren@uchicago.edu.

1 Introduction

Innovation is a critical driver of economic growth and human well-being (Romer, 1990). 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 the uncertain prospect of future regulation, forward-looking firms have a strategic incentive to innovate (should regulation come to pass), but they must trade off this innovation response against their strategic option to lobby the regulator over the probability of regulation. This paper asks what is the direct effect of uncertain regulation on firm innovation and on firm lobbying, and additionally how does the strategic option to lobby affect the strategic incentive to innovate?

The contribution of this paper is to estimate both firm innovation and lobbying responses to an unanticipated shock over the possibility of future regulation, recognizing that forward-looking firms jointly optimize over these strategies. My empirical estimates leverage the randomness of the scientific discovery process of independent researchers: I assume firms do not know when the next discovery of a chemical harm will occur, nor if that discovery will affect them. Unlike much of the extant literature (reviewed by Popp, Newell, and Jaffe (2010) and Bombardini and Trebbi (2020)), my estimates include forward-looking firms’ anticipation effects. These estimates are made possible by creating a novel dataset of over 100 scientific discoveries of previously-unknown chemical harms, spanning 60 industries, 44 years, and over 300 firms. I create this dataset by scraping the epidemiological literature for relevant scientific discoveries and manually reviewing over 7,000 studies to identify the first publications of previously-unknown chemical harms.¹ I pair this dataset with widely-used panels of data on firm innovation expenditures (Compustat) and on firm lobbying expenditures (Center for Responsive Politics). 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’ annual innovation expenditures and biannual lobbying expenditures. Identification rests on within-firm variation in exposure to scientific

¹To be clear, though firm innovation is an outcome I study, all but one of these scientific discoveries of chemical harms come from the epidemiological research community, not private sector researchers.

discoveries affecting its products, controlling for sector-wide temporal shocks common to both treated and untreated firms.

To fix ideas, consider a well-documented example (Benedick, 1991). 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. In 1974, two atmospheric chemists posited that CFCs could *theoretically* catalyze the destruction of Earth’s ozone layer. Thus DuPont was suddenly faced with a new but uncertain possibility of future regulation over CFCs. This possibility of regulation 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 would change the probability of 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.² 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 decades.

I model a simple two-stage game between firms and the regulator to formalize the problem. The first order conditions yield the testable hypothesis that in response to a new possibility of future regulation, firms will increase both their innovation and lobbying investments. 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. This may vary by firm and by market, and must be tested in the data.

I find that in response to a new discovery of a chemical harm, firms slowly increase innovation investments by a total of 2% relative to their market capitalization over a four year period, but they quickly increase lobbying investments by a total of 20% of pre-period lobbying levels over a 12-18 month period. Results are significant at conventional levels, but standard errors are large. Testing for cross-sectional heterogeneity in the dynamics of firms’ responses, I find evidence consistent with innovation and lobbying as substitutes in firm strategies on average. Two separate approaches yield results consistent with innovation and lobbying as substitutes,

²In 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.

though the estimated effect is noisy in both. However, 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 \$170 billion.³

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 and the regulation of new pharmaceutical drugs found to have previously-unknown side-effects, or finance and the regulation of high-frequency trading over the price volatility it introduces, or network economics and the regulation of technology firms such as Facebook over questions regarding the spread of misinformation. Further, this type of analysis can readily be extended to other forms of innovation (e.g. patents) and other forms of regulatory influence, all in both the developing- and developed-world contexts.

The remainder of the paper proceeds as follows: Section 2 reviews 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

³For context, firms affected by a scientific discovery at any point in my sample represent \$2.4 trillion in market capitalization today.

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 (Acemoglu et al., 2012).⁴ A large literature has examined the effect of environmental regulation on firm innovation (see Popp, Newell, and Jaffe (2010) for a survey). 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, 2021), 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.⁵ 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).

⁴The idea of directed technological change (Acemoglu, 1998, 2002) reintroduced the Hicksian concept of “induced innovation”: technological change that is biased towards certain factors of production (Hicks, 1932).

⁵While Kamenica and Gentzkow (2011) list lobbying as an example, the paper is more general.

Importantly, none of these works in either literature considers that innovation and lobbying 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 \(2021\)](#); [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 avoid bias towards zero in my estimated effects due to firm anticipation.

The literature that comes closest to recognizing the innovation and lobbying tradeoff I address is 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. Work by [Lenway, Morck, and Yeung \(1996\)](#) is the empirical application closest to this paper. 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. If the firm perceives a competitive advantage under regulation⁶ it will lobby for regulation, increasing the expected payoff to innovation (which pays off in the regulated state) and implying that innovation and lobbying are complements in firm strategies). Conversely, if the firm perceives a net harm under regulation it will lobby against regulation, decreasing the expected payoff to innovation and implying that innovation and lobbying are substitutes in firm strategies.

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

⁶Regulation may offer a competitive advantage because the firm can be first to market with a substitute or because regulation increases barriers to entry or otherwise more negatively harms competitors, for example (Salop and Scheffman, 1983).

spending by each firm.⁷

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

Firms play a two-stage game. In stage 1, firms invest in innovation and lobbying. Then legislators vote on the tax. In stage 2, Cournot firms set their 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:

$$\begin{aligned}
q_i^* &= \operatorname{argmax}_{q_i} \pi_i(q_i, t, e(\rho_i), e(\rho_{-i})) \\
\implies \pi_i^* &= \pi_i^*(t, e(\rho_i), e(\rho_{-i}))
\end{aligned}$$

where $e(\rho_{-i}) \equiv \{e(\rho_j)\} \quad \forall j \neq i$.⁸

In the case of taxation at $t = \tau$, denote the firm's payoff as $\pi_i^*(\tau, e(\rho_i), e(\rho_{-i}))$. 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

⁷In 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 \(2019\)](#)).

⁸Note 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.

payoff as $\pi_i^*(0)$. Then, the firm's first stage problem is:

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

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:⁹

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

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

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.¹⁰ Equation 1 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](#)

⁹As 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 .

¹⁰To 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. 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](#)).

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. This is because 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.

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

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 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, 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.

Finally, 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.

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

This paper seeks to estimate how the threat of regulation affects firm innovation and lobbying, and how firms jointly optimize over these two responses. The literature frequently examines firm activities as regulation unfolds (Popp, Newell, and Jaffe, 2010; Bombardini and Trebbi, 2020), yet forward-looking firms may respond years in advance once the possibility of regulation is known. 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. 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.

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 review new chemicals to determine whether *“the relevant chemical substance... presents an unreasonable risk of injury to health or the environment, without consideration of costs or other nonrisk factors...”* (15 USC 2604 a.3.A).¹¹ However, several shortcomings in the TSCA new chemical review process rendered it ineffective (Schmidt, 2016; Gerlach, 2016). First, the 62,000 chemicals already in 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. 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.

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). Thus, 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.

¹¹<https://www.epa.gov/sites/default/files/2018-02/documents/chapter5350usc.pdf>

4.2 Scraping and evaluating scientific publications

Scraping scientific publications As an independent list of chemicals to scrape, I draw on the chemicals listed on the Toxics Release Inventory (TRI). The TRI was created after two widely publicized incidents of acute poisoning from chemical releases at Union Carbide manufacturing plants, and it lists chemicals that cause, “cancer or other chronic human health effects, significant adverse human health effects, or significant adverse environmental effects” (EPA, 2021). I begin by scraping the internet for scientific publications of low-dose, chronic human health effects from exposure to commonly used industrial chemicals. My objective is to find discoveries of low-dose, chronic harms because these are less likely to have been known at the time of chemical approval for widespread use. 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,¹² 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

¹²For reproductive studies in rats and mice only one week of exposure is permitted, since total gestation is only about 21 days.

discoveries compiled from early work was also included (see Appendix D),¹³ for a total of 108 initial scientific discoveries of previously unknown health effects from commonly used industrial chemicals.

