Caring but sharing unintentionally: Lobbying for innovations and the leakage of knowledge*

Michael Park

January 21, 2022

Abstract

When firms engage in lobbying, their *intended* outcome is a regulatory change that benefits them. However, prior literature suggests that there may also be an *unintended* outcome of lobbying—the leakage of knowledge to competitors. In this paper, I explore when the intended and the unintended outcomes are more likely by theorizing about the relationship between lobbying and innovation. I predict that innovations that are novel are more likely to benefit from the intended regulatory changes. However, innovations that use knowledge uniquely possessed by a few firms are more likely to be compromised by the leakage of knowledge that happens during lobbying. I use new data from 1999-2013 on public U.S. firms that engaged in lobbying to federal agencies, the regulatory changes made by federal agencies, and the 16,000 patents applied for by those firms. I employ unsupervised machine learning (Doc2Vec) to measure knowledge leakage and an instrumental variable 2SLS mediation analyses to test the theory. The results suggest that the intended regulatory changes that follow lobbying can benefit innovations by facilitating wider adoption. However, unique technological knowledge that only a few firms possess may be expropriated by competitors during the process of lobbying. Overall, this paper demonstrates that fundamental aspects of innovation—such as institutional change, knowledge transfer, and technology adoption—are closely related to lobbying, a form of nonmarket activity.

*E-mail park1892@umn.edu. I would like to thank Mary Benner, Russell Funk, Aks Zaheer, and participants of the workshops in University of Minnesota's Strategic Management and Entrepreneurship department and Strategic Management Society Annual Conference 2021 for their feedback on earlier versions of the draft. Furthermore, I greatly appreciate In Song Kim's generosity regarding the LobbyView data and Patrick McLaughlin's guidance on RegData. Finally, I would like to thank all the people who agreed to be interviewed for this project.

Introduction

The nonmarket literature has noted that firms often try to change their institutional environments (Hillman & Hitt 1999, Dorobantu, Kaul & Zelner 2017, Funk & Hirschman 2017). One of the activities through which firms attempt to directly shape the regulatory environment is lobbying (De Figueiredo & Richter 2014, Jia 2018, Bombardini & Trebbi 2020). Lobbying can be understood as the act of strategically providing information to policymakers in order to be involved in the process of changing regulations (Grossman & Helpman 2001, Baumgartner, Berry, Hojnacki, Leech & Kimball 2009, De Figueiredo 2002). Policymakers are often unable to keep up with the latest developments in the industry (e.g., Funk & Hirschman 2014). By delivering key information to policymakers, firm lobbying can direct their attention to current issues of an industry (Bernstein 1955, Jones 1999, 2003). Prior literature has shown that regulatory changes made by policymakers can facilitate the introduction of new firms and products into an industry (Haveman 1993, Russo 2001, Hiatt & Park 2013). Therefore, firms from time to time likely engage in lobbying with the intention of changing regulations in ways that support their new activities.

However, there is also likely to be an unintended outcome of lobbying. Lobbying involves firms exchanging various information with competitors, lobbyists, as well as policymakers. During this process, firms may unintentionally leak knowledge to competitors (De Figueiredo & Tiller 2001, De Figueiredo & Kim 2004). Prior research has observed the unintentional leakage of knowledge in other related contexts where various parties within an industry interact and share information, such as venture capital fundraising (Cox Pahnke, McDonald, Wang & Hallen 2015), corporate investments in entrepreneurial ventures (Katila & Chen 2008, Dushnitsky & Shaver 2009), and mergers and acquisitions (Rogan 2014, Rogan & Sorenson 2014). In particular, the leakage of knowledge to competitors has been closely documented in the standard setting organizations literature ("SSO") (Toh & Miller 2017, Ranganathan, Ghosh & Rosenkopf 2018, Jones, Leiponen & Vasudeva 2020)—where firms also try to collectively change institutions as they do in lobbying. When valuable knowledge is leaked, competitors may expropriate the knowledge which can be costly for the compromised firm (Cox Pahnke et al. 2015, Toh & Miller 2017). Therefore, knowledge leakage is a potential cost for firms that engage in lobbying.

This suggests that lobbying is a double-edged sword. On the one hand, lobbying could lead to the intended regulatory changes that support firms' new activities. On the other hand, knowledge may unintentionally leak to competitors during lobbying. Overall, it remains unclear when firms are more likely to gain from the intended benefit and when they may lose out from the unintended cost of lobbying.

In order to explore these dual effects of lobbying, I develop a theoretical framework on lobbying and innovation. I choose to focus on innovation because both the intended and unintended effects of lobbying are likely to be consequential for firms' new technologies. Prior literature has identified institutional change as a significant factor for the extent to which an innovation becomes widely adopted in the market (Anderson & Tushman 1990, Tushman & Rosenkopf 1992, Hargadon & Douglas 2001). Likewise, existing literature has recognized that knowledge leakage can be critical for the potential value of a new technology (Katila & Chen 2008, Katila, Rosenberger & Eisenhardt 2008, Toh & Miller 2017). Given such close associations between the two lobbying outcomes and innovation, I investigate lobbying and its effects on firm innovations. I develop a theory to identify innovations that are likely to benefit from regulatory changes and other innovations that may lose out from knowledge leakage.

Based on this theory, I make two predictions. The first prediction is about which innovations are likely to gain from the intended outcome of regulatory change. Regulatory changes can reduce the uncertainty of new technologies (e.g., Hargadon & Douglas 2001, Sine & Lee 2009, Navis & Glynn 2010). New technologies that are more novel are likely to face higher uncertainty (Fleming 2001). Accordingly, I predict that novel innovations are more likely to benefit from lobbying and the resulting regulatory changes. The second prediction is about which innovations are likely to be compromised by the unintended outcome. Technological knowledge that is already widely used throughout the industry, when leaked during lobbying, is unlikely to be of much interest to competitors. In contrast, when technological knowledge possessed only by a few exclusive firms is leaked, because the knowledge is likely new to most competitors, many may seek to expropriate it. Therefore, I hypothesize that innovations that are technologically unique are more likely to be negatively affected from lobbying. Overall, I argue that the characteristics of an innovation is likely to determine if lobbying will tend to increase or decrease the extent to which an innovation becomes adopted in the market.

I organize the presentation in the following manner. I first provide some background information on lobbying and regulatory changes that are relevant to the framework. Subsequently, I unpack the intended and unintended outcomes of lobbying. Based on these theoretical insights as well as interviews I conducted with lobbyists, federal agency employees, and policy experts, I present the two hypotheses. Finally, I test these hypotheses using a novel, complete dataset from 1999-2013 on public U.S. firms that engaged in lobbying to federal agencies, the regulatory changes made by the corresponding federal agencies, and the 16,000 patents granted to those firms. Using text analysis and machine learning (Doc2Vec), I directly measure regulatory change and knowledge leakage. I find general support for my predictions based on a 2SLS instrumental variable approach.

This paper makes theoretical contributions to both the nonmarket and innovation literatures. I demonstrate that fundamental aspects of innovation—institutional change, knowledge transfer, and technology adoption—can be instigated by lobbying, a nonmarket activity. This insight establishes a direct connection between the innovation and corporate political action literatures for the first time.[1] Furthermore, this paper contributes to each of the the nonmarket and innovation literature in additional ways. With respect to the nonmarket

[1] Flammer & Kacperczyk (2019) could be considered a paper that previously explored the relationship between nonmarket activities and innovation outcomes. However, the paper focused on firms' corporate social responsibility (e.g., stakeholder orientation) rather than corporate political activity (e.g., lobbying).

literature, this paper sheds light on the process of lobbying, which had been treated as a "blackbox." Prior work on lobbying in the management field tended to focus only on the inputs (e.g., lobbying expenditure or the decision to hire lobbyist) and firm outcomes (e.g., contract value or regulatory decisions) (e.g., Raffiee 2017, Barber IV & Diestre 2019, Kim 2019, Diestre, Barber IV & Santaló 2020, Shi, Gao & Aguilera 2021). In contrast, I directly explore the inputs (lobbying amount), intermediary processes (regulatory changes and knowledge leakage), and the firm outcomes (innovation adoption). These insights are essential to advance our understanding of the mechanisms underlying corporate political activities. I also contribute to the innovation literature by showing that different characteristics of new technologies influence how they are affected by political processes firms choose to engage in. Overall, this paper represents one of the first attempts to directly explore the politics of innovation—the intersection between nonmarket strategy and innovation (e.g., Park & Funk 2020).

Theory

Before exploring in detail how the outcomes of lobbying are associated with the adoption of different types of innovations, I need to establish that there are lobbying and regulatory changes relevant for innovation adoption. The lobbying and regulatory change relevant for the adoption of innovations would define the scope of my theory. Therefore, I now present the arguments as to why this particular type of lobbying and regulatory change are likely to exist and thereby spell out the scope of my theory.

Lobbying can be undertaken for a variety of different purposes other than to change regulation. For example, lobbying can lower effective tax rates (Richter, Samphantharak & Timmons 2009), increase the value of government contracts (Ridge, Ingram & Hill 2017), as well as influence judicial confirmations (Caldeira & Wright 1998). Similarly, regulations change for reasons other than due to lobbying. For example, regulations change because of environmental concerns (Porter & Van der Linde 1995), in order to increase competition in a stale industry (Haveman, Russo & Meyer 2001), or to reign in excessive risk taking (Jia, Gao & Julian 2020). The aim of this paper is to understand the relationship between lobbying and the adoption of innovations. In order to do so, based on the following prior literatures I make the claim that there exists a particular set of lobbying efforts and regulatory changes that are closely related to each other due to their shared intention to influence the adoption of innovations.

Existing literature on lobbying suggests that firms often engage in lobbying with the intention to positively influence their innovations. For example, pharmaceutical companies have been shown to lobby federal agencies to advocate for favorable regulatory decisions on new drugs (e.g., Barber IV & Diestre 2019, Diestre et al. 2020). The innovation literature also suggests that regulatory changes can benefit innovation in different ways. In particular, regulations can positively influence the extent to which a particular new technology becomes widely adopted in the market over competing ones. (Hargadon & Douglas 2001, Anderson & Tushman 1990, Tushman & Rosenkopf 1992, Teece 1986). These findings from prior literature imply that there are lobbying efforts intended to change regulation in ways that are supportive of innovations and at the same time that there are regulatory changes that can be made to be supportive of the adoption of certain innovations. Put together, it seems reasonable to assume that at least some of the lobbying efforts undertaken by firms is to change regulation in ways that increase the adoption of their innovations in the market. Consequently, in this paper I theorize specifically about this type of lobbying and regulatory changes intended to increase the adoption of innovations.

In what follows, I explain how firms are able to change regulation in such ways via lobbying. I then argue which type of innovations are more likely to benefit from such lobbying efforts. Then I consider which innovations are likely to unintentionally lose out from those same lobbying efforts.

The intended outcome of lobbying

Regulatory changes

In order to change regulations in intended ways, firms must enlist the help of policymakers who have the ability to change regulations. The first step to recruiting policymakers is drawing their attention to the firm's issues. However, given the limited cognitive capacity of policymakers, their attention is known to be very scarce (Baumgartner et al. 2009, Simon 1947), especially with the increase in the size of the government in recent years (e.g., Lott 2000) as well as the growing complexity of regulations (e.g., McLaughlin, Sherouse, Febrizio & King 2020, Davis 2017, Drutman 2015). Furthermore, regulations tend to evolve around ongoing activities of firms rather than new ones such as innovations (Stigler 1971, Hargadon & Douglas 2001). Therefore, even if policymakers are interested in supporting new technologies of the industry, they are unlikely to understand new technologies well enough to know what regulatory changes may be needed.

Conversely, firms are involved in the nexus of industry regulations and their own new technologies on a daily basis. Furthermore, they are often willing and ready to conduct additional research to gather additional information that can be helpful for shaping the regulatory environment in ways that are beneficial for themselves (Grossman & Helpman 2001, Drutman 2015). In this way, firms are likely better-equipped with information about what regulation is potentially beneficial to their own new technologies. Lobbying is a direct way in which firms can relay this type of information to policymakers (Grossman & Helpman 2001, Baumgartner et al. 2009, Bombardini & Trebbi 2020). Therefore, firms will often lobby for a set of regulatory changes by delivering key pieces of information to policymakers. This informational role of lobbying is illustrated in the following quote from an interview with a lobbyist in the energy sector.

"New technologies, products, and business models often do not fit within the existing regulations. Therefore, regulations tend to lag behind. Lobbying is one of the ways in which the industry can help inform policymakers about new regulatory models needed for changing business models."

In this way, firms of an industry use lobbying to relay information to policymakers with the intention of changing regulations. Many other lobbyists relayed similar ideas that informational lobbying is often used to change regulations to help accommodate their new activities. Therefore, the intention behind some of the lobbying that firms undertake is likely closely tied to new technologies.[2]

However, at times within an industry, firms may not always want to provide consistent information to the policymakers of the industry. For example, firms may want different regulatory changes because competitors are pursuing innovations that would benefit from different policy changes. In such cases, individual firms may possibly want to simultaneously ask for conflicting regulatory changes from a policymaker. However, there is little empirical evidence to suggest that individual lobbying can effectively change regulation that applies to a collection of firms such as an entire industry.[3] This is largely because firms are more likely to be successful in garnering the attention of policymakers for a regulatory change when a collection of lobbying interests provide consistent information to policymakers. (e.g., Yackee & Yackee 2006, McKay & Yackee 2007, Hojnacki 1997, Bonardi, Holburn & Vanden Bergh 2006). For this reason, even lobbying groups that have competing interests have been shown to compromise and cooperate to bring upon regulatory changes that those organizations can collectively benefit from (e.g., Salisbury 1990, Hula 1999, Holyoke 2009). Therefore, when firms of an industry seek regulatory changes that can benefit their innovations, they tend to seek a consistent set of regulatory changes most firms can benefit from.

[2] Appendix A row 2 and 3 provide additional examples of this specific relationship between lobbying and new pproducts shared by a policy expert from Silicon Valley and lobbyist in the defense industry, respectively.

[3] Appendix A rows 1 provides an anecdote shared by an insurance industry lobbyist which describes how individual lobbying efforts to change regulations can often fail.

Novelty of innovations

The preceding argument suggests that firms lobby for regulatory changes that benefit their innovations by providing information to policymakers. Yet, it is unclear what type of regulatory change can support firms' new technologies. Thus, I now focus on a dimension of new technologies that may be important to consider for thinking about what type of regulatory changes can benefit innovations.

Innovations vary in the extent to which they are novel (Rosenkopf & Nerkar 2001, Fleming 2001, Kaplan & Vakili 2015). The novelty of an innovation is closely associated with the combination of its technological components—conceptual ideas or physical objects used to create the innovation (Fleming 2001). Specifically, innovations that use more new components are likely to be higher in novelty (Fleming 2001, Rosenkopf & Nerkar 2001). This characteristic of innovations may be an important factor in determining what type of regulatory change can be beneficial for new technologies.

As mentioned, regulatory environments evolve around things that exist rather than things that have not come into existence yet (Stigler 1971, Hargadon & Douglas 2001). Existing technological components used in less novel innovations are therefore likely to be already incorporated into the regulatory environment to some degree. Conversely, new technological components of novel innovations are unlikely to be incorporated. Prior literature suggests that novel innovations are inclined to initially face uncertainties that inhibit their adoption (Hargadon & Douglas 2001, Navis & Glynn 2010, Sovacool 2009, Sine & Lee 2009). Consequently, a regulatory change that incorporates new components into the regulations for the first time potentially represents an important type of regulatory change for novel innovations. I argue that novel innovations are likely to face uncertainty across three dimensions: political, economic, and technological. Then I illustrate how a subsequent regulatory change can alleviate these uncertainties and increase the adoption of these innovations.

Because new technological components fall outside the scope of existing regulation, there is likely to be political uncertainty regarding the technologies that use the new components.

Initially, innovations using new technological components, such as satellite radio, lack legitimacy (Navis & Glynn 2010). Political recognition in the future may help the technology gain some legitimacy. However, until then, potential adopters of the technology will be unsure if the technology will ever become institutionalized (DiMaggio & Powell 1983). Given the importance of the political environment for new technologies (Tushman & Rosenkopf 1992, Anderson & Tushman 1990), without the assurance that the technology may gain legitimacy in the future, adoption of the technology is likely to be low. Novel innovations may also initially face economic uncertainty. For example, when renewable energy technologies were being introduced for the first time, the components of renewable technologies did not fit neatly into the existing regulatory framework (Sine & Lee 2009, Sovacool 2009). This created uncertainty about the economics of renewable technologies. For example, there were inconsistent tax credit policies and a lack of regulatory guidance on what price the renewable energy should be sold for (Sovacool 2009). This contributed to the failure of renewable energy technologies becoming widely adopted early on (Sovacool 2009, Sine & Lee 2009). In addition to political and economic uncertainty, there may also be technological uncertainty. For example, at the time when the innovation of electricity light was being introduced, there was uncertainty about what configuration of the technology would be available in New York City. Edison wanted to install the new technological components—such as electric wires—underground, just like other existing utility technology components—such as gas pipes and water pipes (Hargadon & Douglas 2001). However, only components of technology that were considered to be a utility according to the existing regulations were allowed to be installed underground in New York City (Hargadon & Douglas 2001, Silverberg 1967). Initially, it was unknown if and when the novel innovation of electricity light would be considered a utility by the regulators (Hargadon & Douglas 2001, Silverberg 1967). This created uncertainty regarding what exact form of electricity light technology would be available for adoption. Edison believed the possibility that the electric wires may not be able to be installed underground would be a key barrier against the wide adoption of electricity light technology (Hargadon & Douglas

2001, Silverberg 1967). Thus, overall novel innovations can face political, economic, and technological uncertainties due to the technological components lying outside the existing regulatory framework, and this tends to have a negative influence on the market's adoption of the new technologies.

However, subsequent regulatory changes may reduce the uncertainty by introducing the new technological components into the regulatory framework. Regulatory changes can expand the scope of the regulatory environment by bringing new components into the regulatory fold. Such changes can reduce the uncertainty of novel innovations providing political, economic, and technological information. For example, regarding political uncertainty, the regulatory change of the Federal Communications Commission (FCC) around 1995 to allocate frequency spectrum for a new digital audio radio service (DARS) contributed significantly to the legitimization of satellite radio technology. The regulatory change was specifically designed to regulate satellite radios for the first time (Navis & Glynn 2010). DARS set the regulatory criteria on things such as the broadcasting frequency coordination, bringing into the regulatory fold a new combination of S-band antennas, receivers, and chipsets that handle frequencies between 2-4 gigahertz, which the regulation set aside for satellite radios (Navis & Glynn 2010). In doing so, DARS added satellite radio technology into the existing regulatory framework of the FCC, giving it some legitimacy, reducing satellite radio's political uncertainty stemming from novelty. Regarding economic uncertainty, the passage of the National Energy Act in 1978—created to directly regulate the emerging technology of renewable energy in a consistent manner—provided much needed clarity on how renewable energy would be taxed and priced, reducing economic uncertainty (Sine & Lee 2009, Sovacool 2009). In addition, the legislation resulted in the addition of the provisions around "qualified facilities" into the regulatory environment, which provided for the first time, detailed regulatory guidance on how renewable energy facilities should operate (Sine & Lee 2009, Sovacool 2009). Finally, regulatory changes can also reduce uncertainty about the technological form. For example, the regulatory change which recognized the electric light technology as one of the utility technologies eventually helped to standardize the technology, leading to a reduction in uncertainty regarding what form of electricity light technology would be available (Granovetter & McGuire 1998).

In this way, the uncertainty that innovations face can be reduced by regulatory changes, subsequently increasing the adoption of those innovations. Novel innovations tend to face higher uncertainty because they are more likely to use new technological components that are not incorporated in the existing regulations. Conversely, less novel innovations are less likely to use new components. Therefore, novel innovations, when compared to less novel ones, are more likely benefit from regulatory changes which incorporate new technological components into the regulatory environment. Informational lobbying is a way in which firms can bring upon such changes to the regulatory environment. This discussion leads to the following prediction about the intended outcome of lobbying:

Hypothesis 1. *Regulatory change that results from lobbying is more likely to have a positive effect on the adoption of innovations that are higher in novelty compared to innovations lower in novelty.*

The unintended outcome of lobbying

Leakage of knowledge

I have argued that regulatory changes that result from lobbying can have a positive association with the adoption of novel innovations that use new technological components. However, when firms lobby with the intention of benefiting their new technologies, there may be other outcomes that transpire in addition to the supportive regulatory change. Lobbying entails interaction among lobbyists and policymakers to facilitate industry lobbying efforts that can deliver desired regulatory changes (Carpenter, Esterling & Lazer 2004, Fernandez & Gould 1994). The interactions among lobbyists and policymakers may have some other consequences for firms and their innovations. In order to assess this possibility, I first explore the dynamics underlying lobbying efforts. Then I consider if and how these dynamics may ultimately influence specific innovations.

For firms of an industry to successfully gain the attention of policymakers, they need to lobby for a consistent set of regulatory changes as an industry. In order to do so, firms need to understand how others' innovations will benefit from specific regulatory changes. Individual firms in the industry are likely not completely aware of what innovations others in the industry are pursuing (Jones 1999). Therefore, initially firms are likely unclear about which innovations they should lobby for as an industry. Prior literature has suggested that the interactions among lobbyists from different firms and policymakers are a key process through which firms can cooperatively share information (Carpenter et al. 2004, Fernandez & Gould 1994). These interactions can largely be characterized as a *cooperative* effort to identify a set of regulatory changes that would benefit a large number of firms in the industry (Laumann & Knoke 1987).

However, a *competitive* dynamic will likely exist in addition to the cooperative one during the same lobbying process. Prior literature suggests that even though the formal procedures are designed to encourage competitors of an industry to cooperatively establish rules, competitive dynamics persist during the lobbying process (Drutman 2015, Holyoke 2009). In SSOs, although a cooperative procedure exists to establish technological standards, firms often act on their competitive nature by exploiting opportunities that arise during the process (Ranganathan et al. 2018, Jones et al. 2020). For example, when firms are required to disclose technological knowledge, competitors may expropriate the new knowledge that is revealed (Toh & Miller 2017). Recent research has highlighted a similar process of expropriation during litigation where competing firms have to disclose detailed knowledge about their patented technologies (Awate & Makhija 2021). For such competitive dynamics to play a significant role during lobbying as it does in SSOs and litigation, the lobbying process must also provide opportunities for firms to act on their competitive nature. Lobbying does not typically require disclosure of technological knowledge. However, firms have expressed that mere verbal interactions with competitors are effective ways in which firms can gain knowledge about competitors' innovations (Levin, Klevorick, Nelson, Winter, Gilbert & Griliches 1987, p806). Also, the inadvertent leakage of knowledge to competitors during intraindustry interactions has been documented in a variety of other settings—during fundraising among startups in the medical device industry (Pahnke, Katila & Eisenhardt 2015), corporate investments in entrepreneurial ventures (Dushnitsky & Shaver 2009), and mergers and acquisitions between advertising firms (Rogan 2014)—where disclosures to competitors are not formally required. Firms have also expressed that the leakage of knowledge during lobbying is a concern for them (De Figueiredo & Tiller 2001, De Figueiredo & Kim 2004). Therefore, sensitive technological knowledge may leak during the lobbying process. In this vein, a lobbyist I interviewed explained the different channels through which leakage often occurs during lobbying.

"You can learn a lot about other companies in the industry through lobbying. Just by talking with lobbyists from other companies, you learn about their longterm strategies. There are other ways in which even more sensitive information may travel. For example, I may choose to share more sensitive things with other lobbyists that I trust deeply. In addition, in-house lobbyists often employ the same political consultants. We have to occasionally share sensitive details with outside consultants in order to enlist their help with advocacy campaigns on the Hill, and that comes with the risk that these political consultants may choose to reveal some of these details to their other clients depending on the circumstances."

Alongside evidence from prior literature, this quote suggests that opportunities for firms to act on their competitive nature—in the form of knowledge leakage—may present themselves in the lobbying process.[4] Although lobbyist will likely not voluntarily reveal sensitive knowledge and are not required to, the lobbying process often entails sharing some sensitive

[4] Appendix A row 4 and 5 provide additional examples of the risk of knowledge leakage based on interviews with a policy expert from Silicon Valley and a federal agency employee in the transportation industry, respectively.

knowledge, especially with consultants that are hired for outsourced lobbying efforts. The lobbyist implies that oftentimes, this can lead to the sensitive knowledge leaking eventually to other competitors the lobbyist did not intend to reveal the knowledge to. Therefore, unintended knowledge leakage represents an opportunity that competitors may expropriate during lobbying. [5]

Uniqueness of innovations

I now consider which innovations may be more susceptible to the expropriation that can result from the leakage of knowledge. Just as innovations vary in the extent to which they use new technological components, innovations may vary in the extent to which they use unique technological components. For example, certain technological components may be widely understood within the industry and thus commonly used by many firms in their innovations. Other technological components may be understood only by a few firms and thus uniquely used in a few innovations of those firms. The latter represents innovations that are higher in uniqueness than the former. I now argue that such variation in the uniqueness of innovations is likely to play a critical role in determining which innovations' are expropriated by competitors.