Matching discoveries to firms Unfortunately, I cannot match chemicals directly to firms in the innovation or lobbying data. I do not observe which firms use or produce each chemical. However, I can match chemicals to industries. The chemicals database PubChem¹⁴ contains an 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 most granular industry identifiers available in the innovation data (four-digit SIC codes) 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 (Jaffe and Trajtenberg, 2002; Bombardini and Trebbi, 2020). 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 annual¹⁵ observations of key financial data from 1950–2020. 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

¹³In Appendix D I also show robustness of results to dropping these non-scraped discoveries.

¹⁴<https://pubchem.ncbi.nlm.nih.gov>

¹⁵Quarterly data are also available, starting in 1989. However, quarterly reporting of innovation expenditures is not required on the SEC 10-Q form and so quarterly innovation reporting is both sparse and inconsistent.

firm’s innovation spending by its pre-period market capitalization. To avoid capturing changes in firm market capitalization when estimating treatment effects, I normalize treated firms by their 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 contemporaneous market capitalization. There is not an established approach to normalizing for scale effects in firm innovation.¹⁶ I would like to normalize by the value at stake to the firm, however I do not observe firm sales by sector. I do observe firm market value, which is the closest proxy available, following [Chan, Lakonishok, and Sougiannis \(2001\)](#). I winsorize outlier years in which firms have innovation expenditures in excess of two times their market capitalization (this represents less than 0.5% of firm-year observations).

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.¹⁷ 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 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

¹⁶Normalizing innovation expenditures by sales is common, but compellingly criticized by [Jaffe and Palmer \(1997\)](#) when used across industries with varying levels of market power. [Jaffe and Palmer \(1997\)](#) studies industry aggregate innovation and so is able to take logs; the many zeros in the firm-level panel prevent me from doing this.

¹⁷One could imagine merging the CRP and Compustat datasets to obtain covariatio information in the lobbying data. However, firm names are the only common identifier across the two datasets and are not standardized. Efforts to find similar matches across the datasets have not proved to be promising.

so are normalized by their panel-average lobbying expenditures. Figure 2e indicates that this normalization is stable over time.

5 The completed datasets of 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 across over 300 firms.

Cross sectional variation in innovation and lobbying Across both treated (blue) and untreated (red) industries in my sample, innovation and lobbying are not strongly correlated in the cross section (Figure 1a). I describe an industry (SIC 4-digit code) as “treated” if it is affected by at least one of the discoveries of previously-unknown chemical harms that I compile. I describe as “untreated” any industry that is in the same SIC division as a treated industry but is never affected by any of the discoveries I compile. Some industries (e.g. the pharmaceuticals industry) have very high average investments in both innovation and lobbying. Some industries have relatively high levels of innovation but low levels of lobbying (e.g. semiconductor manufacturers), and others show the reverse pattern (e.g. oil and gas extractors). Note that the lobbying data are not normalized to account for scale effects because the lobbying data contain little covariate information about firms. The time series of discoveries of chemical harms displays clear temporal structure (Figure 1b), with the rate of discoveries generally increasing over time. With economic growth driving increased consumer demand and firm valuations over this same period, addressing temporal confounds will be important and is discussed further below.

Time series variation in innovation and lobbying Over the time series, both innovation and lobbying are increasing in real terms for both treated and untreated firms (Panels (a) and (d) of Figure 2), and the rate of scientific discoveries of previously-unknown chemical harms

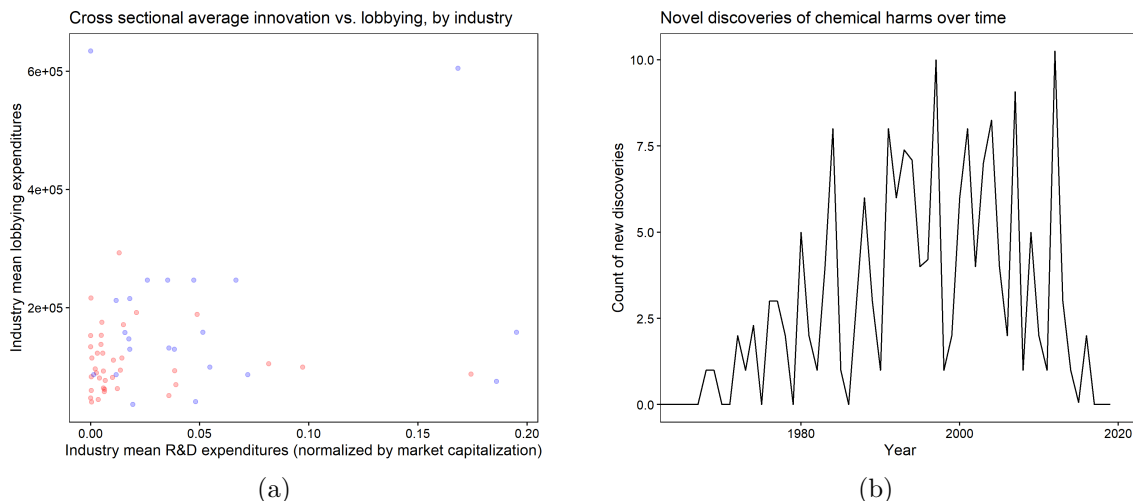


Figure 1: (a) Cross sectional variation in firms’ innovation and lobbying investments. Blue corresponds to treated industries and red to untreated industries. An industry is labeled as “treated” if it is affected by at least one discovery of a previously-unknown chemical harm, and it is labeled as “untreated” otherwise. Innovation expenditures are annual, and normalized by firm market capitalization to account for scale effects. Lobbying expenditures are biannual and are not normalized for scale effects because no normalization covariate is reported in the data. Treated industries are averages over firms at the SIC-4-digit level; untreated industries are averages over firms at the SIC-2-digit level. (b) Time series of the count of discoveries of previously-unknown chemical harms in each year.

is also increasing over time (Figure 2g). 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 firms’ innovation investments. The same pattern of upward bias would hold for firms’ lobbying investments for identical reasons.

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 expenditure¹⁸ (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

¹⁸As noted previously, the lobbying data does not contain any firm covariate information that can be used as a measure of firm scale.

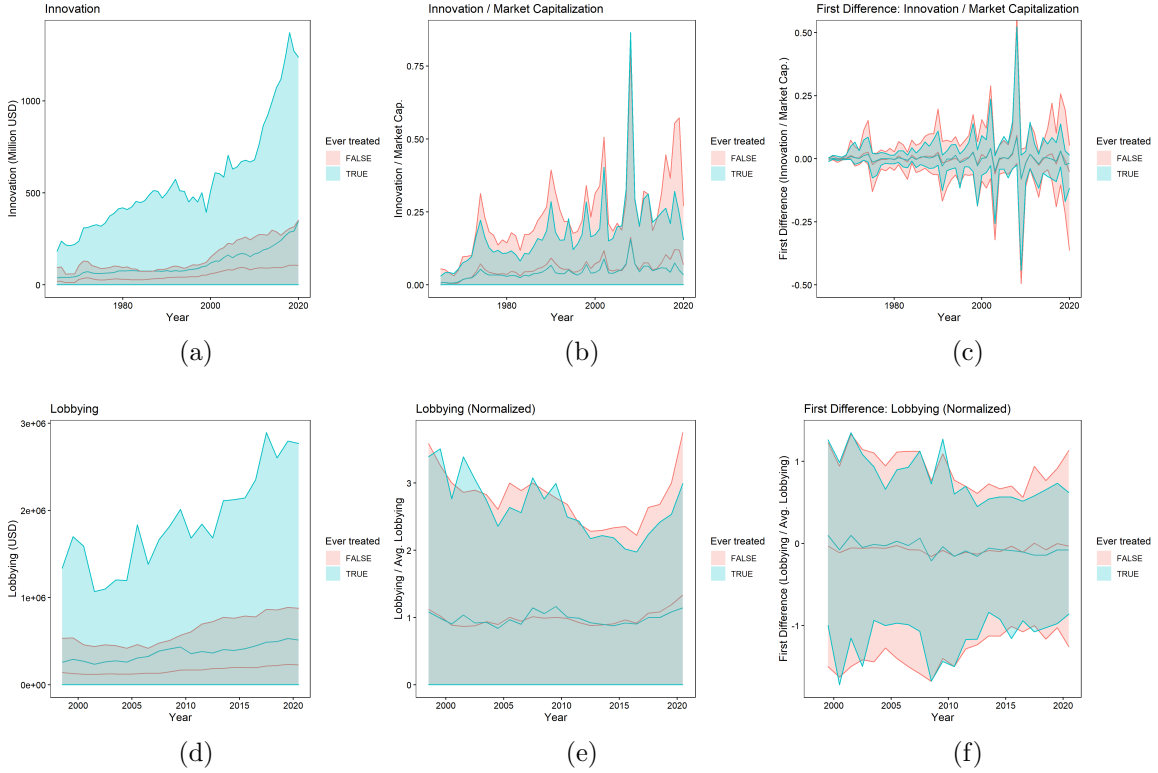


Figure 2: Time series variation in innovation (a-c) and lobbying (d-f), across treated firms (blue) and untreated firms (red). Innovation and lobbying expenditures are shown in real terms. Levels of firm innovation and lobbying investments exhibit average differences in the cross section and over time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

parallel across treated and untreated firms (see e.g. the end of both panels). First differencing each normalized outcome (Panels (c) and (f) of Figure 2), while employed due to

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 maintains balance over treated and untreated firms on average, and eliminates average time trends as well (Panels (c) and (f) of Figure 2). This results in a panel of outcomes which are balanced across treated and untreated firms both in the cross section and over time. Of course, it may be the case (for example) that some treated 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 employ both cross-sectional and temporal fixed effects to address these concerns, and isolate variation in my treatment (the random timing of scientific discoveries) that is uncorrelated with economic trends in the data.

One 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

I seek to estimate how firms jointly optimize over innovation and lobbying investments in response to the uncertain threat of future regulation. The first order conditions derived above

(Equations 1 and 2) predict that firms will increase both their innovation and lobbying expenditures in response to the new possibility of future regulation. I first test these two predictions, using as identifying variation the first scientific discovery of previously unrecognized harms from commonly used industrial chemicals.

Throughout the literature on firm innovation expenditures it is typically 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.¹⁹ This leads me to first-difference normalized innovation and lobbying expenditures in the estimation, to eliminate potential confounding from a unit root in the undifferenced data (Greene, 2003).

For both outcomes I estimate a distributed lag model, which flexibly allows for either transitory or permanent effects of information shocks on firm innovation or lobbying investments. The main specification for firm i in industry sector s and period t is

$$\Delta y_{it} = \sum_{\ell \in b} \beta_{\ell} \sum_{\tau_i} \mathbb{1}\{\ell_i\} + \phi_i + \phi_{st} + \epsilon_{it} \quad (4)$$

where y_{it} either represents firm-level annual innovation or biannual lobbying investments (normalized), b represents the bandwidth (set of leads and lags) over which lagged effects β_{ℓ} are estimated, $\mathbb{1}\{\ell_i\}$ is an indicator for whether period t is ℓ lags from a treatment in the set τ_i ,²⁰ and s indexes SIC divisions in the innovation data and sector codes (similar to SIC divisions) in the lobbying data. Firm fixed effects in levels of the outcome are implicitly captured by first-differencing the outcome variable, purging the data of any constant differences between firms. Thus, in the first differenced model ϕ_i absorbs the average difference in period innovation or lobbying expenditures, equivalent to a firm-specific time trend. 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 marketing of lobbying services, or the cyclical nature of the economy, any of which may have differential effects on innovation and lobbying across sectors of the economy. The coefficients of interest, β_{ℓ} , are estimated off of differences between

¹⁹An 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.

²⁰Formally $\mathbb{1}\{\ell_i\} = t - \ell \in \tau_i$. Note that τ_i denotes the set of treatment years for firm i .

the trajectories of within-firm deviations in innovation and lobbying investments, comparing the trajectories of treated firms to those of similar, untreated firms before and after an information shock. This implies that, in any period in which a given dummy $\mathbb{1}\{\ell_i\}$ is activated for a given discovery, a treated firm is compared to itself in other periods of time, and also to other firms in the same period of time in the same industry division that are not affected by that discovery. Standard errors are clustered 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_t of a shock on an outcome at some time t periods after the shock occurred is given by

$$\Omega_t = \sum_{\ell=0}^t \beta_{\ell} . \quad (5)$$

The identifying assumption I rely on is that the timing of scientific discovery of a new health concern is an exogenous information shock. That is, from the perspective of firm innovation and lobbying investments, the timing of these discoveries is as good as random. This assumption would be violated if firms that will be affected by potential regulation associated with an as-yet unpublished health concern have private information about this possibility of regulation and optimize their innovation or lobbying behavior accordingly. This assumption would also be violated if secular time trends in innovation or lobbying among treated firms resulted in a non-zero correlation between one of my outcomes and the state of being in the post-period, on average across my sample. Both of these threats to identification can be partially tested by a test for parallel trends, which I discuss in the Results section.

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 (e.g. solar panel producers and climate change 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 their capacity for innovation is comparatively limited or if the market for substitutes under regulation is small relative to the incumbent market. Perceived winners under regulation will lobby for regulation, enhancing their innovation payoffs and further increasing the optimal level of innovation; i.e. innovation and lobbying are complements for these firms. The reverse holds for perceived losers under regulation. Thus, since the theoretical result is ambiguous, I turn to the data to test which effect dominates in the markets covered by the present analysis.

To test whether innovation and lobbying are complements or substitutes, one might think to interact lobbying with the innovation response and interact innovation with the lobbying response. Positive interaction effects, for example, would 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 market at stake is 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.

To address this simultaneity bias, I would ideally observe exogenous cost-shifters at the firm level for both innovation and lobbying. This would enable me to instrument for a firm's post-period lobbying, for example, and test whether firms' innovation responses increase or decrease when their lobbying costs increase; and *vice versa* for firms' lobbying responses under shifting innovation costs. However, I cannot match a large number of firms across the innovation and lobbying datasets. As noted in Section 4, firm names (the only common identifier) are not standardized across these data and efforts to find similar matches across the datasets have not proved to be promising. Therefore I cannot instrument for a given firm's lobbying as an interaction term in the innovation regression, for example.