Innovations lower in uniqueness that rely on commonly used technological components are less likely to be susceptible to expropriation. Because most firms are already familiar enough with the technological components to use them, there is likely to be little knowledge associated with these innovations that other firms do not already possess (March 1991, Katila & Chen 2008). Therefore, even though knowledge leakage about these innovations can happen during the lobbying process, it is unlikely to lead to expropriation. Conversely, most in the industry are likely to lack knowledge about unique innovations because the technological components used in these innovations are only familiar enough to the few firms who are

[5] In the same spirit, prior literature has considered lobbyists to be conduits of information for firms (Raffiee 2017, Lambert 2019, Drutman 2015, Potters & Van Winden 1992, Austen-Smith 1994, Bennedsen & Feldmann 2002, Bertrand, Bombardini & Trebbi 2014).

able to incorporate the components into new technologies. In general, firms are likely more interested in obtaining knowledge that they do not already possess than ones that they do (e.g., Rosenkopf & Almeida 2003, Katila et al. 2008) because firms can use new knowledge to create novel combinations for subsequent innovations (Fleming 2001, Fleming & Sorenson 2004). This implies that if leakage regarding unique innovations were to happen, competitors are more likely to try to expropriate the knowledge (Toh & Miller 2017). Such unintended leakage and expropriation by competitors can have a negative impact on the innovation that is compromised (Cox Pahnke et al. 2015, Myles Shaver & Flyer 2000). In particular, competitors may use the knowledge to create substitute innovations or improved versions of the same innovation. This will likely have negative consequences for the focal innovation because potential users of the focal innovation may now simply adopt the competitors' version of the technology instead. This implies that the unintended effect of lobbying is more likely to have a detrimental effect on the adoption of innovations higher in uniqueness compared to ones lower in uniqueness.

Hypothesis 2. *Knowledge leakage that results from lobbying is more likely to have a negative effect on the adoption of innovations that are higher in uniqueness compared to innovations lower in uniqueness.*

Appendix B provides a visualization of the theoretical framework and its hypotheses presented in this paper.

Methods

Data

The data used to test the hypotheses consists of three major parts: lobbying, regulation, and patent data. I obtained the lobbying variables from the lobbying report-level data available on Lobbyview (Kim 2018). This includes all details from 1999 to 2020 that registered lobbyists are required to file on a quarterly basis (Kim 2018). In essence, this contains all of the information on federal lobbying activities in the United States during this time period. With respect to the aims of this paper, the Lobbyview data provides information on the client of the lobbyist (e.g., firms that engaged in lobbying) as well as the specific branches of the federal government that was targeted by the lobbying client (e.g., specific federal agencies). I focused on lobbying reports that had been assigned a GVKEY in the client field, which indicates that a U.S. public firm was behind the lobbying effort.

RegData includes all data on federal regulations in the *Code of Federal Regulations* ("CFR") from 1970 to 2020, a document published yearly by the federal government which lists federal regulations from all agencies and thus covers every industry (Al-Ubaydli & McLaughlin 2017, McLaughlin & Sherouse 2018, 2019, McLaughlin et al. 2020). Although the CFR is exhaustive in its listing of federal regulations, it does not list the regulations according to which federal agency is behind the particular regulation. The CFR is organized according to a set of "titles" and subcomponents of titles called "parts." Al-Ubaydli & McLaughlin (2017), McLaughlin & Sherouse (2018, 2019), McLaughlin et al. (2020) matched each part within titles to federal agencies depending on how relevant that regulation is to the official purview of different federal agencies.[6] This results in a dataset that tracks the regulation set by each regulatory agency over time. I manually crosswalked the federal agency information provided in RegData with the Lobbyview data. This allows me to observe how much lobbying was directed towards specific federal agencies and the changing regulatory environment under the purview of those agencies across time.

In addition, I supplemented RegData with hand collected data from the Federal Register API ("FR"). This contains information on what category of regulatory change was made by each agency from 1994 to 2020 based on its own classification system. For example, using the data from the FR, I am able to distinguish federal agencies that made regulatory changes regarding "Science and Technology" from others that did not make any regulatory

[6]The matching of parts of the CFR to federal agencies is not a one-to-one match, but a probabilistic match. In other words, a part of the CFR is often associated with multiple agencies. This makes sense because the titles and parts often correspond to a specific industry, and an individual industry is likely to be influenced by the regulation of multiple federal agencies.

changes relevant to "Science and Technology" in a particular year.[7] Therefore, using RegData alongside FR, I am able to identify which agencies at which point in time made regulatory changes relevant to the industry's innovations.

With the lobbying and regulation data in hand, I now added patent data. I obtained data on U.S. patents from Funk & Owen-Smith (2017) and the United States Patent and Trademark Office's Patents View database. The Funk & Owen-Smith (2017) data contains the CD index measure for patents granted between 1976 to 2013 (additional information about the CD index is provided in the independent variable section as well as Appendix F). The CD index is used to measure the novelty of patents. The patent office data includes the abstract text of each patent and citations among all patents from 1976 to the present. The abstracts of patents are used to measure knowledge leakage and the uniqueness of patents. The citation data is used to track the number of forward citation a patent receives. In order to focus only on patents that had been granted to public U.S. firms engaged in lobbying, I focused on the subset of patents that had been matched to a public U.S. firm (based on PERMNO) by Kogan, Papanikolaou, Seru & Stoffman (2017).

In order to combine the agency-by-year data (i.e., lobbying and regulation data) with the patent data, I first transformed the agency-by-year data to a NAICS six-digit-by-year structure using the crosswalk provided in RegData (McLaughlin, Sherouse, Francis, Gasvoda, Nelson, Strosko & Richards 2017). Then based on the crosswalk between NAICS and CPC patent classification found in Lybbert & Zolas (2014), I matched the NAICS six-digit-by-year data to the patent data based on each patent's CPC classification and application year.[8] I crosswalked the data at the most granular level possible, which is the CPC group level, to preserve variation in the lobbying and regulatory environment that likely exists across

[7]The other categories are "Business and industry," "Environment", "Health and public welfare", "Money", and "World".

[8]Both matches here are again probabilistic matches. This adds credence to the dataset overall since a group of technologies are likely influenced not only by regulation from a single agency but by regulations from multiple agencies. In other words, these probabilistic matches are more likely to result in a dataset that captures all the relevant regulatory changes for a group of patents than a deterministic one-to-one matching scheme.

different industries. This implies that I observe variation in lobbying and regulatory changes at the CPC-group-by-year level. I also added firm-level controls available for public companies from Compustat (e.g., revenue). This final dataset captures, for each patent for each year, how much lobbying was done by the industry and the firm that was granted the patent, the resulting regulatory changes, and the number of forward citations the patents receive. Given that the lobbying data is available from 1999 and that the CD index is available up to patents granted in 2013, my dataset ranges from 1999 to 2013. Appendix C provides a visualization of how the dataset was constructed.

Dependent Variable

Adoption (log) In order to measure the extent to which an innovation is adopted in the market, I index how many new forward cites a patent receives each year. In the innovation literature, forward citations is a commonly used metric to proxy the degree to which the patent is adopted by future technologies (e.g., Fleming 2001, Hall, Jaffe & Trajtenberg 2005, Jaffe, Trajtenberg & Henderson 1993, Funk & Owen-Smith 2017). Prior literature suggests that the adoption of innovation is closely related to the regulatory environment (Anderson & Tushman 1990, Tushman & Rosenkopf 1992, Teece 1986), suggesting that I should be able to test my hypotheses regarding lobbying and innovation using the number of forward citations. I log transform the measure to account for distributional skewness and diminishing effects of large counts.

Independent Variables

Lobbying to agency (log) The report-level data on Lobbyview includes information on the client of the lobbyist (e.g., the firm that employees the lobbyist). Lobbyview has matched the client information, where applicable, to a GVKEY (Kim 2018). I focus on the set of lobby reports that include a GVKEY, which should cover all the lobbying that is done by public firms. In addition, the lobbying reports provide information on which branch of the government was lobbied during the year. I focus only on lobbying that targeted federal agencies (e.g., Barber IV & Diestre 2019, Ridge et al. 2017). This is because the federal agencies are the ones that are responsible for instituting changes in the federal regulations that are most directly relevant for industries (e.g., Ozcan & Gurses 2018, Igan, Mishra & Tressel 2012, Ridge et al. 2017). I only consider the lobbying efforts directed to agencies that made regulatory changes relevant to "Science and Technology" according to the FR. This is done in order to focus on lobbying efforts relevant for firms' innovations. I sum the dollar amount each agency was lobbied within a year at the CPC group level.[9] This results in a dataset that indexes how much each federal agency was lobbied by groups of public U.S. firms that patent in a CPC group. Although the amount of time it takes for lobbying to lead to regulatory changes is likely to vary, the general consensus is that it is not immediate (Baumgartner et al. 2009, Drutman 2015). Therefore, most studies that use lobbying as an explanatory variable lag the variable one year (e.g., Richter et al. 2009, Diestre et al. 2020), which I follow. I also log transform the measure to account for distributional skewness and diminishing effects of large values.

Regulatory change (log) RegData provides textual measures for each part of the CFR to capture changes in the regulations (McLaughlin et al. 2017). In order to track the regulatory change for each industry, I focus on the number of words in the parts of the CFR linked to each industry. For the regulatory environment to incorporate new technological components and reduce uncertainty, regulations will have to provide details on how these new technological components will be regulated. Therefore, as regulatory changes reduce uncertainty of innovations that use new technological components, they will tend to become longer. However, longer regulatory text may also capture changes in the regulations that are not necessarily expanding the scope of regulated technological components, but also restrictions that prohibit certain technologies (e.g., environmentally detrimental technologies).

[9] Note that a lobbyists may lobby multiple federal agencies in a given year for a client, which means that a single lobbying report may be associated with lobbying across multiple agencies. In such cases, I divide the dollar amount by the number of targeted entities that the report lists.

These restrictive regulatory changes may also require wordy explanations about which technologies are prohibited. In order to exclude such narrowing of regulatory scope from this measure, I exclude from the count specific words that have been flagged by RegData as vocabulary used in restrictive regulations.[10] This count of words should capture the extent to which the scope of regulations are becoming broader in general. Thereafter, to focus on the broadening of regulations relevant for new technologies, I only count parts of the CFR relevant to agencies that make regulatory changes categorized as "Science and Technology" (according to the FR) in a particular year. This should correspond closely to the degree to which the regulatory change is encompassing of new technological components of the industry each year. I log transform the measure to account for distributional skewness and diminishing effects of large counts.

Knowledge leakage (log) To measure the extent to which knowledge about an innovation is leaked, I utilize a novel deep learning methodology, Doc2Vec.[11] Based on the abstract of each patent, I use the deep learning method to count how many patents of other lobbying firms expropriate an extremely large portion of the knowledge embodied in the text of a focal patent.

According to prior literature, the technological knowledge used to create innovations reside in the technological components used (Fleming 2001). This implies that if two patents are based on a similar set of technological knowledge, they likely use a similar set of technological components. Abstracts of patents have been used in prior studies to understand the underlying technologies of patents (e.g., Kaplan & Vakili 2015, Arts, Cassiman & Gomez 2018). Therefore, I measure the similarity in technological components used in patents by comparing the similarity of their abstracts. In particular, I focus only on the nouns found

[10] RegData identifies instances of "prohibited," "shall," "may not," "must," and "required." These words represent the standard set of legal verbs and adjectives used in constraining regulatory text within the CFR (Al-Ubaydli & McLaughlin 2017, McLaughlin & Sherouse 2018, 2019).

[11] I conducted most of the computations for Doc2Vec using the Gensim package on Python (Rehurek & Sojka 2010).

in the abstracts because technological components are physical objects or conceptual ideas, which are most likely to be represented by nouns (Arthur 2021).[12]

Intuitively, using the deep learning method, I am able to measure the meaning of each noun based on the nouns that appear close to the particular noun across all patent abstracts. Inferring the meaning of a word based on other words that are used close to it is grounded in an established theoretical perspective in linguistics (Miller & Charles 1991). Using this approach I am able to capture the meaning of each noun with vectors—or the technological knowledge embodied in each technological component—-as they appear in the original text. I accomplish this without imposing any presumptions, such as the number of categories of text I expect to find, which is required for other text procedures (e.g., Latent Dirichlet Allocation). This represent a significant advantage of Doc2Vec in assessing the similarity across a large number of technical documents since it may be hard to discern even basic characteristics of the documents—such as the number of expected categories—without expert knowledge of many specialized fields.

After obtaining the vectors of technological knowledge embodied in each patent, I compute its pairwise Cosine similarity to all other patents that are applied in subsequent years by other lobbying firms within the same CPC group (i.e., excluding pairs of patents where both are applied for by the same firm and only including pairs of patents from different firms). This is done to capture knowledge flow across different firms that lobby together. Thereafter, I compute the 90^{th} percentile pairwise Cosine similarity for the CPC group each year. This threshold represent a very high level of technological knowledge similarity between two patents within a CPC group applied for one after the other by competitors. Finally, for each patent, I count the number of patents applied for by other firms each year that are higher in similarity than the 90^{th} percentile threshold. This measure captures the count of subsequent patents that are extremely similar in the knowledge to the focal patent and are applied for by other competitors in the industry who the focal firm lobbies with. Therefore,

[12] I provide more information on the preprocessing steps as well as further details in Appendix D.

it should capture the number of patents that represent substitute or competing patents to the focal patent. I log transform the measure to account for distributional skewness and diminishing effects of large counts.