I can, however, match industry sectors across the innovation and lobbying datasets. Doing so, I estimate firm-average lobbying and innovation in the pre-period before the first discovery affecting a sector s in my data. In the presence of fixed costs over new innovation and lobbying projects that decrease with past innovation and lobbying expenditures,²¹ pre-period expendi-

²¹There is good reason to believe this is the case: staffing an unanticipated research project is easier if scientists

tures will positively correlate with post-period expenditures while avoiding the endogeneity problem associated with firm optimization. These regressions are shown in Equations 6 and 7.

$$\Delta y_{it}^{R\&D} = \sum_{\ell \in b} \beta_{\ell}^{R\&D} \mathbb{1}\{\ell_i\} + \sum_{\ell \in b} \gamma_{\ell}^{R\&D} \mathbb{1}\{\ell_i\} \times \overline{\text{Pre-Lobby}}_s + \phi_i + \phi_{st} + \epsilon_{it} \quad (6)$$

$$\Delta y_{it}^{Lobby} = \sum_{\ell \in b} \beta_{\ell}^{Lobby} \mathbb{1}\{\ell_i\} + \sum_{\ell \in b} \gamma_{\ell}^{Lobby} \mathbb{1}\{\ell_i\} \times \overline{\text{Pre-R\&D}}_s + \phi_i + \phi_{st} + \epsilon_{it} \quad (7)$$

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 . $\overline{\text{Pre-Lobby}}_s$ and $\overline{\text{Pre-R\&D}}_s$ are pre-period lobbying and innovation, defined as the average expenditure across firms in each of the most highly resolved industry sector designations available in the data sets.²² The pre-period is the three periods which precede the first discovery an SIC code or lobbying 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 sector and thus represent a cross-sectional source of variation.

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 $\sum_{\ell} \gamma_{\ell}^{R\&D}$ and $\sum_{\ell} \gamma_{\ell}^{Lobby}$ are the estimates of interest. Positive cumulative effects would be consistent with innovation and lobbying as complements in firm strategies (that is, the option to lobby increases firm innovation), and negative cumulative effects 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, they need not agree on sign and may support opposing conclusions.

can be shuffled within the existing pool of employees vs. having to hire and train new scientists. Undertaking a new lobbying effort is easier the greater one's existing lobbying network.

²²Four-digit SIC codes in the innovation data, and "industries" in the lobbying data.

7 Results

7.1 Testing first-order predictions

A simple toy 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 future payoff to each investment (discounted to the present) equals the cost of a dollar to the firm today. Because I do not have data on the market value at risk to each firm affected by any given discovery and I do not have data on firms' subjective probabilities over future regulation, I cannot directly test whether firms correctly optimize innovation and lobbying expenditures on the margin. However, I can 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 firms' response. Each point in Figure 3 represents the cumulative average treatment effect of a scientific discovery of a possible and 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, and 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 2% of a firm's market capitalization, and this response is realized over a period of about four years. This four-year cumulative effect reflects an average annual increase in innovation investments of about 0.5% 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 (i.e. treated firms do not select into treatment). 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.

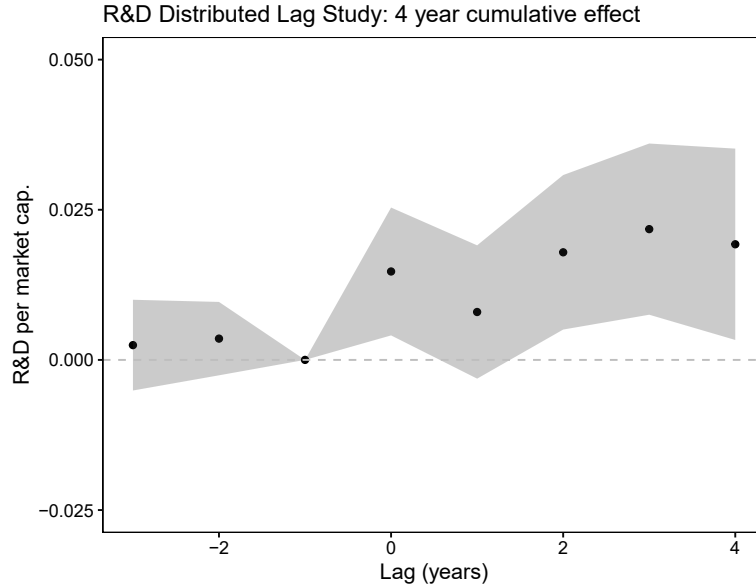


Figure 3: The cumulative effect of the discovery of a previously unknown potential externality on the innovation investments of affected firms. Standard errors are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure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

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 their slower, four-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 20% 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. The exception is the first lead (lag $\ell = -0.5$), which appears to indicate a lobbying response that may begin in the 6-month period before publication of a scientific discovery. While the scientific publication process is generally rapid, with less than a year between the date of submission and article acceptance and printing, it may be that on a six-month time frame the difference between the submission and publication dates of a finding is materially important.

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 and a half years of leads, respectively. Here I show results for six and five years 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 12 and 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

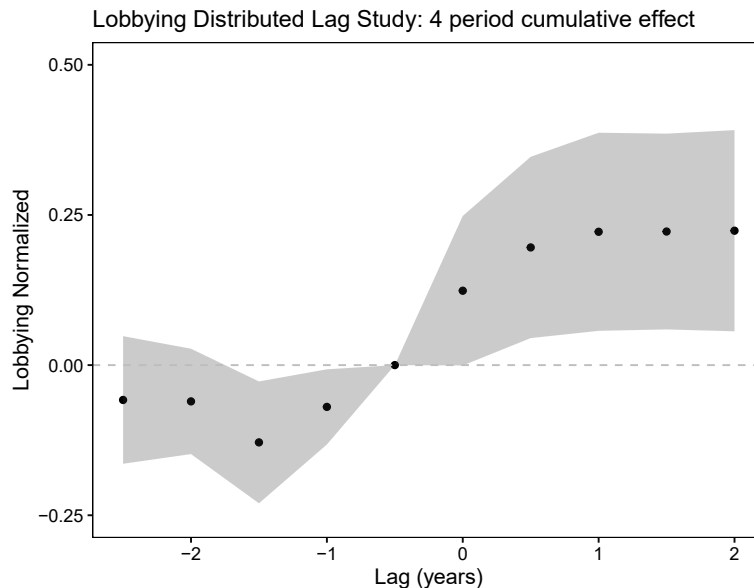


Figure 4: The cumulative effect of the discovery of a previously unknown potential externality on the lobbying investments of affected firms. Standard errors are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of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

estimated in the main models, as is the temporal structure of the responses. Importantly, leads are flat over the longer pre-period for both models, again with the exception of the first lead ($\ell = -0.5$) in the lobbying response. 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 values.

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 – D.

7.2 Evidence of innovation and lobbying specialization

Having estimated that firms increase both their innovation and lobbying investments in response to the scientific discovery of a previously unknown externality, I now ask whether firms “specialize” in innovation or lobbying. That is, do firms with a history of innovation tend to exhibit a stronger innovation response in the post period, and do firms with a history of lobbying tend to exhibit a stronger lobbying response in the post period? Such a pattern would be consistent with, for example, adjustment costs to the expansion of innovation and lobbying efforts that decline with the pre-existing scale of those efforts. For example, firms needing to initiate new research projects or hire new researchers might face lower search frictions if their existing knowledge stock and/or network of researchers is large.

For both the innovation and lobbying responses, I find suggestive evidence of firm specialization (Figure 5). It should be noted that the variation used to estimate this heterogeneity in firm innovation and lobbying responses is cross sectional variation. Thus, omitted firm characteristics that mediate the firm’s response to a discovery and are correlated with pre-period innovation or lobbying could bias the estimates of heterogeneity in treatment effects.

In the left panel of Figure 5, estimates suggest that treated firms with a higher level of pre-period innovation on average exhibit a larger innovation response than do treated firms with a lower level of pre-period innovation, though confidence intervals are widely overlapping. Firms at the 90th percentile of pre-period normalized innovation increase innovation by 2–2.5% of their market capitalization, whereas firms at the 10th percentile of pre-period normalized innovation only increase innovation by about 1% of their market capitalization.

The same holds true of the lobbying response, though some interpretation is required. Figure 5 suggests that historically “low” lobbying firms increase their lobbying more *as a fraction of pre-period expenditures*. However, when accounting for the level difference in pre-period lobbying expenditures, it is the case that firms with a higher level of pre-period lobbying exhibit a higher level of lobbying investment in their estimated lobbying response. So, the right panel of Figure 5 suggests that the post-period lobbying response does increase with pre-period lobbying

expenditures, it just does not increase proportionally. Again, like the left panel, the confidence intervals in the lobbying treatment effect heterogeneity estimates are widely overlapping.

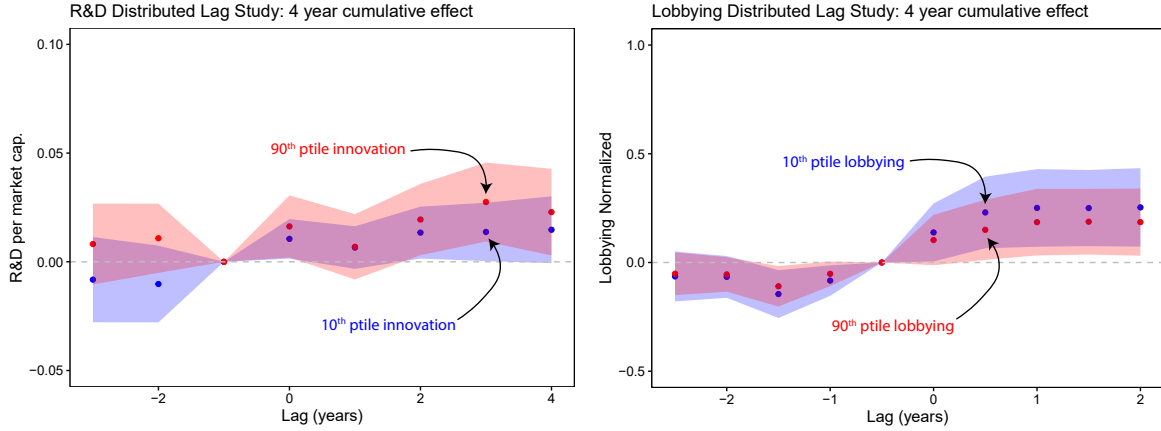


Figure 5: Heterogeneity over pre-period specialization in the cumulative effect of the discovery of a previously unknown potential externality on the innovation (left) and lobbying (right) investments of affected firms. Firms with higher levels of innovation in the pre-period tend to innovate more in the post period (left), and firms with higher levels of lobbying in the pre-period tend to lobby more in the post-period (right). In the case of lobbying, the increase is not proportionately more, which is why the 90th percentile lobbying response as a share of pre-period average lobbying lies below the 10th percentile response. Pre-period innovation and lobbying are cross-sectional interaction terms, thus the heterogeneity in effects shown here should be considered suggestive rather than causal. Both sets of estimates are consistent with fixed costs of new innovation and lobbying efforts that decline with the scale of pre-existing innovation and lobbying capabilities. Standard errors 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

Using the findings that pre-period innovation and lobbying are positively correlated with the magnitude of their respective post-period responses, I implement a suggestive test of whether innovation and lobbying are complements or substitutes in firms' strategic responses to the threat of future regulation. This enables me to provide suggestive evidence in response to the question of 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

In the model of Section 3, innovation pays off in the regulated state, the probability of which is shifted by lobbying. If firms perceive a competitive advantage under regulation (perhaps because they have an advantage in the market for substitutes or because they are hurt less by regulation than their competitors) they may lobby in favor of regulation, increasing the expected payoff to innovation and implying innovation and lobbying are complements in firm strategies. If firms perceive a net harm under the regulated state they may lobby against regulation, implying innovation and lobbying are substitutes in firm strategies. Thus the theoretical result is ambiguous, and the effect must be measured in data to determine which holds on average.

We would like to test whether a firm’s lobbying in response to the scientific discovery of a new externality increases or decreases its innovation investments in response to that same discovery. However an immediate simultaneity problem arises: if the firm value at stake under regulation is large, then the firm will both lobby more and innovate more, regardless of whether the two are complements or substitutes in firm strategies. Thus simultaneity bias will induce the estimation of a complementary relation even if the opposite were in fact the case.

As discussed in Section 6, one potential strategy to break the simultaneity bias would be to instrument for firms’ post-period lobbying responses using their pre-period lobbying investments since these are predetermined and, as demonstrated in the previous section, are positively correlated with firms’ post-period lobbying behavior. Unfortunately, firm names are not standardized across the innovation and lobbying data sets and efforts at inexact matching of firms across the data sets have not proven fruitful. I can, however, match 4-digit SIC codes in the innovation data to the roughly equivalent industry fields (“Catcodes”) in the lobbying data. Doing so leads to the estimator given in Equation 6, where pre-period industry average lobbying is interacted with each lead and lag estimate of a firm’s innovation response.

As in the previous section, the variation used to estimate heterogeneity in firm responses here (pre-period industry average lobbying) is cross sectional, and as such results should be taken as suggestive of the indicated effects.

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, though confidence intervals overlap. 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 in the post-period this is what I find; for every post-period in the cumulative effects estimate, firms in industries with less pre-period lobbying exhibit a larger innovation response, and firms in industries with more pre-period lobbying exhibit a diminished innovation response.

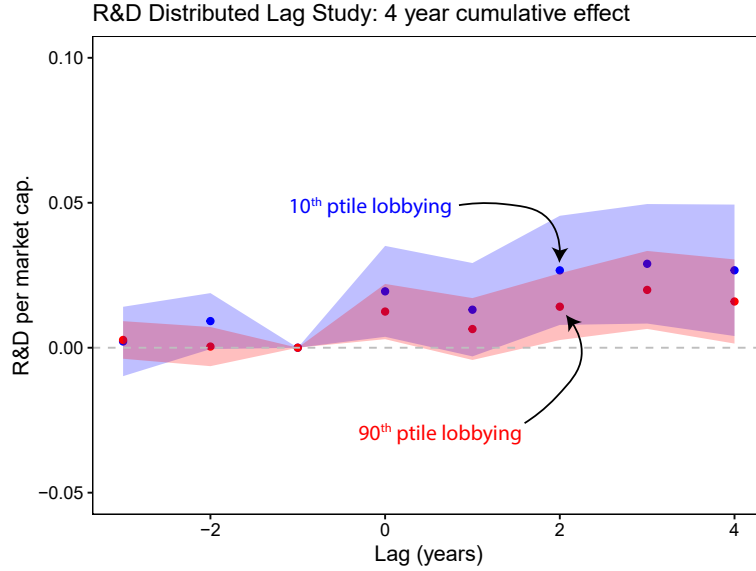


Figure 6: Heterogeneity over pre-period lobbying in the cumulative effect of the discovery of a previously unknown potential externality on the innovation investments of affected firms. Pre-period lobbying is a cross-sectional interaction term, thus the heterogeneity in effects shown here should be considered suggestive rather than causal. These estimates show that firms that had been lobbying more in the pre-period tend to innovate less in the post-period. Given the estimates shown in Figure 5, the estimates here are 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. Standard errors are robust to arbitrary correlations in innovation investments within firm over time, and capture the full intertemporal covariance structure of the estimated dynamic effect.