Given that the measure is new and central to the claims of this paper, I now provide some validity of the measure. Recall that my argument was that when technological knowledge about the focal patent leaks, competitors can expropriate the knowledge by creating highly similar competing patents which can lead to negative consequences for the focal patent. However, one possibility is that because this measure is based on the similarity of technological knowledge used in patents, it may be capturing the extent to which a subsequent patent simply builds on a focal patent. In other words, the proposed measure could be a proxy for how much future technologies use and build on the knowledge in the focal patent, which would imply that a high value of *Knowledge leakage* actually corresponds to the focal patent becoming more foundational and thus more valuable. If this were the case, I would expect the following two empirical relationships. First, *Knowledge leakage* should be highly correlated with forward citations, since the latter is the canonical measure for the extent to which an innovation is adopted and built upon (Fleming 2001, Hall et al. 2005). Second, the measure of *Knowledge leakage* should be positively associated with the value of a patent, in direct contrast to what I have argued about the concept of knowledge leakage.

However, I find that in the sample's panel of patents, *Knowledge leakage* is very weakly correlated to forward citations (0.04, p-value < 0.001). Furthermore, in regression results presented in Appendix E, I find that *Knowledge leakage* has a consistent negative association with the market value of patents (Kogan et al. 2017), especially unique ones as implied in Hypothesis 2, even after controlling for established predictors of market value such as the number of forward citations the patent receives within five years of grant year. This provides support for my claim that *Knowledge leakage* captures the degree to which knowledge is expropriated by competitors to a point that it has material negative consequences for the focal patent.

Novelty of patents Destabilizing innovations are new technologies that disrupt existing technology often by using new technological components. In contrast, consolidating innovations tend to build on existing technology using familiar components (Funk & Owen-Smith 2017, Fleming 2001). The CD index proposed in Funk & Owen-Smith (2017) captures the extent to which a patent is destabilizing by characterizing the backward and forward citations of a focal patent. The intuition is that for destabilizing patents, future patents that cite the focal patent are likely to find the patents referenced by the focal patent to be irrelevant. This is because for subsequent patents that utilize destabilizing patents, the components that went into producing the destabilizing patents are not relevant since the destabilizing innovation introduced new technological components that did not exist previously. For consolidating patents, future patents that cite the focal patent are likely to find the patents referenced by the focal patent to be relevant. This is because for subsequent patents that utilize consolidating patents, the technological components that went into producing the consolidating patents are relevant since it is building on those same ones.

Overall, this suggests that the CD index is a suitable proxy for the extent to which patents are likely to use new technological components. Therefore, it should also correspond closely to the novelty of patents. Patents that are more destabilizing in nature should also be ones higher in novelty, while patents that are more consolidating should be ones that are lower in novelty. In this study, I measure the CD index as of five years after the year in which the focal work was published (indicated by CD_5), consistent with prior work (e.g., Park & Funk 2020, Funk & Owen-Smith 2017). I consider patents with $CD_5 > 0$ to be patents higher in novelty while I consider patents with $CD_5 < 0$ to be patents lower in novelty. Appendix F provides additional information on the CD index.

Uniqueness of patents I measure the uniqueness of patents based on the full-text abstracts of patents. Patents that are unique are likely to use unique technological components. Therefore, to focus on technological components I again limit my attention to the nouns found in patent abstracts. I assess if the set of nouns found in a patent is unique compared to the sets of nouns found in other patents of the same CPC group and application year. This comparison should inform me about the uniqueness of technological components that are used by a patent with respect to other patents that are all being discussed within the same time period by lobbying firms in an industry. I assess the uniqueness of the set of nouns in a focal patent to the sets of nouns found in other patents using Cosine distances. For a patent whose set of nouns is more distant than the CPC group-year mean Cosine distance, I consider it to be a patent higher in uniqueness. For a patent whose set of nouns is less distant than the CPC group-year mean Cosine distance, I consider it to be a patent lower in uniqueness. Further details regarding the process of obtaining this measure can be found in Appendix G.

Controls

I include key controls in all my models. I control for *Regulatory restrictiveness* because the restrictiveness of the regulatory environment can create uncertainty about technological component availability (e.g., Park & Funk 2020), which is likely to influence the extent to which a new technology is adopted. I also include the number of patents applied for each year in the CPC group, *Industry patent count*, as the number of total patents in the industry will correspond to the overall innovation activity of the industry which may influence the number of forward citations a focal patent receives. Furthermore, I control for *Industry average* CD_5 because technological trajectories of the industry can significantly influence the adoption of innovations (Anderson & Tushman 1990). To account for the variation in firms' ability to generate widely adopted new technologies over time, I include *Firm patent average forward citation*. I also include the level of *Firm revenue* since the size of firms are closely related to innovation outcomes. Finally, I control for *Firm lobbying to related agencies*, which is measured by the amount of inhouse lobbying activity directed at agencies that instituted regulations categorized as "Business and industry" by the FR. This is because all firms have limited resources, especially inside the firm, they can direct towards lobbying. Likewise, prior papers on lobbying have also controlled for lobbying activities of firms other than the main lobbying activity of interest (e.g., Diestre et al. 2020, Barber IV & Diestre 2019).

Estimation strategy

The framework presented in this paper suggests that lobbying for regulatory changes can influence innovation adoption through two different channels. Lobbying can influence innovation through the intended benefit of lobbying—by inducing changes in the regulation that are encompassing of the new technological components. Lobbying can also influence innovation through the unintended cost of lobbying—the knowledge leakage and expropriation that can happen among firms. Both of these outcomes of lobbying can be tested by using a mediation model, where *Regulatory change* and *Knowledge leakage* is tested as the mediator between lobbying and innovation adoption (Figure A1). I conduct the best practice two-step least squares (2SLS) mediation tests as suggested by Shaver (2005). In the first step, I predict *Regulatory change* and *Knowledge leakage* based on lobbying, using an instrument for *Adoption*. In the second step, I use the predicted values of *Regulatory change* and *Knowledge leakage* from the first step to estimate innovation adoption. The OLS estimates are conducted in a patent by year panel data with patent and year fixed effects, where I observe changes in the degree to which a firm's patent is adopted each year as a firm's industry engages in lobbying. Note that this is a much more conservative data structure than traditional patent studies that rely on cross-sectional patent data (Kovács, Carnabuci & Wezel 2021). This allows for the estimates to control for many time-invariant characteristics of a patent, including time-invariant firm and industry characteristics. As an example, the final step of the estimation model for *Regulatory change* takes the form, for each patent i in year t,

$$Adoption_{it} = \beta_0 + \beta_1\, Lobbying\ to\ agency_{it-1} + \beta_2\, Regulatory\ change_{it} +$$
$$\beta_3\, Controls_{it} + \alpha_i + \mu_t + \epsilon_{it},$$

where β_1 captures the association between *Lobbying to agency* and *Adoption*, β_2 captures the association between *Regulatory change* and *Adoption*, and β_3 captures the association between *Controls* and *Adoption*. Additionally, α_i represents patent fixed effects, μ_t represents year fixed effects, and ϵ_{it} represents the error term.

Results

Table 1 shows the summary statistics for key variables. Given that the sample consists of public firms engaged in lobbying and patenting activity, it contains larger public firms that have the resources to be active in politics consistently and at the forefront of new technological innovations (Drutman 2015). To be more specific, the 64 unique firms in the sample have a revenue of about 24 billion USD and spend over 4 million USD on lobbying.

Main results

Instrumental variable

In order to obtain an unbiased estimate of the mediation effect, I employ an instrumental variable, *Firm lobbying to others*, to test the mediation effect of *Regulatory change* and another instrumental variable, *Firm novel patent count*, to test the mediation effect of *Knowledge leakage* (Shaver 2005). In this approach, both instruments must not be directly associated with *Adoption*. In addition, the first instrument should be associated with *Regulatory change* while the second instrument should be associated with *Knowledge leakage*. I now explain how each instrumental variable meets these conditions.

Firm lobbying to others In Hypothesis 1, I proposed that the *Regulatory change* which results from *Lobbying to agency* is more likely to be positively associated with novel patents. Recall that *Lobbying to agency* focuses on the industry lobbying targeting federal agencies because federal agencies are responsible for changing regulations that are directly relevant for the day-to-day activities of firms (e.g., Ozcan & Gurses 2018, Igan et al. 2012, Ridge et al. 2017), including firm innovation (Diestre et al. 2020, Barber IV & Diestre 2019). While agencies are responsible for the regulations, other branches of the government such as the house, senate, and the executive office are mainly responsible for legislating acts and statues which serve as overarching guidance as to how federal agencies should regulate (Bernstein 1955). When firms lobby the other branches, it may affect the adoption of innovations, but only through changes in industry regulations that are influenced by the acts and statues. *Firm lobbying to others* is computed by counting how many times the focal firm lobbies to other branches of the government in a given year. Therefore, I should expect a positive association between *Firm lobbying to others* and *Regulatory change* as long as firms are lobbying for their innovations to the house and senate as they do to the agencies. However, there should be no direct association between *Firm lobbying to others* and *Adoption*. This implies that *Firm lobbying to others* is a valid instrument to test Hypothesis 1.

Firm novel patent count In Hypothesis 2, I predicted that the *Knowledge leakage* which results from *Lobbying to agency* is more likely to be negatively associated with unique patents. This is because unique patents are more likely to use knowledge that competitors do not already possess. However, not all of the knowledge competitors do not have may be useful for the competitors. For example, knowledge about technological components that are completely new to the industry may be hard for competitors to expropriate. Competitors may not be able to obtain those technological components or understand the components well enough to start using. Therefore, when a firm produces patents that are more novel, even if the focal firm uses exclusive technological knowledge, it maybe difficult for competitors to expropriate. As a result, I should expect to see a negative association between the number of patents with $CD_5 > 0$ applied for by a firm—*Firm novel patent count*—and *Knowledge leakage*. I also do not expect *Firm novel patents count* to influence the number of new forward citations a patent receives each year because the novelty of a patent is established based on the conditions of the technological landscape when the patent is granted. In other words, the novelty of a patent (i.e., CD_5) does not vary over time. It is possible that firms in the industry acquaint themselves with certain novel technological components as time passes, leading to them expropriating patents that were once novel. However, this should be controlled for by *Industry average* CD_5. Therefore, *Firm novel patent count* should not be directly associated with *Adoption*, especially statistically since I include patent fixed effects which controls for each patent's CD_5 value. This suggests that *Firm novel patent count* is a valid instrument to test Hypothesis 2.

The intended benefit of lobbying

Table 2 presents the results regarding the intended benefit of lobbying. Recall that Hypothesis 1 predicted that the regulatory changes that result from lobbying is more likely to have a positive effect on the adoption of innovations higher in novelty compared to those lower in novelty. Therefore, Table 2 tests the mediating role of *Regulatory change* across subsamples of patents higher in novelty and lower in novelty. Models 1 and 2 show the mediation analysis for the subsample of patents higher in novelty ($CD_5 > 0$). Model 1 shows that the relationship between the main explanatory variable of *Lobbying to agency* and *Regulatory change* is positive and significant, as lobbying to federal agencies about technology issues should lead to regulatory changes that introduce new technological components in to the regulatory framework as discussed previously. In addition, the instrument of *Firm lobbying to others* is positive as expected since firm lobbying to other branches of the government can influence industry regulations. In Model 2, I use the predicted values of *Regulatory change* from Model 1—which should be exogenous to the dependent variable given the instrument (*Firm lobbying to others*)—to predict *Adoption*. The coefficient for *Regulatory change* is positive and significant ($\beta = 2.5266$, p-value < 0.05), suggesting that the regulatory changes that result from lobbying has a positive effect on the adoption of innovations higher in novelty. The magnitude of the coefficient implies that on average when the length of regulatory text of an industry increases by 1%, yearly forward citations to novel patents are likely to increase

2.53%. In Models 3 and 4, the same set of analyses are done as Models 1 and 2 but on the sample of patents lower in novelty ($CD_5 < 0$). In contrast to patents higher in novelty, I note that in Model 4 the coefficient for *Regulatory change* negative and significant ($\beta =$ -4.5941, p-value < 0.001). The magnitude of the coefficient suggests that on average when the length of regulatory text of an industry increases by 1%, yearly forward citations to less novel patents are likely to decrease by 4.59%. This suggests that regulatory changes which introduce new technological components increase the adoption of patents higher in novelty, but decreases the adoption of patents lower in novelty. These patterns provide support for Hypothesis 1.[13]

The unintended cost of lobbying

Table 3 presents the results regarding the unintended cost of lobbying. Recall that Hypothesis 2 predicted that lobbying for regulatory changes is more likely have a negative effect on the adoption of unique innovations because competitors are more interested in expropriating knowledge they do not already possess. In order to test this claim, Table 3 estimates the mediating role of *Knowledge leakage* across subsamples of patents that are higher in uniqueness and lower in uniqueness. Models 1 to 2 show the analysis for the subsample of patents higher in uniqueness—or those whose Cosine distance form other patents in the CPC group is higher than the mean. Model 1 shows that the relationship between the main explanatory variable of *Lobbying to agency* and *Knowledge leakage* is positive and significant. This provides support for the claim that lobbying can lead to the leakage of knowledge. In addition, the instrument of *Firm novel patent count* is negatively associated with *Knowledge leakage*, which provides support for the prior argument that competitors may find it difficult to expropriate knowledge when the focal firm uses new technological components. In Model 2, I use the predicted values of *Knowledge leakage* from Model 1—which should be exogenous to the dependent given the instrument (*Firm novel patent*)—to predict *Adoption*. The coefficients for *Knowledge*

[13] The first stage Cragg-Donald Wald F statistics for both subsamples in Table 2 are greater than 15, indicating that the instrument is strong.

leakage is negative and significant (β = -0.2301, p-value $<$ 0.01), suggesting that the leakage of knowledge that results from lobbying has a negative effect on the adoption of innovations higher in uniqueness. The magnitude of the coefficient suggests that on average, when leakage increases by 1%, the number of yearly forward citations to the focal patent decreases by 0.23%. In Models 3 and 4, the same set of analyses are done as Models 1 and 2 but on the subsample of patents lower in uniqueness—or those whose Cosine distance from other patents in the CPC group is lower than the mean. Importantly, in Model 4 I note that the coefficient for *Knowledge leakage* is not significantly different from zero. This suggests that while the leakage of knowledge that results from lobbying is detrimental to the adoption of innovations higher in uniqueness, it has no significant relationship with innovations lower in uniqueness.[14] This is consistent with the arguments of Hypothesis 2.