This result implies an additional suggestive 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), 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.

I implement this test as Equation 7, interacting industry average pre-period innovation with the firm's lobbying response. Again, consistent with innovation and lobbying as substitutes, I find that firms in industries with higher pre-period innovation tend to lobby less in response to the threat of future regulation, and firms in industries with lower pre-period innovation tend to lobby more in response to the threat of future regulation (Figure 7), though confidence intervals overlap. As before, the interaction effect estimated here is highly flexible: each lead and lag indicator in my estimator has a separate interaction with pre-period innovation and nothing constrains the cumulative effects to be consistently above or below each other in every period.

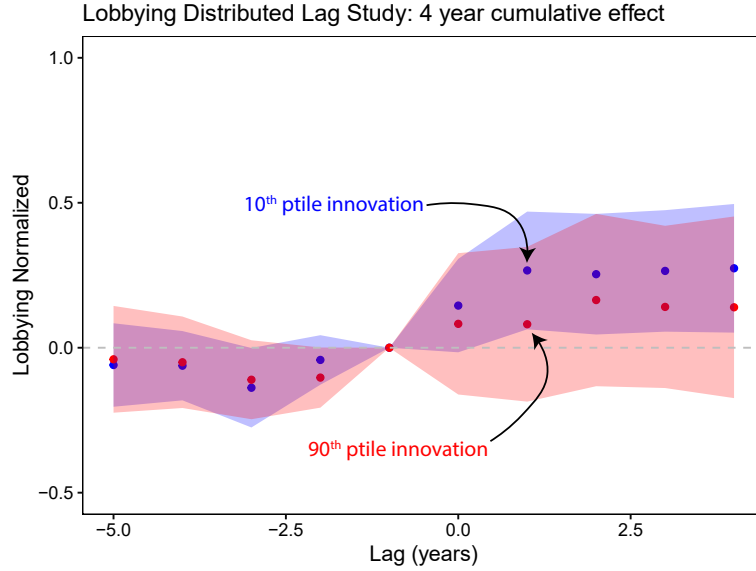


Figure 7: Heterogeneity over pre-period innovation in the cumulative effect of the discovery of a previously unknown potential externality on the lobbying investments of affected firms. Pre-period innovation is a cross-sectional interaction term, thus the heterogeneity in effects shown here should be considered suggestive rather than causal. Firms that had been innovating more in the pre-period tend to lobby less in the post-period. Given the estimates shown in Figure 5, the estimates here are consistent with innovation and lobbying as substitutes in firm strategies. This estimate supports that conclusion by testing its symmetric implication, that the option to innovate may induce firms to lobby less than they otherwise would have. Standard errors 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 firms' 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 firms' private information (e.g. a tax²³), then lobbying by firms may be nothing more than economically wasteful expenditures in 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 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), and the estimated heterogeneity in effects is noisy (Figures 6 and 7). 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 \$250,000. I estimate an average reduction in firms' four-year cumulative innovation responses from 3.2% to 1.8% of firm market capitalization for firms in industries with \$0 in pre-period lobbying versus those with \$250,000 in pre-period lobbying. The total pre-period market capitalization of treated

²³The Pigouvian tax is set at the level of marginal damages, which is not something that firms generally have private information about.

firms in my sample is \$2.4 trillion, which implies a reduction in total innovation investments of \$35 billion. Taking a value of the private returns to innovation from (Hall 1999) of 5x implies the value of forgone innovation is \$174 billion, excluding positive 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 the uncertain prospect of future regulation, forward-looking firms face a strategic incentive to innovate (should regulation come to pass), but they must trade off this innovation response against their strategic option to lobby the regulator. Lobbying changes the probability of future regulation, affecting the firm's payoff to its own innovation, as well as the innovation payoffs of other potentially regulated firms.

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 uncertain regulatory future, and does the option to lobby induce more or less innovation from firms on average.

This analysis requires me to overcome several limitations in the data. 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. Further, I would ideally observe the market structure of each affected industry, however I must pool across industries for statistical power. This means my estimates should be interpreted as average treatment effects of a scientific discovery of a chemical harm, where the industry affected is drawn at random from the industries in my sample. 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 the estimated interactions are noisy, and thus these estimates should be held with caution.

I find that firms slowly increase innovation investments by a total of 2% of their market capitalization over a four year period, but they quickly increase lobbying investments by a

total of 20% of pre-period lobbying levels over a 12-18 month period. Results are significant at conventional levels, but standard errors are large. Testing for cross-sectional heterogeneity in firms’ dynamic responses, I find evidence consistent innovation and lobbying as substitutes in firm strategies. 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. Further, both tests are highly flexible, yet the sign of the interaction is consistent in each period 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 lost market capitalization value of about \$170 billion. 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.

Multiple extensions to this work are possible; I highlight a few of them. 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 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. 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 stock of knowledge could provide a more detailed characterization of the temporal evolution of a given firms’ perceived competitive advantage. 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, which could 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.
- . 2002. “Directed technical change.” *The review of economic studies* 69 (4):781–809.
- 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, 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Dugoua, Eugenie. 2021. “Induced innovation and international environmental agreements: evidence from the Ozone regime.” *Working paper* .
- 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.
- 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.”
- 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.
- 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."
- 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 six and five years 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 12 and 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.

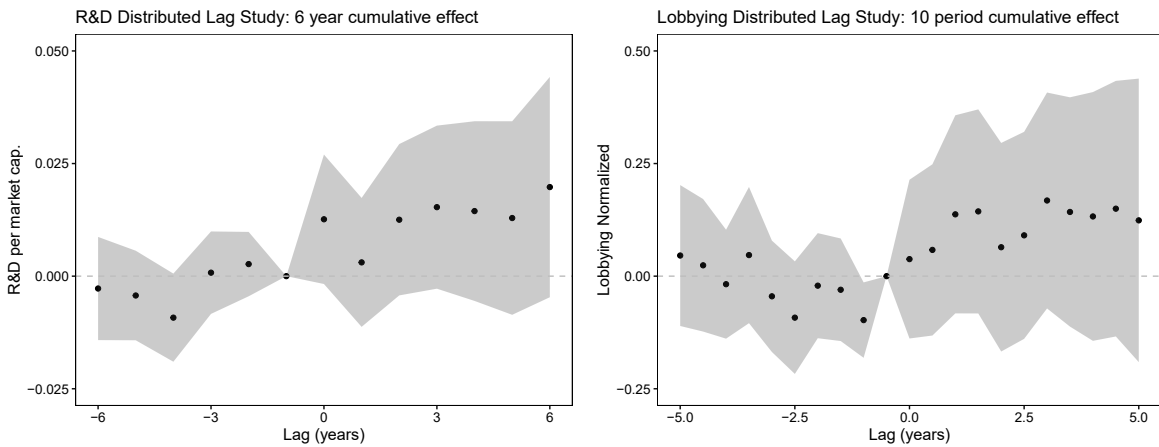


Figure A.1: The cumulative effect of the discovery of a previously unknown potential externality on the innovation (right) and lobbying (left) investments of affected firms. These estimates include an extended set of leads and lags in firm responses, increasing the noise in the estimated effects considerably. 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 six years of leads and lags for the innovation response and five years of leads and lags for the lobbying response. Standard errors 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.2), to alternative lag specifications (Figure B.3), and to dropping individual discoveries (Figure B.4, only the most influential dropped discoveries shown). Robustness of

both the innovation and lobbying responses to dropping all non-scraped studies is shown in Appendix D.

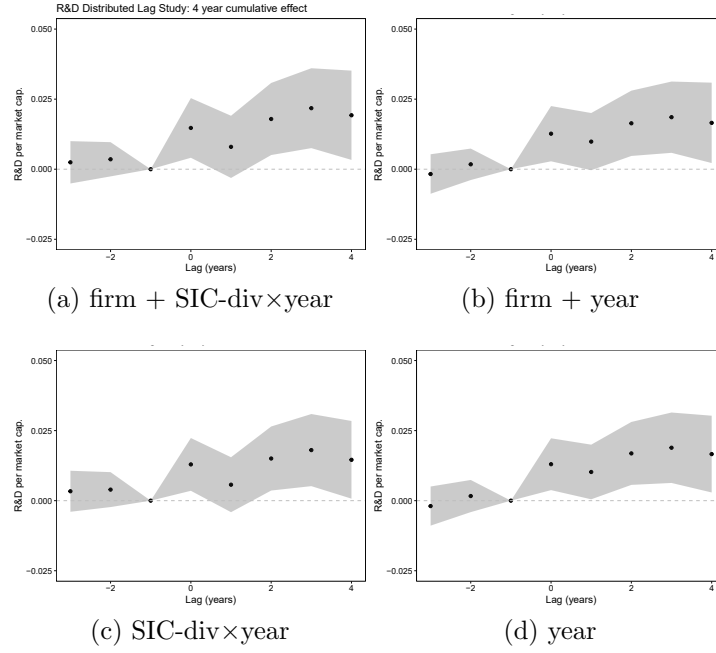


Figure B.2: Robustness of the innovation response to alternative fixed effects specifications. Panel (a) is the preferred specification presented in the main text, included here for comparison.

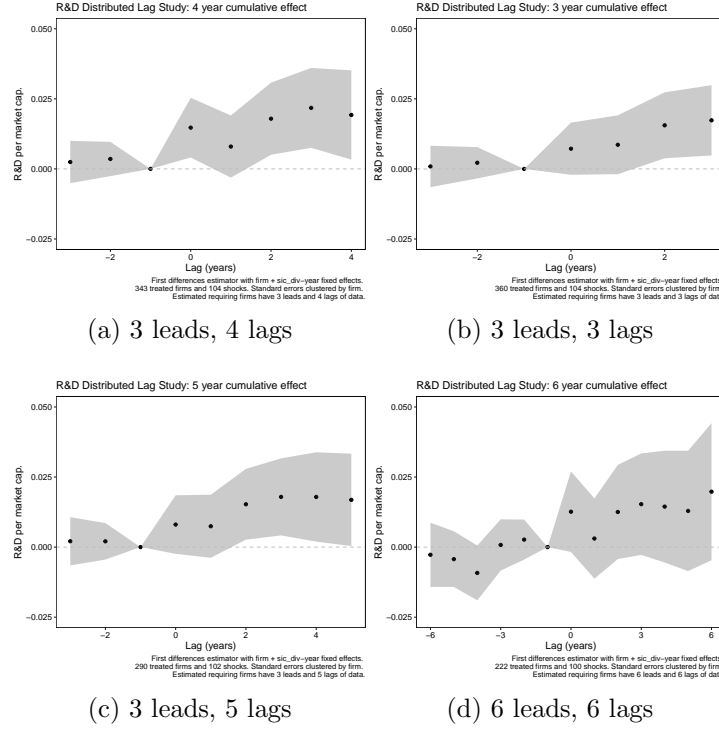


Figure B.3: Robustness of the innovation response to alternative lag specifications. Panel (a) is the preferred specification presented in the main text, included here for comparison.

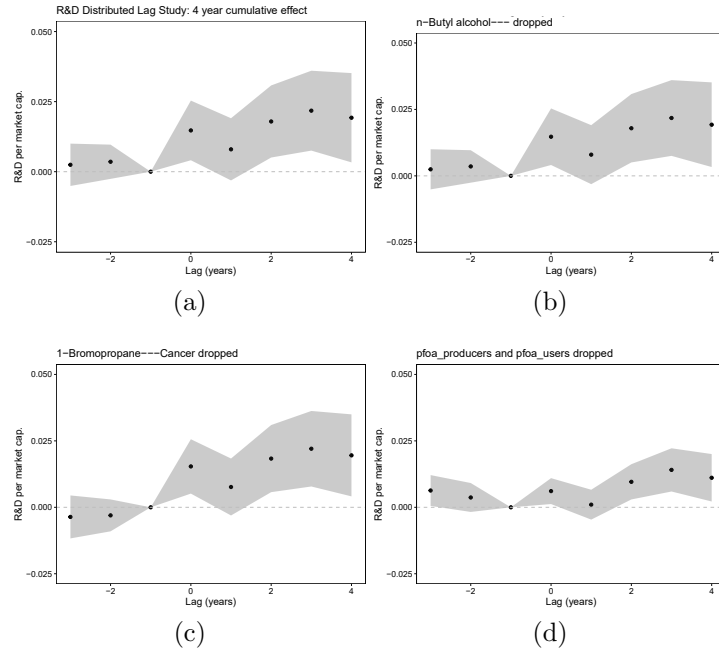


Figure B.4: Robustness of the innovation response to dropping individual discoveries. Panel (a) is the preferred specification presented in the main text, included here for comparison. Panel (b) is representative of dropping a random discovery. Panels (c) and (d) are the two discoveries that *most* affect the estimates when they are dropped.

C Robustness of the lobbying response

This section demonstrates robustness of the lobbying response to alternative fixed effects specifications (Figure C.5), to alternative lag specifications (Figure C.6), and to dropping individual discoveries (Figure C.7, only the most influential dropped discoveries shown). Robustness of both the innovation and lobbying responses to dropping all non-scraped studies is shown in Appendix D.