Discussion

Although prior literature suggests that firms do engage in lobbying to benefit their innovations (e.g., Barber IV & Diestre 2019, Diestre et al. 2020), many aspects of lobbying and innovation remained unexplored. This paper is the first attempt to simultaneously understand the intended and unintended outcomes of lobbying. Based on existing literature, interviews, and a novel dataset, I develop a theory on how lobbying is associated with the adoption of different types of new technologies. In doing so, I highlight the importance of both the regulatory changes as well as the leakage of knowledge that occur during lobbying. I also show the importance of novelty and uniqueness of innovations in this political process. More specifically, I show that the outcome of lobbying is more likely to have a positive association with novel innovations through regulatory change. However, lobbying is more likely to have a negative relationship with unique innovations due to knowledge leakage. Overall, this research

[14]The first stage Cragg-Donald Wald F statistics for both subsamples in Table 3 are greater than 15, indicating that the instrument is strong.

seeks to expand the field's understanding of the relationship between business and politics by examining the politics of innovation (e.g., Park & Funk 2020).

When evaluating the results shown here, some limitations should be kept in mind. First, I am unable to observe the heterogeneity in lobbying that might exist other than what is recorded in the lobbying reports. For example, the effectiveness of lobbying may differ depending on the experience of the lobbyist or the importance of the policymaker who was targeted by the lobbyist. These details are not reported in a systematic source available to the public. However, they may play a significant role in the efficacy of lobbying. In other words, my measure of lobbying is likely to be coarse in some ways. Nonetheless, the data utilized in this study is the most exhaustive source of lobbying publicly available today. Second, I rely heavily on patent data. Although patent data have been used widely for innovation studies in the past (e.g., Fleming 2001, Funk & Hirschman 2017), they are not a perfect measure of innovation. Many innovations may not be patented (Levin et al. 1987). I am unable to take those innovations into account. Finally, the dataset relied on multiple crosswalks. I manually crosswalked the federal agency data to combine the lobbying and regulatory change data. I used the crosswalk between federal agencies and NAICS six-digit industries provided by RegData (McLaughlin et al. 2017). I also utilized the crosswalk between NAICS six-digit industries with CPC patent classification (Lybbert & Zolas 2014). I am not able to verify that the non-patent data have been appropriately transformed to match with the CPC patent classification through this series of crosswalks. However, the coherence in the results suggests that the crosswalks are largely reliable.

Notwithstanding these limitations, this study makes contributions to both the innovation and nonmarket literature in interconnected ways. Although innovation is a widely studied topic across many fields, to my knowledge no existing study has directly explored the relationship between corporate political activity and innovation. This marks a significant contribution as I show that institutional changes that are purposefully instigated by firms can have consequences for new technologies—in direct contrast to other existing innovation studies that have mostly focused on exogenous institutional changes (e.g., Arora, Belenzon & Sheer 2021, Marx, Strumsky & Fleming 2009, Balsmeier, Fleming & Manso 2017). Given the increasing importance that firms are placing on lobbying (Drutman 2015), there was a substantive need for innovation scholars to understand what kind of relationship exists between lobbying and innovation. This paper represents one of the first steps in that direction. Furthermore, it adds to the existing literature that has explored factors external to the firm which can impact innovation such as technological trajectories (Anderson & Tushman 1990), analysts (Benner 2010), institutional environments (Vasudeva, Zaheer & Hernandez 2013), and networks (Kumar & Zaheer 2019, Funk 2014).

In related ways, the paper also contributes to the nonmarket literature. Most importantly, it raises the significance of nonmarket literature overall by showing that nonmarket activities of firms can directly impact fundamental processes of innovation such as institutional change, knowledge leakage, and technology adoption. Furthermore, it sheds light on unexplored aspects of lobbying. Although industry and firm lobbying have been a focus of lobbying studies (e.g., Barber IV & Diestre 2019, Diestre et al. 2020, Ridge et al. 2017), prior literature has overlooked how specific lobbying outcomes, such as regulatory change and knowledge leakage (e.g., Carpenter et al. 2004, Laumann & Knoke 1987, Fernandez & Gould 1994), can be consequential for firms. In addition, existing literature had not considered what type of regulatory change firms may seek through lobbying for specific purposes.

Of the many different ways in which future scholars can build on the framework developed in this paper, two stand out as especially promising. First, further efforts to more closely understand the ways in which lobbying can lead to leakage of knowledge would be helpful. Based on empirical results, this paper shows that the leakage occurs with respect to technological knowledge. However, interviews with lobbyists, policy experts, and agency employees suggest that this intriguing phenomenon is a common occurrence, which likely has consequences beyond innovation. Future research should perhaps examine the network among lobbyists and policymakers to try to quantify the degree of leakage under different contexts and conditions. Second, exploring other unintended consequences of lobbying may be interesting. For example, interviews revealed that small start up firms gain significant information from joining trade associations and participating in industry lobbying. Therefore, probing the potential for "learning-by-lobbying" mechanism may further advance the field's understanding about the externalities of lobbying.

References

Al-Ubaydli, O. & McLaughlin, P. A. (2017), 'Regdata: A numerical database on industry-specific regulations for all united states industries and federal regulations, 1997–2012', *Regulation & Governance* **11**(1), 109–123.

Albert, M. B., Avery, D., Narin, F. & McAllister, P. (1991), 'Direct validation of citation counts as indicators of industrially important patents', *Research policy* **20**(3), 251–259.

Anderson, P. & Tushman, M. L. (1990), 'Technological discontinuities and dominant designs: A cyclical model of technological change', *Administrative Science Quarterly* pp. 604–633.

Andrade, F. F. N., Figueiredo, F., Silva, D. & Morais, F. (2020), Measuring disruption in song similarity networks, *in* 'Proc. of the 21st Int. Society for Music Information Retrieval Conf'.

Arora, A., Belenzon, S. & Sheer, L. (2021), 'Knowledge spillovers and corporate investment in scientific research', *American Economic Review* **111**(3), 871–98.

Arthur, W. B. (2021), 'Economics in nouns and verbs', *arXiv preprint arXiv:2104.01868* .

Arts, S., Cassiman, B. & Gomez, J. C. (2018), 'Text matching to measure patent similarity', *Strategic Management Journal* **39**(1), 62–84.

Austen-Smith, D. (1994), 'Strategic transmission of costly information', *Econometrica* pp. 955–963.

Awate, K. S. & Makhija, M. V. (2021), 'A trojan horse inside the gates? knowledge spillovers during patent litigation', *Academy of Management Journal* **X**(In press).

Azoulay, P., Fons-Rosen, C. & Graff Zivin, J. S. (2019), 'Does science advance one funeral at a time?', *American Economic Review* **109**(8), 2889–2920.

Balsmeier, B., Fleming, L. & Manso, G. (2017), 'Independent boards and innovation', *Journal of Financial Economics* **123**(3), 536–557.

Barber IV, B. & Diestre, L. (2019), 'Pushing for speed or scope? pharmaceutical lobbying and food and drug administration drug review', *Strategic Management Journal* **40**(8), 1194–1218.

Baroni, M., Dinu, G. & Kruszewski, G. (2014), Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors, *in* 'Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)', pp. 238–247.

Baumgartner, F. R., Berry, J. M., Hojnacki, M., Leech, B. L. & Kimball, D. C. (2009), *Lobbying and policy change: Who wins, who loses, and why*, University of Chicago Press.

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

Benner, M. J. (2010), 'Securities analysts and incumbent response to radical technological change: Evidence from digital photography and internet telephony', *Organization Science* **21**(1), 42–62.

Bernstein, M. H. (1955), *Regulating business by independent commission*, Princeton University Press.

Bertrand, M., Bombardini, M. & Trebbi, F. (2014), 'Is it whom you know or what you know? an empirical assessment of the lobbying process', *American Economic Review* **104**(12), 3885–3920.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), 'Latent dirichlet allocation', *the Journal of machine Learning research* **3**, 993–1022.

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

Bonardi, J.-P., Holburn, G. L. & Vanden Bergh, R. G. (2006), 'Nonmarket strategy performance: Evidence from us electric utilities', *Academy of Management Journal* **49**(6), 1209–1228.

Caldeira, G. A. & Wright, J. R. (1998), 'Lobbying for justice: Organized interests supreme court nominations, and united states senate', *American Journal of Political Science* pp. 499–523.

Carpenter, D. P., Esterling, K. M. & Lazer, D. M. (2004), 'Friends, brokers, and transitivity: Who informs whom in washington politics?', *The Journal of Politics* **66**(1), 224–246.

Choudhury, P., Wang, D., Carlson, N. A. & Khanna, T. (2019), 'Machine learning approaches to facial and text analysis: Discovering ceo oral communication styles', *Strategic Management Journal* **40**(11), 1705–1732.

Cox Pahnke, E., McDonald, R., Wang, D. & Hallen, B. (2015), 'Exposed: Venture capital, competitor ties, and entrepreneurial innovation', *Academy of Management Journal* **58**(5), 1334–1360.

Davis, S. J. (2017), 'Regulatory complexity and policy uncertainty: headwinds of our own making', *Becker Friedman Institute for Research in Economics Working Paper* .

De Figueiredo, J. M. (2002), 'Lobbying and information in politics', *Business and Politics* **4**(2), 125–129.

De Figueiredo, J. M. & Kim, J. J. (2004), 'When do firms hire lobbyists? the organization of lobbying at the federal communications commission', *Industrial and Corporate Change* **13**(6), 883–900.

De Figueiredo, J. M. & Richter, B. K. (2014), 'Advancing the empirical research on lobbying', *Annual Review of Political Science* **17**, 163–185.

De Figueiredo, J. M. & Tiller, E. H. (2001), 'The structure and conduct of corporate lobbying: How firms lobby the federal communications commission', *Journal of Economics & Management Strategy* **10**(1), 91–122.

Diestre, L., Barber IV, B. & Santaló, J. (2020), 'The friday effect: Firm lobbying, the timing of drug safety alerts, and drug side effects', *Management Science* **66**(8), 3677–3698.

DiMaggio, P. J. & Powell, W. W. (1983), 'The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields', *American Sociological Review* pp. 147–160.

Dorobantu, S., Kaul, A. & Zelner, B. (2017), 'Nonmarket strategy research through the lens of new institutional economics: An integrative review and future directions', *Strategic Management Journal* **38**(1), 114–140.

Drutman, L. (2015), *The business of America is lobbying: How corporations became politicized and politics became more corporate*, Oxford University Press.

Dushnitsky, G. & Shaver, J. M. (2009), 'Limitations to interorganizational knowledge acquisition: The paradox of corporate venture capital', *Strategic Management Journal* **30**(10), 1045–1064.

Fernandez, R. M. & Gould, R. V. (1994), 'A dilemma of state power: Brokerage and influence in the national health policy domain', *American Journal of Sociology* **99**(6), 1455–1491.

Figueiredo, F. & Andrade, N. (2019), Quantifying disruptive influence in the allmusic guide., *in* 'ISMIR', pp. 832–838.

Flammer, C. & Kacperczyk, A. (2019), 'Corporate social responsibility as a defense against knowledge spillovers: Evidence from the inevitable disclosure doctrine', *Strategic Management Journal* **40**(8), 1243–1267.

Fleming, L. (2001), 'Recombinant uncertainty in technological search', *Management Science* **47**(1), 117–132.

Fleming, L. & Sorenson, O. (2004), 'Science as a map in technological search', *Strategic Management Journal* **25**(8-9), 909–928.

Funk, R. J. (2014), 'Making the most of where you are: Geography, networks, and innovation in organizations', *Academy of Management Journal* **57**(1), 193–222.

Funk, R. J. & Hirschman, D. (2014), 'Derivatives and deregulation: Financial innovation and the demise of glass–steagall', *Administrative Science Quarterly* **59**(4), 669–704.

Funk, R. J. & Hirschman, D. (2017), 'Beyond nonmarket strategy: Market actions as corporate political activity', *Academy of Management Review* **42**(1), 32–52.

Funk, R. J. & Owen-Smith, J. (2017), 'A dynamic network measure of technological change', *Management Science* **63**(3), 791–817.

Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–74.

Granovetter, M. & McGuire, P. (1998), 'The making of an industry: electricity in the united states', *The Sociological Review* **46**(1_suppl), 147–173.

Grossman, G. M. & Helpman, E. (2001), *Special interest politics*, MIT press.

Hall, B. H., Jaffe, A. & Trajtenberg, M. (2005), 'Market value and patent citations', *RAND Journal of Economics* pp. 16–38.

Hargadon, A. B. & Douglas, Y. (2001), 'When innovations meet institutions: Edison and the design of the electric light', *Administrative Science Quarterly* **46**(3), 476–501.

Haveman, H. A. (1993), 'Follow the leader: Mimetic isomorphism and entry into new markets', *Administrative science quarterly* pp. 593–627.

Haveman, H. A., Russo, M. V. & Meyer, A. D. (2001), 'Organizational environments in flux: The impact of regulatory punctuations on organizational domains, ceo succession, and performance', *Organization Science* **12**(3), 253–273.

Hiatt, S. R. & Park, S. (2013), 'Lords of the harvest: Third-party influence and regulatory approval of genetically modified organisms', *Academy of Management Journal* **56**(4), 923–944.

Hillman, A. J. & Hitt, M. A. (1999), 'Corporate political strategy formulation: A model of approach, participation, and strategy decisions', *Academy of Management Review* **24**(4), 825–842.

Hojnacki, M. (1997), 'Interest groups' decisions to join alliances or work alone', *American Journal of Political Science* pp. 61–87.

Holyoke, T. T. (2009), 'Interest group competition and coalition formation', *American Journal of Political Science* **53**(2), 360–375.

Hula, K. W. (1999), *Lobbying together: Interest group coalitions in legislative politics*, Georgetown University Press.

Igan, D., Mishra, P. & Tressel, T. (2012), 'A fistful of dollars: lobbying and the financial crisis', *NBER Macroeconomics Annual* **26**(1), 195–230.

Jaffe, A. B., Trajtenberg, M. & Henderson, R. (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citations', *The Quarterly Journal of Economics* **108**(3), 577–598.

Jia, N. (2018), 'The "make and/or buy" decisions of corporate political lobbying: Integrating the economic efficiency and legitimacy perspectives', *Academy of Management Review* **43**(2), 307–326.

Jia, Y., Gao, X. & Julian, S. (2020), 'Do firms use corporate social responsibility to insure against stock price risk? evidence from a natural experiment', *Strategic Management Journal* **41**(2), 290–307.

Jones, B. D. (1999), 'Bounded rationality', *Annual Review of Political Science* **2**(1), 297–321.

Jones, B. D. (2003), 'Bounded rationality and political science: Lessons from public administration and public policy', *Journal of Public Administration Research and Theory* **13**(4), 395–412.

Jones, S. L., Leiponen, A. & Vasudeva, G. (2020), 'The evolution of cooperation in the face of conflict: Evidence from the innovation ecosystem for mobile telecom standards development', *Strategic Management Journal* .

Kaplan, S. & Vakili, K. (2015), 'The double-edged sword of recombination in breakthrough innovation', *Strategic Management Journal* **36**(10), 1435–1457.

Katila, R. & Chen, E. L. (2008), 'Effects of search timing on innovation: The value of not being in sync with rivals', *Administrative Science Quarterly* **53**(4), 593–625.

Katila, R., Rosenberger, J. D. & Eisenhardt, K. M. (2008), 'Swimming with sharks: Technology ventures, defense mechanisms and corporate relationships', *Administrative Science Quarterly* **53**(2), 295–332.

Kelly, B., Papanikolaou, D., Seru, A. & Taddy, M. (2018), Measuring technological innovation over the long run, Technical report, National Bureau of Economic Research.

Kim, I. S. (2018), 'Lobbyview: Firm-level lobbying & congressional bills database', *Unpublished manuscript, MIT, Cambridge, MA. http://web. mit. edu/insong/www/pdf/lobbyview. pdf Google Scholar Article Location* .

Kim, J. H. (2019), 'Is your playing field unleveled? us defense contracts and foreign firm lobbying', *Strategic Management Journal* **40**(12), 1911–1937.

Kogan, L., Papanikolaou, D., Seru, A. & Stoffman, N. (2017), 'Technological innovation, resource allocation, and growth', *The Quarterly Journal of Economics* **132**(2), 665–712.

Kovács, B., Carnabuci, G. & Wezel, F. C. (2021), 'Categories, attention, and the impact of inventions', *Strategic Management Journal* **42**(5), 992–1023.

Kumar, P. & Zaheer, A. (2019), 'Ego-network stability and innovation in alliances', *Academy of Management Journal* **62**(3), 691–716.

Lambert, T. (2019), 'Lobbying on regulatory enforcement actions: Evidence from us commercial and savings banks', *Management Science* **65**(6), 2545–2572.

Laumann, E. O. & Knoke, D. (1987), *The organizational state: Social choice in national policy domains*, University of Wisconsin Press.

Leahey, E. (2016), ‘From sole investigator to team scientist: Trends in the practice and study of research collaboration’, *Annual Review of Sociology* **42**, 81–100.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R. & Griliches, Z. (1987), ‘Appropriating the returns from industrial research and development’, *Brookings Papers on Economic Activity* **1987**(3), 783–831.

Lott, Jr, J. R. (2000), ‘A simple explanation for why campaign expenditures are increasing: The government is getting bigger’, *The Journal of Law and Economics* **43**(2), 359–394.

Lybbert, T. J. & Zolas, N. J. (2014), ‘Getting patents and economic data to speak to each other: An ‘algorithmic links with probabilities’ approach for joint analyses of patenting and economic activity’, *Research Policy* **43**(3), 530–542.

March, J. G. (1991), ‘Exploration and exploitation in organizational learning’, *Organization Science* **2**(1), 71–87.

Marx, M., Strumsky, D. & Fleming, L. (2009), ‘Mobility, skills, and the michigan non-compete experiment’, *Management Science* **55**(6), 875–889.

McKay, A. & Yackee, S. W. (2007), ‘Interest group competition on federal agency rules’, *American Politics Research* **35**(3), 336–357.

McLaughlin, P. A. & Sherouse, O. (2018), ‘Regdata us 3.1 annual (dataset)’, *QuantGov, Mercatus Center at George Mason University, Arlington, VA* .

McLaughlin, P. A. & Sherouse, O. (2019), ‘Regdata 2.2: A panel dataset on us federal regulations’, *Public Choice* **180**(1), 43–55.

McLaughlin, P. A., Sherouse, O., Febrizio, M. & King, M. S. (2020), ‘Is dodd-frank the biggest law ever?’, *Journal of Financial Regulation* .

McLaughlin, P. A., Sherouse, O., Francis, D., Gasvoda, M., Nelson, J., Strosko, S. & Richards, T. (2017), ‘Regdata 3.0 user’s guide’, *Available at SSRN 3044802* .

Miller, G. A. & Charles, W. G. (1991), ‘Contextual correlates of semantic similarity’, *Language and cognitive processes* **6**(1), 1–28.

Myles Shaver, J. & Flyer, F. (2000), ‘Agglomeration economies, firm heterogeneity, and foreign direct investment in the united states’, *Strategic Management Journal* **21**(12), 1175–1193.

Navis, C. & Glynn, M. A. (2010), ‘How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005’, *Administrative Science Quarterly* **55**(3), 439–471.

Niu, L.-Q. & Dai, X.-Y. T. (2015), ‘Learning distributed representations of topics’, *CoRR, abs/1506.08422* .

Ozcan, P. & Gurses, K. (2018), ‘Playing cat and mouse: Contests over regulatory categorization of dietary supplements in the united states’, *Academy of Management Journal* **61**(5), 1789–1820.

Pahnke, E. C., Katila, R. & Eisenhardt, K. M. (2015), 'Who takes you to the dance? how partners' institutional logics influence innovation in young firms', *Administrative Science Quarterly* **60**(4), 596–633.

Park, M. & Funk, R. (2020), 'Regulation and innovation revisited: How restrictive regulations can promote destabilizing new technologies', *Working Paper* .

Park, M., Leahey, E. & Funk, R. (2021), 'The decline of disruptive science and technology', *https://arxiv.org/abs/2106.11184* .

Porter, M. E. & Van der Linde, C. (1995), 'Toward a new conception of the environment-competitiveness relationship', *Journal of Economic Perspectives* **9**(4), 97–118.

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

Raffiee, J. (2017), 'Employee mobility and interfirm relationship transfer: Evidence from the mobility and client attachments of united states federal lobbyists, 1998–2014', *Strategic Management Journal* **38**(10), 2019–2040.

Ranganathan, R., Ghosh, A. & Rosenkopf, L. (2018), 'Competition–cooperation interplay during multifirm technology coordination: The effect of firm heterogeneity on conflict and consensus in a technology standards organization', *Strategic Management Journal* **39**(12), 3193–3221.

Rehurek, R. & Sojka, P. (2010), Software framework for topic modelling with large corpora, *in* 'In Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks', Citeseer.

Richter, B. K., Samphantharak, K. & Timmons, J. F. (2009), 'Lobbying and taxes', *American Journal of Political Science* **53**(4), 893–909.

Ridge, J. W., Ingram, A. & Hill, A. D. (2017), 'Beyond lobbying expenditures: How lobbying breadth and political connectedness affect firm outcomes', *Academy of Management Journal* **60**(3), 1138–1163.

Rogan, M. (2014), 'Too close for comfort? the effect of embeddedness and competitive overlap on client relationship retention following an acquisition', *Organization Science* **25**(1), 185–203.

Rogan, M. & Sorenson, O. (2014), 'Picking a (poor) partner: A relational perspective on acquisitions', *Administrative Science Quarterly* **59**(2), 301–329.

Rosenkopf, L. & Almeida, P. (2003), 'Overcoming local search through alliances and mobility', *Management Science* **49**(6), 751–766.

Rosenkopf, L. & Nerkar, A. (2001), 'Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry', *Strategic management journal* **22**(4), 287–306.

Russo, M. V. (2001), 'Institutions, exchange relations, and the emergence of new fields: Regulatory policies and independent power production in america, 1978–1992', *Administrative Science Quarterly* **46**(1), 57–86.

Salisbury, R. H. (1990), *The paradox of interest groups in Washington—More groups, less clout*, Springer.

Shaver, J. M. (2005), 'Testing for mediating variables in management research: Concerns, implications, and alternative strategies', *Journal of Management* **31**(3), 330–353.

Shi, W., Gao, C. & Aguilera, R. V. (2021), 'The liabilities of foreign institutional ownership: Managing political dependence through corporate political spending', *Strategic Management Journal* **42**(1), 84–113.

Silverberg, R. (1967), *Light for the world: Edison and the power industry*, Princeton, NJ: Van Nostrand.

Simon, H. A. (1947), *Administrative behavior*, Simon and Schuster.

Sine, W. D. & Lee, B. H. (2009), 'Tilting at windmills? the environmental movement and the emergence of the us wind energy sector', *Administrative Science Quarterly* **54**(1), 123–155.

Sovacool, B. K. (2009), 'The importance of comprehensiveness in renewable electricity and energy-efficiency policy', *Energy Policy* **37**(4), 1529–1541.

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

Teece, D. J. (1986), 'Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy', *Research Policy* **15**(6), 285–305.

Toh, P. K. & Miller, C. D. (2017), 'Pawn to save a chariot, or drawbridge into the fort? firms' disclosure during standard setting and complementary technologies within ecosystems', *Strategic Management Journal* **38**(11), 2213–2236.

Tushman, M. L. & Rosenkopf, L. (1992), 'Organizational determinants of technological change: toward a sociology of technological evolution', *Research in Organizational Behavior* **14**, 311–347.

Vasudeva, G., Zaheer, A. & Hernandez, E. (2013), 'The embeddedness of networks: Institutions, structural holes, and innovativeness in the fuel cell industry', *Organization Science* **24**(3), 645–663.

Wu, L., Wang, D. & Evans, J. A. (2019), 'Large teams develop and small teams disrupt science and technology', *Nature* **566**(7744), 378–382.

Yackee, J. W. & Yackee, S. W. (2006), 'A bias towards business? assessing interest group influence on the us bureaucracy', *The Journal of Politics* **68**(1), 128–139.