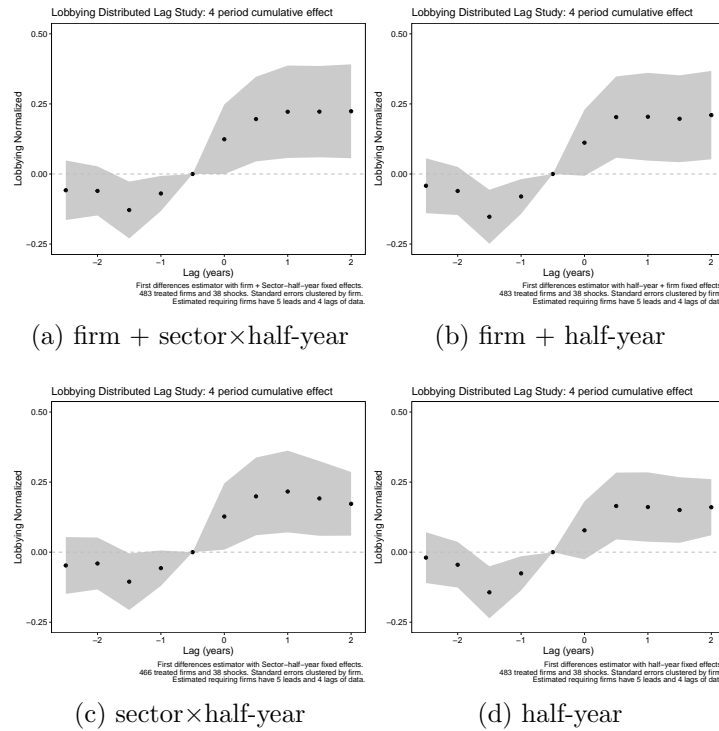


Figure C.5: Robustness of the lobbying response to alternative fixed effects specifications. Panel (a) is the preferred specification presented in the main text, included here for comparison.

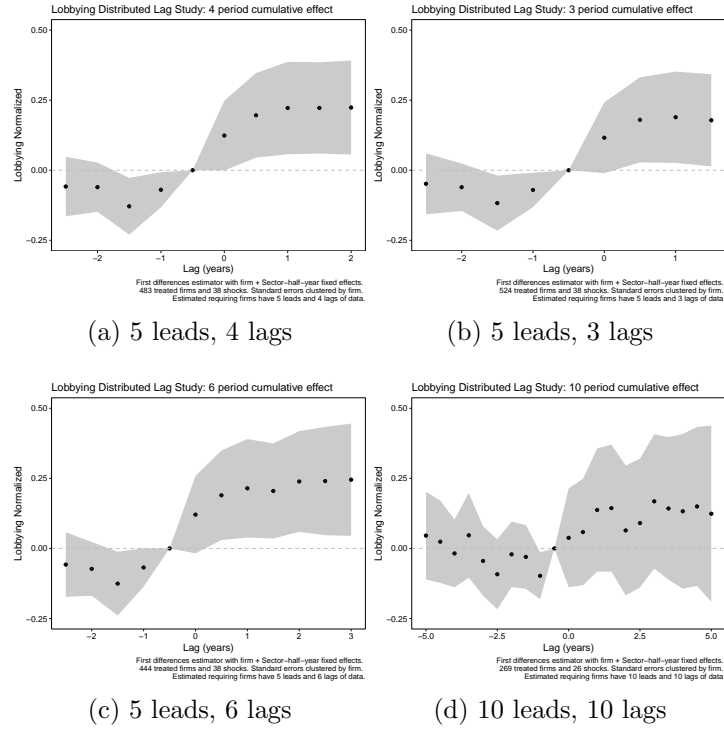


Figure C.6: Robustness of the lobbying response to alternative lag specifications. Panel (a) is the preferred specification presented in the main text, included here for comparison.

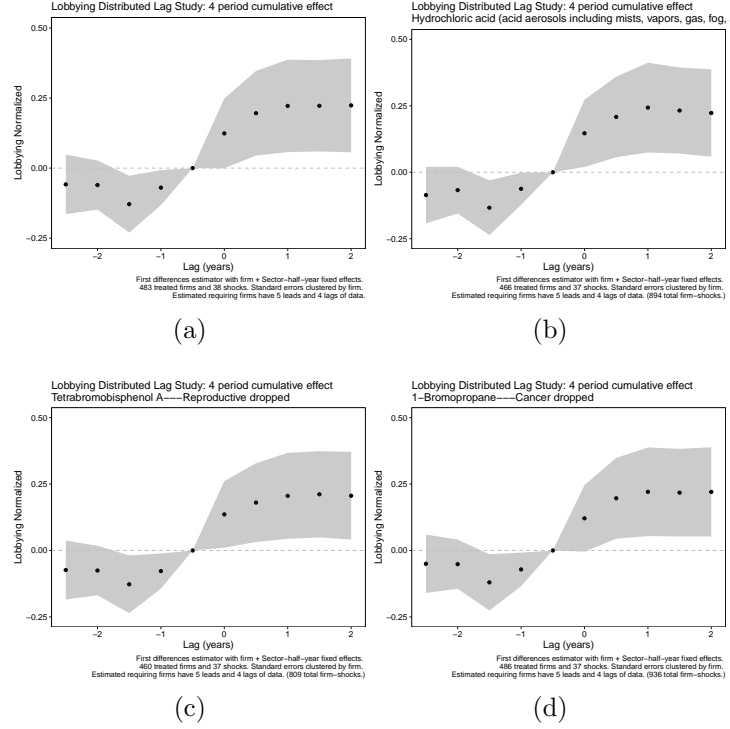


Figure C.7: Robustness of the lobbying response to dropping individual discoveries. Panel (a) is the preferred specification presented in the main text, included here for comparison. Panel (b) is representative of dropping a random discovery. Panels (c) and (d) are the two discoveries that *most* affect the estimates when they are dropped.

D Non-scraped discoveries and robustness to dropping them

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 pharmaceuticals
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 1: 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.

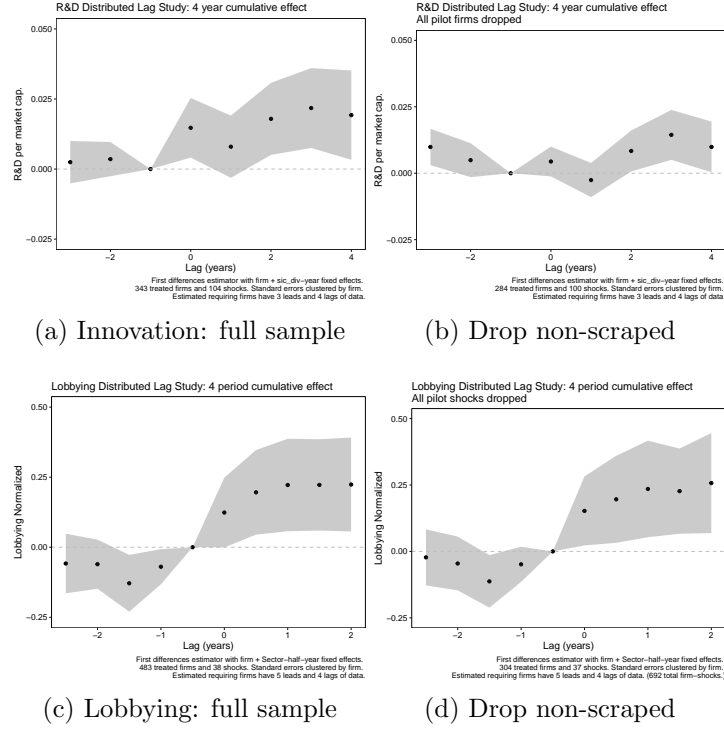


Figure D.8: Robustness of the innovation and lobbying responses to dropping all non-scraped discoveries. Panel (a) is the main-text innovation response based on the full sample, included here for comparison. Panel (b) is the innovation response when all non-scraped discoveries are dropped. Panel (c) is the main-text lobbying response based on the full sample, included here for comparison. Panel (d) is the lobbying response when all non-scraped discoveries are dropped.