Table 1: Summary statistics

Variable	N	Mean	SD	Min	Max	Level of measurement
Dependent variable						
Adoption (log)	144,521	0.35	0.59	0.00	6.00	Patent x Year
Independent variables						
Lobbying to agency lagged (log)	126,985	17.15	1.73	6.70	21.96	CPC group x Year
Regulatory change (log)	142,991	12.43	1.99	0.78	16.91	CPC group x Year
Knowledge leakage (log)	57,541	1.56	1.75	0	7.47	Patent x Year
Patent novelty	144,521	0.02	0.14	-1.00	1.00	Patent
Patent uniqueness	138,185	0.67	0.15	0.00	0.94	Patent
Instrumental variable						
Firm lobbying to others lagged (log)	84,768	3.66	.88	0.00	5.81	Firm x Year
Firm novel patent count (log)	128,329	6.95	1.35	0.00	7.79	Firm x Year
Control variables						
Regulatory restrictiveness (log)	142,911	7.89	2.02	-3.64	12.24	CPC group X Year
Industry patent count (log)	144,521	4.99	1.47	0.00	6.78	CPC group x Year
Industry average CD_5 (log)	138,397	-3.55	0.67	-10.53	0.01	CPC group x Year
Firm patent average forward citations (log)	144,521	1.60	0.23	0.00	3.22	Firm x Year
Firm revenue (log)	144,517	8.48	2.38	1.12	11.00	Firm x Year
Firm lobbying to related agencies lagged (log)	84,768	8.78	4.69	0.00	13.90	Firm x Year

Table 2: The intended benefit of lobbying: Regulatory change

	Sample: Patents higher in novelty		*Sample: Patents lower in novelty*	
	DV: Regulatory change (log)	DV: Adoption (log)	DV: Regulatory change (log)	DV: Adoption (log)
	(1)	(2)	(3)	(4)
Independent variables				
Lobbying to agency lagged (log)	0.0562***	-0.2196**	0.0662***	0.2442**
	(0.0024)	(0.0735)	(0.0040)	(0.0768)
Regulatory change (log)		2.5266*		-4.5941***
		(1.2810)		(1.0974)
Instrumental variable				
Firm lobbying to others lagged (log)	0.0154***		0.0247***	
	(0.0037)		(0.0050)	
Control variables				
Regulatory restrictiveness (log)	1.0390***	-2.6849*	1.0405***	4.7045***
	(0.0034)	(1.3310)	(0.0064)	(1.1419)
Industry patent count (log)	0.0219***	0.0946*	0.0111*	0.3776***
	(0.0035)	(0.0372)	(0.0055)	(0.0460)
Industry average CD_5 (log)	0.0492***	-0.0758	0.0372***	0.1711***
	(0.0066)	(0.0673)	(0.0088)	(0.0486)
Firm patent average forward citations (log)	-0.0012	0.2361	0.0007	0.5230*
	(0.0217)	(0.1389)	(0.0312)	(0.2113)
Firm revenue (log)	0.0026	-0.1232**	0.0126	-0.0583
	(0.0095)	(0.0425)	(0.0117)	(0.0711)
Firm lobbying to related agencies lagged (log)	-0.0005	-0.0047*	-0.0000	0.0053
	(0.0003)	(0.0019)	(0.0005)	(0.0030)
Constant	3.2348***	-5.8026	2.9421***	14.6640***
	(0.1264)	(4.2483)	(0.2015)	(3.4950)
Patent fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.9990	0.4814	0.9991	0.5288
N	33,418	33,418	18,369	18,369
Unique firms	39	39	31	31
Unique patents	3,775	3,775	2,156	2,156

Notes: These models present the mediation analysis of regulatory change based on a two-staged least squares approach (Shaver 2005). The results suggest that the adoption of patents higher in novelty (Models 1 and 2)–compared to ones lower in novelty (Models 3 and 4)—are more likely to be positively mediated by regulatory changes, providing support for Hypothesis 1. Cluster-robust standard errors (on individual patents) are shown in parentheses; p-values correspond to two-tailed tests.

*$p<0.05$; **$p<0.01$; ***$p<0.001$

†The two-stage least squares approach requires a variable that is related to the proposed mediator (*Regulatory change*) but is directly unrelated to the dependent variable (*Adoption*) to be included in the first stage (Shaver 2005). Therefore, Models 1, and 3 include *Firm lobbying to others lagged (log)* as the instrument.

Table 3: The unintended cost of lobbying: Knowledge leakage

	Sample: Patents higher in uniqueness		*Sample: Patents lower in uniqueness*	
	DV: Knowledge leakage (log)	DV: Adoption (log)	DV: Knowledge leakage (log)	DV: Adoption (log)
	(1)	(2)	(3)	(4)
Independent variables				
Lobbying to agency lagged (log)	0.2295***	0.0271	0.2129***	0.0142
	(0.0372)	(0.0272)	(0.0378)	(0.0371)
Knowledge leakage (log)		-0.2301**		-0.0187
		(0.0773)		(0.1446)
Instrumental variable				
Firm novel patent count (log)	-0.4210***		-0.2611***	
	(0.0735)		(0.0775)	
Control variables				
Regulatory restrictiveness (log)	0.1545***	0.0030	0.1280***	-0.0249
	(0.0236)	(0.0188)	(0.0237)	(0.0241)
Industry patent count (log)	-0.8081***	-0.0876	-0.6591***	0.1772
	(0.0667)	(0.0751)	(0.0635)	(0.1043)
Industry average CD_5 (log)	-0.2740***	-0.0203	-0.2211***	0.0259
	(0.0590)	(0.0323)	(0.0637)	(0.0420)
Firm patent average forward citations (log)	-0.0377	0.1941	0.6768*	0.3226
	(0.2968)	(0.1437)	(0.3338)	(0.2007)
Firm revenue (log)	0.1724	-0.0584	0.6449***	-0.1145
	(0.1064)	(0.0532)	(0.1240)	(0.1070)
Firm lobbying to related agencies lagged (log)	-0.0030	-0.0034	0.0004	-0.0037
	(0.0046)	(0.0020)	(0.0047)	(0.0020)
Constant	1.1132	1.0019	-5.8045***	0.2157
	(1.5886)	(0.7052)	(1.7512)	(1.3574)
Patent fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.8522	0.5434	0.8515	0.5455
N	18,367	18,367	16,112	16,112
Unique patents	2,702	2,702	2,432	2,432
Unique firms	27	27	31	31

Notes: These models present the mediation analysis of knowledge leakage based on a two-staged least squares approach (Shaver 2005). The results suggest that the adoption of patents higher in uniqueness (Models 1 and 2)–compared to ones lower in uniqueness (Models 3 and 4)—are more likely to be negatively mediated by regulatory changes, providing support for Hypothesis 2. Cluster-robust standard errors (on individual patents) are shown in parentheses; p-values correspond to two-tailed tests.
*$p<0.05$; **$p<0.01$; ***$p<0.001$
†The two-stage least squares approach requires a variable that is related to the proposed mediator (*Knowledge leakage*) but is directly unrelated to the dependent variable (*Adoption*) to be included in the first stage (Shaver 2005). Therefore, Models 1, and 3 include *Firm novel patent count (log)* as the instrument.

Appendix

Section	Description	Category
Appendix A	Additional quotes from interviews	Supplemental information
Appendix B	Visualization of theoretical framework	Visualization
Appendix C	Visualization of data sources	Visualization
Appendix D	Patent knowledge leakage procedure	Supplemental information
Appendix E	Validation of the measure for knowledge leakage	Supplemental analysis
Appendix F	CD index	Supplemental information
Appendix G	Patent uniqueness procedure	Supplemental information

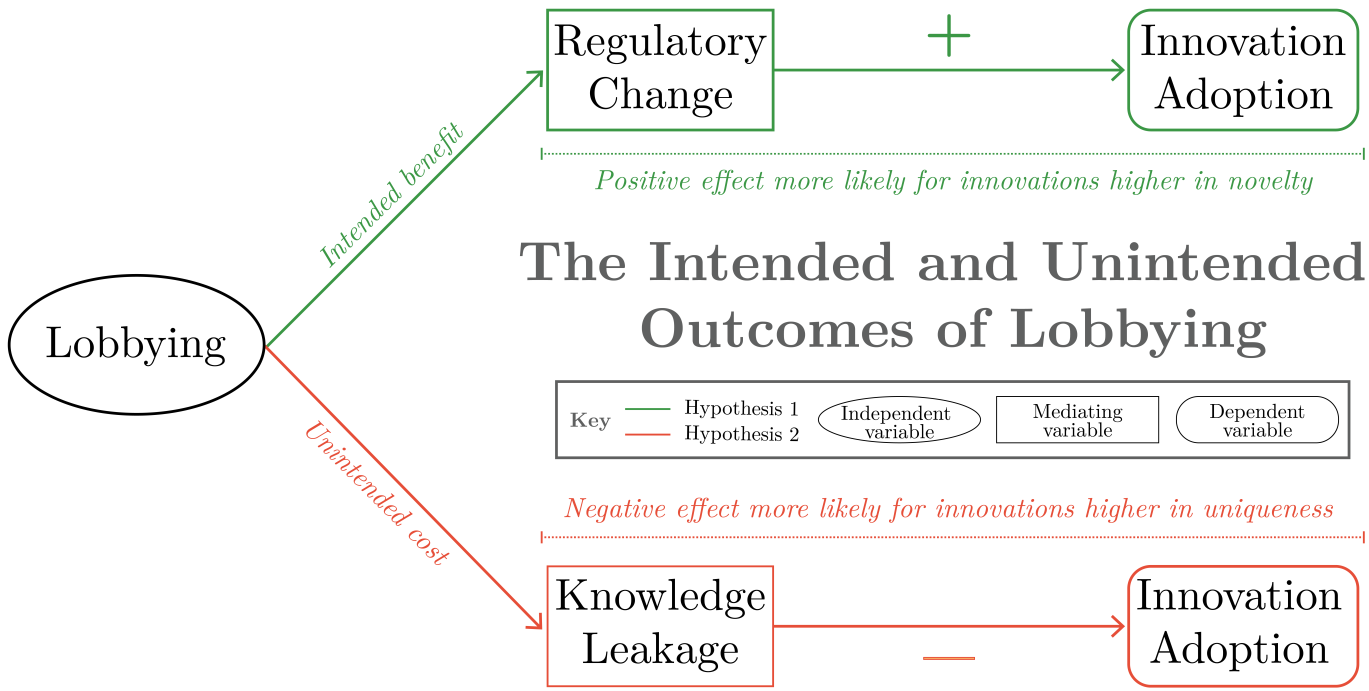
Appendix A Additional quotes from interviews

Table A1: Additional illustrative quotes from the interview

	Quote	Implication
1	*"There was a time when we wanted to take the lead on changing a specific piece of regulation in our industry. We knew that some of our direct competitors would likely want to see a slightly different change. Therefore, we started the lobbying process with a legislator without bringing on board our competitors. However, we discovered that since the lobbying industry is so interconnected, the legislator we were working with unwittingly brought up to our competitor lobbyists that we had been contacting him about the specific changes we were seeking. The cat was out of the bag at this point and we were not able to pursue the specific changes in the way we would have wanted at that point. We were forced to choose a different path towards legislation and changes."*	It can be difficult for individual firms to change regulations in ways that benefit only itself, especially when competitors have a stake in those regulations.
2	*"When we develop new technology, (e.g., AI, facial recognition, encryption), we often need to change existing regulations to make sure that we as well as others can use the technology in the way we intended for it to be used. Lobbying is a way in which we can inform policymakers on how the regulations should change."*	Informational lobbying is often used to help policymakers change regulation in ways that facilitate the adoption of firms' new technologies.
3	*"There was a defense equipment manufacturer that developed a new product— a transportation vehicle that could function both as a commercial and military vehicle. The regulations regarding this type of vehicle at the time was geared towards either a commercial or military version of this particular vehicle. The new product they had developed certainly was useful but the regulations at the time just did not support it. Therefore, they engaged heavily in lobbying to see some changes in the regulation for this type of vehicle."*	Lobbying is often used to change regulation in ways that are helpful for firms' new products, especially those do not fit neatly into the existing regulatory framework due to their novelty.
4	*"Anything we relay to policymakers, we assume it will be leaked. So we are tight-lipped about extremely sensitive information. However, based on the specific policies and issues that a competitor in the industry cares about during industry meetings or conversations with policymakers, you can learn about the technology that the competitor is interested in, especially if you are knowledgeable about the industry."*	Even though leakage of information—especially through politicians—is a widely recognized phenomena among those that are involved in lobbying, there are other less direct ways in which leakage can still occur.
5	*"There was a time when our agency was having a meeting with representatives of manufacturers to update our regulations so that it would include the latest technology in highspeed rail. Each firm was giving a presentation explaining what regulatory changes would be helpful given its latest proprietary technologies. We noticed when one firm handed over their USB to upload the slides to our office computer, malwares were detected on the USB. So we asked the representative to load the shares on our own USB we had. However, to this day, I strongly believe that the particular firm had purposefully loaded malwares onto their USB, which they hoped to load onto our computer, to collect the information on proprietary technologies of the competitors that were presented that day."*	Even policy specialists working at federal agencies are well aware of how firms try to gain as much knowledge about competitors as they can during the lobbying process—sometimes even by taking questionable and extreme measures.

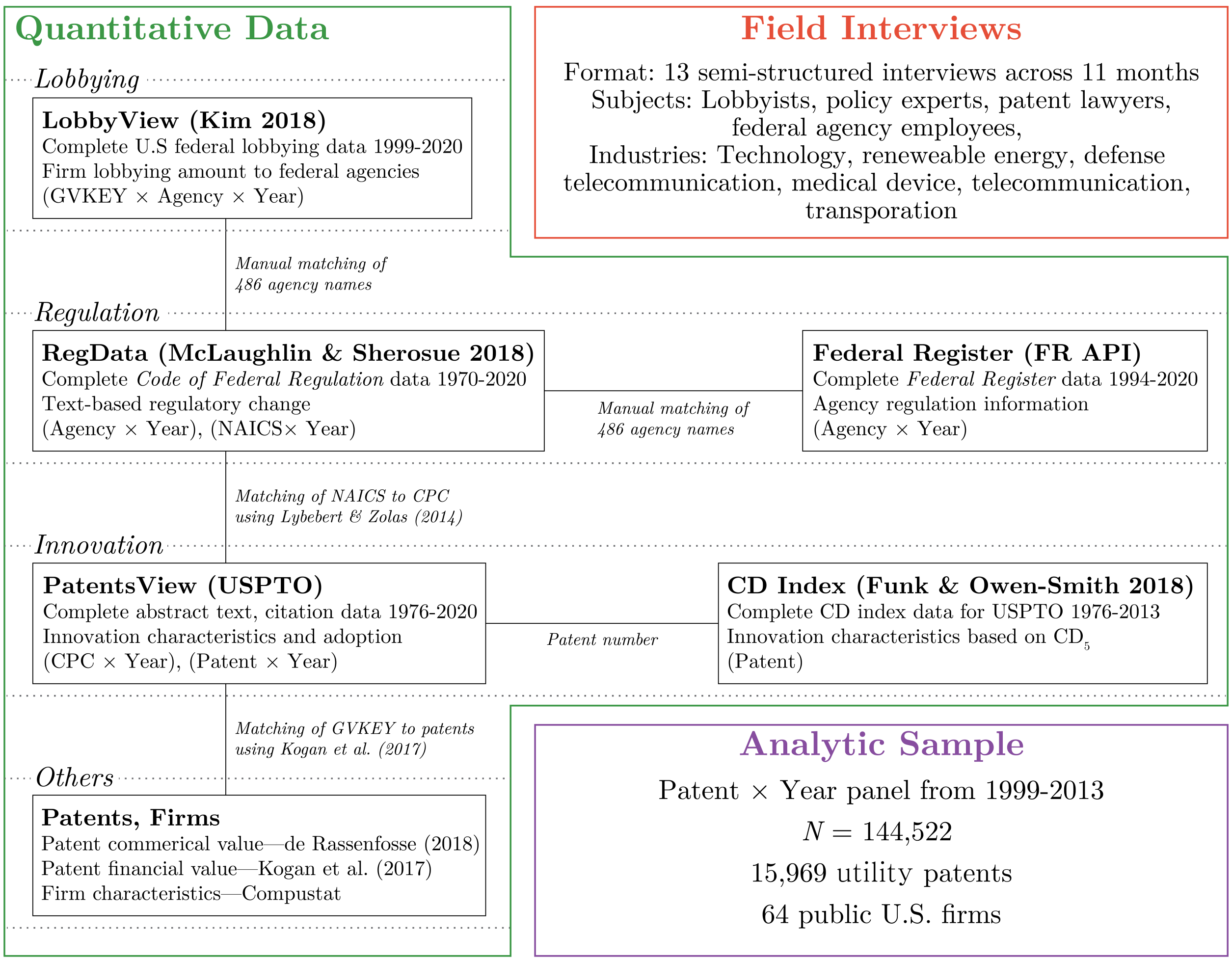
Appendix B Visualization of theoretical framework

Figure A1: Lobbying for innovations: This figure presents a visualization of the theoretical framework. When a firm engages in lobbying, the intended outcome is regulatory change, which is more likely to have a positive effect on innovations higher in novelty (Hypothesis 1, green). Conversely, the unintended outcome of lobbying is knowledge leakage, which is more likely to have a negative effect on innovations higher in uniqueness (Hypothesis 2, red).



Appendix C Visualization of data sources

Figure A2: Collection of empirical data This figure presents a visualization of the empirical data used to test the theory developed in this paper. The quantitative data (green box) consists of novel lobbying, regulation, and innovation datasets that have been combined in a unique fashion to test the hypotheses. This results in an analytic sample (purple box) that consists of a panel of 15,969 patents from 64 firms from 1999 to 2013. In addition, I also conducted field interviews (red box) to better understand the empirical context and establish face validity of the theoretical claims made in the paper.



Appendix D Patent knowledge leakage procedure

Doc2Vec is a recently developed Natural Language Processing techniques that incorporates a “deep learning” procedure to create vector space models. Latent Dirichlet Allocation (LDA) (Blei, Ng & Jordan 2003) is another vector space model that has been used in recent innovation work (e.g., Kaplan & Vakili 2015, Choudhury, Wang, Carlson & Khanna 2019), mainly to categorize the content of documents based on the overall topic represented by each document. LDA infers the content of documents by considering the probability of words occurring together across documents. However, LDA makes two restrictive assumptions. Fist, LDA assumes that the total number of topics appearing across all documents is known. Second, the order in which words appear are not considered. Doc2Vec is a technique developed to take into consideration the semantics of words by capturing where in the document a specific word is embedded with respect to other words (Niu & Dai 2015). This approach is based on an established view in linguistics which suggests that similar words appear in similar contexts (Miller & Charles 1991). Doc2Vec, therefore, makes no assumptions about the number of categories that may exist across all documents but takes into explicit consideration where in the documents specific words tend to appear to capture the meaning of specific words. Overall, while LDA may be suitable for categorizing documents when the number of topics are known and the actual meaning of words are not important, Doc2Vec has significant performance advantages in capturing the similarity of content across documents based on the meaning of words as they are used (Rehurek & Sojka 2010, Niu & Dai 2015, Baroni, Dinu & Kruszewski 2014).

I obtained the full text abstracts of all patents that are in my sample from PatentsView. Using the Python package spaCy I preprocessed and tokenized the text. I first removed all punctuation, digits, and spaces. I also removed words shorter than two characters and longer than 100 characters. My aim is to focus on the knowledge embodied in the set of technological components that are used in the patent. Based on the assumption that components—or physical objects or conceptions—are likely to be represented by nouns than other parts of speech, I focused on words that were tagged as nouns by spaCy in the abstracts.

After tokenzing the abstract text as described above, two tasks remained. I first needed to quantitatively measure knowledge embodied in each patent using the collection of nouns found in the patent abstract. Then based on this quantitative measure of the knowledge used in each patent, I needed to compute the similarity between all the sample abstracts.

For the first step, the standard deep learning procedure of vectorizing each noun requires three neural network that consists of an input, hidden, and an output layer. In this case, the input layer is the set of nouns extracted from the patent abstracts. The hidden layer defines each word from the input layer based on other words that appear close to the focal word. It trains itself to provide the most accurate assignment using the input layer. Then the output layer represents the assignment of nouns as well as each patent abstract (based on the collection of nouns used in the abstract) in a 100-dimension vector space, or “embeddings.” For this first step, I used all patent abstracts that were available on PatentsView, including those outside my sample year, because having a larger input layer is more likely to lead to an accurate embedding of each patent abstract.

Thereafter, with the embedding of each patent abstract in hand, I calculated the pairwise Cosine distance between a focal patent and other patents applied by other firms within the same CPC group. For this calculation, I focused only on other patents that were applied for after the application year of the focal patent in order to make sure that the competitor had time to expropriate the knowledge in the focal patent. Furthermore, I only considered patents that were applied for by other non-lobbying firms because lobbying is the activity through which I argue that knowledge leakage can occur.

Appendix E Validation of the measure for knowledge leakage

Table A2: Validation of knowledge leakage

	Sample: More unique patents			*Sample: Less unique patents*		
	DV: Market value (log)	DV: Market value (log)	DV: Market value (log)	DV: Market value (log)	DV: Market value (log)	DV: Market value (log)
	(1)	(2)	(3)	(4)	(5)	(6)
Knowledge leakage (log)		-0.0542***	-0.0152*		-0.0808***	-0.0077
		(0.0091)	(0.0075)		(0.0096)	(0.0074)
Forward citation within five years (log)	0.1053***	0.0328	0.0234	0.0915***	0.0100	0.0077
	(0.0144)	(0.0192)	(0.0121)	(0.0150)	(0.0201)	(0.0123)
Number of backward citations (log)			-0.0020			-0.0103
			(0.0140)			(0.0143)
Firm average patent count (log)			0.0504			-0.0557
			(0.0552)			(0.0686)
Industry average patent count (log)			-0.0108			-0.0060
			(0.0110)			(0.0115)
Constant	1.4040***	1.4519***	1.0789**	1.4231***	1.5467***	1.8254***
	(0.0215)	(0.0343)	(0.3819)	(0.0230)	(0.0362)	(0.4751)
Firm fixed effects	No	No	Yes	No	No	Yes
Year fixed effects	No	No	Yes	No	No	Yes
R^2	0.0120	0.6594	0.6596	0.0252	0.6518	0.6518
N	2,990	2,984	2,984	2,756	2,749	2,749

Notes: These models provide support for the validity of the measure for knowledge leakage presented in this paper. *Knowledge leakage* seems to be negatively associated with the market value (Kogan et al. 2017) of more unique patents (Models 1 to 3)—unlike that of less unique patents (Models 4 to 6). The results hold when controlling for various patent, firm, and industry characteristics as well as firm and year fixed effects. Results holding when including the number of forward citations it receives within five years of grant year is especially notable as forward citations are considered to be a close determinant of the economic value of patents (e.g., Albert, Avery, Narin & McAllister 1991, Fleming 2001, Hall et al. 2005). Overall, the models suggest that knowledge leakage is capturing the extent to which ideas are taken from the focal patent by subsequent patents to an extent where the newer patents begin to compete with the focal patent and affect the focal patent negatively.

Appendix F CD index

CD_t for a focal patent ("f"), which references a collection of prior patents ("b"), varies over time as subsequent patents ("i") that references only the focal patent, only the prior patents cited by the focal patent, or both are granted. f and b are fixed once the focal patent is granted, however CD_5 varies over time as future patents i's are granted. The exact equation of CD_t is as follows:

$$CD_5 = \frac{1}{n} \sum_{n=i}^{n} 2 f_{it} b_{it} + f_{it},$$

where n is the total number of works that cite either the focal patent, its predecessors, or both. The index ranges from -1 to 1 with values between 0 to 1 indicating destabilizing patents and -1 to 0 indicating consolidating patents. The CD index has been used in a wide variety of prior studies across sociology, science of science, computer science, and even musicology. (Leahey 2016, Azoulay, Fons-Rosen & Graff Zivin 2019, Figueiredo & Andrade 2019, Wu, Wang & Evans 2019, Andrade, Figueiredo, Silva & Morais 2020, Park, Leahey & Funk 2021).

Appendix G Patent uniqueness procedure

I followed the same pre-processing procedure described for the Doc2Vec procedure above to obtain the set of noun tokens for each patent. Thereafter, I vectorized the noun tokens for each patent within its CPC group and application year by calculating the term frequency-inverse document frequency ("TFIDF") for each token. This implies that each patent is considered a document and all the patents applied for in the same CPC group and year represents the corpus of documents. TFIDF is a commonly used vertorizing technique to measure the importance of each token in a body of text (e.g., Kelly, Papanikolaou, Seru & Taddy 2018, Gentzkow, Kelly & Taddy 2019). I use the formula that adjusts for the differences in the length of documents because patent abstracts can vary significantly in length. The reason why I vectorized within each CPC group and application year is because I am interested in calculating the commonality of each technological component with respect to all the patents (i.e., new technologies) that are likely to be lobbied for together by the same industry within the same time period. Patents in different CPC groups maybe lobbied for by different industries. Similarly, a technological component that is unique among patents of one year maybe common in another year. The formula for the term frequency, $tf(t, d)$, and inverse document frequency, $idf(t, D)$ is as follows:

$$tf(t,p) = \frac{f_{t,p}}{max\{f_{t,p} : t \in p\}}$$

$$idf(t,C) = log\frac{N}{|\{p \in C : t \in C\}|}$$

where the TFIDF for a token is the the product of $tf(t, p)$ and $idf(t, D)$. $f_{t,p}$ refers to the frequency of token t in patent p and N is the total number of patents in the corpus C (where the corpus is the patents in the focal CPC group and application year). Based on the vectors of tokens for each patent, I calculate the mean centroid vector for each corpus. Then I calculate the Cosine distance between the focal patent and the centroid vector for each patent. Within each corpus, patents whose distance from the centroid are below the mean are considered to be patents that are low in uniqueness; patents whose distance from the centroid is above the mean are considered to be patents high in uniqueness.